DATA ANALYTICS FOR EDUCATIONAL PROCESSES:
ASSESSING BASIC SKILLS OF ADULT STUDENTS

Autore
Elena Salvatori

Relatori
Prof.ssa Fosca Giannotti
Prof. Dino Pedreschi

10 Settembre 2018
Abstract

The EC publication “Horizon Report Europe: 2014 Schools Edition”, analyzed the emerging technologies and their impact on education. The report identified a number of trends and challenges to be faced by educational institutions in the near future. The educational data, made available thanks to the introduction of web-based learning systems and the large scale adoption of ICT in education, allows the analysis of the named trends and challenges through different types of analytics. These include Educational Data Mining, Learning Analytics and Academic Analytics.

At the same time, the contemporary information society is mostly grounded on text-based technologies where reading, writing and numeracy skills play a fundamental role. International, European and national organizations, recognizing the strategic importance of human capital for economic growth, have placed a lot of efforts in defining core skills. OECD developed a strategy called “Better skills. Better jobs. Better lives” to help countries analyze their national skills systems. The most relevant program launched by OECD has been the Survey PIAAC which started in 2012 and was later on supported by a tool called EsOnline. Furthermore, in Italy, the national school reform identified the certification of basic skills as one of the priorities to be implemented.

This research intends to investigate the possible use of EsOnline precisely to respond to this need. Two case studies were developed, the first intended to test EsOnline in a real adult education setting, in Grosseto, Italy. A second case study was based on the data available through the 2012 PIACC Survey covering more than 30 countries including Italy. Data mining techniques were applied to the Italian dataset and results were used to build a classifier.

The analyses carried out with reference to the first case study allowed to answer several research questions: a first question was related to benchmarking local samples with national and international ones. We investigated whether local samples scored differently than the national and OECD ones. The question was answered positively since the local performance was found to be higher. A second research question was related to the change of core skills of adult students after attending a full school year. However in this case, results were not conclusive. Research question 3 analyzed differences in the mean scores when considering the learning period at school. The analysis revealed that, on average, the students belonging to the second learning period scored significantly higher than those of the first one. A further research question was formulated on
whether EsOnline can be effectively contribute to the Individual Training Pact, and this was answered positively. Finally we investigated conditions for further adoption of EsOnline in the adult school system. A number of drivers and constraints were identified, which could support the adoption process in the future.

Regarding the second case study an analysis was performed on the datasets of the large-scale survey PIAAC, resulting in homogeneous sub-groups of test takers with common features. Visual analysis supported spatially the “digital division” of the samples. Finally, the above data-driven process was applied to 28 OECD national datasets. The Italian dataset made it possible to build a classifier and each student in the Grosseto case was assigned to a cluster. Most of the test takers were assigned to the better performing clusters (where skilled workers prevail) and their average level of proficiency was consistent with the level computed for the clusters of the national dataset. The classification procedure, without having the ambition of replacing EsOnline, could offer an effective and immediate support in the definition of the Individual Training Pact, by indicating a cluster membership and, associated to it, a level of basic skills proficiency.

A limitation of this study is that PIAAC and EsOnline datasets are not fully comparable. This limits the possibility of predicting the level of performance with regards to basic skills. The use of the full set of PIAAC variables offers, in principle, important opportunities to be investigated by the communities of Educational Data Mining, Learning Analytics and Academic Analytics.

Investigating further the relation between skills performance, as defined and measured by PIAAC and skill competences as established by the Italian Ministry of Education, would help evaluating competences at the different levels of the school system.
## Contents

1 **Introduction** .......................... 1  
1.1 Document structure .......................... 3

2 **Setting the stage: educational data**  .... 5  
2.1 EDM, Learning Analytics and Academic Analytics  .......... 5  
2.2 Main driving factors in the research field  .......... 6  
2.3 Educational technologies for e-learning  .......... 7  
2.3.1 A Blended Learning approach to Adult Education  .......... 9  
2.3.2 Computer based assessing systems  .......... 10  
2.3.3 Large-scale Assessments  .......... 11  
2.4 Extracting knowledge from educational data  .......... 13  
2.4.1 Case study 1: the EDM approach  .......... 14  
2.4.2 Case study 2: Social capital and academic outcome  .......... 16  
2.4.3 Case study 3: Resources’ access patterns  .......... 18  
2.4.4 Case study 4: Social capital and language  .......... 20  
2.4.5 Case study 5: Language analysis in the Community of Inquiry framework  .......... 21  
2.4.6 Case study 6: Automatic text grading  .......... 23  
2.4.7 Measuring and understanding learner emotions: evidence and prospects  .......... 24  
2.4.8 Main conclusions  .......... 26

3 **Setting the stage: why skills matter**  .... 29  
3.1 Skills: the international perspective  .......... 29  
3.2 Skills, the Italian perspective  .......... 30  
3.2.1 The PIAAC Results 2012  .......... 30  
3.2.2 An update from 2016 by ISTAT  .......... 31  
3.2.3 The Italian adult education system  .......... 33  
3.3 The PIAAC Methodology  .......... 35  
3.4 How PIAAC works  .......... 35  
3.4.1 Multistage adaptive testing  .......... 36  
3.4.2 Scaling and comparing proficiency  .......... 38
3.5 PIAAC, description of the main findings .......................... 39
3.5.1 Differences across participating countries .................... 39
3.5.2 Use of ICTs ......................................................... 40
3.5.3 Education and proficiency ......................................... 40
3.5.4 Employment and proficiency ...................................... 41
3.5.5 Participating in adult learning and proficiency ................. 42
3.5.6 Age and proficiency ............................................. 44

4 Analyzing and evaluating Education & Skills Online 47
4.1 Introduction ............................................................. 47
4.2 Education & Skills Online, the test ................................ 48
4.2.1 The field work ...................................................... 50
4.3 Data, methods and analysis ........................................... 52
4.3.1 Datasets ............................................................. 52
4.3.2 Methods .............................................................. 53
4.3.3 The samples: socio-demographic features and response rate .. 53
4.3.4 Benchmarking the samples’ performances ..................... 55
4.3.5 Discussion of the results .......................................... 57
4.3.6 Repeated measure .................................................. 60
4.3.7 Discussion of the results .......................................... 61
4.3.8 Learning periods and skills ....................................... 61
4.3.9 Discussion of the results .......................................... 63
4.3.10 To evaluate whether the EsOnline results could be used when defining the Individual Training Pact ..................... 64
4.3.11 To test EsOnline in a real case study with adult students for further adoption by the partnership ..................... 65

5 Extracting knowledge from PIAAC data ............................. 67
5.1 Data preparation .......................................................... 67
5.1.1 The datasets .......................................................... 68
5.2 Methods ................................................................. 69
5.2.1 Decision Trees ....................................................... 71
5.2.2 Clustering ............................................................. 72
5.2.3 Multi Dimensional Scaling ........................................ 73
5.3 The analysis of the Italian dataset .................................... 73
5.3.1 Adding the macro-regions’ variable .............................. 78
5.4 The analysis of the international datasets ............................ 79
5.5 Classifying the Grosseto sample ...................................... 89
5.6 Discussion of the results .............................................. 93

6 Conclusions ................................................................. 97
  6.1 Results from the case study ........................................... 97
  6.2 Results from the data mining exercises ............................... 99
    6.2.1 Limitations and possible developments of this study ........ 100

Appendices ................................................................. 103

A EsOnline evaluation survey ............................................. 105

B Education and Skills Online Score Reports ........................... 107

Bibliography .................................................................. 141
List of Figures

2.1 The Diplomarsionline VLE ................................................. 10
2.2 Classification results using data balancing ............................... 16
2.3 The Read network ......................................................... 17
2.4 The student-resources clusters ........................................... 19
2.5 The Community of Inquire framework .................................. 22

3.1 The flow of the PIAAC Main study ....................................... 37
3.2 The test-lets of the PIAAC Main study .................................. 37
3.3 The PIAAC computer-based adaptive assessment ....................... 38
3.4 Literacy, PIAAC achievement levels by country ....................... 40
3.5 Computer experience of the PIAAC respondents by country ........ 41
3.6 Numeracy, PIAAC average scores by educational level ................ 42
3.7 Numeracy score by employment status ................................ 43
3.8 Numeracy score and participating in education. ....................... 43
3.9 Literacy score by age-group in OECD countries ....................... 44
3.10 Numeracy score by age-group in OECD countries .................... 45

4.1 The case study portal ...................................................... 48
4.2 The Education & Skills Online test flow ................................ 49
4.3 The test home-page ....................................................... 52
4.4 Participants by age-group ............................................... 55
4.5 Discrete Achievement Levels in Numeracy and Literacy ............. 58
4.6 The distribution among age groups in the Italy and in the local sample 59
4.7 Literacy average score by age group, Italy and Grosseto ........... 60
4.8 Box Plot of the Literacy scores by Learning Period .................. 62
4.9 Box Plot of the Numeracy scores by Learning Period ................ 62
4.10 Box Plot of the PSTRE scores by Learning Period .................... 62

5.1 The Multidimensional Scaling of Italian Dataset stratified by cluster id. 77
5.2 Distribution of Numeracy scores by working status ................... 79
5.3 Distribution of Literacy scores of by working status ................. 80
5.4 Literacy scores in the Italian macro region ........................... 80
5.5 Numeracy scores in the Italian macro regions ........................ 81
List of Tables

2.1 The research field .................................................. 6
4.1 Schools and students participating in the testing .................. 51
4.2 The socio-demographic features of the samples ................. 54
4.3 Response Rate of the samples ..................................... 55
4.4 Mean scores of the samples ....................................... 56
4.5 The T-test results .................................................. 57
4.6 Repeated measures, mean scores in 2016 and 2017 ............. 61
4.7 The statistics by Learning Period ................................. 63
5.1 Demographic features of the Italian sample ..................... 70
5.2 Coding of the cluster analysis’ variables ......................... 74
5.3 The centroids matrix ............................................... 75
5.4 Numeracy and Literacy scores by cluster ....................... 76
5.5 Percentages across the five proficiency levels, literacy and numeracy, for Italian workers. ........................................ 78
5.6 Literacy proficiency levels by working status, 5 countries. .... 82
5.7 Numeracy proficiency levels by working status. ................. 83
5.8 Average proficiency level in Literacy by working status, 28 countries . 88
5.9 Average proficiency level in Numeracy by working status, 28 countries . 89
5.10 Numeracy and Literacy average scores by predicted cluster .... 93
A.1 The results of the satisfaction survey ............................. 106
Chapter 1

Introduction

As a result of the massive introduction of ICT, education is facing new challenges. The EC publication “Horizon Report Europe: 2014 Schools Edition”, [16], analysed the emerging technologies and their impact on education. The report identified a number of trends and challenges to be faced by educational institutions at every level (primary, secondary and higher) in the near future.

In the short-term, the trends that will have a significant impact on the ways educational institutions approach teaching will be rethinking the role of teachers, the growing ubiquity of social media and the integration of hybrid learning.

In the mid-term, we will see an increasing focus on Open Educational Resources (OER) and the increasing use of hybrid learning design, while in the long-term we will witness the introduction of data-driven learning assessment, the evolution of online learning and the acceleration of intuitive technologies.

In the short run the most important learning technologies will be cloud computing, tablet computing, while in the mid and long-term, the emerging technologies will be games and gamification, mobile learning, learning analytics, personalized learning, the Internet of things, wearable technology and virtual and remote laboratories.

These changes will create also new challenges for learning in ICT based environment and learners will become co-designers of the contents and will experience blended forms of formal and non-formal learning activities.

In this scenario educational institutions will have to define their policy on data privacy taking into considerations learner’s digital rights such as control on consent for processing individual’s data, the right to be forgotten, and easier access to and protection of one’s personal data.

Also the competences to be acquired are affected by the digitalization of education; the digital fluency of future learners is a variable that must be taken into consideration. If fluency of young learners is not seen as a problem, the same does not hold for teachers, adult learners, as well as faculty members and administrative personnel, whose competences do not necessarily include ICT training.
Furthermore the information society is mostly grounded on text-based technologies where reading, writing and numeracy skills still play a fundamental role. The speed of accessing ICT in today’s world as well as the changes occurred in the social, economic and cultural landscapes require citizens to master an expanding set of skills to cope with new media. Education is already revising its practices to define the “new media literacies” beyond traditional literacy and numeracy skills. The concept of multi-literacy emerges, which extends “the scope of traditional literacy to include the diversity of media and modes of communication that are now available to learners” who are seen both as producers and consumers. Multi-literacy addresses the increasing cultural and linguistic diversity that is prevalent today [42].

International, European and national organizations, recognizing the strategic importance of human capital for economic growth, have already launched large research programs to define the core skills of the 21st century society. The OECD (Organization for Economic Co-operation and Development) developed the strategies “Better skills. Better jobs. Better lives” to help countries to analyze their national skills systems, develop policies aimed at transforming “better skills in better jobs, social inclusion and economic growth”. The most relevant programme launched by OECD in the framework of the Skills Strategy has been the Survey PIAAC (Programme for the International Assessment of Adult Competencies) started in 2012. Its results indicate that skills assessment rather than educational attainment reflects the status of knowledge in a population and the levels within.

The national school’s reform identifies skills’ certification among the activities that the Italian adult education system will have to implement. Furthermore, the document [36] promotes the experimentation of international self-assessing methodologies as a good practice when designing educational projects for adult learners. To conclude, in 2015, OECD released a set of web-tools developed in the framework of PIAAC such as the Computer Based Testing environment called Education&Skills Online (EsOnline), and the PIAAC data Explorer.

The previous considerations emphasize the need, for the Italian adult education system, to assess basic skills through internationally recognized methods. Currently there are no standardized methodologies: each school develops its own procedures and, as a consequence, results are not comparable. Therefore, this research is aiming, above all, at investigating the possible use of PIAAC EsOline precisely to fill this gap. With this in mind, we have designed a first case study where EsOline was tested in a real adult education setting. The objectives underlying the case study were drawn from our personal experience as teachers in adult education. The following were identified:
• to analyze the performance of this sample and benchmark it against national and international figures;

• to analyze whether core skills of adult students change after attending school for a full year and to analyze the performance according to the learning period attended.

• to evaluate whether the EsOnline results could be used when defining the Individual Training Pact;

• to test EsOnline in a real case study with adult students for further adoption by the partnership.

A second case study was developed based on the data available through the 2012 PIACC survey covering more than 30 countries including Italy. Data mining techniques were applied with the objective of discovering patterns and relations in the Italian dataset without formulating any prior hypothesis.

Ultimately this data driven analytical process was intended to shed light on our knowledge on core skills of the Italians.

The knowledge gained from the data mining was used to classify the results of the first case study, establishing meaningful relations between the two.

1.1 Document structure

Following from the observational research approach above, the thesis is organized into 6 chapters. Chapter 1 is the introduction, including objectives and structure of the thesis. Chapter 2 defines the research field of data analysis applied to educational data, giving an overview of educational technologies for on-line learning and reporting a selected number of exemplar case studies in the research field. Chapter 3 introduces the international and European strategies for education, the main results of the 2012 PIAAC Survey in Italy and the innovations taking place in the Italian system of adult education. Secondly, the PIAAC approach at measuring adult competencies is presented along with an analysis of its components, test flow and of the competences evaluated by the test.

Chapter 4 provides a descriptions of the first case study, the partnership of schools, the analysis performed and the discussion of the results. Chapter 5 presents the data driven analytical analysis on the national and international datasets. Chapter 6 discusses the overall results of the work, and the newly gained knowledge.
Chapter 2

Setting the stage: educational data

The research field, i.e. the analysis of educational data, has as its main goal that of making a better use of the large amount of digital data about learners. These data are made available thanks to the introduction of web-based learning systems and the large scale adoption of ICT in education. The field has its roots in business intelligence and data mining. Business intelligence is defined as a set of technologies supporting decision making in organizations using the large volume of business data. Data Mining (DM) or Knowledge Discovery in Databases (KDD), is defined as the process of discovering interesting and useful patterns and relationships in large volumes of data. This field is strongly related to statistics, information retrieval, and artificial intelligence.

2.1 EDM, Learning Analytics and Academic Analytics

In the domain of the analysis of educational data the following subfields can be found: Educational Data Mining (EDM), Learning Analytics and Academic Analytics.

EDM, which is the application of DM techniques to educational data, has the objective of analyzing the data in order to answer educational research issues [39]; the International Society on educational data mining was established in 2008 together with the first EDM conference.

Learning analytics is the measurement, collection, analysis and reporting of data about learners and their contexts, for the purpose of understanding and optimizing learning and the environments in which it occurs. This definition dates back to 2011 and was suggested in the occasion of the First international Conference on Learning Analytics & Knowledge; since then, the Society for Learning Analytics Research (SoLAR) organizes every year the Learning Analytics and Knowledge Conference (LAK). In 2012 [6], Social Learning Analytics (SLA), is defined as “a distinctive subset of learning analytics that draws on the substantial body of work demonstrating that new skills and ideas are not solely individual achievements, but are developed, carried forward, and passed on through interaction and collaboration”.
### Type of Analytics | Level or object of analysis | Who benefits?
--- | --- | ---
Learning analytics and EDM | Course-level: social networks, discourse analysis, intelligent curriculum
Departmental: predictive modeling, pattern of success/failure | Learners, Faculty
Academic analytics and EDM | Institutional: learner profiles, performance of academics,
Regional (state/provincial): comparison between systems
National and international | Administrators, funders, marketing

| Academic analytics and EDM | Funders, administrators
| National governments, Education authorities |

**Table 2.1: The research field**

**Academic analytics** is the application of business intelligence in the educational sector and makes use of institutional data for comparisons at regional, national and international level. Its definition appeared first in 2007.

The table 2.1 by Siemens [43], describes the distinction among the mentioned fields: learning analytics focuses on activities at the learner-educator level, academic analytics targets organizational efficiency and finally EDM has a role in both.

According to [7], the three terms refer to three different research challenges on educational data. Educational data mining is focused on the technical challenge: “how can we extract value from the big sets of educational data?”. Learning analytics is focused on the educational challenge: “how can we optimize opportunities for on-line learning using data generated in the different environments?” Academic analytics is focused on the political/economic challenge: “how can we substantially improve learning opportunities and educational results at national or international levels using data on learning outcomes?”

#### 2.2 Main driving factors in the research field

As Ferguson indicates in [7], the main driving factors in our research field. **BIG Data** represents a first factor: the diffusion of Virtual Learning Environments also referred to as learning management systems (LMS) like Moodle (Modular Object-Oriented Dynamic Learning Environment) or Blackboard, has resulted in an increasing volume of
data handled in the educational information systems. Such systems store everyday interaction data, personal data, system information and academic information. More and more frequently however, learning activities occur externally to the VLE, and across a variety of web resources characterized by different standards, levels of access and owners. A main issue then, is that of the need of new educational datasets integrating data from VLEs and data related to e-learning activities occurring outside the e-environments through different devices.

On-line learning The second factor to be faced is of educational nature and deals with the optimization of opportunities for on-line learning. As on-line learning programs are increasing in number and are made available to a massive number of students, new issues are arising: isolation of learners, technical problems, low motivation and high rate of drop-out students as well as difficulties experienced by the teachers. How can we then improve the individual learning experience of a student? How can teachers, in lack of a face to face contact, monitor the level of motivation and engagement of a Student, as well as his/her eventual confusion? How can the contributions of individual learners and the quality of his/her learning in open on-line environments be evaluated when hundreds of interactions take place daily and hundreds of contributions are uploaded by students?

Political concerns A third factor is related to the fact that governments and institutions show an increased interest in measuring the performance of the education offered. New approaches have to be identified for optimizing learning and educational results at national and international level.

Stakeholders Who benefits? Three different interest groups are involved: teachers/learners, educational institutions, and governments. Depending on the target audience, the way researchers will conceptualize the problems, analyze the data and report their findings will be affected.

2.3 Educational technologies for e-learning

A Virtual Learning Environments (VLE) is a computer based learning system that delivers course-ware plus e-tutoring over the Internet. VLEs are usually closed platforms for registered users only; they provide course management Information such as lists of courses, registration, credit information and syllabus, pre-requisites as well as teaching materials and self-assessment quizzes. Emails, forums, chat, teleconferencing are used for asynchronous and synchronous communications among learners and teachers; students as well teachers have home pages with tools to track their activities. Interactions
Chapter 2. Setting the stage: educational data

among participants may take place also outside the VLE on social media platforms such as Facebook and Twitter, YouTube channels or software platforms allowing video communication (e.g. Skype). Although Computer based learning systems may be also referred as Learning Management System (LMS), Course Management System, Pedagogical Platform or E-Learning Platform, VLEs are "above" LMS since they enable the creation of VLEs.

E-learning can be synchronous, asynchronous, blended or ubiquitous. Synchronous learning systems may include features such as shared applications and whiteboards, emoticons, hand rising, and a class-voting (clicker), as well as direct communication that occurs in real time through instant messaging, chat rooms, and audio and video conferencing. An example of asynchronous learning is that of MOOC, an acronym for 'massive open on-line course’. Since 2008 many universities worldwide offer on-line academic courses to students that can participate through open access on the internet. We refer to “blended-learning” when face-to-face classroom methods are expanded by digital delivery of content; the approach of expanding the course content with on-line resources is referred to also as “hybrid learning”. “Ubiquitous learning”, or u-learning, is when VLEs are supported by ubiquitous computing technologies such as wireless communication equipment, mobile phones, PDAs (Personal Digital Assistants), sensor network nodes and wearable computers.

The term “open”, that is often encountered in educational technology, is used when educational contents are introduced with an open license or are in the public domain and can be used, modified and shared; we can also refer to them as “Open Educational Resources”.

Other educational technologies exploited in education are virtual reality environments, educational game environments, and Intelligent tutoring systems (ITS).

ITSs are automatic system supporting students in a domain of knowledge (mathematics, physics, etc.), in educational games and in game-based learning. They track user inputs and provide the learner with questions and problems. They are also able to understand students’ solutions and to interact with students providing hints and adapting tasks, questions, etc. to the needs of the learners. The adaptation of the learning content and the difficulty level of the assignment given to a student are based on the estimation and prediction of the student’s knowledge with respect to a student model.

Hence, intelligent tutoring systems are built upon a learner model to infer student’s knowledge on a specific topic. ITSs store data about the student model, the domain model and the interactions.
2.3.1 A Blended Learning approach to Adult Education

The project “Diplomarsionline” presented in this section, is an Italian Blended Learning program supported by a network of partners. Such network includes: Ufficio Scolastico Provinciale di Grosseto, SdS - COeSO - Grosseto, Istituto di Istruzione Superiore L. Bianciardi Grosseto, CPIA Grosseto, CPIA Livorno, Istituto Comprensivo di Monte Argentario, Comune di Civitella Paganico, Comune di Isola del Giglio, Comune di M. Argentario, Comune di Monterotondo M.ma, Comune di Follonica.

In 2005, in the framework of the Action Plan of the European Commission, Regione Toscana, Provincia di Grosseto and ISIS Bianciardi (formerly IP Einaudi) promoted a research on adults and education. The research showed that about 50% of the 115,000 residents in the province, in the age range 19-55, had no upper secondary education. About 30% of them were interested in being involved in a distance learning program especially in relation to the fact that they are living in areas far from schools.

The project “Diplomarsionline” (“version 1.0”) was launched in 2006 in the municipalities of Cinigiano, Capalbio and Monterotondo Marittimo with a "blended" learning model, that is a mix of “face to face” school and e-learning experiences addressed to residents of the province.

In 2008, a “version 2.0” of Diplomarsionline was in place: the programme migrated to Moodle and opened to other schools of the region thanks to the project “ACCE.DI. – ACCEsso al Diploma”, promoted by Regione Toscana, and funded under the European programme “F.S.E. Por Ob. Competitività Regionale E Occupazione Asse IV - Capitale Umano”.

Starting from autumn 2015, Diplomarsionline was partially re-designed according to the new secondary education curricula and to the indications in [35]. Furthermore, to respond to the many requests by adults not living in the province, in the same period Diplomarsionline started new "long distance" courses. During the school year 2017/2018, Diplomarsionline counts about 300 students grouped in 16 virtual on-line classes.

Diplomarsionline offers two type of courses: "blended" learning courses and "long distance" courses. In the "blended" courses, lessons are held in physical classes in different towns of the province of Grosseto. Classes are given twice a week with a duration of 8 hours.

In the "long distance" courses, lessons are held by teachers using Skype according to a monthly schedule. Students and teachers meet also three times a year in Grosseto: at the beginning of the school year in September, at the end of January and at the end of May. During the meetings, students have "face to face" lessons and testing sessions.

\[1\]http://www.diplomarsionline.eu/
Chapter 2. Setting the stage: educational data

For on-line activities students and teachers interact using emails, Skype, Whatsapp, the project Facebook’s page and a Moodle Virtual Learning Environment. The VLE home-page is shown in Fig. 2.3.1; it gives access to the message system of the community called "Maremma Impara" and to the on-line classes. Each class has a tutor who acts as a contact point for students and teachers moderating the messages of the community. Tutors manage the monthly calendar of a group and have weekly meetings with students and teachers. They offer individual support to students on topics ranging from academic difficulties to emotional as well as social problems, illness or traumatic life events. They also facilitate the communication between students and the administrative personnel of ISIS Bianciardi whenever needed.

The curriculum is organized in three periods of one year each. Depending on their previous educational experiences, students can complete their education in 1, 2 or 3 years. The decision about the duration is taken at enrollment, it is based on evidences on previous educational experiences, certified work experiences in the field of study as well as personal needs. The school year starts in September and ends in August. At the end of the third year, students take the final state examination (Esame di Stato) to get the upper secondary vocational qualifications as “Tecnico della Gestione Aziendale Informatico”.

2.3.2 Computer based assessing systems

Computer based testing/assessing systems are tools often integrated in VLEs and MOOC that provide immediate feedbacks on the proficiency and knowledge of a student in a

\[2\text{http://www.diplomarsionline.eu/moodle30/}\]
specific subject area. The tests may include multiple choice questions, yes/no question, numeric entry items, clicking items, highlighting items, open-ended answered questions, etc. These tools can be designed to present the same tests to every student or to adapt the assessing procedure to the individual ability of a student. In this second case, we call the systems as adaptive; the test starts presenting a first question of moderate difficulty with respect to a standard threshold; the system then will score the answer and will estimate the student’s ability as higher, if the answer was correct, or lower, if not. Thus, the next question will be selected according to such estimation: it will be less difficult in case of wrong answer or more complex otherwise. At each step of this process, the system will re-estimate the student’s ability until the maximum number of questions have been presented or a statistically accuracy of the estimation is reached. These systems are called multistage adaptive system when at each step of the process the evaluation is made on a group of questions, testlet, and not on a single item.

2.3.3 Large-scale Assessments

A particular category of skills assessment tools is represented by Large-scale Assessments. Such assessments are factually surveys of knowledge, skills, behaviors conducted in one or several specific domains. In contrast to assessment of individuals, they focus on group scores with the aim of describing populations of interest. Such methods are relatively recent. The first work for assessing student population started in the United States in the late 1950s, by the International Association for the Evaluation of Educational Achievement (IEA). The population of interest consisted of 13-year-old students in 12 countries. Participating countries included Belgium, England, Finland, France, Germany, Israel, Poland, Scotland, Sweden, Switzerland, the United States and Yugoslavia [20].

The project, known as the Pilot Twelve-Country Study, resulted in the assessment of academic skills and non-verbal ability, and took place from 1959 to 1962. A key result of the project was that comparing across languages, cultures and educational systems was possible and could be supported by common instruments. This opened the way for new large-scale international assessments targeting both students and adults.

In 1969 in the framework of the National Assessment of Educational Progress (NAEP) project run by the U.S. Department of Education, a first assessment was conducted which involved 17-year-old students in-school.

Since 1983, driven by policy questions focused on a desire to better understand how student competencies related to national concerns, human resource needs, and school effectiveness, the NAEP adopted specific psychometric methodologies. These all aimed
at interpretations of the data going beyond the initial ones, fixed to the individual items used in the assessments and included the Balanced incomplete block (BIB), the Item response theory (IRT), and the marginal estimation procedures.

The introduction of BIB allowed to administer to each student a subset of the item pool while the adoption of IRT based methodologies allowed the creation of scales that could be compared across the forms assumed by the test. In addition, the development of marginal estimation procedures optimized the computation of proficiency scales and their reporting.

Starting from the 1990s the importance of skills and knowledge gained by adults gain through education, on-the-job training and lifelong learning, was widely acknowledged by policy makers. This led to a series of large-scale assessments targeting adults such as the International Adult Literacy Survey (IALS), conducted from 1994 to 1999, and followed by the survey of Adult Literacy and Lifeskills (ALL). IALS and ALL developed their cognitive measures around the domains of prose and document literacy, and with ALL numeracy and analytic problem solving skills were added as additional dimensions.

Since 2009, the NAEP began experimenting computer-based assessment tests in which part of the student sample was administered interactive computer tasks. In 2011 the writing assessment for grades 8 and 12 was administered only through its computer based version.

The evolution of such innovative psychometric methodologies led to several international large-scale studies of student and adults skills. Examples include the Trends in International Mathematics and Science Study (TIMSS), the Progress in International Reading Literacy Study (PIRLS), the OECD’s Programme for International Student Assessment (PISA) and, the latest in the series, the Programme for International Assessment of Adult’s Competencies (PIAAC).

Assessment, as mentioned by the authors in [21], represents a “wicked” problem in educational institutions. Although teachers see assessment as a tool for supporting both learning and teaching, assessment within an educational institution is considered as a powerful tool in the hands of policy makers for measuring and comparing quality standards at regional, national and international level.

Assessment practices are intricately linked both with learning and accountability measures as shown in the NAEP initiative “No Child Behind” (2002). This act was meant to contrast poor literacy and numeracy standards and reshaped teaching contents as well as teaching and assessment methods; teachers’ incentives were linked to performance in the high-stake tests.

An unexpected outcome of the program was that teachers encouraged low-performing
students not to take high-stakes tests and narrowed the contents delivered to meet the requirements of the test. This practice is well illustrated by Goodhart’s law that states that “when a measure becomes a target, it ceases to be a useful measure”.

This example shows also how a change in a specific area of the educational organization may have unpredicted consequences in a subsystem. It also indicates that every policy in education must take into account the fact that educational institutions are an example of complex adaptive system (CAS). They are resilient to changes and require constant energy to maintain their organizational structure. To be successful, any change strategy must take into account interdependencies among subsystems.

2.4 Extracting knowledge from educational data

In this Section a number of case studies, exemplar of relevant problems in the domain of interest, have been selected showing how specific techniques described in 2.1 can be used to extract knowledge from educational data. The selection of such case studies was based on important sector reviews. These included those presented in the Journal of Learning Analytics, [46] by the Society for Learning Analytics Research, SoLAR (editor D. Gasevic); the review on Educational Data Mining by C. Romero et al. (in IEEE Transaction on Systems, Man, and Cybernetics, [38]) and, with specific reference to Italy, the "Dossier Learning analytics", [9], published in the "Italian Journal of Educational Technology" by the Istituto Tecnologie Didattiche (ITD) of the National Research Council. Dragan Gasevic and Cristobal Romero represent among the most authoritative figures in the fields, respectively, of Learning Analytics and Educational Data Mining.

The selection was further guided by a number of open questions which I derived from my personal professional experience: can we measure the interactions among students and teachers and the use of resources in the Moodle based VLE of Diplomarsionline? What can we learn about the Diplomarsionline community from the texts exchanged by the participants? Can we predict students’ success?

To evaluate which techniques were more appropriate to answer the questions above, these were matched against the analytics categorized by Siemens in [42] and summarized in Table 2.1. All questions could be answered within the Learning and EDM types of analytics.

The first case study is an example of application of educational data mining techniques for predicting student’s failure in high school classes. The second and the third case studies describe the application of data analysis techniques such as SNA, statistics, network analysis and visualization techniques to e-learning platform data. One case
study addresses the problem of tracking student’s access to resources in virtual learning environments while the other investigates the predictive value of network parameters. Then a group of three case studies on the application of text analysis techniques and statistics on text data collected in e-learning or social media platforms will follow. One analyses the influence of written text, in on-line community, on social capital accumulation. The second experiment investigates the linguistic features of on-line transcripts according to the Community of Inquire Model. The last case study describes an experiment about automatic grading of digital text artifacts written by students. The Section ends with a subsection investigating how learning analytics have been successfully applied to measure emotions in on-line learning.

2.4.1 Case study 1: the EDM approach

The paper [23] addresses the problem of predicting student failure at high-school applying data mining techniques to educational data. Data used in the research come from a 3 years high school program named “Program II of the Academic Unit Preparation” conducted at the University of Zacateca (UAPAZ, Mexico). It involves student aged 15-18 years old who want to prepare for University.

The dataset consists of 77 attributes of 670 students; the information comes from different sources: registration data, a specific students’ survey about personal and family information and scores obtained by students of the department of school services.

Data were integrated into a single dataset; students’ incomplete records were excluded and continuous variables were transformed into discrete ones; all the information has been saved in the .ARFF format of Weka.

During the pre-processing, the attribute selection was performed in Weka in order to handle the high number of attributes of the dataset; using ten different feature selection algorithms, the best attributes were ranked according to their frequency. Only attributes selected by more than two feature selection algorithms have been chosen; after this step, 15 best attributes out of 77 were selected.

Since the majority of the students passed and a minority failed, the dataset is imbalanced and not suited for traditional classification algorithms which do not take into consideration class distribution.

To balance the data, the SMOOTE algorithm was applied to the training files of the 15 best attributes, obtaining a rebalanced file with 50% “Pass” and 50% “Not Pass” students.

The prediction of student’s failure is addressed as a classification problem over the two classes “Pass” and “Not Pass” class.
The authors introduce a new classification algorithm based on the genetic programming algorithm G3P, also known as grammar-based GP, which has been successfully used in classification problems. The first variant ICRMv1 (Interpretable Classification Rule Mining) allows one rule per class; ICRMv2 allows multiple rules per class; ICRMv3 decides the number of rules needed to predict student failure.

To compare performances among Weka classification algorithms and ICRM, the confusion matrix is used along with the following measures: Accuracy, True Positive rate, True Negative rate, Geometric Mean, Number of rules, #Conditions per rule, #Conditions.

Four experiments have been carried out testing 10 well-known classification algorithms available in Weka DM against the three version of ICRM. White-box models have been chosen since they provide IF-THEN prediction rules that can be directly used for decision making.

The first experiment was performed on the original dataset of 670 students and 77 attributes. The second one on the reduced dataset of the 15 best attributes, and the third on the rebalanced dataset with the best 15 attributes. The fourth and last experiment was performed applying a cost sensitive classification with cost matrix on the best 15 attributes dataset.

The results of the first experiment performed on the whole population of students show high accuracies for almost all classifiers - >85 % -, high TP rate but lower TN rate. ICRM models result in high TN rate, a reduced number of rules and conditions, as well as high accuracy. All classifiers use a reduced number of the 77 attributes available.

The second experiment performed on the reduced dataset of the best 15 attributes show that all the classifiers improve in some quality measure (TN rate and GM) and do slightly better or worse with regard to other measures (TP rate and Accuracy). ICRM v1 obtains the maximum TN rate and GM (93.3% and 92.5%), the minimum number of rules and a low number of conditions.

The experiment performed on the rebalanced training file shows that more than half of the classifiers improved the evaluation measures except accuracy (see Fig. 2.2). The Cost sensitive classification performed in the 4th experiment did not show any general improvement in the evaluation measures.

The results obtained show that classification algorithms can be used in predicting student’s failure or success with high accuracy. Moreover, feature selections resulted in less rules and conditions without a lower classification performance.

The best classification results are obtained by ICRM models on the rebalanced dataset: ICRMv3 obtains the highest TN rate, 98.7%, using five rules and only 1.5 conditions per rule, with accuracy of 92.7%.
Only a small subset of the 77 attributes appear in rules: those appearing more frequently are scores in Math, English, Physics, Humanities with values as “poor”, “very poor” and “not present” when modeling the “Not pass” class.

### 2.4.2 Case study 2: Social capital and academic outcome

The study presented by Garcia in [11], applies Social Network Analysis theory to educational data from the online course “Introduction to financial information” held at the Open University of Catalonia (UOC) in the period September 2013 - January 2014. The online course involved 656 students, organized into 10 classrooms under the supervision of 10 consultant teachers coordinated by one professor.

The objectives of the exploratory study were to observe the relations existing among social network parameters and academic outcomes of each student and of the class as a whole, and to visualize in Gephi the interactions among students, teachers and posts in order to unveil the complex networks built by the course participants during their online activities. The analysis was performed on three datasets from the message board of the course, logs, and final grades.

- **Dataset 1**, the “reply” network; in this network the nodes represent students and teachers and every oriented edge from node \( a \) to node \( b \), has a weight corresponding to the numbers of replies of node \( a \) to a message created by node \( b \).

- **Dataset 2**, the “read” network; it describes the reading of posts by the agents (students and teachers); edge from node \( a \) to node \( b \) has a weight corresponding to the numbers of times that node \( a \) read a message created by node \( b \).

Both datasets contain also other student’s attributes such as the number of messages read, the number of replies, the number of new posts, role, classroom and final grade.
Dataset 3 represents messages as nodes and relations among them as edges; a directed edge from node $a$ to node $b$ means that message corresponding to node $a$ has been written to answer to message of node $b$; for each message, the number of times it has been read, its creation date, and other information are stored.

Centrality measures such as degree centrality, betweenness centrality, closeness centrality, eigenvector centrality etc. have been computed for the three networks. The closeness centrality of a node is based on the natural distance metric between other nodes; the eigenvalue centrality measure indicates the influence of a node on other nodes; the betweenness centrality allows identifying nodes acting as a bridge among other two nodes quantifying the numbers of shortest path passing via a given node.

The Fig. 2.3 of the “read” network visualized by Gephi shows the interactions among students; the dimension of the node is proportional to the messages posted by each student; the classroom subnetworks can be easily identified as well as weakly connected nodes corresponding to students passively engaged in the activities. The authors assessed with the experimental results the potential of SNA analysis as a promising approach to better understand actors and behaviors in an online learning system since they can provide “easy to use” information for decision-making.

The three networks have been analyzed to test the relation between network parameters and academic achievement through parametric correlation analysis (Pearson’s $r$) and non-parametric correlation analysis (Spearman’s rho). The correlation analysis...
confirmed a significant but low positive relation between the different centrality measures and final grade except for eigenvector centrality. Due to the mixed results obtained the authors advises “against relying exclusively on SNA parameters for predictive purposes”.

### 2.4.3 Case study 3: Resources’ access patterns

This case studies addresses the problem of how learning resources’ are used in online courses and presents an analysis of resource access patterns [10]. Experimental data come from the blended learning course “Design of interactive teaching and learning systems”, with 40 students, and the Mooc “Computer-mediated Communication”, with 173 students, University of Duisburg-Essen.

The data sources and dataset are: Moodle log files. Logs identify the student, the resource name and the time when the event occurred and contains all the actions made by the students on the resources in a given time-interval. Sliding such time-interval over time allows to generate a series of actor-resource networks evolving over time. Dataset: actor-resource networks.

The Data analysis techniques used are an adaptation of methods from SNA to dynamic bipartite graphs; descriptive statistics, network analysis and visualization techniques.

Based upon the consideration that agents in a social network tend to organize in subgroups, the authors introduce the concept of “bipartite cluster” corresponding to densely connected areas of a bipartite graph. Bipartite graphs, also called “two-mode” or “affiliation” network, have two types of nodes, called vertexes, that belong to two distinct sets and edges connect vertexes of different types. The authors introduce a new algorithm which is an adaption of the method of Greene, Boyle e Cunningham (2010) designed for one-mode networks, to address the problem of identifying similar “bipartite clusters” in subsequent actor-resource networks.

The method defined to find group of agents - “actor group” - and resources - “resource group of the cluster”-, is based on a subgroup detection method called Biclique Communities method which relies on $K_{a,b}$ biclique notion.

A $K_{a,b}$ biclique is a bipartite subgraph with a nodes of type one and b nodes of the second type, which is also maximal connected; when a set of students – nodes a - are all connected to each resources – node b – we get a $K_{a,b}$ biclique. A biclique community consists of the union of adjacent $K_{a,b}$ bicliques; two subgraphs are adjacent if they share at least (a-1) node of the first mode and (b-1) of the second mode.

Once subgroups are identified, the algorithm is run to identify similar “bipartite clusters” in subsequent actor-resource networks. The algorithm at each step identifies
which “step-group” (in this context, the actor-group and the resource groups at a time slice are called “step-groups”) matches the “step-group” of the previous time slice and it ends when the subgraph corresponding to the last time-slice is processed.

The authors propose a semiautomatic labeling step in order to better understand the content of the “bipartite clusters” identified; for this purpose metadata of the resources or values inserted manually are used.

The experimental results on the two datasets showed peculiarities in the resource access patterns of the two groups. The blended GILLS course showed a regular access of resources and wiki assigned to the topic of each week as well as to wiki articles written by other students. Lecture videos access by small groups of students revealed the presence of groups using such resources “on demand” to rework a presence lecture not attended or to find additional learning resources. During the last weeks of the course, when students were preparing for the exam, it was possible to isolate some “core” learning resources used by more than one group of students. GILLS students in the exam period can be clustered into small overlapping groups affiliated to two types of learning resources: lecture videos and wiki (Fig. 2.4). The CMC course shows different students’ behaviors with respect to the GILLS course where many clusters were found. The tracing of the bipartite clusters indicate that the resources provided at the beginning of the course were used for several weeks and that access of students started to diversify after some weeks of decline.
2.4.4 Case study 4: Social capital and language

In the paper entitled "How do you connect? Analysis of social capital accumulation in Connectivist MOOCs", [18], SNA and automated language and discourse analysis are applied to investigate if language use influences the social capital accumulation of learners in Connectivist MOOCs – CMOOcs. In CMOOcs learners interact “by sharing, aggregating and connecting information through the use of a diverse set of media”.

Experimental data come from the 2011 and 2012 editions of a 12 week CMOOcs course called CCK (“Connectivism and Connective knowledge”); the study analyzes the communications of 1326 “active” learners in the two editions of the course; such data were extracted from Facebook, Twitter and from the message board of the course.

Data analytics used are SNA analysis, linguistic analysis with Coh-Metrics and mixed-effects modeling.

First the network parameters were computed for 72 undirected graphs representing the interaction mediated by the three media (Twitter, Facebook, blogs) for each week of the two courses.

The following measures were computed: degree centrality that takes into account the numbers of edges of a node; the closeness centrality of a node that is based on the natural distance metric between other nodes; the eigenvalue centrality measure that indicate the influence of a node on other nodes; the betweenness centrality that allows to identify nodes acting as a bridge among other two nodes quantifying the numbers of shortest path passing via a given node.

The linguistic analysis was performed using Coh-Metrix on the texts produced by each student in the three networks; the measures adopted were the following Coh-Metrix principal components: narrative, deep cohesion, referential cohesion, syntactic simplicity and word concreteness.

The dataset has been analyzed using R v.3.0.1, with package Ime4, for fitting linear mixed-effects model.

For each centrality measure, the independent fixed effect variables were the five Coh-Metrix principal components. The results of the analysis confirm that linguistic characteristics of the language influence the way a learner interacts in a cMOOcs.

Learners with higher centrality values have also better narrative trend and deep cohesion values, while nodes with lower degree of centrality measures show high degree of text simplicity and referential cohesion meaning that learners with better Coh-Metrics indicators have more links to peer students and teachers. Learners with lower scores for referential cohesion show higher eigenvector centrality values since they are more connected to the most influent nodes. Learners with higher referential cohesion and simple use of text constructs have lower value of betweenness centrality; among those with
higher values of betweenness, learners emerge that enriched the post with new information but low level of cohesion.

The study results allow concluding that learners with better Coh-Metrics indicators develop better their social capital since they are able to build new ties with peers. Therefore an effective use of language and learning settings define the position of the learners in the social network of an on-line learning environment.

Further investigation is needed to assess if the ability to develop social capital could facilitate learning outcomes.

### 2.4.5 Case study 5: Language analysis in the Community of Inquire framework

The paper "Psychological characteristics in cognitive presence of communities of inquiry: A linguistic analysis of online discussions", [19], presents the application of data analysis techniques to the online discussions of an intensive software engineering course in a Canadian online master program. The objective of the work is to investigate the linguistic features of students’ transcripts at different stage of their learning process and their relations to specific individual psychological characteristic related to learning.

The dataset consists of 1747 messages of the asynchronous forums of the online course that involved 82 students. After each message has been coded manually the datasets consists of 308 messages in the triggering phase, 684 messages in the exploration phase, 508 messages in the integration phase, 107 messages in the resolution phase. Finally 140 messages have been coded as others.

The Data analytics used are “Linguistic Inquiry and Word Count” (LICW), a software that calculates several psychological indicators of written text, descriptive statistics and non-parametric analysis.

The research is related to previous works in different areas such as computer-supported collaborative learning (Community of Inquire and cognitive presence), cognitive psychology (cognitive load) and text analysis. The framework used is that of the Community of Inquire (CoI) which was first introduced by Randy Garrison, Terry Anderson and Walter Archer (2000). CoI aims at modeling and enhancing the group learning experience through on-line activities in Computer Supported Collaborative Learning (CSCL) environments. CoI identifies three dimensions in the online learning community: cognitive presence, social presence, and teaching presence (Fig. 2.5). Cognitive presence is related to the ability of learners to build common knowledge through discourse with peers. Social presence describes relationships among peers and social climate in an online community. Teaching presence is finally related to designing, facilitating and supporting the social processes among learners.
Cognitive presence is characterized by the emerging of higher order thinking processes and is operationalized through critical thinking.

The authors define a four steps model of critical thinking: triggering (problem definition), exploration (analysis of different ideas), integration (construction of the meaning of a solution) and resolution phase (how to apply the new knowledge).

The analysis compares at each phases of cognitive presence the usage of several words categories such as “filler”, “tentative”, “causation”, “discrepancy”, “exclusive” words. Other linguistic categories -16-, such as certainty, insight, inhibition, inclusive, cognitive processes, etc. and other well-known language categories from cognitive psychology reported in literature as indicators of specific psychological features are tested at different phases of cognitive presence as well.

Commenting the measures obtained, the authors recognize that the analysis did not reveal exactly what they expected. For instance, the measures obtained in the first experiment do not show any difference in the usage of “filler” words across the 4 phases.

The authors would have expected filler words to characterize the “triggering phase” as these are indicators of ideas which are not completely developed.

The results do not confirm the initial assumptions about the different usage of “discrepancy” and “causal” words across the phases. Such categories were expected to differ between the exploration phase and all other phases.

The only linguistic category showing differences among the pair-wise comparisons is the count of words, while differences of count of words per sentence and causal words between most of the pair-wise comparisons are not significant.

Nevertheless, previous findings about the linguistic categories tested and cognitive presence are confirmed. Such categories, namely cognitive process word, articles,
prepositions, conjunctions, auxiliary verbs, certainty, exclusive, tentative, functional, inhibition, inclusive words, "share a common pattern" across the phases (except between the triggering and the other phases) and are "recognized as indicators of an increase in cognitive load".

2.4.6 Case study 6: Automatic text grading

In their work "Automatic Assessment of Student Reading Comprehension from Short Summaries", the authors present a method for predicting human grading of students’ summaries [22].

The data source and dataset consist of 225 summaries written by students who were asked to read three text containing from 1000 and 1500 words on different topics and to summarize them using between 75 and 100 words. Human experts rated manually the 225 summaries using 4 grades.

Data analytics techniques used: Coh-Metrix: language analysis, descriptive statistics, Pearson-correlation analysis and linear regression analysis.

Information Content (IC) is a measure of text informativeness of a concept in the hierarchical organization of nouns, verbs, adjectives and adverbs such that of WordNet that was first introduced by Resnik P. in 1995.

The authors introduce then a new measure based on Information Content; this new measure is called Information Content of a word and is defined as follows. The value of the measure is given by the IC of the most general concept \( c \) that a given word \( w \) can represent.

IC of a text fragment is then defined as the sum of IC(w) values for the words occurring in that text; the value can be normalized with respect of the total numbers of words in that text.

The text dimension measured with Coh-Metrics are narrativity, syntactic simplicity, word correctness, referential cohesion and deep cohesion.

Grading, performed by experts, has been tested to measure inter-rater agreement using the Cronbach’s alpha test: the value obtained, .802, represented high agreement among the raters. The statistical study performed on the dataset consisted of Person correlation analysis, linear regression and multiple regression analysis. The result suggests that IC is a better predictor for human ratings of summaries but future studies are needed to identify other linguistic indices to predict summary scores; moreover the five dimensions of text complexity were not correlated with summary score as expected.
2.4.7 Measuring and understanding learner emotions: evidence and prospects

Emotions are a critical factor in educational science since they drive attention. This in turn influences learning and memory. Until recent years emotions have not received attention from education theory. If in a face-to-face settings, learners’ emotions such as frustration, anger or boredom can be easily observed by teachers while the same task performed in online learning needs special learning analytics, [37].

The studies presented in [37] have been classified into two big families of analytics: learning analytics that measure emotions using existing data and those using new data.

Learning analytics using existing data: Content analysis was used in an experiment performed in 2012 at University of Limerick, Ireland, in the framework of an electronic mentoring program aiming at facilitating the transition to college for beginner students. The objective of the work was to evaluate whether e-mentoring could support “freshmen” adaptation both emotionally and academically. The dataset contained records on the participation and interaction in the virtual learning environment of a sample of 123 participants (42 mentors and 81 mentees) and a total of 1811 speech acts.

Natural language processing has been successfully applied in many researches to evaluate text in online environment as well as to identify emotions, opinions in written text. In 2010, Dodds and Danforth analyzed 2.4 million blogs to identify the words “I feel . . .” and defined a 9 grade scale for happiness evaluation They finally developed an algorithm to compute a daily “net feel-good factor”.

The iTalk2Learn project at Birbeck College, U.K., analyzed data from speech recognition software to adapt learning environment to the students’ reasoning process. OpenEssayist and OpenMentor have been developed at the Open University, U.K. OpenEssayist is a natural language analytics engine providing feedbacks to students in the writing of essays. OpenMentor is addressed to teachers and supports them in the evaluation of students’ assignments. This is performed through the evaluation of the quality of written feedbacks to students.

Another approach for measuring emotions is through identifying behavioral indicators or extracting them from the transcript of discussion forums. Another source is through tracking learners’ clicking behavior in the virtual learning environment. In an experiment performed in 2007 in online chats, students could use text as well as emoticons in their posts; the analysis showed that in socio-emotional conversations students tended to use more emoticons than in task-oriented ones. Negative conversations and task-oriented discussions reported the least number of emoticons. In 2011, a study
about loafing in small group interactions, showed the positive correlation between negative affect (such as tension or feeling tired) and social loafing, while positive feelings such as happiness or calm were related to positive group interactions.

According to several authors, VLE metrics provide less understanding of the learning dynamics over time than learning motivations and emotions shown by learners. Such understanding could be acted upon by teachers to help weaker students at an early stage.

SNA techniques allow to analyze behavior in the form of interaction patterns among learners. Several studies have been performed in this field often integrating natural language processing of written interactions in order to measure participation. A research performed in 2014 by Rienties, showed different behaviors between autonomous and extrinsically motivated learners. The former related quite easily to other autonomous learners developing written communications, while the latter gradually got isolated. Another experiment in 2014, were “like buttons” were provided, showed that better-written contributions received more attention from the community of learners as well as more “likes”.

Learning analytics new data: a widely used instrument for collecting data about learners’ emotion in blended and online learning is the “Achievement Emotions Questionnaire (AEQ)” developed in 2011.

This questionnaire define a 24 scale of enjoyment, hope, pride, relief, anger, anxiety, shame, hopelessness and boredom and has shown to be highly reliable in identifying self-regulated behavior and in the relationship between emotions and task significance. AEQ is based on the control-value theory of achievement of emotions; this theory posits that achievement emotions are determined by antecedents such as the perceived control of learners over activities and the subjective value given to the activities in the learning process.

Wellbeing word clouds provide the visualization of the words used by students to describe their emotions. The visualization of a word is proportional to its frequency and words may be arranged in clouds of words within a given shape and according to different formats.

Intelligent tutoring systems are automatic system supporting students in a domain of knowledge (math, physics, etc.), in educational games and in game-based learning. They track user inputs and provide the learner with questions and problems. They are also able to understand students’ solutions and to interact with students providing hints and adapting tasks, questions, etc. to the needs of the learners.

AutoTutor, is an intelligent tutoring system where tutoring occurs in the form of a dialogue in natural language. An animated conversational agent with a humanoid
aspect who is able to change facial expression and gestures as well as intonation, interacts through speech. Recent versions of AutoTutor systems also adapt to the learner’s emotional states.

The adaptation of the learning content and the difficulty level of assignment given to a student is based on the estimation and prediction of the student’s knowledge with respect to a student model. Hence, intelligent tutoring systems are built upon a learner model to infer student’s knowledge on a specific topic.

One of the most popular approaches for estimating student knowledge is Bayesian Knowledge Tracing (BKT) [17]. BKT infers students’ latent knowledge from previous results on similar problems and set of skills and recalculates estimations after each attempt. It operates through a hidden Markov model with two hidden states that represent if a knowledge K is mastered or not.

### 2.4.8 Main conclusions

The main take-aways from the case studies were the following. Concerning the interactions among participants of a VLE, from the big amount of data recorded and illustrated in paper [11], SNAs can be built and drawn. The expressive power of SNAs is highlighted when visualizing the interactions that take place within the on-line communities of students and teachers. The density of arcs to/from nodes allows to spot instantaneously those students that interact within the learning community as well isolated nodes corresponding to those students who do not. Furthermore, ways of measuring students’ interactions were treated in paper [18] where metrics were calculated at each learner’s nodes together with linguistic measures computed by specific text analytics.

Papers [18], [19], [22] and [37] all address the second open question. They show how text analytics can help in extracting meaningful indicators from texts exchanged on-line. This refers not only to the formal correctness of texts but also to the evaluation of emotional and psychological aspects of language. These papers also support the opinion that core writing skills are a pre-requisite for communicating within an on-line learning community.

In paper [10], the application of graph theory is demonstrated in relation to the use of educational resources in Moodle based VLEs.

As to the third open question, paper [23] reports on a case where the success of students was accurately predicted by means of EDM techniques. Low scores in math, humanities, physics and English are the best classifiers of those not passing. This confirms something which is already well known in educational practices. It has, however, important implications for preventing school failure on time with targeted interventions.
The above case studies provided useful insights on the application of different types of analytics to educational data. However, in the course of the research, two issues emerged:

a) from the analysis of the Moodle database of the VLE Diplomarsionline it appeared that the system was mainly used as a repository and most interactions occurred outside the VLE, through email, skype and whatsapp. Therefore there were very few transactions which could be used for applying the intended analytics.

b) in the meantime the Italian formal education system underwent a major change. As it will be better illustrated in 3.2.3 the assessment of prior learning and certification of basic skills now becomes a fundamental requirement.

The two points above oriented the research in a different direction, which could better serve the needs of the education community, i.e. towards computer based tools capable of objectively assessing basic skills. A number of take-aways from the case studies still proved to be sources of inspiration for the research: for instance the recognition of the importance of writing skills in digital environments confirmed the need, for educational institutions, to support their development and therefore to measure them effectively.

The case studies also helped in the design of the data mining experimental approach and in providing solutions in the visualization of datasets.
Chapter 3

Setting the stage: why skills matter

3.1 Skills: the international perspective

If the goal of education is to provide students with competences enabling them to keep pace with the rapid changes in the knowledge society, the development of key-skills is fundamental for the success of any formal educational system.

In the European Union we expect that 16 million jobs requiring high qualifications and 12 million requiring lower skill levels will be needed by 2020 but in Europe, around 80 million people have low basic skills and have less access to life-long learning when compared with people with higher levels of education [3].

European policymakers are aware of the importance of skills to sustain economic growth and social cohesion. The EU 2020 strategy for inclusive growth, seeks to promote smart, sustainable, and inclusive growth where a key factor of the strategy concerns literacy, numeracy, science, and technology, so called basic skills.

The book “Better skills. Better jobs. Better lives” contains the strategic approach to skills policies of the OECD to help countries to analyze their national skills systems, to develop policies aiming at translating better skills into better jobs, social inclusion and economic growth. The underlying assumption is that good skills reduce the risk of unemployment, and are positively related to income, health, as well as to social trust. In this framework, OECD has started a comparative survey on education called Programme for the International Assessment of Adult Competencies (PIAAC), [24]. The OECD Report "Skills for a Digital World” published in 2016, analyses the effects on skills for citizens and workers in the digital economy. Authors advise policy-makers to put in place all possible actions which can help citizens mastering digital skills as well as “strong foundation skills, higher order thinking competencies, and emotional skills to respond to greater levels of uncertainty” ([26]).

In an increasingly knowledge-based global society, without the work force having the right skills, technological innovation does not translate into economic growth. This is even more the case when skills are not maintained and upgraded.
Although access to education involves an increasing number of adults all over the world, the impact of skills and education on individuals' opportunities has become stronger.

On average 60% of people with below upper secondary education is employed compared to more than 80% of adults with tertiary education. Furthermore, only 5.8% of adults without upper secondary education but with a moderate level of literacy proficiency were unemployed; this percentage was 2 points higher 8% for adults with similar education attainment but lower levels of literacy proficiency. Across the OECD countries, adults with a tertiary degree earn, on average, 70% more than adults with upper secondary educations [25].

3.2 Skills, the Italian perspective

3.2.1 The PIAAC Results 2012

The results of the OECD analysis are well-known in Italy, as they have been used by national media as a starting point for discussions on education related matters. In 2013, after the publications of the first PIAAC results, the newspaper "La Repubblica" published an interview with Mrs. Carrozza, Ministry of Education, entitled "OECD, Italian "illiterate" of the millennium. Carrozza: a turn-about is needed", Intravaia [12]. Lastly in 2017, the magazine "L’Espresso", section data-journalism, in the article entitled "Italian functional illiterates the Italian drama: who are and why our country is among the worst. The 70 % of Italians is illiterate: read, watch, listen, but do not understand", reported the main findings of the survey with communicative infographics.

In the PIAAC literacy and numeracy rankings published after 2012, Italy alternates between the last and second last position among the 24 OECD countries belonging to the EU. More than 70% of Italians scored at Level 2 or lower and only a minority, 30%, scored at Level 3 or above [4]. This is against the OECD average of 47% where Level 3 is considered the level that matches the minimum necessary skills to live and work in nowadays society, see Section 3.1. The survey revealed that 24.6% of the Italians had no experience with a computer; among those having used a computer, 2.5% did not have the minimal digital skills to perform the computer version of the test, 14.7% opted for the paper version and only 58.2 % of the sample performed the computer version of the test. Additionally, if macro regions are considered, it emerges that Italy is divided in two. In the domain of literacy, for instance, in the North East and Center 39.3%and 36.8% of the sample scored at Level 3 or above while in the South and Islands, these percentages were much lower: 22.1% and 18.0% respectively [4]. Moreover, the figures
about education attainment of Italy in 2012, age group 18-64, showed that 54% had no secondary education attainment, 34% has a secondary diploma and the remaining 12% has a tertiary qualification in contrast with the OECD average of 27%, 43%, and 29% respectively, [4].

3.2.2 An update from 2016 by ISTAT

The “Bes Report 2017: equitable and sustainable well-being in Italy” by ISTAT (National Institute of Statistics), is an yearly report analyzing the main economic, social and environmental phenomena characterizing Italy, through a set of 12 indicators including “Education and Training”. According to the data presented in the Report, [14], in 2016 Italy has managed to reduce, but not to fill, the gap accumulated in previous decades with regards to other European countries as reported by the PIAAC Survey in 2012.

In 2016 in the age group 15-64, the percentage of those without a secondary diploma diminished to 47%: it was 54% in 2012. About 36% of Italians reached a secondary diploma: they were 34% in 2012. Furthermore, the percentage of those with a tertiary qualification showed an increase, from 12% in 2012 to 16% in 2016.

The school dropout rate remains higher than the EU average, although decreasing at 13.8%; it was 14.7% in 2015 and the rate achieved is better than the one predicted by Europe 2020 for Italy (16%). In addition, difficulties in the integration of students born outside Italy have to be highlighted. To date, their rate of early leavers is equal to 30% (19.7% the EU average). In 2016 the rate of NEET (Not in Education, Employment or Training) is equal to 24.3% in the group-age 15 - 29: it was 25.7% in 2015.

In 2016, the number of students going to university registered an increase of 4.5% with about 50.1% of the Italian students choosing university after secondary education. In the same year, an increased number participated in training (8.3%), 1 percent more than in 2015 but far from the EU 2020 objective at 15% and the EU average which is 10.8%. Significant gender and territorial gaps in participation in the training system remain, and in some cases, are increased. The school dropout rate among boys is always higher than that of girls. In the last five years the gap between the North and the South has widened, in terms of both participation and performance, including the acquisition of basic skills.

The tertiary education rate of young people aged between 30 and 34, continues to be the lowest in the EU and below the national target set by Europe 2020. Third-level vocational training (Higher Technical Institutes), in fact, is still not widespread and is struggling to become the alternative channel to university degree courses for those who want to continue their studies after graduation without enrolling in university.
The analysis on digital skills in the Bes Report 2017 is based on data from the yearly survey called “Indagine Aspetti della vita quotidiana” [13]. In this survey ISTAT collects information on the digital activities that people have actually carried out. This is done through home interviews structured according to the “Digital competence framework” system. This framework identifies 4 domains: information, communication, content creation, problem solving. For each domain, three proficiency levels are defined, ranging from 0 to 2; users with level 2 of digital competences in the four domains are classified as high skilled, [2].

Even with this framework, it can be noted that Italy’s position is far from the European average: only 19.5% of the Italian population aged 16-74 years claims to have a high level of digital skills, compared to 28% of the EU average. A low level of skills is also found among the youngest, the so-called "digital natives", born and grown over the years of the dissemination of new information and communication technologies: in the 20-24 age group, young Italians with high skills are 36.5% compared to 52% of the European average. Digital skills are also higher among residents of the Center-North regions (about 22%), in particular of Valle D’Aosta and Friuli-Venezia Giulia, and lower among those living in the South (14.1%), in particular in Campania and Puglia (12.2% and 13.1% respectively). Distances between generations and due to gender are more evident if the levels of access and use of new technologies are compared. People who are able to use a computer with the necessary skills are more than 34% between 16 and 34 years but they reduce to 3% among people aged 65-74. Computer skills are more common among males (22%) than among women (17%), and the gender gap is particularly evident in older age groups.

In 2017 the share of families having access to a broadband Internet connection (with preference to a land-line connection such as adsl or fiber-optic) grew to 69.5% from 67.4% in 2016.

Within Italy differences between regions are still remarkable until 2017, with the center and northern parts of the country having an advantage; regions with a lower degree of penetration of broadband among families are Calabria and Molise.

65.3% of the population of 6 years and above connected to the Internet in the last 12 months (it was 63.2% in 2016, and only 52.5% in 2012), while around 47.6% gets connected on a daily basis. Age is still the main discriminating factor in the use of Internet: the young segments of the population are using it most (more than 92% in the age range 15-24) although there is a remarkable increase also among people in the age group 55-59 (from 62.7% to 68.2%: it was only 45.3% in 2012).

Smartphone and cloud services allow to connect to Internet at anytime, anywhere: 44.6% of Internet users utilizes a Smartphone while away from home or the workplace;
32% uses cloud services to store documents or other files for personal usage.

### 3.2.3 The Italian adult education system

Over the past years Italy has transposed the European Directives stemming from the strategy for inclusive growth, in its national legislation targeting adult education.

In 2015, the above resulted in specific guidelines [35] reshaping the Italian formal adult education system as it was established with the 2007 reform. The new system consists of CPIAs (Provincial Centers for Adult Education) and their service networks of technical, vocational and artistic secondary schools, offering second-level education. The education offer includes the following levels:

**First-level education.** The courses are managed by the provincial CPIA and are organized in two teaching periods. The first teaching period ends with a final examination which grants a lower secondary certification. At the end of the second teaching period, the student receives a certification of intermediate secondary level. This document certifies the end of compulsory education (10 years) and the levels of achievement in those skills characterizing the first and the second year of secondary education.

**Literacy and Italian language courses.** These courses are managed by the provincial CPIA and are open to foreigners. The courses provide language skills that make it possible to reach the A2 proficiency level in Italian according to the Common European Framework for Languages by the Council of Europe.

**Second-level education.** Adults lacking a second-level education can access adult classes in secondary schools. Courses are subdivided into three learning periods. With respect to the ordinary secondary education which consists of five school years, the first learning period corresponds to the first and second school year. At the end of this period, secondary schools, as CPIAs, provide the documentation certifying the accomplishment of compulsory education and the levels of achievement in those skills characterizing the first and the second year of secondary education. The second learning period covers the third and the fourth school year; the third learning period corresponds to the fifth which is also the final school year. Formal and informal credits acquired by the students in previous learning and working experiences are evaluated at admission and students may be assigned to the first, second or third learning period. At the end of the third period, students participate in the final examination to obtain the upper secondary technical, vocational or artistic qualifications.

According to [35], CPIA and the secondary schools offering adult education, have
to subscribe a network agreement and to establish a technical commission for the definition of the so called “Individual Training Pact” 1, a common way to recognize and assess prior learning and to certify the accomplishment of compulsory education in the framework of the national guidelines [33]. The commission should provide guidelines for the design of educational courses as well as for supporting students to continue their learning along the different levels of the adult education system.

As mentioned before, the activities of the provincial CPIA are organized in a provincial network. The central node of the network is the CPIA itself and the other nodes are the physical locations where courses take places. These nodes, referred to as points of service, are identified by regional authorities and may include upper secondary schools, lower secondary schools as well as prisons. CPIA can also collaborate with local public authorities, certified training centers, universities and regions and can perform R&D activities in the field of adult education. The identified research fields are the assessment of skills and the recognition of prior learning based on previous work or study experience, as well as guidance for disadvantaged adults.

In the 2015 reform named "La Buona Scuola", [34], the strategic role of CPIAs is confirmed. This entails supporting education in adulthood, reinforcing basic competencies for life-long learning as well as new skills needed in the labor-market. According to [34], CPIAs can contribute to tackle the so called NEET (not in education, employment or training) problem and can improve the knowledge about Italian culture and language among foreigners.

The reform introduces e-learning modules for the first time in Italy. Two types of e-learning activities are identified:

- **asynchronous e-learning**: up to 20% of the learning period is covered by self-taught activities entailing the use of digital educational resources;

- **synchronous e-learning**, where 100% of the learning period is supported by VLE. These are integrated by face-to-face activities such as guidance at enrollment, specific teaching sessions and assessments during the school year.

The document “Improving competences of Italian adults” [36], states that foundation skills (core skills: reading, writing, oral communication and numeracy) should be assessed and certified as it is already done for ICT (e.g. through the ECDL) and languages in the broader logic of supporting transition to the labor market.

Notwithstanding the central role of the adult education system envisaged in the "La Buona Scuola" reform and by the authors of the document “Improving competences of Italian adults” [36], the latest reforms reduced the duration of the secondary adult

---

1The Training Pact, Patto Formativo, is a document discussed and signed by the student and the school at the beginning of the school year. It contains the educational plan of the student: list of learning modules, skills developed, desired outcomes, learning times, ecc
education. The number of school-years was reduced from five to three and the teaching hours of the school-year were reduced by 30% with respect to ordinary education.

3.3 The PIAAC Methodology

Over the last decades, as globalization and new technologies affirmed, policy makers have concentrated their focus on those skills and knowledge sustaining economic growth and competitiveness. As we have seen in Section 2.3.3, the OECD’s PIAAC survey is a large-scale assessment on adults’ skills, designed to answer such new questions. Since 2012 it collected data on about 200,000 working age respondents in more than 30 countries.

The objective of a large-scale assessment is to gather comparable data on competences and knowledge of the population of interest across different countries through common instruments. The information gathered should provide policy makers and other stakeholders with information on the effectiveness of educational systems in different domains and populations: pupils, students and adults.

3.4 How PIAAC works

Even if PIAAC was originally related to IALS and ALL in terms of the overall design, it has broadened and refined the existing assessment domains and introduced new ones. Altogether it focusses on four cognitive domains: literacy, numeracy, reading components and problem solving in technology-rich environments.

In respect to the previous assessments, PIAAC has made substantial progresses both on what is measured and on how a large-scale assessment is designed, implemented, and also administered. Already from its first round PIAAC introduced a number of innovations:

- a platform delivering both the cognitive instruments and nationally-specific versions of a background questionnaire;
- an integrated assessment design including both computer- and paper-based instruments;
- items designed to mirror the technology-based tasks increasingly required in workplaces and everyday life;
- items which are automatically scored across around 50 language versions of the cognitive instruments;
Chapter 3. Setting the stage: why skills matter

- multistage computer-adaptive testing algorithms which are incorporated into the large-scale assessment;

- data processing, in particular timing information, to both enhance the interpretation of performance and evaluate the quality of the assessments.

We should also mention the PIAAC background questionnaire (BQ), which is a significant component of the survey. It has the scope of relating skills to a number of demographic and explanatory variables. The information collected via the background questionnaire adds to the interpretation of the assessment, and helps investigating how the distribution of skills is associated with variables including educational attainment, gender, employment and immigration status.

Background information also supports the psychometric analysis of the data by providing auxiliary information that can improve the precision of the skills measurement. Moreover the use of background data permits flexible routing so that parts of the questionnaire can be skipped in order to tailor it to the taker.

Ultimately variables can be derived that control the flow of the questionnaire as shown in Fig. 3.1. According to the answer given in the BQ, three groups are identified: i) those that did not use a computer at home or work; ii) those who “opted out”, or refused to take the computer-based version of the assessment; iii) those who reported that they had computer experience but failed the basic computer skills test. These groups were administered paper booklets with literacy or numeracy tasks followed by the assessment of Reading Component skills. The Reading Component module, which was designed to add a better measurement for those who scored at the lower end of the scales, is presented also to adults who failed the Core test of the Cognitive Screener.

3.4.1 Multistage adaptive testing

In PIAAC, a Field Test was carried out which, among others, provided the initial IRT parameters used to construct the multistage adaptive testing algorithms, then implemented in the Main Study. In the surveys before PIAAC, item-level adaptive testing was performed, where the response to a single item determines the next item presented. In PIAAC instead, the multistage design algorithms work on responses to a number of items (a test-let). These in turn determine the next test-let, which will be adapted to the level of the taker. In other words more able respondents will receive a more difficult set of items.

The adaptive design optimizes the match between item difficulty and the ability of the taker, providing a more reliable assessment of the respondent’s skills within the specified testing time. As shown in Fig. 3.2 from [20], the cognitive assessment tools
Chapter 3. Setting the stage: why skills matter

Figure 3.1: The flow of the PIAAC Main study

were designed around two stages with a total of seven test-lets. The set of items presented to a given respondent in Stage 1 was based on a limited set of initial routing information collected in the BQ, as well as the score received on the cognitive screener. Stage 1 included only four blocks of items that broadly covered the range of item difficulty.

Stage 2 included seven blocks and covered a narrower range of the difficulty spectrum. The item selection was based on routing information from the background variables, the cognitive screener and the respondent’s performance in Stage 1. To optimize the match between respondent ability and item difficulty, more able respondents received a more difficult set of tasks than less able respondents.

Figure 3.2: The test-lets of the PIAAC Main study
In Module 1 of the cognitive assessment, as in Fig. 3.3 ([20]) respondents were randomly assigned to either the literacy, numeracy or PSTRE test. Those who were redirected to the literacy or numeracy test, received a total of 20 items. Later, those who received the literacy test were randomly assigned to either the numeracy or PSTRE tests.

### Figure 3.3: The PIAAC computer-based adaptive assessment

#### 3.4.2 Scaling and comparing proficiency

Across the different instruments (computer-based and paper-based), quite a number of items (a total of 58 for literacy, 56 numeracy and 14 PSTRE) are administered to representative samples of adults in each participating country. Comparing performance across this set of tasks posed a challenge: indeed each participant receives and responds to a subset of tasks from each of the three cognitive domains. Items have then to be assembled into test-lets to establish a common scale for each of the domains. Once the data are collected, the pool of tasks within each domain is analyzed in a way that would array the set of tasks along a continuum. This reflects both the proficiency of adults in a particular domain as well as the level of skill and knowledge associated with a correct response.

For dichotomously scored responses PIAAC used the two-parameter logistic model and for items with more than two response categories the generalized partial credit model. The 2PL model was initially used to calibrate the items for each domain and link items across surveys. Once a fixed set of international and national item parameters
was established, a latent regression model was fitted to the data and plausible values were derived.

### 3.5 PIAAC, description of the main findings

The PIAAC survey assesses three domains of cognitive skill: literacy (including reading components), numeracy and problem solving in technology-rich environments (PSTRE). The Literacy and Numeracy scores can range from 0 to 500; five achievement levels are computed from the score obtained in the test ranging from below Level 1 to Level 4/5. For OECD, proficiency is reached with Level 3 in both the case of literacy and numeracy, and with Level 2 for PSTRE. These correspond to the levels that match the minimum necessary skills to live and work nowadays, always according to OECD. Among the relevant findings of the survey, the following can be listed:

- There are considerable differences in the distribution of skills across participating countries.
- A large proportion of the population is not able to use ICTs.
- Better education does not necessarily imply better proficiency.
- Employed have on average higher skills than those looking for a job or inactive.
- Low performers do participate less in adult learning.
- Skills, if not used, tend to deteriorate over time.

#### 3.5.1 Differences across participating countries

Considering as a benchmark the OECD average percentage of the working-age population scoring at Level 3 in literacy (35.4%), we can identify three groups of countries. A group of low performer countries, which includes two important European countries such as Spain and Italy where less than 30% of the respondents scored at Level 3. A second group of countries where this percentage approaches the OECD average; this group includes some of the most developed world economies such as Canada, United States, United Kingdom (UK) and Germany. Finally, we have the group of top performers countries, including Austria and Japan, with more than 40% of the population scoring at Level 3 as shown in Fig. 3.4. For numeracy, the same patterns can be found across most countries.
3.5.2 Use of ICTs

From the bar diagram in Fig. 3.5 that shows the shares of respondents who took the Computer Based Assessment (CBA) of the PIAAC test, we can observe that ICT skills are unevenly distributed in the participating countries. There are countries such as Turkey, Slovak Rep., Cyprus, and Italy were more than 20% of the population declared no computer experience. Furthermore, across almost every country, among respondents declaring that they have used a computer before, some failed the ICT core test or refused to take the computer version of the test opting for the paper version. Notwithstanding, there are countries were more than 80% of the working-age population took the computer version of the assessment.

3.5.3 Education and proficiency

Thanks to the data collected by PIAAC, it is possible to analyze the relation between average scores in numeracy or literacy of respondents and their educational attainment expressed as the highest level of education obtained. In Fig. 3.6, we can notice differences in numeracy proficiency not only across countries but within countries as well. While in Singapore and Chile the difference in performance between upper secondary graduates
Chapter 3. Setting the stage: why skills matter

3.5.4 Employment and proficiency

Fig. 3.7 shows average scores in numeracy of employed, unemployed and those out of the labor force. We can observe that those employed have on average higher skills than those looking for a job or inactive. The difference between average scores of employed and unemployed is less than 10 score points in Japan, in the Republic of Korea, Flanders (Belgium), Greece, Israel and Chile. In UK, Germany, Sweden the gap is very

and tertiary graduates is significantly large, more than 50 score points, in Italy, Spain, Slovak Republic and the Russian Federation it is of less than 20 score points. When comparing mean scores in numeracy across countries, upper secondary graduates in the Netherlands and Japan roughly perform at the same level of post-secondary educated in United States, Israel, France, Greece, Italy and Russian Federation. The impact of initial education on skills maintenance is not as strong as one could expect. Some of the differences across countries may be due to the quality and efficiency of the initial educational systems as well as to individual opportunities to maintain and develop skills throughout the working life.
pronounced and it reaches 30 score points. The positive correlation between better performance and employment suggests that higher skilled have more chances to maintain their skills being part of the active work force than lower skilled.

### 3.5.5 Participating in adult learning and proficiency

When looking at Fig. 3.8 reporting the average scores in numeracy of those who participated in formal education during the 12 months before the survey and those who didn’t, we can notice a number of countries where there is a gap of more than 20 score points between the two groups. This includes Chile, Turkey, Spain, Italy, Israel, Slovenia, Poland, Rep. of Korea, and UK. This difference is less pronounced in the remaining countries. These results confirm that the highly skilled tend to participate more in formal learning activities thus reinforcing and maintaining their skills better than lower skilled individuals.
Figure 3.7: Numeracy score by employment status.

Figure 3.8: Numeracy score and participating in education.
3.5.6 Age and proficiency

PIAAC data reveal significant age-related differences in proficiencies, strongly suggesting that proficiency tends "to naturally" decline with age. Age differences in proficiency are, at first sight, substantial. As depicted in Fig. 3.9 and Fig. 3.10 on average across the OECD countries participating in PIAAC, adults aged 55 to 65 score some 30 points less than adults aged 25 to 34 on the PIAAC literacy scale. This difference is slightly smaller than the score difference between tertiary educated and individuals with less-than-upper-secondary education.

The higher proficiencies of younger individuals may be determined by an increased quality of education while the decline of skills over time may depend on a lack of opportunities to use skills at the workplace or to participate in adult education.
Chapter 3. Setting the stage: why skills matter

Figure 3.10: Numeracy score by age-group in OECD countries
Chapter 4

Analyzing and evaluating Education & Skills Online

4.1 Introduction

The general aim of the case study was to experiment with the derivative product of the PIAAC Survey, called Education & Skills Online (EsOnline in the following).

The research objectives of the case study were:

- to test EsOnline in a real case study with adult students for further adoption by the partnership;
- to analyze the performance of this sample and benchmark it against national and international figures;
- to analyze whether core skills of adult students change after attending school for a full year and to analyze the performance according to the learning period attended.
- to evaluate whether the EsOnline results could be used when defining the Individual Training Pact.

The case study was carried out in partnership with ISIS Bianciardi, the CPIA and other schools of the Province of Grosseto.

This chapter is organized as follows. An introductory section about the testing environment of EsOnline is followed by an account of the field-work and the school partnership. Section 4.3 covers first the datasets then the methodologies and the samples. Afterwards three research questions stemming from the initial objectives are addressed and the respective data analyses and results are discussed. The two last sections of the chapter discuss how the EsOnline results were used by schools and whether the test could be adopted by the partnership.
4.2 Education & Skills Online, the test

The case study portal in Fig. 4.1 was opened in May 2016 thanks to the fundings from the Knowledge Discovery and Data Mining Laboratory (KDD Lab), ISTI, Institute of CNR and Department of Computer Science of the University of Pisa 1.

EsOnline is the commercial product derived from the computer based assessment tool developed in the PIAAC Main Study, see Section 3.3. It provides individual-level results that can be benchmarked against those of participating countries to the PIAAC Survey and it is the result of a joint project by OECD and the European Commission, that involved Italy and other 8 countries. It is available on the web since summer 2015 and is managed by Educational Testing Service ETS, Princeton, USA.

EsOnline assesses the three domains of cognitive skill: literacy (including reading components), numeracy and problem solving in technology-rich environments (PSTRE), and returns as output a score and a level of proficiency for each domain, where:

**Literacy.** “Literacy is the ability to understand and use information from written texts in a variety of contexts to achieve goals and develop knowledge and potential”. This ability is considered a requirement for developing higher-order skills and for “positive economic and social outcomes”.

**Numeracy** “Numeracy is the ability to use, apply, interpret, and communicate mathematical information and ideas”. This is still an essential skill in the information society where a wide range of mathematical and quantitative information is accessible in the

---

1The datasets belong to KDD Lab. and were managed under the CNR Privacy System and according to the current Code of conduct on personal data treatment.
daily life. Literacy and numeracy are distributed differently across subgroups of the population and the survey assesses how these competences interact.

**Problem solving in technology-rich environments (PSTRE)** “This refers to the ability in using technology to solve problems and accomplish complex tasks”. The test does not measure “computer literacy”, but rather how information is used and evaluated to solve a problem through higher-order skills.

![Education & Skills Online test flow diagram](image)

**Figure 4.2: The Education & Skills Online test flow**

EsOnline is delivered in two versions: the "core assessment package" and the "core and non-cognitive bundled package", the latter including three optional non-cognitive modules. The testing environment flow of the "core and non-cognitive bundled package" is depicted in Fig. 4.2. After completing the background questionnaire, three preliminary literacy and numeracy items are presented to the test-taker. Users with very low literacy scores on these core cognitive items are routed directly to the Reading components module, else to the core-assessment which includes the literacy and numeracy test. Next two optional cognitive modules follow: the PSTRE test and the optional component called Reading Module. Three optional non-cognitive modules are available: the "Skill Use Module", the "Career interest and intentionality Module" and the "Subjective wellbeing Module". The latter two modules were not included in the Main PIAAC Survey of 2012. The core cognitive tests and the non-cognitive modules take approximately 120 minutes.
Background Questionnaire This module collects data about the employment status, current activity and income, health status, volunteering activities, political efficacy and social trust. These factors are considered to be influencing the development and maintenance of skills as well as education and social background.

The reading component module, delivered only in case of low level of proficiency, allows to collect data about “the basic set of decoding skills that enable individuals to extract meaning from written texts: knowledge of vocabulary, ability to process meaning at the level of the sentence, and fluency in reading passages of text”.

Module on Skills Use This module collects data about how frequently and how intensively, ICT, reading, writing and numeracy skills are used in the workplace by employed adults. It collects also information about four broad categories of generic work skills that consist of cognitive skills, interaction and social skills, physical skills, and learning skills.

Module Subjective well-being collects data about health, life satisfaction, sleep, fitness, smoking, physical activity, ability to cooperate with others, generosity, etc.

Module Career interest and intentionality collects information on career interests, satisfaction with the current occupation, list of occupations. For those in search of new employment, it evaluates of the level of motivation when seeking employment.

EsOnline scores may vary from 150 to 400: after completing the tests, the user is shown his/her scores together with an analytical report for each cognitive and non-cognitive module completed; these report can be downloaded by the user (see Appendix B). For each domain a proficiency level is computed as follows: Below Level 1 (score less than 175), Level 2 (score from 176 to 225), Level 2 (score from 226 to 275), Level 3 (score from 276 to 325), Level 4/5 (score > 326). For PSTRE, level values are four: Below Level 1 (score < 240), Level 1 (score from 241 to 290), Level 2 (score from 241 to 340), Level 3 (score > 340).

4.2.1 The field work

The testing of the EsOnline methodology started in autumn 2016, ended in autumn 2017 and involved about 150 adult students and four schools of the province of Grosseto. (see Table 4.1 for details).

The data-collection activities, i.e. the testing sessions, were carried out over two rounds. The first round, held in 2016, involved 75 students from a group of four schools. In autumn 2017 the second session involved 77 students coming from the course Diplomarsionline, ISIS Bianciardi; among them, a group of 23 students participated in both
In 2016, the testing was organized in 7 sessions that took place from September to November 2016 in the computer laboratories of the partner schools. In 2017, 3 sessions were held at ISIS Bianciardi in October 2017 during the first weeks of the school year. In 2017 a group of adults was allowed to carry out the activity from home. The 2017 round was meant to collect a second set of measures from the 42 students of the ISIS Bianciardi who participated in 2016. Unfortunately, 16 abandoned the course and only 23 decided to be involved in the second edition of the experimentation.

Each student used a computer connected to Internet and the browser Mozilla Firefox; after inserting the test authorization code in the web-page of Fig. 4.1, the test-taker could start the test selecting the click button next to the test name, see Fig. 4.3. Similarly to test-takers participating in the 2012 Survey, students could use dictionary, calculator, paper and pencil during the testing session.

Before joining the experiment, each test-taker was informed about the objectives of the study and privacy related issues. Furthermore, to receive the test authorization code, the test-takers undersigned a Letter of intent and the data-privacy Informed Consent. The individual results of the Core and PSTRE assessment tests were communicated to the schools.

---

2 The two versions of the EsOnline test were used: the "core and non-cognitive bundled package" version in 2016 and the "core assessment package" version in 2017.

3 For the students attending the course Diplomarsionline for the first time, this activity was mandatory.
Chapter 4. Analyzing and evaluating Education & Skills Online

4.3 Data, methods and analysis

4.3.1 Datasets

The datasets used to perform the analysis are based on information provided by the students or downloaded from the administration portal. The dataset stored in EsOnline, "core and non-cognitive bundled package" version, consists of 133 variables while it consists of 50 variables in the "core assessment package" version. The variables of the complete version of the test, i.e. "core and non-cognitive bundled package" version, can be grouped into 5 domains, each corresponding to a Module of the test. Group 1: Information on age, sex, level of education, mother tongue, employment situation and, for the employed, ISCO (International Standard Classification of Occupations) identification code.

Group 2: Information on the execution time of the literacy, numeracy and problem solving test, scores obtained in the three areas, execution order of the tests; information about the accuracy of vocabulary and text comprehension for those who have scored low in Data literacy, and thus had to undergo the Reading Component test.

Group 3: Information collected through the non-cognitive "Module subjective well being": about health, life satisfaction, sleep, fitness, smoking, physical activity, ability to cooperate with others, generosity, etc.

Group 4: Information collected through the non-cognitive Module "Skill use", on the usage of writing, reading and new technologies at home and at work.
Group 5: Information collected through the non-cognitive Module "Career interest and intentionality" of the test-taker, on their career interests, satisfaction with the current occupation, list of occupations ISCO that are on the one hand closest and on the other less suited to their expressed interests, information about the will to improve their training and finally, for those in search of new employment, the evaluation of the level of motivation when seeking employment.

For each test-taker, two variables were defined: school, indicating the name of the school and the learning period attended. The learning period variable can assume a value from 0 to 3. When its value is 0, it means that the student is enrolled in a short learning programme such as an ECDL course. When the values are 1, 2 or 3, it means that the student is attending the first, the second or the third year of the three years educational programme leading to an upper secondary qualification.

To sum up, the datasets resulting from the experiments were:

- EsOnline Dataset Round 2016, consisting of 75 records, and 133 variable of the EsOnline dataset + 2 user defined variables.
- EsOnline Dataset Round 2017, consisting of 77 records and 50 variable of the EsOnline dataset + 2 user defined variables.
- Repeated measures sample dataset, consisting of 23 records and 6 variables storing the pair of literacy, numeracy and PSTRE scores in 2016 and 2017.
- Whole sample dataset, consisting of 126 records, 50 variable of the EsOnline dataset + 2 user defined variables. It is the union of the 2016 and 2017 samples minus the record-set of the measurements of test-takers who repeated the test twice.

4.3.2 Methods

In the following analysis methods from descriptive statistics and inferential statistics were used. The aim of descriptive statistics is to present data in their most direct meaning, and make them more readable in terms of tables, plots, and with characteristics measures. Inferential statistics are based on probability theory and make rigorous assumptions about data: the conclusions drawn are valid only when these assumptions are satisfied.

4.3.3 The samples: socio-demographic features and response rate

As shown in Table 4.2, the study participants in 2016 were 75 of which 50.6% were female, in 2017 the sample consisted of 77 individuals with almost the same numbers
of male and female participants. As shown in Fig. 4.4, all class-ages were represented in the samples: the average age was 38 years old in 2016 and 36 in 2017.

Most participants (81%) in 2016 and (93%) in 2017 had a Low level of education (primary education, or secondary education without a diploma), 14% and 4% had a Medium level of education (secondary education with a diploma). In 2016 5% of the sample had a High level of education (4-year College or University degree or some post-secondary education), in 2017 this percentage was 3%.

The majority were employed, full-time or part-time workers including self-employed, 60% in 2016 and 78% in 2017. About 30% were unemployed and looking for a job in 2016 while this percentage was about 17% in 2017; the remaining 10% and 5% were not employed and were not looking for a job in 2016 and 2017 respectively.

<table>
<thead>
<tr>
<th></th>
<th>2016</th>
<th>2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size</td>
<td>75</td>
<td>77</td>
</tr>
<tr>
<td>Age, average</td>
<td>38</td>
<td>35.7</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>49.4%</td>
<td>51%</td>
</tr>
<tr>
<td>Female</td>
<td>50.6%</td>
<td>49%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below secondary</td>
<td>81%</td>
<td>93%</td>
</tr>
<tr>
<td>Secondary</td>
<td>14%</td>
<td>4%</td>
</tr>
<tr>
<td>Above secondary</td>
<td>5%</td>
<td>3%</td>
</tr>
<tr>
<td>Occupation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed</td>
<td>60%</td>
<td>78%</td>
</tr>
<tr>
<td>Unemployed</td>
<td>30%</td>
<td>17%</td>
</tr>
<tr>
<td>Out of the labor's market</td>
<td>10%</td>
<td>5%</td>
</tr>
<tr>
<td>Learning period in secondary school</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Learning period 1</td>
<td>36%</td>
<td>16%</td>
</tr>
<tr>
<td>Learning period 2</td>
<td>44%</td>
<td>56%</td>
</tr>
<tr>
<td>Learning period 3</td>
<td>15%</td>
<td>27%</td>
</tr>
<tr>
<td>Learning period 0</td>
<td>5%</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.2: The socio-demographic features of the samples

In 2016 about 36% of the students were attending their first learning period of secondary education while in 2017 the students of this learning period were only 16% of the sample. The second year was attended by the 44% of the sample in 2016 and 56% in 2017 and the third year, the last one of the secondary adult education, was attended by the 15% of students in 2016 and 27% in 2017. In 2016 a small percentage of participants, 5%, attended an ECDL courses (Learning period 0).
Chapter 4. Analyzing and evaluating Education & Skills Online

The response rates of 2016 and 2017 are given in Table 4.3. In 2016, 100% of the students concluded the Background Questionnaire, 96% the Core Test of Literacy and Numeracy, and 84% of them concluded the PSTRE Test; 21% of the sample was redirected to the Reading components Module.

Also in 2017, 100% of the students concluded the Background Questionnaire; 97% the Core Test of Literacy and Numeracy, and 84% of them concluded the PSTRE Test; 13% of the sample was redirected to the Reading components Module.

<table>
<thead>
<tr>
<th>Module</th>
<th>Response Rate 2016</th>
<th>Response Rate 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td>Background Questionnaire</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Literacy and Numeracy Test</td>
<td>96%</td>
<td>97%</td>
</tr>
<tr>
<td>PSTRE</td>
<td>84%</td>
<td>91%</td>
</tr>
<tr>
<td>Reading components</td>
<td>21%</td>
<td>13%</td>
</tr>
</tbody>
</table>

Table 4.3: Response Rate of the samples

4.3.4 Benchmarking the samples’ performances

The Research Question (RQ) driving this analysis was:

Are the average scores of the local samples higher than national and international ones?
The analysis were performed in SPSS on the datasets Round 2016 and Round 2017, described in Section 4.3.1.

Table 4.4 shows the mean scores in Literacy, Numeracy, and PSTRE in Grosseto, Italy and in the OECD countries. Considering that, in the 2012 survey, those who opted for the computer based version of the test had better average scores than those who did not ([4]), their statistics, indicated by CBA in the table, are reported together with the averages of the OECD countries and Italy.

<table>
<thead>
<tr>
<th>Jurisdiction</th>
<th>Literacy</th>
<th></th>
<th>Numeracy</th>
<th></th>
<th>PSTRE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Aver.</td>
<td>S.E.</td>
<td>S.D.</td>
<td>Aver.</td>
<td>S.E.</td>
</tr>
<tr>
<td>OECD Average</td>
<td>268</td>
<td>(0.2)</td>
<td>47</td>
<td>263</td>
<td>(0.2)</td>
</tr>
<tr>
<td>OECD CBA</td>
<td>276</td>
<td>(0.2)</td>
<td>42</td>
<td>275</td>
<td>(0.2)</td>
</tr>
<tr>
<td>Italy</td>
<td>250</td>
<td>(1.1)</td>
<td>45</td>
<td>247</td>
<td>(1.1)</td>
</tr>
<tr>
<td>Italy CBA</td>
<td>261</td>
<td>(1.4)</td>
<td>41</td>
<td>264</td>
<td>(1.3)</td>
</tr>
<tr>
<td>Grosseto 2016</td>
<td>289</td>
<td>(4.9)</td>
<td>43</td>
<td>273</td>
<td>(4.3)</td>
</tr>
<tr>
<td>Grosseto 2017</td>
<td>295</td>
<td>(4.3)</td>
<td>37.8</td>
<td>273</td>
<td>(4.5)</td>
</tr>
</tbody>
</table>

Table 4.4: Mean scores of the samples

To test whether the mean scores of our samples were significantly different from those coming from the PIAAC survey, we applied the independent-samples T-Tests. It is based on the assumptions that the samples are independent and data are approximately normal distributed. This test is used when dealing with the comparison between sample and census means because it tolerates violations to its normality assumption rather well [8]. Having verified the T-Test assumptions, we performed the independent-samples T-tests whose results are given in Table 4.5; when t-values are greater than 1.96, we can reject the null hypothesis that mean values are equal, with 0.95 of confidence interval, [32]. The local samples did score significantly higher than the other ones in almost every comparison, with the exception of some cases, indicated by the symbol **. Therefore, we can conclude that adult students of the local samples did score significantly higher than the Italian sample in literacy and numeracy; this is also true, although only in the case of literacy, when considering the average scores reported by Italians who did the computer version of the test. The comparison with the average scores by adults who opted for the computer version of the test in the OECD countries, allows us to say that the local samples performed significantly better in literacy while they had a non significant difference in performance with regards to numeracy (in both years) and PSTRE (in 2016).

Another way to compare the PIAAC results is to analyze the distribution among achievement levels in a population as we have see in Section 3.5. To do so, we computed
### Table 4.5: The T-test results

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>T-value 2016</th>
<th>T-value 2017</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Literacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OECD</td>
<td>268</td>
<td>47</td>
<td>4.06</td>
<td>6.27</td>
</tr>
<tr>
<td>OECD CBA</td>
<td>276</td>
<td>42</td>
<td>2.47</td>
<td>4.41</td>
</tr>
<tr>
<td>Italy</td>
<td>250</td>
<td>45</td>
<td>7.46</td>
<td>10.15</td>
</tr>
<tr>
<td>Italy CBA</td>
<td>261</td>
<td>41</td>
<td>5.25</td>
<td>7.52</td>
</tr>
<tr>
<td><strong>Numeracy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OECD</td>
<td>263</td>
<td>52</td>
<td>2.15</td>
<td>2.28</td>
</tr>
<tr>
<td>OECD CBA</td>
<td>275</td>
<td>46</td>
<td>0.61**</td>
<td>0.40**</td>
</tr>
<tr>
<td>Italy</td>
<td>247</td>
<td>50</td>
<td>5.65</td>
<td>5.70</td>
</tr>
<tr>
<td>Italy CBA</td>
<td>264</td>
<td>44</td>
<td>1.84**</td>
<td>1.98</td>
</tr>
<tr>
<td><strong>PSTRE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OECD CBA</td>
<td>279</td>
<td>44</td>
<td>1.28**</td>
<td>2.96</td>
</tr>
</tbody>
</table>

The distributions among the achievement levels in Italy, OECD countries, and Grosseto. Results are shown in Figure 4.5. Most of the local sample in 2016 and in 2017 scored at Level 3 or above in Numeracy and in Literacy. In Italy, the percentages of population scoring at Level 3 or above is less than 30% in both the domains. Moreover, around 50% of the samples reached Level 2 or above in PSTRE against the 41% in OECD countries.

Hence, based on the analysis performed, the answer to Research Question is affirmative. The average scores in the three domains of adult students who participated in the experiment in 2016 and 2017 were significantly higher than those attained in the PIAAC survey in 2012. The analysis of the distribution among achievement levels showed that more than 50% of the participants performed at Level 3 or above in the literacy and numeracy domains. In Italy this percentage is around 30% both in literacy and numeracy.

#### 4.3.5 Discussion of the results

The results of the tests performed comparing the mean scores of the local samples and those of the Italians or OECD test-takers, indicated that the mean scores of the samples were statistically significantly higher in literacy and numeracy. The comparison with the average scores of those who opted for the computer version of the test in Italy and

---

4 statistics based on http://piaacdataexplorer.oecd.org/ide/idepiaac/
internationally, showed that the local samples scored significantly higher in literacy; in numeracy and PSTRE no statistically significant differences were found.

Additionally, the analysis of the distributions of the discrete achievement levels in literacy, highlighted that in Grosseto only 37% and 28% of the test takers scored lower than level 2 respectively in 2016 and 2017; in Italy the percentage was 70% and in the OECD countries 54% with reference to the year 2012. Moreover as far as numeracy goes, 49% of the 2016 and 46% of the 2017 sample in Grosseto scored lower than level 2; these percentages were 71% in Italy and 56% in the OECD countries. Furthermore, the majority of adult students (52% of the sample in 2016 and 51% in 2017), scored at Level 2 and above in PSTRE, against 41% in the OECD countries.

There are several factors which could explain the performances of the students despite their low level of education.

Firstly all the students had the ICT basic skills to take the computer version of the PIAAC test. According to the national PIAAC report published in 2012, the Italians who took the computer based version scored higher than those who took the paper version. Their average scores were 10+ points higher than the national averages: 261 vs 250 in literacy and 264 vs. 247 in numeracy.

Secondly, both in 2016 and in 2017, the majority of the participants were employed, (60% and 78% respectively) and, according to PIAAC data, employment conditions and proficiency are positively correlated, as reported in Section 3.5.4.

In addition, the samples consisted of adult students that is individuals participating
in formal education. As shown in Section 3.5.5, maintaining the skills over time is important: those who participated in formal education managed to achieve better average results compared with those who did not participate.

Furthermore, when looking at the age of participants we found a significant difference between the distributions of the variable age in the local and Italian samples. In 2012, the best performance in literacy is found for the age groups 16-19 and 25-29, and tended to decrease in older age groups. As shown by the histogram in Figure 4.7, the average scores in literacy of the two groups were around 265 points, while individuals aged 55+ scored about 30 points less. In the local sample, which by construction was not representative of the national population, individuals aged 50+, are underrepresented, i.e. 13% against 30% in the Italian population. On the contrary, individuals below 30 are overrepresented, i.e. 30% against 24% as shown in Fig. 4.6.

In the local sample, the age group of students aged less than 19, outperformed the others with an average score of 323 points in literacy. The second best performance, which is 307 points, was achieved by students aged 30-34 and not, as for Italy as a whole in 2012, by younger individuals aged 25-29. As for the Italian population, the average performance tended to decrease in older age groups.
4.3.6 Repeated measure

The objective of this analysis was to test whether the core-skills changed after one school year. To do this, we compared the mean scores for the same population at 2016 and 2017, the population of interest being composed by adult students of ISIS Bianciardi. Unfortunately, among the 42 students who participated in 2016, only 62% of them were still attending classes regularly in 2017; 23 students, (88%) volunteered in the second round and were allowed to do the test from home.

The Research Question driving this analysis was:
Did the average scores in literacy, numeracy and PSTRE of the sample differ after one year at school?

The mean scores in the three domains of the sample are displayed in Table 4.6. Although the average scores in the 2016 and 2017 were quite similar, we ran the Wilcoxon Signed Rank Test on the dataset, [32]. This method belongs to the non-parametric statistics and is based on the assumptions that data are paired and come from the same population, [8]. The test confirmed that there were not statistically significant differences in the mean scores after one year of schooling.
4.3.7 Discussion of the results

As anticipated the performance of students was measured twice, firstly in October 2016 and then in October 2017. We did not find any significant difference in the average scores in the three domains. Considering the small dimension of the sample and that a number of students did the test in two different settings, that is at school in 2016 and from home in 2017, results do not seem to be conclusive. The RQ therefore, remains unanswered.

4.3.8 Learning periods and skills

In this section we report the results of a further analysis performed on the sample named Whole sample. The objectives in this case, were to test whether there were differences in scores when considering the learning period of the test-takers. The Research Question driving the analysis was:

Is there any difference in the mean scores when considering the learning period attended at school?

We first created the box-plots of the test-scores in the three groups (Learning period 1, 2 and 3). This visualization subdivides the data into 4 subsets Quartiles, each of them containing the 25% of the sample. The values separating each group are called the first, the second, and the third quartiles and are denoted by $Q_1$, $Q_2$, and $Q_3$; the second interquartile $Q_2$ divides the sample in two equal parts and coincides with the median (see Table 4.7 for the statistics of the literacy score).

The interquartile range is defined as the difference between the third and the first quartiles ($Q_3 - Q_1$). The maximum length of the whisker is proportional to the interquartile range (1.5 time) and extreme values are plotted outside the whiskers. Box-plots allows to easily evaluate the dispersion and symmetry of the data and the presence of outliers.

In the visualization given in 4.8, 4.9 and 4.10, we could spot some outliers, at the extremes of the whiskers, in the three domains that were subsequently evaluated to be genuine. The box-plot of the third learning period showed very large interquartile ranges in literacy and numeracy. This implies a large variability in data. Furthermore,
Chapter 4. Analyzing and evaluating Education & Skills Online

**Figure 4.8**: Box Plot of the Literacy scores by Learning Period

**Figure 4.9**: Box Plot of the Numeracy scores by Learning Period

**Figure 4.10**: Box Plot of the PSTRE scores by Learning Period
we observed a difference in the mean scores in literacy numeracy and PSTRE among the three learning periods.

To analyze the significance of such differences we ran a One-way between-groups Anova with post-hoc test. The Anova found a significant difference among the mean scores of the three groups; to find among which groups the differences were, we ran the Tukey post hoc test, [32].

<table>
<thead>
<tr>
<th>Learning Period 1</th>
<th>Learning Period 2</th>
<th>Learning Period 3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Literacy score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data points</td>
<td>39</td>
<td>65</td>
</tr>
<tr>
<td>Median</td>
<td>280</td>
<td>300</td>
</tr>
<tr>
<td>Mean</td>
<td>275</td>
<td>300</td>
</tr>
<tr>
<td>$(Q_3 - Q_1)$</td>
<td>50</td>
<td>40</td>
</tr>
<tr>
<td><strong>Numeracy score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>250</td>
<td>290</td>
</tr>
<tr>
<td>Mean</td>
<td>255</td>
<td>279</td>
</tr>
<tr>
<td>$(Q_3 - Q_1)$</td>
<td>60</td>
<td>55</td>
</tr>
<tr>
<td><strong>PSTRE score</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>280</td>
<td>300</td>
</tr>
<tr>
<td>Mean</td>
<td>274</td>
<td>296</td>
</tr>
<tr>
<td>$(Q_3 - Q_1)$</td>
<td>50</td>
<td>32</td>
</tr>
</tbody>
</table>

Table 4.7: The statistics by Learning Period

Based on the results of the post-hoc test, the mean scores in literacy, numeracy and PSTRE were significantly different for the comparison with the Learning period 1 and Learning period 2 and not for the comparison with Learning period 3.

### 4.3.9 Discussion of the results

The analysis presented above revealed a significant difference in literacy, numeracy and PSTRE when considering the first and second learning period. On average, the students belonging to the second learning period, scored significantly higher than those of the first one. The mean scores of the students of the first learning period were 275 in literacy, 255 in numeracy and 274 in PSTRE while adults attending the second period, reported a mean score of 300 in literacy, 279 in numeracy and 296 in PSTRE respectively.

The students attending the lessons of the first learning period of the secondary adult education system, are usually adults with the lower secondary school diploma, who have never attended the secondary school in their youth neither have had any work experiences in the domain related to their studies.
Students belonging to the second learning period, are usually coming from the first period of secondary adult education. There may also consist of adults that have direct access to the second year. These students, besides the lower secondary school diploma, must have also successfully completed at least two school years of secondary education. However the latter condition can be replaced by a certified work experiences in the field of study. Overall, the performance of the students in the second year is higher. This is the results of both the better proficiency attained by those coming from the first year and of the very criteria for accessing directly the second year.

One could also expect the students of the third learning period to perform better than the others. Nevertheless we did not observe a further increase in proficiency from the second to the third year. It should be noted, however, that due to the small size of the group attending the third learning period (18 students, about 14% of the whole sample), we are not able to draw any statistically significant conclusions in this respect.

We have further investigated the possible underlying reasons and, as observed in the previous section, found that the distributions of scores of the third learning period have a very large variability. This, besides from being related to the small sample size, may be caused by the origin of the students. These are in fact coming from four types of technical and vocational secondary schools where the ICTs are not always taught. Thus the lower level of digital skills of some students could have contributed to diminish the overall performance of the group.

4.3.10 To evaluate whether the EsOnline results could be used when defining the Individual Training Pact

As we have already mentioned in Section 4.2.1, undertaking the test has been made compulsory for students who enrolled for the first time in the ISIS Bianciardi courses in 2016 and 2017. Moreover, the results obtained integrated the curricular information of the students.

Access to the second educational period was granted to students having the necessary school qualifications for the first learning period and, at the same time, a level 3 proficiency in literacy and numeracy and a level 2 proficiency in PSTRE. Therefore we can conclude that the EsOnline results were instrumental in the definition of the Individual Training Pact.
4.3.11 To test EsOnline in a real case study with adult students for further adoption by the partnership

The following factors behind adoption by the different organization in the partnership have been tested:

The assessment is entirely computer based. This means first of all easiness of use, then the potential to delivery the test to a large number of students, flexibility in testing, authenticity, efficiency, and accuracy. It adapts to the abilities of each candidate and it delivers real-world tasks and the performances are automatically graded by computer. According to ETS, the core cognitive tests and the non-cognitive modules should take approximately 120 minutes to be completed. It was designed to meet ETS’s high standards for quality and fairness and it is based on an international, consistent and objective scoring system. This allows the comparison of the results across test-takers, courses, schools and their benchmarking against national and international ones.

The high response rates indicate that students found EsOnline to be "usable": about 97% of the sample completed the core questionnaire of literacy and numeracy and 84% and 91% of them completed the test PSTRE, in 2016 and in 2017 respectively. We observed that a number of students took more than the estimated time (120 minutes) to complete the test and some of them requested a break.

However a number of issues emerged in the actual use of the test. EsOnline is easy to administer: a good user guide and support by ETS, both in terms of documentation and on-line assistance, are available. However the assessment, which is carried out online, is delivered solely for the Mozilla Firefox browser. Moreover the user interface of the test is not really up-to-date, since it was designed before 2012.

Although the test is adaptive, that is each person is presented with a unique subset of the total pool of test items, the number of testlets is limited. This may be conducive to cheating, given that two nearby test takers can receive the same set of testlets. Minor translation mistakes were reported by students during the first round in 2016, which were solved by ETS based on our indications.

It may be relevant also to evaluate how would the adoption of EsOnline help mostly. As done by ISIS Bianciardi, that used it as a tool for the recognition of prior learning of new students or it could be used by CPIAs and adult secondary schools to certify fundamental skills at the end of the compulsory education cycle.

What are the costs involved in the use of this test?

The price of the "core assessment package" and the "core and non-cognitive bundled package" varies depending on the quantity of tests purchased. The "core assessment package", contains the core background questionnaire, the literacy and numeracy
assessment, and the optional problem solving and reading components measures. The cost of each test-code is of 9.00 Euro for up to 5,000 test-codes.

The "core and non-cognitive bundled package" contains the core background questionnaire, the literacy and numeracy assessment, the optional problem solving and reading components measures, and the four non-cognitive modules. It costs 12.50 Euro for up to 5,000 test-codes.

Finally, the purchase of the EsaOnline test by an Italian public administration, as in the case of any international procurement, involves a number of declaratory and accounting obligations.
Chapter 5

Extracting knowledge from PIAAC data

The objective of the data-driven analytical process, was to shed light on our knowledge on fundamental skills of adults based on the PIAAC datasets.

The chapter is organized as follows: after a section about the data, the adopted methodologies will be described in Section 5.2. Section 5.3 shows the results of the analytical process performed on the national dataset and Section 5.4 presents the analysis of the international datasets. Section 5.5 displays how we used the results of the previous analysis to build a predictor to classify new instances in the Grosseto sample. In Section 5.6 the results of the analytical process are discussed. The first version of the analyses presented in this chapter was published in [40]. It was then part of an "invited talk" at the satellite seminar "Università e Ricerca: costruire una Research Unit sull’educazione nel 21mo secolo" on the state of the art of the Italian research on digital innovations and education. The seminar took place during the event "Il Piano Scuola Digitale incontra il Paese", organized by the Italian Ministry of Education and Research in July 2017, Rome.

5.1 Data preparation

The preparation of the OECD PIAAC dataset and the new data collected in the case-study was performed with the KNIME Analytics Platform and in SPSS.

The PIAAC files, named PUFs (Public Usage Files), store the results of more than two hundreds thousands individual tests performed by adults in the framework of the 2012 PIAAC. For each test, 1.328 variables are stored; the variables are organized in 16 domains. Some of them are recorded during the testing session, others are derived. The domains are: Background (coded), Background Questionnaire, Background Questionnaire (derived), Background Questionnaire (trend), Literacy (computer), Literacy (paper), Not assigned, Observation module (ZZ questions), Numeracy (computer),
Numeracy (paper), Problem solving (computer), Reading components (paper), Sampling/weighting, Sampling/weighting (derived), Workflow/logistics, Scale scores.

As shown in Section 4.3.1, EsOnline stores only 133 variables which consists of roughly 10% of the OECD dataset; among these variables two groups are not part of the PIAAC dataset. This is due to the fact that the EsOnline Test includes two new non cognitive modules (Career interest and intentionality Module and Subjective Wellbeing Module) which were not available in the main PIAAC study of 2012.

The two datasets not only have different numbers of variables, but also differ in their attribute names, types, and scales. For instance, the variables age group is missing in the EsOnline dataset while the variable occupation uses a different coding with respect the OECD dataset.

Out of the 1,300 variables of the Italian dataset and the 130 of EsOnline, the common variables selected for the analysis were only 26.

5.1.1 The datasets

Dataset1 was used in the first analysis and consists of 4,600 records and 20 attributes. It stores demographic status, education, participation in formal education, and Skills Use of the Italians 1. The Skill Use variables, which are 14, code the levels for Reading, Writing, Numeracy, and ICT at work and at home (8 scores) and the responses to 6 questions about computer use. The skills use at home and at work variables contain values that vary from 1, lowest level of use, to 5, highest level. The ICT related questions are the following: "Have you ever used a computer?", "Do you use a computer at work?", "What level of computer use is needed to perform your job?", "Do you think you have the computer skills you need to do your job well?", "Has a lack of computer skills affected your chances of being hired for a job or getting a promotion or pay raise?", "Do you use a computer in your everyday life now (outside work)?".

Dataset2 extends Dataset1 with four new variables quantifying individual proficiency in Literacy and Numeracy. Two variables, called PVLitMedian and PVNumMedian, contain the median scores in the two domain and were computed from the ten Plausible Values for Literacy and Numeracy 2. The last two variables, PVLitLevel and PVNumLevel, contain the achievement levels in the two domains and were computed from the PVLitMedian and PVNumMedian. Their values vary from 0 to 4 as follows: 0, median score < 175; 1, median score from 176 to 225; 2, median score from 226 to 275; 3, median score from 276 to 325; 4, median score > 326.

---

1Data source: Data Agreement with ISFOL - Programme for the International Assessment of Adult Competencies (PIAAC), File ISF57.

2Plausible values (PVs) are used to estimate population parameters in large-scale assessment programs.
The Dataset3, which was used in the classification analysis in Section 5.5, is the extension of Dataset2 with a new variable identifying the geographical area of origin of the test-takers; five Italian macro-regions are defined: north west, north east, center, south, and islands.

Other analyses were performed on 28 OECD Datasets. The datasets of 25 countries are based on the OECD Public Use Files (PUFs) and have been built with the same procedure followed for the Italian dataset. The Finnish, Austrian and German datasets have been released by the national PIAAC authorities after the signing of a distribution contract.

The DatasetGrosseto1, is based on the Dataset 2016 of the case study held in Grosseto. It contains 65 records and the same variables of Dataset1.

Datasetclustered consists of Dataset1 extended with the new variable called "cluster id". For each instance of the dataset, this variable contains a value varying from 0 to 8 corresponding to one of the 9 clusters identified by the analysis in Section 5.3.

In Table 5.1 the demographic, geographical origin, educational and occupational statistics of the Italian sample are shown; we report in brackets the values stored in the dataset.

5.2 Methods

The data-driven analytical process was carried out using Data Mining techniques in order to discover interesting patterns and relationships in our data without formulating any prior hypothesis. These techniques do not make preventive assumptions on data and are actually aimed at extracting new knowledge from them for generating further hypothesis [1]. Data Mining methods can be applied to several problem’s categories such as:

- Classification problems, when there is a categorical variable to be predicted.
- Regression problems, when the target variable to be predicted is numerical.
- Clustering and Segmentation, when there is the need to find groups in data according to similarity.
- Association Analysis, when there is the need to analyze the relationships between the attributes.

---

Chapter 5. Extracting knowledge from PIAAC data

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male (1)</td>
<td>50.0</td>
</tr>
<tr>
<td>Female (2)</td>
<td>50.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age group</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>16-24 (1)</td>
<td>14.4</td>
</tr>
<tr>
<td>25-34 (2)</td>
<td>18.9</td>
</tr>
<tr>
<td>35-44 (3)</td>
<td>24.4</td>
</tr>
<tr>
<td>45-54 (4)</td>
<td>21.8</td>
</tr>
<tr>
<td>55-65 (5)</td>
<td>20.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Macro region</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>North West (1)</td>
<td>21.8</td>
</tr>
<tr>
<td>North East (2)</td>
<td>21</td>
</tr>
<tr>
<td>Center (3)</td>
<td>18.3</td>
</tr>
<tr>
<td>South (4)</td>
<td>27</td>
</tr>
<tr>
<td>Islands (5)</td>
<td>11.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Education</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below secondary (1)</td>
<td>53.4</td>
</tr>
<tr>
<td>Secondary (2)</td>
<td>33.8</td>
</tr>
<tr>
<td>Above secondary (3)</td>
<td>12.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employed (1)</td>
<td>55.8</td>
</tr>
<tr>
<td>Unemployed (2)</td>
<td>9.0</td>
</tr>
<tr>
<td>Out of the labor’s market (3)</td>
<td>34.5</td>
</tr>
</tbody>
</table>

Table 5.1: Demographic features of the Italian sample

- Deviation Analysis, when there is the need to find subgroups behaving exceptionally with respect to a predictor.

Classification and regression methods aim to model a dependency among input attributes and the response variable, and are often referred to as supervised learning techniques. Supervised learning relies on labeled datasets, i.e. each instance in the input table has a value/label in the column corresponding to the response variable. The response variable may be called the dependent variable and the other variables can be referred to as independent variables or predictor variables. Methods such as clustering, which do not concentrate on a response variable but aim at discovering unknown patterns in the dataset, are referred to as unsupervised learning techniques.

As shown in Section 2.4, DM methods find their application in several fields related to education. A typical classification problem is the prediction of students’ success (failed or not failed), while the prediction of students’ scores is a typical regression problem.
5.2.1 Decision Trees

Decision trees belong to the category of supervised learning techniques. They can be easily understood by humans and represent popular and powerful techniques for classification and prediction. A decision tree is a classifier in the form of a tree structure; there is one categorical response variable labeling data; instances or examples are classified by starting at the root of the tree and moving through it until a leaf node is reached. Each internal node is a decision node and specifies a test on a single attribute; arcs/edges represent the split of one attribute; a path is the disjunction of test to make the final decision and the Leaf nodes correspond to the possible values of the response variable. The decision tree is first trained on a labeled set of instances, Training Set, and then is evaluated on a Test Set of unlabeled instances. The learning phase of the decision tree ends with a classifier, that is a function that maps the set of values of attributes into the set of classes. The decision tree learning algorithms are known as recursive partitioning methods because at each split they divide the data into disjoint sub-partitions. In the decision tree construction process, the main task is to determine which is the best attribute to split a given node. Splits should be made in such a way that the resulting attribute partitions are as pure as possible. To calculate the impurity of a sample, Entropy is often used.

Let \( S \) denote the set of data elements in a node, \( n(S) \) be the cardinality of \( S \) and \( n^k(S) \) the number of elements of set \( S \) from the \( k \)th class. For each attribute indexed by \( i, i = 1, ..., D \), the data set \( S \) is divided into disjoint \( Q \) subsets \( S_{i,q}, q = 1, ..., Q \), where \( Q = 2 \) for binary trees and \( Q = v_i \) for non-binary trees. For convenience, one can define the fraction of elements from the \( k \)th class in set \( S \) as a ratio between \( n^k(S) \) and \( n(S) \):

\[
p_k(S) = \frac{n^k(S)}{n(S)}
\]

Using the notation of fractions \( p_k(S) \), the Information Entropy is given by

\[
E(S) = -\sum_{k=1}^{K} p_k(S) \log p_k(S)
\]

The Entropy ranges from 0 to 1 and is measured in bits of information. When the subset contains many different elements which are unique, the variation is maximal, and it takes many bits to encode the individual instances. On the contrary, an Entropy equal to 0 indicates that all elements in the subset are the same, and therefore no bits are needed to encode the individual instances.

If we denote the entropy of set \( S \) by \( g(S) \), then, the split measure function can be expressed as

\[
\Delta g_i(S) = g(S) - \sum_{q=1}^{Q} \frac{n(S^q_i)}{n(S)} g(S^q_i)
\]

The attribute providing the highest value of the split measure function, which is called Information Gain, is used to split the given node.
The result of the classification model is summarized in a table, called confusion matrix, that contains the counts of the instances that are correctly and incorrectly predicted. Based on the values in the confusion matrix, it is possible to compute the total numbers of correct and incorrect predictions and the following performance metrics:

- **Accuracy**, which is the number of correct predictions divided by the total number of predictions.

- **Error rate**, which is the number of wrong predictions divided by the number of predictions.

### 5.2.2 Clustering

Cluster analysis belongs to unsupervised learning methods which do not focus on a target attribute but aim at finding groups of similar data within the whole dataset. Instances in any given group should be similar to each other and dissimilar to instances in other groups according to some dissimilarity measure, [1].

One of the most common clustering algorithms is K-means. Each instance is considered as a n-dimensional vector where \( n \) is the number of variables. Each cluster has a centroid denoting its center that is computed by taking the average of the coordinates of the instances in the cluster. All instances are assigned to the closest cluster according to the Euclidean distance between the instance and the centroid of the cluster. Based on the average distance of an instance to its centroid, it is possible to define the quality of a particular clustering.

In the K-means algorithm, the number of clusters \( k \) is fixed from the beginning. Thus, when using k-means clustering, there is the need to determine whether the algorithm is initialized with the right value of \( k \). An approach which addresses this problem is to start with a reduced number of clusters and then increase \( k \) if there are significant improvements. One method to validate the number of clusters \( k \) is the **elbow** method. The idea behind this method is to run k-means clustering on the dataset for a range of values of \( k \) and for each value of \( k \) calculate the sum of squared errors (SSE). SSE is the sum of the squared differences between each observation and its group’s mean and it can be used as a measure of variation within a cluster. If all cases within a cluster are identical the SSE would then be equal to 0. Then for each value of \( k \), a line chart of the SSE is plotted. Whenever the line chart looks like an arm, then the elbow on the arm is best value of \( k \). SSE tends to decrease toward 0 when \( k \) is increased: the SSE is 0 when \( k \) is equal to the number of data-points in the dataset, because each data point is its own cluster, and there is no error between it and the center of its cluster. The goal
of the method is to choose a small value of \( k \) that still has a low SSE, and the \textit{elbow} usually represents the point where there will be diminished returns by increasing \( k \).

### 5.2.3 Multi Dimensional Scaling

MDS is a visualization technique that helps to observe similarities or dissimilarities between the investigated objects by providing a geometrical representation of their relations. MDS maps high-dimensional datasets into lower-dimensional ones preserving pairwise distances and addresses the problem of constructing a configuration of \( n \) points in Euclidean space by using information about the distances between the original instances.

For a dataset of \( n \) instances, MDS builds a \( n \times n \) matrix \( D \) called Distance matrix where \( d_{i,j} \) is the distance between data object \( i \) and \( j \) for all \( i, j \). The distance matrix is usually based on the Euclidean distance between the multidimensional data objects. A simplified version of the algorithm is as follows:

1) Assign each point to arbitrary coordinates in a \( p \)-dimensional space;
2) Build a new Distance matrix, computing the Euclidean distances among all pairs of points in the \( p \)-dimensional space;
3) Compare the new Distance matrix with the input Distance matrix \( D \) by evaluating a loss function called \textit{Stress}: the smaller the value, the greater the correspondence between the two;
4) Adjust coordinates of each point to minimize Stress;
5) Repeat steps 2 through 4 until stress values will not diminish any longer.

Two thing to be observed in an MDS plot are the directions within the space and the grouping/separation of data-points as they reveal which samples are more likely to be similar/dissimilar one another. The distances among data-points are a distorted representation of the original distances, the orientation of the picture is arbitrary and the coordinate axes are a device to "hang" the points within the \( m \)-dimensional space, [15].

### 5.3 The analysis of the Italian dataset

In our first analysis, we applied the K-Means clustering model to the Italian dataset named Dataset1; the analysis was performed in Python applying the K-means algorithm of the Scikit-learn library ([41]). This was done after computing the optimal number of cluster \( k \) (9 in our case), with the Silhouette method.
All the variables of the dataset were coded into integer values as shown in Table 5.2. Table 5.3 shows the centroids matrix, where the columns are the means vectors of the nine clusters, and the rows are the variables of the observation.

<table>
<thead>
<tr>
<th>Name</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>&gt; 16</td>
</tr>
<tr>
<td>Gender</td>
<td>male=1, female=2</td>
</tr>
<tr>
<td>Occupation</td>
<td>employed=1, unemployed=2, out of the labor’s market=3</td>
</tr>
<tr>
<td>Age group</td>
<td>16-24=1, 25-34=2, 35-44=3, 45-54=4, 55-65=5</td>
</tr>
<tr>
<td>Education</td>
<td>participated in formal education=1, did not=0</td>
</tr>
<tr>
<td>Computer at work ?</td>
<td>yes=1, no=2, no answer=9</td>
</tr>
<tr>
<td>ICT level job?</td>
<td>straightforward=1, moderate=2, complex=3, no answer=9</td>
</tr>
<tr>
<td>Right ICT skills ?</td>
<td>yes=1, no=2, no answer=9</td>
</tr>
<tr>
<td>Lack of ICT skills ?</td>
<td>yes=1, no=2, no answer=9</td>
</tr>
<tr>
<td>Ever used a computer ?</td>
<td>yes=1, no=2, no answer=9</td>
</tr>
<tr>
<td>Computer at Home ?</td>
<td>yes=1, no=2, no answer=9</td>
</tr>
<tr>
<td>ICT Home / Work</td>
<td>low use=1,..., high use=5, no answer=9</td>
</tr>
<tr>
<td>Numeracy Home / Work</td>
<td>low use=1,..., high use=5, no answer=9</td>
</tr>
<tr>
<td>Reading Home / Work</td>
<td>low use=1,..., high use=5, no answer=9</td>
</tr>
<tr>
<td>Writing Home / Work</td>
<td>low use=1,..., high use=5, no answer=9</td>
</tr>
</tbody>
</table>

**Table 5.2: Coding of the cluster analysis’ variables**

Analyzing the Table 5.3, we can observe some of the determinant variables which were used by the K-means algorithm to assign a record to a specific cluster. These were the answers to the following questions “Have you ever used a computer?”; “Do you use a computer at work?”; “Do you use a computer in your everyday life now (outside work)?”. Based on the analysis of the centroids matrix, we have characterized the nine clusters as follows.

- **Cluster 0** can be defined as the *middle aged unskilled workers* cluster because it is mainly composed of workers, and the average age is 41.95 years. It has almost the same number of males and females, with education below diploma; respondents falling in this cluster do not use a computer at work and outside work. The skills usage at home and at work is low.

- **Cluster 1**, can be defined as the *mature unskilled workers*. The average age is about 55 and its centroid is quite similar to that of cluster 0.

- **Cluster 2**, can be named as the *young skilled workers* cluster. The average age is 29.32; it groups mostly employed people, with education above diploma; they
use a computer at work and outside work. The mean values for skills usage at home and at work are around 3 on a scale from 1 to 5.

**Cluster 3**, *young adults*, is composed mainly by individuals out of the labor market, whose age is on average 21. Most participated in formal education during the 12 months before the interview and do not have a secondary diploma. They all use a computer and make an intensive use of ICT skills at home (4 out 5) as well as of reading, writing, and numeracy skills.

**Cluster 4**, is composed of those who do not work anymore and can be defined as *elderly people* cluster, the average age is 60, and the level education is below secondary. The majority has never used a computer but some of them use ICT at home. The use of writing, reading and numeracy skills at home is low (mean of 1 on a scale from 1 to 5).

**Cluster 5**, is made of *mature skilled workers*. The average age is around 54 years and for the remaining variables, its centroid is quite similar to that of cluster 2.

**Cluster 6**, *unemployed women*, is mainly composed of women looking for a job; the average age is 41, education is below secondary; the majority have used a computer but do not use it in the daily life. The mean values for skills use at home are around 2 while the skills use at work values are missing.
Cluster 7, can be named as the *middle aged skilled workers* cluster. The average age is 41.35; it groups mostly employed people, the education is above diploma, they use a computer at work and outside work. The mean values for skills usage at home and at work are around 3 on a scale from 1 to 5.

Cluster 8, is made of *young unskilled workers*. The average age is around 28 years and for the remaining variables, its centroid is quite similar to that of cluster 0 and 1.

To generate an estimate of the proficiency of each cluster, we created Dataset2 which extends Dataset1 with four new variables quantifying individual proficiency in Literacy and Numeracy. As shown in Table 5.4, our estimations of proficiency scores are above the Italian average which is 250 for literacy and 247 for numeracy on a scale from 0 to 500 for cluster 2, 5, and 7. The definition mature/middle aged/young skilled workers can be further refined in mature/middle aged/young skilled and literate workers. Young unskilled-workers, middle aged unskilled workers, and mature unskilled workers have, on average, values which are less than the Italian average both for literacy and numeracy scores. For cluster 4 and 6, proficiency scores are below the national averages as well, while the average scores for cluster 3, young adults are both above 250 so that we can better qualify this group as *literate young adults*.

We have then performed a visual analysis of the dataset to inspect basic relations among instances. The analysis was performed in Python applying the Manifold algorithm of the Scikit-learn library ([41]) on Dataset1; the dissimilarity measure used was the Euclidean distance.

Fig. 5.1, shows the scatter-plot resulting from the processing of the original dataset stratified by cluster id.

We can easily spot three separated and parallel large groups or clouds, and each is further separated into subgroups. The central cloud is divided in three and is composed of data-points belonging to cluster 0, 1, and 8, that is *middle aged unskilled workers*, *mature unskilled workers*, and *young unskilled workers*. The right plot is separated into three parts (i.e. the data-points rendered in pink, pale orange, and blue) belonging to cluster 5, 7, and 2. These are *mature and literate skilled workers*, *middle aged skilled and literate workers*, and *young skilled and literate workers* cluster respectively.
The left cloud displays three separated clusters (4, 6, and 3) and represents individuals belonging to the group of elderly people with dark green data-points; red data-points for unemployed women, and pale green data-points for the literate young adults’ cluster.

The picture produced by MDS confirmed the patterns found in the dataset by the previous cluster analysis. The similarities of the instances belonging to the skilled workers, unskilled workers and out of the labor market clusters are made clear by the positioning of each item along three parallel oblique axes. The distances among the three axes confirm the dissimilarities among the three groups. After the stratification by cluster id and the coloring of the data-points accordingly, we could appreciate the ordering by age along the three oblique axes.

After the visual analysis performed applying the Multidimensional Scaling algorithm, we computed the distribution of Numeracy and Literacy scores in the three clouds of skilled workers, unskilled workers and out of the labor market, see Fig.5.3. Box plots allow to display the data distribution based on the five numbers summary (minimum, maximum, first quartile, median, and third quartile) and to identify outliers values that are above or below the third or the first quartile. Scores are not evenly split at the median and we can also notice the presence of outliers, especially below the minimum, in every box plots. This indicates the presence of relatively low performer test-takers in every group.
Chapter 5. Extracting knowledge from PIAAC data

<table>
<thead>
<tr>
<th></th>
<th>&lt;Level 1</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4/5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Literacy, all workers</td>
<td>5</td>
<td>21</td>
<td>41</td>
<td>29</td>
<td>4</td>
</tr>
<tr>
<td>Numeracy, all workers</td>
<td>6</td>
<td>21</td>
<td>38</td>
<td>29</td>
<td>6</td>
</tr>
<tr>
<td>Literacy, skilled</td>
<td>0.51</td>
<td>8.87</td>
<td>37.56</td>
<td>46.04</td>
<td>7.02</td>
</tr>
<tr>
<td>Literacy, unskilled</td>
<td>5.09</td>
<td>26.52</td>
<td>46.89</td>
<td>20.15</td>
<td>1.35</td>
</tr>
<tr>
<td>Literacy, out</td>
<td>5.84</td>
<td>20.50</td>
<td>48.13</td>
<td>23.96</td>
<td>1.56</td>
</tr>
<tr>
<td>Numeracy, skilled</td>
<td>0.67</td>
<td>8.53</td>
<td>36.33</td>
<td>43.46</td>
<td>11.01</td>
</tr>
<tr>
<td>Numeracy, unskilled</td>
<td>6.97</td>
<td>27.64</td>
<td>44.64</td>
<td>19.55</td>
<td>1.20</td>
</tr>
<tr>
<td>Numeracy, out</td>
<td>9.44</td>
<td>25.59</td>
<td>43.25</td>
<td>19.69</td>
<td>2.04</td>
</tr>
</tbody>
</table>

Table 5.5: Percentages across the five proficiency levels, literacy and numeracy, for Italian workers.

In Table 5.5, we show the distribution among the achievement levels in Literacy and Numeracy of Italian workers, [4], and in the three groups of test-takers identified by our analysis. According to national data, only 33% of Italian workers scored at Level 3 or above in Literacy and 35% in Numeracy.

The statistics of the three groups of skilled workers, unskilled workers and out of the labor market, show that the percentage of workers belonging to clusters of skilled workers scoring at Level 3 or above, is 53% for Literacy and 54.5% for Numeracy while these percentages fall down to 21.50% and 20.75% for unskilled workers. Also, among those stating their use ICT at work, the percentage of low scorers, people scoring at Level 1 or below, is less 9.38% in literacy and 9.21% in numeracy, while for those workers doing jobs where computers are not used, the percentage of low scorers is 31.61% in literacy and 34.61% in numeracy.

According to Falck in [5], the better results of those declaring to use ICT at work may be due to the kind of tasks performed in their jobs: the use of computer at work characterizes professions with high abstract task intensity and not professions with high routine or manual task intensity. Furthermore, our cluster analysis confirmed the statistics presented in [4]: workers declaring to use ICT at work, achieved also high values in the use of reading, writing and numeracy skills both at work and at home. This also contributes to keeping and developing skills and thus to higher proficiency levels.

5.3.1 Adding the macro-regions’ variable

In Fig. 5.4 and 5.5, we plot the distribution of literacy and numeracy scores in the the five Italian macro-regions and at national level: the north west, north east, and the center regions have higher scores than the Italian average, while the south and the islands show a lower performance.
To understand the behavior of skilled workers, unskilled workers and out of the labor market in the different areas of the country, we added the macro-region variable to the clustered dataset and we computed the proportion of test-takers in the three groups.

In Fig. 5.6, we can observe that while in the north and center of Italy about 45% of test-takers belong to the group of skilled workers, this share goes down to 24% in the islands and 29% in the south.

The share of the unskilled workers is quite similar, about 30%, both nationally and in the five macro-regions.

Those test-takers who are out of the labor market represent 46% of the whole population in the islands and the 41% in the south. This share is 32% nationally while it is around 25% in the northern and center of the Italy. This confirms that the differences in the socio economic structure of parts the country have a great impact on the labor market; jobs requiring high skilled professionals are concentrated in the center and north of Italy and the unemployment rate is higher in the southern areas and in the islands. The differences in the distribution of scores in literacy and numeracy in the five areas of the country go together with the different shares of test-takers belonging to the three groups of clusters.

### 5.4 The analysis of the international datasets

In this section we will show the results of the application of the data-driven analytical process to 28 countries among the 34 countries that participated in the PIAAC Survey since 2012.
Chapter 5. Extracting knowledge from PIAAC data

Figure 5.3: Distribution of Literacy scores of by working status.

Figure 5.4: Literacy scores in the Italian macro region.
Figure 5.5: Numeracy scores in the Italian macro regions.

Figure 5.6: Distribution by working status, Italian macro regions.
After building the datasets from OECD PUFs, we performed the cluster analysis following the same procedure described in Section 5.3 and applied to the Italian dataset. The datasets were then extended with the variables PVLitLevel and PVNumLevel to estimate the proficiency of each cluster, and finally the clustered datasets were visualized applying MDS.

Although the numbers of clusters identified by the silhouette methods may vary in the countries, the patterns found in the datasets are quite similar to those found in the Italian dataset. In every country, test-takers are grouped according to occupational status, ICT usage at work, and age. Based on these premises, we will comment on the results of a number of countries selected according to their average scores in Numeracy. We selected Japan and Denmark that scored above the OECD average (263 points), then we chose UK that scored next to the OECD average, and Spain and Turkey that scored below the OECD average.

As in Tables 5.6 and 5.7, the percentage of high scorers, test-takers scoring at Level 3 or above, is higher in clusters grouping workers using ICT at work than in those clusters of people not using digital technologies at work or of people out of the labor market.

We could name the three groups as *out of the labor market*, *skilled workers*, and *unskilled workers*. This is despite the fact that in Japan one can observe high performers throughout the groups.

As shown in the scatter-plot of Japan after MDS stratified by cluster-id in Fig. 5.7, we can easily identify the three well separated clouds of data-points already found for

<table>
<thead>
<tr>
<th>Country-group</th>
<th>&lt; Level 1</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4/5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan, skilled</td>
<td>0.03</td>
<td>1.36</td>
<td>14.23</td>
<td>52.16</td>
<td>32.22</td>
</tr>
<tr>
<td>Japan, unskilled</td>
<td>1.32</td>
<td>5.52</td>
<td>29.84</td>
<td>49.63</td>
<td>13.69</td>
</tr>
<tr>
<td>Japan, out</td>
<td>0.49</td>
<td>3.94</td>
<td>25.12</td>
<td>49.16</td>
<td>21.28</td>
</tr>
<tr>
<td>Denmark, skilled</td>
<td>1.34</td>
<td>6.38</td>
<td>30.86</td>
<td>48.38</td>
<td>13.04</td>
</tr>
<tr>
<td>Denmark, unskilled</td>
<td>8.35</td>
<td>21.55</td>
<td>38.33</td>
<td>27.77</td>
<td>4.00</td>
</tr>
<tr>
<td>Denmark, out</td>
<td>9.20</td>
<td>21.07</td>
<td>39.99</td>
<td>26.78</td>
<td>2.97</td>
</tr>
<tr>
<td>UK, skilled</td>
<td>0.42</td>
<td>5.36</td>
<td>26.76</td>
<td>47.72</td>
<td>19.74</td>
</tr>
<tr>
<td>UK, unskilled</td>
<td>2.99</td>
<td>19.78</td>
<td>47.26</td>
<td>26.65</td>
<td>3.31</td>
</tr>
<tr>
<td>UK, out</td>
<td>3.68</td>
<td>21.19</td>
<td>40.73</td>
<td>28.83</td>
<td>5.58</td>
</tr>
<tr>
<td>Spain, skilled</td>
<td>0.60</td>
<td>10.00</td>
<td>35.35</td>
<td>45.07</td>
<td>8.98</td>
</tr>
<tr>
<td>Spain, unskilled</td>
<td>8.66</td>
<td>26.61</td>
<td>43.97</td>
<td>19.33</td>
<td>1.43</td>
</tr>
<tr>
<td>Spain, out</td>
<td>12.25</td>
<td>22.16</td>
<td>41.67</td>
<td>21.48</td>
<td>2.44</td>
</tr>
<tr>
<td>Turkey, skilled</td>
<td>2.10</td>
<td>17.70</td>
<td>52.60</td>
<td>26.40</td>
<td>1.20</td>
</tr>
<tr>
<td>Turkey, unskilled</td>
<td>10.90</td>
<td>38.30</td>
<td>42.00</td>
<td>8.80</td>
<td>0.00</td>
</tr>
<tr>
<td>Turkey, out</td>
<td>7.80</td>
<td>36.70</td>
<td>44.10</td>
<td>11.30</td>
<td>0.10</td>
</tr>
</tbody>
</table>
### Table 5.7: Numeracy proficiency levels by working status.

<table>
<thead>
<tr>
<th>Country-group</th>
<th>&lt;Level 1</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4/5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan, skilled</td>
<td>0.10</td>
<td>1.87</td>
<td>18.71</td>
<td>50.46</td>
<td>28.86</td>
</tr>
<tr>
<td>Japan, unskilled</td>
<td>1.90</td>
<td>10.39</td>
<td>36.52</td>
<td>42.62</td>
<td>8.57</td>
</tr>
<tr>
<td>Japan, out</td>
<td>1.18</td>
<td>6.90</td>
<td>31.63</td>
<td>47.78</td>
<td>12.51</td>
</tr>
<tr>
<td>Denmark, skilled</td>
<td>1.15</td>
<td>5.27</td>
<td>25.06</td>
<td>45.42</td>
<td>23.09</td>
</tr>
<tr>
<td>Denmark, unskilled</td>
<td>6.98</td>
<td>16.01</td>
<td>39.95</td>
<td>30.07</td>
<td>6.98</td>
</tr>
<tr>
<td>Denmark, out</td>
<td>7.57</td>
<td>18.99</td>
<td>39.39</td>
<td>28.04</td>
<td>6.01</td>
</tr>
<tr>
<td>UK, skilled</td>
<td>0.81</td>
<td>9.45</td>
<td>31.25</td>
<td>42.01</td>
<td>16.48</td>
</tr>
<tr>
<td>UK, unskilled</td>
<td>7.06</td>
<td>27.42</td>
<td>43.64</td>
<td>19.27</td>
<td>2.61</td>
</tr>
<tr>
<td>UK, out</td>
<td>10.86</td>
<td>28.05</td>
<td>36.76</td>
<td>19.58</td>
<td>4.75</td>
</tr>
<tr>
<td>Spain, skilled</td>
<td>0.83</td>
<td>9.58</td>
<td>39.75</td>
<td>41.55</td>
<td>8.28</td>
</tr>
<tr>
<td>Spain, unskilled</td>
<td>12.37</td>
<td>27.40</td>
<td>43.76</td>
<td>15.40</td>
<td>1.06</td>
</tr>
<tr>
<td>Spain, out</td>
<td>16.40</td>
<td>24.49</td>
<td>41.05</td>
<td>16.61</td>
<td>1.45</td>
</tr>
<tr>
<td>Turkey, skilled</td>
<td>2.39</td>
<td>19.50</td>
<td>41.40</td>
<td>31.90</td>
<td>4.90</td>
</tr>
<tr>
<td>Turkey, unskilled</td>
<td>20.10</td>
<td>33.90</td>
<td>35.80</td>
<td>9.80</td>
<td>0.40</td>
</tr>
<tr>
<td>Turkey, out</td>
<td>19.00</td>
<td>34.50</td>
<td>34.10</td>
<td>11.60</td>
<td>0.70</td>
</tr>
</tbody>
</table>

**Figure 5.7:** The Multidimensional Scaling of the Japanese Dataset
Chapter 5. Extracting knowledge from PIAAC data

Figure 5.8: The Multidimensional Scaling of the Danish Dataset

Figure 5.9: The Multidimensional Scaling of the UK Dataset
Chapter 5. Extracting knowledge from PIAAC data

Figure 5.10: The Multidimensional Scaling of the Italian Dataset

Figure 5.11: The Multidimensional Scaling of the Spanish Dataset
Each cloud is further separated into subgroups of different colors indicating cluster’s membership; a cloud of data-points corresponds to clusters of test-takers out of the labor market, the central cloud identifies the clusters of people in the labor market not using ICT at work, and a cloud of data-points is made up by clusters of workers using ICT at work. We can observe that for Japan, Denmark, and UK the numbers of clusters is 10 while for Italy, Spain, and Turkey it is 9.

We then computed the percentage of test-takers belonging to the three clouds for each country, see Fig. 5.13. The skilled workers, groups more than 50% of test-takers in Japan, Denmark, and UK and less than 30% in Spain and Turkey.

Our analysis confirms that, at global level, Japan has the highest numbers of people performing at Level 3 or above, and the smallest numbers of low performers both in literacy and numeracy. Also, the gaps among skilled, unskilled and out of the labor market test-takers are quite small: in every group, the shares of those performing at Level 3 or above are greater than 50%. Those out of the labor market outperform unskilled workers: this can be due to the presence of students and inactive tertiary educated women in this group [28]. For Denmark, those performing at Level 1 or below, are around 20% among the unskilled workers and those out of the labor market. One of the most important factors characterizing low performers in Denmark is immigrant status. Nevertheless, a large proportion of poor performers are Danish and they can be found throughout socio-economic groups in the country [30].

The UK confirms to be a country with a large proportion of highly skilled adults (those performing at Levels 4 or 5) in literacy, [27]; they represent about the 20% of the share among skilled workers. Nevertheless there are large proportions of adults and
Chapter 5. Extracting knowledge from PIAAC data

Figure 5.13: The distribution by working status, Japan, Denmark, UK, Italy, Spain, and Turkey.

young adults with poor numeracy skills. The shares of those performing at Level 1 or below among unskilled workers are around 23% in literacy and 34% in numeracy. These shares are 25% in literacy and 39% in numeracy for those test-takers out of the labor market.

Spain is a country of which a large proportion of the population has low skills and a small share of people scoring at Level 3 or above as in the case of Italy. Around 35% of unskilled workers and inactive people scored at Level 1 or below in literacy while this share is about 10% among skilled workers. The clusters representing people out of the labor market have a slightly larger share of people at Level 3 both in literacy and in numeracy. This can be due to the presence of young adults among inactive people: in Spain 16-24 year-olds had the best performance in literacy and were outperformed in numeracy only by 25-34 year-olds, [29].

Our analysis shows that 27.6% of Turkish skilled workers scored at Level 3 or above in literacy while the shares among unskilled workers and inactive people go down to 8.80% and 11.4% respectively.

Those out of the labor market have a better performance than the unskilled workers; as for other countries, young adults belong to the group of inactive people and they scored better than other age groups in Turkey, [31].

The results of the analysis for all the 28 OECD countries are summarized in Table 5.8 and 5.9. Rankings of the average literacy/numeracy level are reported by country and group in descending order. In the second column, the ranking of the whole population is presented. The third, fourth, and fifth column display the ranking of skilled
Chapter 5. Extracting knowledge from PIAAC data

--

```
<table>
<thead>
<tr>
<th>rank</th>
<th>all</th>
<th>skilled</th>
<th>unskilled</th>
<th>out of the labor market</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
<td>Japan</td>
</tr>
<tr>
<td>3</td>
<td>Netherlands</td>
<td>Netherlands</td>
<td>Finland</td>
<td>Poland</td>
</tr>
<tr>
<td>4</td>
<td>Sweden</td>
<td>Sweden</td>
<td>Slovak Rep.</td>
<td>Korea</td>
</tr>
<tr>
<td>6</td>
<td>Russian Fed.</td>
<td>Belgium (Flanders)</td>
<td>Estonia</td>
<td>Estonia</td>
</tr>
<tr>
<td>7</td>
<td>New Zealand</td>
<td>Norway</td>
<td>Netherlands</td>
<td>Belgium (Flanders)</td>
</tr>
<tr>
<td>8</td>
<td>Belgium (Flanders)</td>
<td>UK</td>
<td>Sweden</td>
<td>Estonia</td>
</tr>
<tr>
<td>11</td>
<td>Poland</td>
<td>Germany</td>
<td>Germany</td>
<td>Austria</td>
</tr>
<tr>
<td>12</td>
<td>Germany</td>
<td>United States</td>
<td>Norway</td>
<td>Germany</td>
</tr>
<tr>
<td>13</td>
<td>Slovak Rep.</td>
<td>Poland</td>
<td>Poland</td>
<td>New Zealand</td>
</tr>
<tr>
<td>14</td>
<td>Austria</td>
<td>Czech. Rep.</td>
<td>Canada</td>
<td>Slovenia</td>
</tr>
<tr>
<td>15</td>
<td>UK</td>
<td>Korea</td>
<td>Austria</td>
<td>Norway</td>
</tr>
<tr>
<td>16</td>
<td>Korea</td>
<td>Austria</td>
<td>United States</td>
<td>Norway</td>
</tr>
<tr>
<td>17</td>
<td>United States</td>
<td>Ireland</td>
<td>Belgium (Flanders)</td>
<td>United States</td>
</tr>
<tr>
<td>18</td>
<td>Ireland</td>
<td>Canada</td>
<td>Canada</td>
<td>France</td>
</tr>
<tr>
<td>19</td>
<td>Canada</td>
<td>Russian Fed.</td>
<td>Denmark</td>
<td>UK</td>
</tr>
<tr>
<td>20</td>
<td>Denmark</td>
<td>Denmark</td>
<td>United States</td>
<td>Denmark</td>
</tr>
<tr>
<td>21</td>
<td>France</td>
<td>France</td>
<td>Greece</td>
<td>Greece</td>
</tr>
<tr>
<td>22</td>
<td>Singapore</td>
<td>Singapore</td>
<td>France</td>
<td>Singapore</td>
</tr>
<tr>
<td>23</td>
<td>Slovenia</td>
<td>Spain</td>
<td>Italy</td>
<td>Sweden</td>
</tr>
<tr>
<td>24</td>
<td>Italy</td>
<td>Italy</td>
<td>Slovenia</td>
<td>Canada</td>
</tr>
<tr>
<td>25</td>
<td>Greece</td>
<td>Israel</td>
<td>Spain</td>
<td>Turkey</td>
</tr>
<tr>
<td>26</td>
<td>Israel</td>
<td>Slovenia</td>
<td>Spain</td>
<td>Turkey</td>
</tr>
<tr>
<td>27</td>
<td>Spain</td>
<td>Greece</td>
<td>Denmark</td>
<td>Spain</td>
</tr>
<tr>
<td>28</td>
<td>Turkey</td>
<td>Turkey</td>
<td>Spain</td>
<td>Japan</td>
</tr>
</tbody>
</table>
```

*Table 5.8: Average proficiency level in Literacy by working status, 28 countries*

workers, unskilled workers and out of the labor market.

With the exception of Japan having the highest ranking in literacy and in numeracy, the position in the ranking of Denmark, UK, Spain, Italy and Turkey vary when we consider the three groups identified by our analysis.

As in Table 5.8, UK, which has a large share of skilled workers, reaches the 8th position in the skilled workers ranking while it has the 14th and 18th place when we consider unskilled test-takers or those out of the labor market respectively.

Denmark maintains its position around the 20th place in the first three literacy rankings while it goes down to the 25th position in the ranking of those out of the labor market.

Spain, Italy, and Turkey roughly maintain their places in the four rankings both for literacy and numeracy.
Table 5.9: Average proficiency level in Numeracy by working status, 28 countries

5.5 Classifying the Grosseto sample

The analysis presented in this section aimed at building a classifier based on Dataset3 which is the clustered Italian dataset labeled by the categorical variables cluster id. The objective was to classify the new instances of the Grosseto sample based on the model learned.

The analysis was performed based on two environments: Knime and Clus. Knime is an open source platform that processes complex datasets where users can create their analysis procedures in the form of a work-flow [44]. Each step in the analysis is executed by a node. A node reads a dataset as input, and after processing it makes its results available at its output port. The analysis flow is the pipeline of the analytical process which usually includes reading of the data, cleaning and filtering them and training a
model. Clus is a decision tree and rule learning system implementing the predictive clustering framework which is a method containing elements of both predictive modeling and clustering, [45]. The overall accuracy of the cross validated model by Knime proved to be higher than that of Clus. We have therefore retained the first model whose work-flow and results will be described below.

Figure 5.14 shows the work-flow of the analytical process. A File Reader node reads the clustered dataset used to feed the classifier. Then, the Column Filtering node selects the variables for the training set. Next, a node transforms variable type "cluster id" from text to number.

The Partitioning node reads the previously filtered table and randomly splits the set of samples creating a partition: one representative of the Training Set with 50% of the input dataset, and the other representative of the Test Set which consists of the remaining 50% of the input data.

The Decision Tree Learner and the Weka predictor nodes are the heart of the true classifier. The Decision Tree Learner node is trained first and then, thanks to the training, the Weka predictor is able to make predictions on the Test Set. The Scorer node contains the evaluation of the classifier which is presented in the confusion and the accuracy matrix. The confusion matrix shows how many instances have a given value in
the first column (reference column) and another value in the second column (predictions column). The values on the diagonal of the matrix show the numbers of correct classification. As shown by the values in the confusion matrix in Fig. 5.15, our classifier was able to correctly classify almost every instance of the Test Set. The accuracy matrix of the experiment is shown in Fig. 5.16; it contains, among others, the statistics of True-Positives, False-Positives, True-Negatives, False-Negatives, Recall (the proportion of actual positives that was identified correctly) as well as the overall Accuracy and the Cohen’s kappa. Accuracy is computed across all classes and when it has the value equal to 1 it means that the classifier has put every instance in its original class. The overall accuracy of our classifier was near to 1 (0.99) which implies that the classifier was able to assign correctly almost every instance of the Test Set.

![Confusion matrix](image)

**Figure 5.15: The Confusion matrix**

![Accuracy statistics](image)

**Figure 5.16: The Accuracy statistics**

The stability of the model was tested using the Knime meta node Cross Validation. This node encapsulates an inner work-flow which is executed several times on different partitions of the input dataset; the error rates of each iterations is reported on the second output port. The error rates for each of the 10 iterations performed by the Cross
Validation node are shown in Fig. 5.17. These values are the percentages of the wrong classified instances with respect to the total number of the instances of the Training Set. Being such values quite small, less than 1% for each iteration, we could conclude that the stability of the model was high.

As shown at the bottom of the work-flow in Fig. 5.14, a File Reader reads the unlabeled DatasetGrosseto1 and then the classifier generated previously is run to predict the cluster of each instance in the Grosseto sample. In Fig. 5.18, we present the results of the classification. Cluster 0 consists of "Middle aged unskilled workers" which makes 11% of the sample. Cluster 1 represents the "Mature unskilled", accounting for 7% of the sample. Cluster 2 is made of "Young skilled workers", 29% of the sample. The "Mature skilled workers" are represented in Cluster 5, 14% of the sample. "Middle aged skilled workers" are in cluster 7, which accounts for 23%, while Cluster 8 consists of "Young unskilled workers", 16% for the Grosseto sample. Cluster 3 represents the "Young adults", Cluster 4 the "Elderly people" and finally Cluster 6 the "Unemployed women": no instances were allocated to these last three clusters.

Finally, for the Grosseto sample, we computed the average scores in Literacy and Numeracy of each cluster identified, see Table 5.10. This was done to verify whether the proficiency in the local clusters was consistent with the estimation of proficiency in the Italian clusters (see Table 5.4). According to the latter estimation, clusters 2, 5 and 7 in the national analysis, which include skilled workers, had average scores above the national average. On the contrary clusters formed by unskilled workers (i.e. cluster 0, 1, and 8), or people out of the labor market (cluster 3, 4, and 6), had average scores below the Italian average. The same applies for the local sample: in Table 5.10, we observe that the identified clusters 2, 5 and 7 scored above the average of the local sample which was 273 in Numeracy, 289 in Literacy, and 287 in PSTRE. The average scores of the
clusters 0, 1 and 8 were lower than the average of the local sample.

![Figure 5.18: The classification results](image)

<table>
<thead>
<tr>
<th>Predicted cluster</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Numeracy average score</td>
<td>269</td>
<td>268</td>
<td>279</td>
<td>283</td>
<td>280</td>
<td>247</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Literacy average score</td>
<td>276</td>
<td>272</td>
<td>298</td>
<td>288</td>
<td>298</td>
<td>272</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 5.10: Numeracy and Literacy average scores by predicted cluster**

### 5.6 Discussion of the results

The results of the data-driven analytical process integrate those presented in the report “PIAAC-OCSE Rapporto Nazionale sulle Competenze degli Adulti”, as far as Italy is concerned [4], and in the "OECD Skills Outlook", as to the international context. [25].

DM techniques were successfully applied to the datasets of the large-scale survey PIAAC. Thanks to the first cluster analysis performed on the Italian dataset, test-takers were grouped in homogeneous sub-groups with common features. More specifically the most important features in the grouping were occupation, the ICT usage at work
and at home, and age. Adding the estimation of proficiency to the clustered dataset, the inspection of the centroids’ matrix confirmed the positive correlation between individual proficiency levels and skills use, occupation, age, and education. The visual analysis confirmed spatially the “digital division” of the samples between those at work, those using and those not using ICT at work. Data-points were plotted according to use of ICT at work and age. Better achievement levels can be found in clusters corresponding to test-takers declaring to use ICT at work and at home, or, if unemployed and young, having high levels of skills’ use at home.

When the macro-region variable is added to the Italian clustered dataset, it can be observed that the proportion of individuals belonging to the clusters of skilled workers is larger in the central and northern part of Italy. In the southern areas and in the islands, there is a larger share of test-takers belonging to the clusters of those who are out of the labor market. The percentage of individuals in the unskilled workers clusters is almost the same across the five macro-regions.

In the second analysis, the data-driven process was applied to 28 OECD national datasets. The clustering and MDS on these datasets revealed the same patterns found at national level. Test-takers were grouped in clusters according to occupation, the ICT usage at work and at home, and age. Countries with better average scores have also a larger proportion of test-takers belonging to the clusters of skilled workers.

We observed that Italian macro-regions and OECD countries with higher proficiency in literacy and numeracy, have also a larger number of individuals belonging to the clusters of skilled workers.

Based on the clustered Italian dataset, we built a classifier predicting the cluster. On this basis, we assigned each student in the local dataset to one of the clusters.

Since each cluster was evaluated in terms of its PIAAC performance, we were able to observe that as much as 66% of the instances in the Grosseto dataset belonged to the better performing clusters (namely those with skilled workers). The remaining 34% of takers were assigned to clusters which we have previously qualified as "unskilled", either because of their age or because they do not make use of ICT.

One of the implications is that, for any new student entering the adult education system, a cluster membership can be indicated and, associated to it, a level of proficiency. Therefore an organization in the adult education system, in case more complex assessment are difficult to undertake, could still rely on this more simple classification procedure.

What are, once more, the information required for the classification? There are questions which can be easily answered by a student upon enrollment, and more precisely in the definition of the Individual Training Pact, and would not require complex
and expensive tests such as EsOnline. These are background variables such as age, gender, age group, education level, occupation and furthermore the answers to questions such as: have you attended formal adult education in the previous 12 months? Have you ever used a computer? Do you use a computer at work? Do you use numeracy skills at work? Moreover, it would be necessary to add other variables related to the level of use of basic skills at home and at work. These would allow us to provide the full range of variables included in the classifier.
Chapter 6

Conclusions

6.1 Results from the case study

Research question n. 1 was elaborated in response to the first research objective of benchmarking local samples with national and international ones. We investigated whether local samples scored higher than the national and international ones, and the question was mostly answered positively. Mean scores were statistically significantly higher in literacy and numeracy. When considering only those who opted for the computer version of the test, the local samples scored significantly higher in literacy while in numeracy and PSTRE no significant differences were found. The distributions of the achievement levels in literacy highlighted a similar pattern, with the test taker from Grosseto scoring better than the national and international samples. For instance in literacy only 28\% of the test takers scored lower than level 2 in 2016 while, with reference to the year 2012, in Italy the percentage was 70\% and in the OECD countries 54\%. Similar results apply to numeracy and PSTRE.

Among the factors underlying the good performances despite the low education level of the students, we identified the possession of ICT basic skills. Age was also relevant: in the case of Grosseto both samples were made of relatively young adults, aged on average 37. Being employed also had a positive relation with proficiency. Finally, the local samples were made of adult students who have been able to maintain their skills over time participating in formal education.

Research question 2 addressed the second research objective, and was related to the change of core skills of adult students after attending a full school year. However, due to the small dimension of the sample and the different settings in which the test was conducted, the results were not conclusive and the question remains unanswered.

Research question 3, always in relation to the second research objective on skills, analyzed differences in the mean scores when considering the learning period at school. The analysis revealed that, on average, the students belonging to the second learning period scored significantly higher than those of the first one: i.e. mean scores of 300 vs.
275 in literacy, 279 vs. 255 in numeracy and 296 vs. 274 in PSTRE. On the contrary we did not observe an improvement when comparing scores of the second and third training periods, although we were not able to draw any statistically significant conclusions in this respect.

However, an even more important aspect is that results from EsOnline prove to be coherent with the placement performed by the schools to the different learning periods.

A third research objective was formulated on whether EsOnline can be effectively used for defining the Individual Training Pact. This is proven by the fact that access to the second educational period was granted by ISIS Bianciardi to students having the necessary school qualifications for the first learning period and, at the same time, sufficient proficiency levels in literacy, numeracy, and in PSTRE as certified by EsOnline. Further evidence is provided by the fact that the test was compulsory for students at their enrollment and that the results obtained were integrated in their curricular information.

The final research objective was to identify the conditions for further adoption of EsOnline by the organizations composing the partnership. A number of drivers behind adoption have been tested.

Scalability to large numbers of takers, authenticity, efficiency, and accuracy were among the factors identified. Also the ability to deliver real-world tasks and automatic grading emerged as important factors as well as the compliance with ETS’s standards for quality and fairness. Comparability of results across test-takers, courses, schools and their benchmarking against national and international tests are also important drivers for adoption. From the viewpoint of the test takers easiness of use emerged as an important element favoring adoption. 97% of the sample completed the core questionnaire of literacy and numeracy and 91% completed the test PSTRE, in 2017.

There are, on the one hand, certain drawbacks which appeared when using the test, ranging from limitations of the tool (e.g. the fact that EsOnline is delivered solely for the Mozilla Firefox browser and has limits in its user interface), to the fact that it can be conducive to cheating due to the limited number of testlets.

On the other hand the are a number of reasons why adopting EsOnline would be advantageous. For instance, as ISIS Bianciardi did, it could be used as a tool for providing evidences during the recognition of prior learning in newly enrolled students. Or applied by CPIAs and adult secondary schools at the end of the compulsory education cycle to certify core skills.

Costs involved and constraints in procurement by the public administration were also assessed as part of the analysis of adoption factors.
Chapter 6. Conclusions

6.2 Results from the data mining exercises

Results from this research integrate what was reported for Italy in the “PIAAC-OCSE Rapporto Nazionale sulle Competenze degli Adulti”, [4] and internationally in "OECD Skills Outlook 2013: First Results from the Survey of Adult Skills", [25]. The novelty of this research is also in the application of DM techniques and in the visual representations of the patterns obtained through clustering and multidimensional scaling techniques.

The first, cluster analysis performed on the datasets of the large-scale survey PIAAC, resulted in homogeneous sub-groups of test takers with common features. More specifically the determinants found were occupation, the usage of ICT at work and at home, and age. When adding proficiency to the clustered dataset, the positive relation between individual proficiency levels and skills use, occupation, age, and education, is confirmed. The visual analysis supports spatially a “digital division” of the samples between those at work, those using ICT at work and those that do not. Thanks to this visualization, we made it possible to portray the information 20 variables in one bi-dimensional plot.

Better achievement levels correspond to test-takers declaring that they use ICT at work and at home, or, in case of unemployed and young, having high levels of skills’ use at home.

When the macro-region variable is added to the Italian clustered dataset, it can be observed that the proportion of individuals belonging to the clusters of skilled workers is larger in the central and northern part of Italy. In the southern areas and in the islands, there is a larger share of test-takers belonging to the clusters of those who are out of the labor market. The percentage of individuals in the unskilled workers clusters is almost the same across the five macro-regions.

In the second analysis, the data-driven process was applied to 28 OECD national datasets. The clustering and MDS on these datasets revealed the same patterns found at national level. Test-takers were grouped in clusters according to occupation, the ICT usage at work and at home, and age. Countries with better average scores, have also a larger proportion of test-takers belonging to the clusters of skilled workers.

The clustered Italian dataset made it possible to build a classifier and each student in the Grosseto case was assigned to a cluster. Most of the test takers were assigned to the better performing clusters (the skilled workers) which proves the validity of the first analysis conducted on the local dataset.

While acknowledging the importance of introducing the EsOnline test in all Italian schools, we also reckon that this perspective encounters for now practical difficulties,
namely related to its high cost. Our classification procedure, without having the ambition of replacing EsOnline, could offer an effective and immediate support to the Individual Training Pact, identifying fundamental skills in a rather simple way. A precondition, however, is that an appropriate data collection mechanism is in place in order to gather the full range of variables needed for the classifier. With this objective in mind we planned to develop a survey with a dedicated questionnaire for the school involved in the Grosseto case study.

As a concluding remark we would like to acknowledge the fact that, despite our availability, no teachers participated in the case study. This, unfortunately, did not allow us to appreciate the eventual positive effects of the test on the teaching practices.

Although it was not possible to meet the students involved in the case study, they received a link to a short video on the results and a link to an evaluation survey. More details are given in Appendix A.

As a next step we would like to present the results of the case studies to students and teachers involved. To increase awareness especially on the side of teachers, we will also introduce teachers to EsOnline through the demo test, now available in Italian on the OECD portal.

6.2.1 Limitations and possible developments of this study

As to the limitations of this study, we can highlight the following ones. As mentioned in Section 5.1 PIAAC and EsOnline datasets are not fully comparable, i.e. PIAAC has a much richer dataset. This did not allow the use of the full set of PIAAC variables which in principle, could have helped in predicting the level of performances with regards to basic skills. Another limitation of our work is that it was not possible investigate the relation between skills performance, as defined and measured by PIAAC and skill competences as established by the Italian Ministry of Education. This is becoming an even more interesting perspective as the same Ministry, through the Istituto Nazionale per la Valutazione del Sistema Educativo di Istruzione e di Formazione (INVALSI), is currently introducing computer based assessment tools for evaluating competences at the different levels of the school system.

In Section 6.1 we tried to indicate possible positive developments of this research for the Italian adult-education system. This applies for instance to the recognition of prior learning skills in newly enrolled students, or when carried-out at the end of the education cycle in order to certify basic skills.

As to further possible advancements in the Educational Data Mining, Learning Analytics and Academic Analytics field, a major takeaway from the case studies relates to

1http://www.oecd.org/skills/ESonline-assessment/takethetest/?
the OECD surveys (PIAAC and PISA). Datasets generated by the named surveys are extremely important for analyzing and comparing skills worldwide. Nevertheless, they are totally ignored by the communities of learning analytics, educational data mining and academic analytics. It appears that only economists and statisticians are aware of them. On the contrary, they do not seem to apply any data mining technique in this specific domain. Bringing together the communities would open the doors to the investigation of a very rich data base with many skills related variables (PIAAC alone has over 1,300 variables and over 200,000 persons covered). A concrete opportunity towards the general understanding, and possibly the prediction of basic skills.
Appendices
Appendix A

EsOnline evaluation survey

The survey was meant to collect the opinions of the students about their experience with EsOnline; it was held in June 2018 and 29 students participated. The answers were collected using the Likert scales with a five point scale, where the middle option was given to allow a neutral opinion. The survey was designed in collaboration with Dott. A. Del Cimuto of the Statistics Unit of the National Agency on Labor Market policies.

Through the first question, "According to your experience, the PIAAC Online Test was", we wanted to analyze whether the students found difficult to answer the test. As shown in row 1 of Table A.1, 11% of the sample considered the test very difficult/difficult, about 38% did not find the test either difficult or easy, while more than 50% found the test not difficult/not difficult at all.

The second question, "According to your experience, understanding how to answer the test questions was", was meant to understand whether students found intuitive to respond to EsOnline. About 38% of the sample considered the interaction very intuitive/intuitive, 37.9% expressed a neutral opinion while about 24% of the sample considered the test not very intuitive.

In the third row of the table, students expressed their opinion on the duration of the test; about 24% of the sample found the duration of the test very long/long, 37.9% expressed a neutral opinion while 27.5% of the sample did not think that the test was too long.

The fourth row shows the opinions given on the following statement "The test was useful for understanding your strengths and/or weaknesses". The majority of the sample, about 65%, considered the test very useful/useful for self evaluation, 13.8% expressed a neutral opinion and 20.7% said that it was not useful. Nobody considered the test totally not useful to evaluate personal strengths/weaknesses.

"The score report was easy to understand" for about 90% of the sample as reported in the fifth row of the table. The operators’ support during the test was considered not useful/not useful at all, by about 11% of the sample, 27.6% of the students found the
support of the operators very useful, 41.4% reported that the support was useful and 20.75% expressed a neutral opinion.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>very difficult</td>
<td>3.4%</td>
<td>6.9%</td>
<td>37.9%</td>
<td>44.8%</td>
</tr>
<tr>
<td>2</td>
<td>very intuitive</td>
<td>6.9%</td>
<td>31%</td>
<td>37.9%</td>
<td>17.2%</td>
</tr>
<tr>
<td>3</td>
<td>very long</td>
<td>6.9%</td>
<td>27.6%</td>
<td>37.9%</td>
<td>24.1%</td>
</tr>
<tr>
<td>4</td>
<td>very useful</td>
<td>13.8%</td>
<td>51.7%</td>
<td>13.8%</td>
<td>20.7%</td>
</tr>
<tr>
<td>5</td>
<td>very easy</td>
<td>31%</td>
<td>37.9%</td>
<td>20.7%</td>
<td>6.9%</td>
</tr>
<tr>
<td>6</td>
<td>very useful</td>
<td>27.6%</td>
<td>41.4%</td>
<td>20.75%</td>
<td>3.4%</td>
</tr>
</tbody>
</table>

Table A.1: The results of the satisfaction survey
Appendix B

Education and Skills Online Score Reports
Appendix B: Sample Score Reports
Literacy

Your Results:

What do the Literacy Questions Measure?
The Literacy questions measure how well you understand and use information found in materials such as newspapers, brochures, manuals or websites. They also measure how well you find and use information in forms, schedules, charts or tables.

Most adults use their literacy skills to answer questions, to learn how to do something, or for entertainment. For example, you are using literacy skills when you:

- Look for a job online
- Learn about quitting smoking from a brochure at your health clinic
- Take aspirin as directed on the package
- Find out when you must wear safety glasses at work from a health and safety manual
- Use a bus schedule to figure out the time of your next bus
- Read and complete a form giving permission for your child to go on a school field trip
- Use an inventory sheet to check warehouse stock at work

Adults use literacy skills at work, at home, and in the community.

What does my Literacy score mean? [text displays for Below Level 1]

Based on your score, there are some everyday literacy tasks that you can likely do very well. Adults with scores similar to yours are typically able to locate specific information from a text with a few sentences or paragraphs about familiar topics. For example, they are likely able to:

- Locate a phone number or address of a store from a newspaper advertisement
- Locate the date and time of a community art show from a flyer
- Identify the winner of an employee contest from a company announcement
- Identify key ingredients from a food package label

While you have demonstrated some basic literacy skills you need to handle the demands of everyday life, you might sometimes have trouble using your literacy skills to understand longer unfamiliar texts or to complete a form. For example, some adults with scores similar to yours might find it challenging to:

- Complete a short form to order a magazine subscription
- Submit a vote for or against a new workplace dress code on an employer’s Web page
- Locate the link on a theater’s website that would be used to find information about the theater
- Use a table in a newspaper article to identify the top three companies with the most employees
- Name two reasons stated in a newspaper article for an increase in local food prices
What does my Literacy score mean? [text displays for Level 1]

Based on your score, there are some everyday literacy tasks that you can likely do very well. Adults with scores similar to yours are typically able to understand longer texts about familiar topics. For example, they are likely able to:

- Identify key ingredients from a food package label
- Complete a short form to order a magazine subscription
- Submit a vote for or against a new workplace dress code on an employer’s Web page
- Locate the link on a theater’s website that would be used to find information about the theater
- Use a table in a newspaper article to identify the three companies with the most employees

While you have demonstrated some literacy skills you need to handle the demands of everyday life, you might sometimes have trouble using your literacy skills to understand longer and more complicated texts. For example, some adults with scores similar to yours might find it challenging to:

- Determine what forms are needed to return a damaged telephone according to instructions in the warranty brochure
- Identify information in a camera store’s single Web page that explains how this year’s photo contest rules differ from those in previous years
- Name two reasons stated in an employee newsletter for an increase in company sales
- Use a music store’s Web page to compare and contrast several reviews to determine which song to download based on the price and the type of music you like

What does my Literacy score mean? [text displays for Level 2]

Based on your score, there are some everyday literacy tasks that you can likely do very well. Adults with scores similar to yours are typically able to understand longer and more complicated texts about unfamiliar topics. For example, they are likely able to:

- Submit a vote for or against a new workplace dress code on an employer’s Web page
- Determine what forms are needed to return a damaged telephone according to instructions in the warranty brokerage
- Identify information in a camera store’s single Web page that explains how this year’s photo contest rules differ from those in previous years
- Name two reasons stated in an employee newsletter for an increase in company sales

While you have demonstrated the literacy skills you need to cope with the demands of everyday life, you might sometimes experience frustration if you need to use your literacy skills to understand longer and more complicated digital and printed texts with a variety of text features. For example, some adults with scores similar to yours might find it challenging to:

- Find out whether a utility company accepts the same type of payment if paid by mail or online using information from a monthly billing statement
- Use a music store’s Web page to compare and contrast several reviews to determine which song to download based on the price and the type of music you like
- Search several Web pages of a national health organization for evidence supporting the claim that exercise can lead to greater work productivity
- Determine which parents in a childcare discussion forum share a similar viewpoint by comparing their comments
What does my Literacy score mean? [text displays for Level 3]
Based on your score, there are some everyday literacy tasks that you can likely do very well. Adults with scores similar to yours are typically able to understand longer and more complicated digital and print texts with a variety of text features. For example, they are likely able to:

- Name two reasons stated in an employee newsletter for an increase in company sales
- Find out whether a utility company accepts the same type of payment if paid by mail or online using information from a monthly billing statement
- Use a music store’s Web page to compare and contrast several reviews to determine which song to download based on the price and the type of music you like
- Search several Web pages of a national health organization for evidence supporting the claim that exercise can lead to greater work productivity

While you have demonstrated the literacy skills you need to cope with the demands of everyday life, you might sometimes experience frustration if you need to use your literacy skills to understand longer and more complicated texts from a number of different sources. For example, some adults with scores similar to yours might find it challenging to:

- Use online search results for books on alternative energy to identify a book that includes arguments both for and against solar energy
- Evaluate posts in a discussion forum on health remedies by comparing the information they provide against that in a website from a well-known medical center
- Use several links in a city’s transportation Web page to locate information about special fares or services on holidays
- From a list of workplace safety suggestions, determine which a company will be likely to adopt based on a complex chart showing the company’s existing policies and procedures

What does my Literacy score mean? [text displays for Level 4/5]
Based on your score, you can likely do most everyday literacy tasks very well. You are able to use your literacy skills to understand longer and more complicated texts from a number of different sources. Adults with scores similar to yours are typically able to do things like:

- Find out whether a utility company accepts the same type of payment if paid by mail or online using information from a monthly billing statement
- Use a music store’s Web page to compare and contrast several reviews to determine which song to download based on the price and the type of music you like
- Search several Web pages of a national health organization for evidence supporting the claim that exercise can lead to greater work productivity
- Evaluate posts in a discussion forum on health remedies by comparing the information they provide against that in a website from a well-known medical center
- Use several links in a city’s transportation Web page to locate information about special fares or services on holidays
- Determine which claims in a newspaper article about the benefits of sleep are supported by information and graphs in two long research articles
How does my Literacy score compare?
The graphs below show how your Literacy test score compares with the average test scores of adults (ages 16-65) in your country and internationally (OECD PIAAC average) by their education level, occupation and age group. As you look at the graphs, remember that you use other types of skills besides literacy in school or on the job. Therefore, you may find that people with scores similar to your own have different levels of education, work in different occupations, and are in different age groups than you.

Some Facts
In general, adults with higher skill levels are more likely than adults with low skill levels to:
- Be employed regularly
- Earn a higher wage
- Have opportunities for job mobility
- Be more efficient in the workplace
- Have greater opportunities for furthering their education and training
Numeracy

What do the Numeracy Questions Measure?
The numeracy questions measure how well you are able to interpret, communicate, or use mathematical information to solve a problem or understand a situation. You may find mathematical information in materials such as tables, graphs, maps, product labels or advertisements.

Most adults use numeracy skills to answer a question or make a decision in a variety of everyday situations. For example, you are using numeracy skills when you:

▲ Estimate the distance between two cities on a map
▲ Follow a recipe when cooking
▲ Manage your personal finances
▲ Interpret a weather report
▲ Keep a timesheet of hours worked
▲ Keep track of the inventory at a store

Adults use numeracy skills at work, at home, and in the community.

What does my Numeracy score mean? [text displays for Below Level 1]
Based on your score, there are some numeracy tasks that you can likely do very well. Adults with scores similar to yours are typically able to do simple arithmetic in familiar situations. For example, they are likely able to:

▲ Figure out how much money it will cost to buy a few common items in the grocery store
▲ Identify the amount that corresponds to an unlabeled mark on a measuring cup
▲ Find the range in daily temperatures by subtracting the lowest from the highest temperature

While you have demonstrated the numeracy skills you need to handle certain demands of everyday life, you may have trouble using numeracy skills that require computing with percents and decimal numbers, or understanding mathematical information in a table. For example, some adults with scores similar to yours might find it challenging to:

▲ Figure out the price of a shirt that will be discounted by 25 percent
▲ Determine the price of a single bottle of water when given the cost of an entire case of bottles
▲ Look at a weekly timesheet to find out which employee worked the most hours in a single day
What does my Numeracy score mean? [text displays for Level 1]
Based on your score, there are some numeracy tasks that you can likely do very well. Adults with scores similar to yours are typically able to compute with percents and decimal numbers, or understand mathematical information in a table. For example, they are likely able to:

- Identify the amount that corresponds to an unlabeled mark on a measuring cup
- Find the range in daily temperatures by subtracting the lowest from the highest temperature
- Figure out the price of a shirt that will be discounted by 25 percent
- Determine the price of a single bottle of water when given the cost of an entire case of bottles

While you have demonstrated the numeracy skills you need to handle certain demands of everyday life, you may have trouble using numeracy skills that require performing an intermediate computation before being able to answer a question, having to interpret a graph, or using ratios. For example, some adults with scores similar to yours might find it challenging to:

- Determine how many months in a year had sales above the mean sales for the year from a table of monthly sales
- Identify which predicted monthly gasoline price was most accurate based on line graphs of predicted and actual gasoline prices for a year
- Determine the amount of concentrated lemonade flavoring and water needed to make a large container of lemonade that is in the same ratio of flavoring to water as a smaller amount of lemonade

What does my Numeracy score mean? [text displays for Level 2]
Based on your score, there are some numeracy tasks that you can likely do very well. Adults with scores similar to yours are typically able to perform an intermediate computation before being able to answer a question, understand mathematical information in a table, or interpret a simple graph. For example, they are likely able to:

- Figure out the price of a shirt that will be discounted by 25 percent
- Determine the price of a single bottle of water when given the cost of an entire case of bottles
- Determine how many months in a year had sales above the mean sales for the year from a table of monthly sales
- Identify which predicted monthly gasoline price was most accurate based on line graphs of predicted and actual gasoline prices for a year

While you have demonstrated the numeracy skills you need to handle certain demands of everyday life, you may have trouble using numeracy skills that require using ratios, reading a complex graph, or comparing changes in percentages. For example, some adults with scores similar to yours might find it challenging to:

- Determine the amount of concentrated lemonade flavoring and water needed to make a large container of lemonade that is in the same ratio of flavoring to water as a smaller amount of lemonade
- Read a complex graph, comparing the amount of salt, sugar, and fat in a typical diet for men and a typical diet for women, to determine the amount of sugar consumed by men
- Convert the number of students enrolled in a university each year into percentages, and then compute the change in the percentage of students enrolled each year
What does my Numeracy score mean? [text displays for Level 3]
Based on your score, there are some numeracy tasks that you can likely do very well. Adults with scores similar to yours are typically able to use ratios, understand mathematical information in a table, or read a complex graph. For example, they are likely able to:

- Determine the price of a single bottle of water when given the cost of an entire case of bottles
- Determine how many months in a year had sales above the mean sales for the year from a table of monthly sales
- Identify which predicted monthly gasoline price was most accurate based on line graphs of predicted and actual gasoline prices for a year
- Determine the amount of concentrated lemonade flavoring and water needed to make a large container of lemonade that is in the same ratio of flavoring to water as a smaller amount of lemonade
- Read a complex graph, comparing the amount of salt, sugar, and fat in a typical diet for men and a typical diet for women, to determine the amount of sugar consumed by men

While you have demonstrated the numeracy skills you need to handle certain demands of everyday life, you may have trouble using numeracy skills that require using percentages, using rates, or understanding how quantities are related. For example, some adults with scores similar to yours might find it challenging to:

- Convert the number of students enrolled in a university each year into percentages, and then compute the change in the percentage of students enrolled each year
- Determine how much medicine to give to a child when the dosage is based on the child’s body weight
- Calculate profit from a table containing lists of income and expense sources

What does my Numeracy score mean? [text displays for Level 4/5]
Based on your score, you are generally able to handle most of the mathematical demands of everyday tasks. Adults with scores similar to yours are typically able to use percentages and rates, interpret information presented in various ways, or understand how quantities are related. For example, they are likely able to:

- Identify which predicted monthly gasoline price was most accurate based on line graphs of predicted and actual gasoline prices for a year
- Determine the amount of concentrated lemonade flavoring and water needed to make a large container of lemonade that is in the same ratio of flavoring to water as a smaller amount of lemonade
- Convert the number of students enrolled in a university each year into percentages, and then compute the change in the percentage of students enrolled each year
- Read a complex graph, comparing the amount of salt, sugar, and fat in a typical diet for men and a typical diet for women, to determine the amount of sugar consumed by men
- Determine how much medicine to give to a child when the dosage is based on the child’s body weight
- Calculate profit from a table containing lists of income and expense sources
How does my Numeracy score compare?
The graphs below show how your numeracy test score compares with the average test scores of adults (ages 16-65) in your country and internationally (OECD PIAAC average) by their education level, occupation, and age group. As you look at the graphs, remember that you use other types of skills besides numeracy in school or on the job. Therefore, you may find that people who have scores similar to your own score have different levels of education, work in different occupations, and are in different age groups than you.

National Comparison
International Comparison

Some Facts
In general, adults with higher skill levels are more likely than adults with low skill levels to:
- Be employed regularly
- Earn a higher wage
- Have opportunities for job mobility
- Be more efficient in the workplace
- Have greater opportunities for furthering their education and training
What are Reading Components?

Reading Components are the skills that work together to help you understand what you read. There are many different reading components. In this test, you took 3 of them: vocabulary, sentence comprehension, and passage comprehension. Each of these components represents an important skill that supports the ability to read well.

- **Vocabulary**: The more words you know, the better you can understand what you read. The vocabulary exercise you completed measures your ability to understand the kinds of words that you will find in a wide range of printed materials in your everyday life.

- **Sentence Comprehension**: The sentences we read can be very short, or they can go on for a long time, with many phrases and clauses. The sentence comprehension exercise you completed measures your ability to understand sentences of different lengths and difficulty levels.

- **Passage Comprehension**: The reason we read is to understand news articles, e-mails, books, and so on. The passage comprehension exercise you completed measures your ability to understand a variety of types of reading materials like the ones you might encounter in your daily life.
How do my Reading Components scores compare?

Below you will see your results on the 3 reading components as compared to people who have Literacy scores that are like yours. You will see your results in terms of **accuracy** and **rate**.

- **Accuracy** means how many you got right.
- **Rate** means how quickly you could do the exercises, whether you got the answer right or not.

You will see graphs that show your accuracy and rate results for each component. Your results will be in one of three groups: Low, Medium, or High. The closer your result is to the High group, the better you are likely to be.

If your result is:

- **Low**: Continue to work on that skill to improve it.
- **Medium**: Your performance on the skill is adequate, but you need to improve it.
- **High or Fast**: You likely have the skill in place.

Your results on each reading component fall into one of these patterns:

<table>
<thead>
<tr>
<th>Results</th>
<th>What Should You Do?</th>
</tr>
</thead>
<tbody>
<tr>
<td>High accuracy and fast rate</td>
<td>You have good basic reading skills and can focus on building your comprehension skills.</td>
</tr>
<tr>
<td>High accuracy and low or medium rate</td>
<td>You have good basic reading skills; you can work on building your comprehension skills AND increasing your rate.</td>
</tr>
<tr>
<td>Low or medium accuracy and fast rate</td>
<td>You might be trying to go too fast. You need to build your basic skills.</td>
</tr>
<tr>
<td>Low or medium accuracy and low or medium rate</td>
<td>You should work on building your basic skills AND getting faster.</td>
</tr>
</tbody>
</table>
Your Reading Components Results:
Remember: if your results fall into the Low or Medium categories, you should continue building the skill so you can get MORE ACCURATE and FASTER.
Passage Comprehension Results

Accuracy

Low  Medium  High

More Accurate ➔

Rate

Low  Medium  High

Faster ➔

Your Accuracy

Your Rate
Problem Solving in Technology-Rich Environments

Education & Skills Online: Problem Solving in Technology-Rich Environments

Date: 5/22/14   Authorization Code: ####   Location: Institution

Your Results:

What do the Problem Solving in Technology-Rich Environments Questions Measure?

The Problem Solving in Technology-Rich Environments questions measure how well you use different types of technology to solve everyday problems and complete tasks to successfully meet your goals. They also measure how well you understand and use information in different environments, such as e-mail, Web pages, or spreadsheets. In this test, a problem is any situation where you don’t already have a good idea about how to achieve a goal. This may be because the strategy to use is not obvious to you or because you have never tried such a task in the past. As you have more practice in meeting different goals using technology, those tasks that were once problems will become automatic and routine for you.

Most adults use problem-solving in technology-rich environments skills to find information or answer questions, use online tools and functions that can make tasks easier, and communicate with others. For example, you are using these skills when you:

- Read and answer emails from friends or co-workers
- Search for a website with information about treatment for a medical issue
- Use a spreadsheet to set up a budget and keep track of spending
- Help a friend figure out how to install a new software program
- Set up folders on your computer to organize your emails or files
- Evaluate whether information on a Web page comes from a reliable source

Adults use problem solving skills for various technology-related tasks at work, at home, and in the community.

What does my Problem Solving in Technology-Rich Environments score mean? [text displays for Below Level 1]

Based on your score, there are some everyday problem-solving tasks that you can likely do very well using technology. Adults with scores similar to yours are typically able to complete tasks that are quite routine for them using familiar technology programs. For example, they are likely able to:

- Use a familiar email program to open and read emails
- Write a short summary of a club meeting using a word processing program they know well
- Enter the name of a local store into a search engine they have used in the past to find the store’s phone number

While you have demonstrated basic skills you need to handle the demands of everyday tasks such as these, you might sometimes have trouble using technology to solve more complex problems. For example, some adults with scores similar to yours might find it challenging to:

- Open and read email using an unfamiliar email program similar to one they regularly use
What does my Problem Solving in Technology-Rich Environments score mean? [text displays for Level 1]
Based on your score, there are some everyday problem-solving tasks that you can likely do very well using technology. Adults with scores similar to yours are typically able to use unfamiliar software programs that work like ones they have used in the past to solve problems where the goal is clear and a limited number of steps are required. For example, they are likely able to:

- Open, read, and respond to email using an unfamiliar email program
- Locate specific information on the homepage of a website that a friend has recommended
- Set up a system of folders that allow files or emails to be organized and easily retrieved

While you have demonstrated basic skills you need to handle with the demands of everyday tasks such as these, you might sometimes have trouble using technology to solve more complex problems. For example, some adults with scores similar to yours might find it challenging to:

- Figure out how to send an email message to a number of contacts using an unfamiliar bulk email function
- Use a sorting tool to make it easier to locate sales numbers for a specific product in a company spreadsheet
- Conduct a web search to find out how to solve a problem with other software, such as how to view a column that won’t display properly in a spreadsheet
- Find an email message or file that has been “lost” somewhere on a computer hard drive

What does my Problem Solving in Technology-Rich Environments score mean? [text displays for Level 2]
Based on your score, there are some everyday problem-solving tasks that you can likely do very well using technology. Adults with scores similar to yours are typically able to use software they have never seen before to solve more complex problems, even when unexpected impasses/outcomes occur. For example, they are likely able to:

- Figure out how to send an email message to a number of contacts using an unfamiliar bulk email function
- Use a sorting tool to make it easier to locate sales numbers for a specific product in a company spreadsheet
- Conduct a web search to find out how to solve a problem with other software, such as how to view a column that won’t display properly in a spreadsheet
- Find an email message or file that has been “lost” somewhere on a computer hard drive

While you have demonstrated basic skills you need to handle with the demands of everyday tasks such as these, you might sometimes have trouble using technology to solve more complex problems. For example, some adults with scores similar to yours might find it challenging to:

- Establish criteria for narrowing a Web search, documenting results using a spreadsheet, and communicating the results to others through email
- Evaluate a number of Web search results to determine which has the most relevant and reliable information. Part of this process includes evaluating and refining a search to determine if additional or different types of websites should be considered
- Use a software program that they have never seen before with limited or unclear directions based on general experience with technology or by consulting other online resources including
What does my Problem Solving in Technology-Rich Environments score mean? [text displays for Level 3]

Based on your score, there are many everyday problem-solving tasks that you can likely do very well using technology. Adults with scores similar to yours are typically able to use one or more complex software programs to solve ill-defined problems with multiple goals. For example, they are likely able to:

- Conduct a web search to find out how to solve a problem with other software, such as how to view a column that won’t display properly in a spreadsheet
- Figure out how to send an email message to a number of contacts using an unfamiliar bulk email function
- Evaluate a number of web search results to determine which has the most relevant and reliable information. Part of this process includes evaluating and refining a search to determine if additional or different types of websites should be considered
- Use a software program that they have never seen before with limited or unclear direction. Success may be based on a user’s general experience with technology or information may be gathered by consulting other online resources including websites or user blogs
- Select from among a number of choices the best software to use for a particular task
How does my Problem Solving in Technology-Rich Environments score compare?

The graphs below show how your Problem Solving in Technology-Rich Environments test score compares with the average test scores of adults (ages 16-65) in your country and internationally (OECD PIAAC average) by their education level, occupation and age group. As you look at the graphs, remember that you use other types of skills besides problem solving in school or on the job. Therefore, you may find that people with scores similar to your own have different levels of education, work in different occupations, and are in different age groups than you.

### National Comparison

**Problem Solving Score by Education Level**

- **Tertiary**
- **Upper secondary**
- **Below upper secondary**

### International Comparison

**Problem Solving Score by Education Level**

- **Tertiary**
- **Upper secondary**
- **Below upper secondary**

### Problem Solving Score by Occupation

- **Skilled**
- **Semi- Skilled white collar**
- **Semi- Skilled blue collar**
- **Elementary**

### Problem Solving Score by Age Group

- 16-24
- 25-34
- 35-44
- 45-54
- 55-65

Some Facts

In general, adults with higher skill levels are more likely than adults with low skill levels to:

- Be employed regularly
- Earn a higher wage
- Have opportunities for job mobility
- Be more efficient in the workplace
- Have greater opportunities for furthering their education and training
### Skill Use

**Education & Skills Online: Skill Use**

| Date: 5/22/14 | Authorization Code: #### | Location: Institution |

**What do the Skill Use questions measure?**

The questions in this section focus on skills associated with reading, writing, numeracy, and information and communication technology (ICT). They were designed to collect information about how often you use these skills as well as the variety of your activities in each of these areas:

- **Reading**: Reading documents (directions, instructions, letters, memos, emails, articles, books, manuals, bills, invoices, diagrams, maps)
- **Writing**: Writing documents (letters, memos, emails, articles, reports, forms)
- **Numeracy**: Calculating prices, costs or budgets; using fractions, decimals or percentages; using calculators; preparing graphs or tables; algebra or formulas; using advanced math or statistics
- **ICT skills**: Using email, Internet, spreadsheets, word processors, programming languages; conducting transactions online; participating in online discussions (conferences, chats).

These activities are important for building and maintaining skills in literacy, numeracy and problem solving in technology environments. Practicing skills in a range of environments has many benefits. By practicing your skills and expanding your experiences, you are likely to continue to improve your skills and be able to use them in new situations.
How do my Skill Use scores compare?

Below are your results on the 4 skill use areas compared to other people internationally.¹ Your results are reported separately in terms of your skill use at **home** and at **work**.

Your results will be shown as a shaded section in one of four groups: N/A, Low, Moderate, or High. The closer your result is to the High group, the more often you are likely to use skills in a greater variety of activities.

If your result is:
- **N/A**: You reported that you never engaged in any of the activities involving this skill.
- **Low**: You reported that you rarely engaged in most of the activities involving this skill.
- **Moderate**: You reported that your engagement in activities varied in terms of how many activities you did and how often you did them.
- **High**: You reported that you engaged in most activities most days or every day.

### Reading Skill Use

<table>
<thead>
<tr>
<th></th>
<th>At Home</th>
<th>At Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Low</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

### Writing Skill Use

<table>
<thead>
<tr>
<th></th>
<th>At Home</th>
<th>At Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Low</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

### Numeracy Skill Use

<table>
<thead>
<tr>
<th></th>
<th>At Home</th>
<th>At Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Low</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

### ICT Skill Use

<table>
<thead>
<tr>
<th></th>
<th>At Home</th>
<th>At Work</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>Low</td>
<td>Moderate</td>
</tr>
</tbody>
</table>

**Some Facts**

In general, adults who use a variety information processing skills both at home and at work are likely to:

- Continue developing and maintaining proficiency in literacy, numeracy, and problem solving in technology environments
- Be employed regularly
- Earn a higher wage
- Have opportunities for job mobility
- Be more efficient and productive in the workplace
- Have greater opportunities for furthering their education and training

---

¹ Based on responses from participants from 24 countries in the OECD Programme for the International Assessment of Adult Competencies (PIAAC).
Here are your Career Interest and Intentionality Results!

The Career Interest and Intentionality module measures your preferences for different types of work activities and environments, how well your interests match your current or intended job and the level of your intention to seek out new job opportunities and career- and job-related training.

The more your intended career and job-related training match your career interests, the greater your career fit. The greater your career fit, the more likely your job will be satisfying and rewarding to you.

**CAREER INTEREST**

Identifying your interests, or the work you like to do, can help you find careers you might enjoy. The more a career meets your area of interests, the more likely it will be satisfying and rewarding to you.

Below are your career interest results. Your score for each area of interest can tell you more about what you like to do. You will find the areas with higher scores more interesting to you than areas with lower scores. Career categories are described in the table following your results.

<table>
<thead>
<tr>
<th>Areas of Career Interest</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Realistic (R)</td>
<td>5</td>
</tr>
<tr>
<td>Investigative (I)</td>
<td>10</td>
</tr>
<tr>
<td>Artistic (A)</td>
<td>15</td>
</tr>
<tr>
<td>Social (S)</td>
<td>20</td>
</tr>
<tr>
<td>Enterprising (E)</td>
<td>25</td>
</tr>
<tr>
<td>Conventional (C)</td>
<td>30</td>
</tr>
</tbody>
</table>

This Product uses the O*NET™ Interest Profiler Short Form. U.S. Department of Labor, Employment and Training Administration is the exclusive owner of all rights under U.S. copyright laws and international treaty provisions in the O*NET™ Interest Profiler Short Form. Any other copyright notices refer only to Educational Testing Service’s original work in the product.
**AREAS OF CAREER INTEREST**

Below are descriptions of people with high scores in each interest area and examples of work those people like to do.

<table>
<thead>
<tr>
<th>People with this interest typically:</th>
<th>Examples of Work:</th>
</tr>
</thead>
</table>
| REALISTIC (R) “The Do-ers” – independent, practical, enjoy the outdoors, and prefer working with their hands | • Working with plants and animals  
• Work that involves real-world materials like wood, tools, and machinery  
• Outside work |
| Like: Work that includes practical, hands-on problems and answers.  
Dislike: Careers that involve paperwork or working closely with others. | |
| INVESTIGATIVE (I) “The Thinkers” – curious, analytical, logical, and enjoy problem solving | • Searching for facts  
• Figuring out problems |
| Like: Work that has to do with ideas and thinking  
Dislike: Work that has to do with physical activity or leading people | |
| ARTISTIC (A) “The Creators” – creative, expressive, imaginative, and like to work with ideas | • Creativity in their work  
• Work that can be done without following a set of rules |
| Like: Work that deals with the artistic side of things, such as acting, music, art, and design  
Dislike: Work that is routine and follows set rules and methods | |
| SOCIAL (S) “The Helpers” - generous, helpful, enjoy teamwork, and helping others | • Teaching  
• Giving advice  
• Helping and being of service to people |
| Like: Working with people to help them learn and grow.  
Dislike: Working with objects, machines, or information; not working with people | |
| ENTERPRISING (E) “The Persuaders” – ambitious, extroverted, confident, and enjoy leading | • Persuading and leading people  
• Making decisions  
• Taking risks for profits |
| Like: Work that has to do with starting up and carrying out business projects; taking action  
Dislike: Work that involves thinking about things | |
| CONVENTIONAL (C) “The Organizers” – logical, organized, detail-oriented, and prefer structured environments | • Working with clear rules  
• Following a strong leader |
| Like: Work that follows set procedures and routines; working with information and paying attention to details  
Dislike: Working with ideas and work without rules and established methods | |
CAREER FIT

Based on your responses, you will find indicators of how well your interests align with your current and desired jobs. Also, you will find 20 jobs that most align with your interests.

<table>
<thead>
<tr>
<th>FIT BETWEEN INTEREST PROFILE &amp; CURRENT JOB</th>
<th>FIT BETWEEN INTEREST PROFILE &amp; DESIRED JOB</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;BQ_9&gt;</td>
<td>&lt;CII_90&gt;</td>
</tr>
<tr>
<td>LOW FIT</td>
<td>LOW FIT</td>
</tr>
<tr>
<td>MODERATE FIT</td>
<td>MODERATE FIT</td>
</tr>
<tr>
<td>GOOD FIT</td>
<td>GOOD FIT</td>
</tr>
<tr>
<td>NOT APPLICABLE</td>
<td>NOT APPLICABLE</td>
</tr>
</tbody>
</table>

Top 20 Jobs That MOST Fit Your Interest

<table>
<thead>
<tr>
<th>Job Title</th>
<th>Level of Fit Based on Responses (100 = Best Fit with Interests)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job Title 1</td>
<td></td>
</tr>
<tr>
<td>Job Title 2</td>
<td></td>
</tr>
<tr>
<td>Job Title 3</td>
<td></td>
</tr>
<tr>
<td>Job Title 4</td>
<td></td>
</tr>
<tr>
<td>Job Title 5</td>
<td></td>
</tr>
<tr>
<td>Job Title 6</td>
<td></td>
</tr>
<tr>
<td>Job Title 7</td>
<td></td>
</tr>
<tr>
<td>Job Title 8</td>
<td></td>
</tr>
<tr>
<td>Job Title 9</td>
<td></td>
</tr>
<tr>
<td>Job Title 10</td>
<td></td>
</tr>
<tr>
<td>Job Title 11</td>
<td></td>
</tr>
<tr>
<td>Job Title 12</td>
<td></td>
</tr>
<tr>
<td>Job Title 13</td>
<td></td>
</tr>
<tr>
<td>Job Title 14</td>
<td></td>
</tr>
<tr>
<td>Job Title 15</td>
<td></td>
</tr>
<tr>
<td>Job Title 16</td>
<td></td>
</tr>
<tr>
<td>Job Title 17</td>
<td></td>
</tr>
<tr>
<td>Job Title 18</td>
<td></td>
</tr>
<tr>
<td>Job Title 19</td>
<td></td>
</tr>
<tr>
<td>Job Title 20</td>
<td></td>
</tr>
</tbody>
</table>
**CAREER FIT**

In addition to jobs that are most like you, based on your responses, below are 10 jobs you would likely not be interested in pursuing.

<table>
<thead>
<tr>
<th>Job Title</th>
<th>Level of Fit Based on Responses (-100 = Least Fit with Interests)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Job Title 452</td>
<td></td>
</tr>
<tr>
<td>Job Title 453</td>
<td></td>
</tr>
<tr>
<td>Job Title 454</td>
<td></td>
</tr>
<tr>
<td>Job Title 455</td>
<td></td>
</tr>
<tr>
<td>Job Title 456</td>
<td></td>
</tr>
<tr>
<td>Job Title 457</td>
<td></td>
</tr>
<tr>
<td>Job Title 458</td>
<td></td>
</tr>
<tr>
<td>Job Title 459</td>
<td></td>
</tr>
<tr>
<td>Job Title 460</td>
<td></td>
</tr>
<tr>
<td>Job Title 461</td>
<td></td>
</tr>
</tbody>
</table>
### CAREER INTENTIONALITY*

<table>
<thead>
<tr>
<th>Job Seeking</th>
<th>Additional Training</th>
<th>Self-Efficacy</th>
<th>Taking Active Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intention and benefit from finding and securing a new job.</td>
<td>Intention and benefit of seeking out and completing new or additional training.</td>
<td>Personal intention and internal motivation for pursuing a new job.</td>
<td>Actions and steps taken in pursuit of a new job.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Job Seeking</th>
<th>Additional Training</th>
<th>Self-Efficacy</th>
<th>Taking Active Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low: Locating and securing a new job is not important to you right now. You believe that finding a new job is not a top priority.</td>
<td>Low: You have little intention of seeking additional job training within the next year.</td>
<td>Low: You may lack confidence in locating and securing a new job or successfully seeking additional training.</td>
<td>Low: You have taken few active steps in seeking a new job and may not be motivated to do so.</td>
</tr>
<tr>
<td>Moderate: You are aware that locating and securing a new job is important, but are not necessarily motivated to do so.</td>
<td>Moderate: You have moderate intention of seeking additional job training within the next year.</td>
<td>Moderate: You are moderately confident in locating and securing a new job or seeking additional training.</td>
<td>Moderate: You have demonstrated some initiative in seeking a new job.</td>
</tr>
<tr>
<td>High: You see the importance of and value in locating and securing a new job and recognize the positive impact a new job will have both personally and to those important to you.</td>
<td>High: You have high intention of seeking additional job training within the next year.</td>
<td>High: You are highly confident in locating and securing a new job or seeking additional training.</td>
<td>High: You have demonstrated strong initiative and taken active steps in seeking a new job.</td>
</tr>
</tbody>
</table>

*Note: These scores are compared to individuals internationally.

Compare your scores in the above areas. The level of your intention to seek a new job and/or additional training, your confidence in successfully doing so, and your actions should be relatively balanced.

---

Education & Skills Online Technical Report

Appendix B
If the test taker answered BQ_Q8 = Unemployed, not looking for work, then the following text will display under the Career Intentionality heading instead of the charts and high/moderate/low descriptive text:

Because you indicated that you were currently “Unemployed, Not Looking for Work,” it was determined that the Career Intentionality survey was not applicable to your future occupational goals.
Subjective Well-Being and Health

This report explains:

- What Subjective Well-Being is and why it is important
- Your Life Satisfaction, Positive Affect, and Negative Affect scores
- A summary of your health attitudes and behaviors

Your individual report follows below.

What Subjective Well-Being Is and Why It Is Important

Subjective Well-Being refers to how people think and feel about their lives. In general, people who have a positive view of their own lives are hopeful for the future and have more positive experiences. They are likely to have higher Subjective Well-Being. Researchers have found that people with higher subjective well-being are usually more productive and more successful in their personal and professional lives. They are more likely to live longer, healthier lives. The questions in this test have been used in many research studies worldwide. This score report describes your results in three major areas of Subjective Well-Being that are described below: Life Satisfaction, Positive Affect, and Negative Affect.

The Life Satisfaction aspect of Subjective Well-Being refers to how people think about their lives overall. How people feel about their lives is shown by such things as moods and emotions experienced day to day. The experience of positive emotions, such as joy or excitement, is called Positive Affect. The experience of negative emotions, such as anger, distress, or shame, is called Negative Affect. The Positive and Negative Affect scores represent your experience of positive and negative emotions in the past week. They make up the emotional components of Subjective Well-Being.

Your Life Satisfaction Score: High, Moderate, or Low

High – Your score shows that you are very satisfied with your life and feel good about how it is going. Generally, people who score in this range take on life’s challenges without feeling overwhelmed.

Moderate – Your score shows that you are somewhat satisfied with your life. You may feel as though you are doing well in some areas while feeling other areas need improvement. People who report having a moderate level of life satisfaction for long periods of time may want to think about why this is. After reflecting, it is important for them to try to make positive changes in their lives.

Low – Your score shows that you are not very satisfied with your life. When possible, changes in circumstances (e.g., schedule, activities), attitudes, and behaviors are recommended for people with a low score. These changes may result in positive ways of dealing with difficult situations and improvement in life satisfaction.
Your Positive Affect Score: High, Moderate, or Low

**High Positive Affect** – Your score shows that you had positive moods and emotions in the past week. People who usually score high in this category feel happiness and are often quick to smile, are energetic, and enjoy their work.

**Moderate Positive Affect** – Your score indicates that you experienced moderate positive moods and emotions in the past week. People who score in this range can appear emotionally controlled while being hard to read due to a lack of obvious enthusiasm.

**Low Positive Affect** – Your score shows that you had low levels of positive moods and emotions in the past week. People who score in this range have had fewer positive experiences and felt sadness in the past week, which sometimes results in feeling tired and engaging in little activity.

Your Negative Affect Score: Low, Moderate, or High

**Low Negative Affect** – Your score shows that you experienced low levels of negativity in the past week. People who score in this range appear calm and composed.

**Moderate Negative Affect** – Your score shows that you experienced moderately negative moods and emotions in the past week. People who score in this range appear somewhat angry, annoyed, and tense.

**High Negative Affect** – Your score shows that you experienced negative moods and emotions multiple times in the past week. People who score in this range experience negative feelings more often than others. They are often frustrated and depressed.
The Importance of a Healthy Lifestyle

Leading a healthy lifestyle is important to reaching your best physical and mental health and well-being. A healthy lifestyle includes having a positive outlook on your health, maintaining a healthy weight, and adopting healthy behaviors and habits. A positive health outlook includes focusing more on what you can do and less on what you cannot do. This can help you handle and recover from physical and mental health challenges.

A healthy lifestyle can increase your energy levels. It can improve your mood and physical and mental performance. It may also reduce your risk for illness, increase the length of your life, and improve your overall quality of life. Four of the most important health behaviors include eating a balanced diet, not smoking, getting regular exercise, and getting enough quality sleep. While a positive health outlook and a healthy lifestyle are important for your personal well-being, no single aspect should be used as an overall assessment of your health.

<table>
<thead>
<tr>
<th>Your Health Outlook</th>
<th>Being optimistic about your health given your particular situation is important to managing illness. Focusing more on what you can do, and less on what you cannot do in terms of your health, can positively impact your ability to cope with and recover from health challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>You have a positive view of your current health.</td>
<td></td>
</tr>
<tr>
<td>You have a fair view of your current health.</td>
<td></td>
</tr>
<tr>
<td>You have a poor view of your current health</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Body Mass Index (BMI)</th>
<th>Body Mass Index (BMI) is a simple index of weight for height and is commonly used to classify individuals as Underweight, Normal Weight, and Overweight. Maintaining a healthy weight is important for your overall health. It can lower your risk for many illnesses and conditions while increasing your energy level.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your BMI indicates that you are Underweight.*</td>
<td></td>
</tr>
<tr>
<td>Your BMI indicates that you are a Normal Weight.*</td>
<td></td>
</tr>
<tr>
<td>Your BMI indicates that you are Overweight.*</td>
<td></td>
</tr>
<tr>
<td>Your BMI indicates that you are Overweight.*</td>
<td></td>
</tr>
</tbody>
</table>

*Note: While the cutoffs used in this assessment correspond to the World Health Organization (WHO) (2014) classifications for Underweight, Normal Weight, and Overweight, these cutoffs may not be indicative of health risk for all populations.
### Health Behaviors

<table>
<thead>
<tr>
<th>Health Behavior</th>
<th>Description</th>
<th>Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Diet/Nutrition</strong></td>
<td>A healthy diet includes eating fruits and vegetables every day. This can reduce your risk for illnesses such as heart disease, cancer, and diabetes. Current recommendations suggest that we eat 400 g (approximately 3 cups or 5 servings) of fruits and vegetables per day to maintain good health.</td>
<td></td>
</tr>
<tr>
<td><strong>Smoking</strong></td>
<td>Smoking is a large risk factor for serious illnesses such as heart attack, stroke, and cancer. Avoid smoking and secondhand smoke to positively impact your health.</td>
<td></td>
</tr>
<tr>
<td><strong>Exercise</strong></td>
<td>Regular physical exercise is important to reducing stress, managing your weight and maintaining health. The current recommendation is 150 minutes of moderate or 75 minutes of intense exercise each week. Walking regularly, taking the stairs, and starting a new sport are ways to increase physical activity.</td>
<td></td>
</tr>
<tr>
<td><strong>Sleep</strong></td>
<td>Good sleep habits lead to better mood and functioning and reduce the risk of illness. Too little sleep can lead to illness, irritability, and difficulty concentrating. It is best to get 7 to 9 hours of quality sleep per night. A regular sleep schedule, including habits that encourage uninterrupted sleep such as limiting caffeine and alcohol, is helpful.</td>
<td></td>
</tr>
</tbody>
</table>
**Behavioral Competencies (Forthcoming)**

<table>
<thead>
<tr>
<th>Behavioral Competencies</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Conscientiousness</strong></td>
</tr>
<tr>
<td>The following skills are reflective of <strong>Conscientiousness</strong> in the workplace. People who are conscientious are usually thorough, organized and efficient as well as committed to doing a good job.</td>
</tr>
</tbody>
</table>

**Diligence** describes behaviors associated with working towards objectives. Individuals who are high in diligence tend to be described as hard working, ambitious and confident.

| Percentile Rank* | 66 | 100 |

**Organization** describes behaviors associated with maintaining a sense of order as well as an ability to plan work tasks and work activities.

| Percentile Rank* | 66 | 100 |

**Dependability** describes behaviors related to a sense of personal responsibility. Individuals who are high in dependability tend to be reliable and make every effort to keep promises.

| Percentile Rank* | 66 | 100 |

**Self Discipline** indicates an ability to be patient, cautious and level-headed. People who are high in self discipline tend to maintain control at work.

| Percentile Rank* | 66 | 100 |

*Percentile ranks are based on international data from Educational Skills Online*

Items included in the Behavioral Competencies module are Copyright © Drasgow Consulting Group, LLC and licensed and used with permission by Educational Testing Service.
## Behavioral Competencies

### Extraversion

The following skills are reflective of Extraversion in the workplace. People who are extraverted are often described as social, talkative and assertive.

**Assertiveness** indicates an ability to take charge at work. People who are assertive are often described as direct, decisive and “natural leaders.”

| Percentile Rank* | 0 | 66 | 100 |

**Friendliness** indicates an interest in social interactions. People high in friendliness are often interested in meeting new people at work and using this skill for the betterment of the organization.

| Percentile Rank* | 0 | 66 | 100 |

### Agreeableness

The following skills are reflective of Agreeableness in the workplace. People who are agreeable are often perceived as good natured and cooperative.

**Generosity** describes individuals who are willing to offer their time and resources in support of others. People high in generosity tend to be helpful to others at work.

| Percentile Rank* | 0 | 66 | 100 |

**Collaboration** describes individuals who are viewed as trusting and cooperative. People high in collaboration are often easy to get along with and work well on teams.

| Percentile Rank* | 0 | 66 | 100 |

### Emotional Stability

The following skills are reflective of Emotional Stability in the workplace. People who are emotionally stable tend to be even tempered, composed and maintain a positive attitude.

**Stability** describes individuals who are relaxed and worry free. People high in stability work well with changing work priorities and manage stress well.

| Percentile Rank* | 0 | 66 | 100 |

**Optimism** describes individuals who have a positive outlook and cope well with setbacks. People who are optimistic tend to incorporate feedback well at work.

| Percentile Rank* | 0 | 66 | 100 |
| Percentile Rank* | 0 | 66 | 100 |
The following skills are reflective of *Openness* in the workplace. People who are open to experience tend to be creative, interested in learning and have an intellectual approach.

<table>
<thead>
<tr>
<th>Behavioral Competencies</th>
<th>Percentile Rank*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Creativity</strong> describes behaviors that are inventive and imaginative. People high in creativity tend to be innovators at work.</td>
<td>0</td>
</tr>
<tr>
<td><strong>Intellectual Orientation</strong> is indicative of an ability to process information and make decisions quickly. People high in intellectual orientation are often viewed as knowledgeable by others.</td>
<td>0</td>
</tr>
<tr>
<td><strong>Inquisitiveness</strong> describes behaviors that relate to being perceptive and curious. People high in inquisitiveness tend to be interested in learning more by attending workshops at work.</td>
<td>0</td>
</tr>
</tbody>
</table>
Bibliography


