International Journal of Statistics and Probability; Vol. 4, No. 3; 2015
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Students' Emerging Reasoning about Data Tables of Large-Scale Data

Theodosia Prodromou[1]

[1] School of Education, University of New England, Armidale, Australia

Correspondence: Theodosia Prodromou, School of Education, University of New England, Armidale, NSW, 2351, Australia. E-mail: theodosia.prodromou@une.edu.au

Received: June 23, 2015 Accepted: July 12, 2015 Online Published: July 24, 2015

doi:10.5539/ijsp.v4n3pxx URL: http://dx.doi.org/10.5539/ijsp.v4n3pxx

Abstract

This study investigated thirty-two Year 9 secondary school students' (15 year olds) reasoning about data tables of large-scale data. Eight groups of four students, drawn from six classes, participated in a workshop that examined the components of population change for EU and candidate countries, namely natural increase of population, net overseas migration for Europe and their country, and total population growth. Students investigated trends in real data displayed in tables, and responded to a set of reflective questions. Analysis of the reasoning used by the students revealed four levels of data-table comprehension—reading the data, reading within the data, reading beyond the data, and reading behind the data—similar to the levels described for students working with smaller data sets.

Keywords: Statistics education, large-scale data, open data, big data, data tables of large-scale data

1. Introduction

Developments enabled by novel technologies have completely altered the ways that citizens can access data. Indeed, citizens can access an enormous amount of numerical information that is even greater nowadays with the data revolution that gives rise to emerging data sources that are providing new sorts of evidence used to influence public opinion. Three emerging trends have impacted this revolution in our increasingly data-driven society. These trends include: 1) the increasing use of large-scale databases within the open data movement, 2) the growing use of big data, and 3) novel ways of visualising data.

The open data movement supports the availability of high quality data sets collected by national statistics offices and non-government organisations for a specific purpose. These data are characterised by several features: the data are multivariate, consist of clearly defined measures, the population is known, and the data generation and presentation have been subjected to extensive scrutiny. These data are made available to all citizens. The revolution of the open data movement has had significant success in recent years in persuading major data providers, such as Eurostat and national statistics offices, to give citizens access to huge databases in order to create new variables, and explore new relationships.

Big data, in contrast to open data, is not publicly available. In contrast to open data generated by national statistics offices, big data is only available through proprietary sources, and is owned by companies that take financial advantage of using big data.

Big data is extremely large and complex collections of data that do not necessarily have a numerical form. The emergence of big data spans four dimensions: variety, volume, velocity, and veracity. By variety, we mean that big data involves many sources and types of data both structured and unstructured. Nowadays data comes in the form of audio, video, click streams, log files, emails, photos, videos, monitoring devices, PDFs, etc. By volume, we mean that big data implies enormous quantities of data generated by machines, networks, and human interaction on systems like social media. By velocity, we refer to the time sensitive nature of much data: big data can increase its value when delivered rapidly. Finally, by veracity, we refer to the biases, noise, and abnormalities in big data due to data inconsistency.

Novel ways of visualising data using highly advanced computational methods with sophisticated graphics engines aim to tap the surprising ability of humans to see patterns and structure in even the most complex visual presentations. Currently major data providers make assessible powerful visualisations of dynamic datasets (e.g., GapMinder and eXplorer) that allow users to zoom in to data displays and interactively manipulate multivariate datasets. New research is now beginning to apply these sorts of tools to the social sciences and humanities as well, and the techniques offer considerable promise in helping us understand complex social processes like learning, political and organizational change, and the diffusion of knowledge.

The three emerging trends that characterise the revolution of our increasingly data-driven society offer considerable promise in enhancing people's understanding of complex scientific and societal issues, such as political and organizational change, population immigration. The revolution of data in our society is having a profound impact on teaching statistics.

The statistics community is not faced merely with the challenge of educating students to become competent explorers of large-scale authentic data on a huge variety of social important topics, it is faced with educating an entire population about difficult statistical tasks, including interpreting multivariate data sets, and drawing conclusions about samples of large-scale data.

My purpose in this article is to report an empirical study to more clearly identify the patterns of reasoning used by students interacting with lasrge-scale data tables. It brings together key ideas from various perspectives, going beyond several earlier reviews of the literature, to identify critical factors that influence secondary school students' reasoning about large-scale data organised in tables, to enable an in-depth analysis of key aspects of participants' reasoning. I also provide some recommendations for instruction and future research.

2. Theoretical Framework

The expanding use of large-scale data for prediction and decision-making in almost all domains of life makes it a priority for mathematics school curricula worldwide to help students develop their understanding of key statistical ideas prior to entering college. This includes understanding of data-tables, which is a core aspect of statistics, essential to conducting meaningful data analysis. Tables, graphs, and other data displays are used broadly in the media to present, disseminate, and explain information, thus students need to be able to read and interpret them in meaningful ways.

A number of research studies about the difficulties that learners have with drawing inferences from tables and graphs showed that students have particular difficulty in drawing inferences from tables and graphs in order to interpret the data, and make predictions (e.g., Bright and Friel, 1998; Estrepa, Batanero, and Sanchez, 1999; Pereira-Mendoza and Mellor, 1991; Sharma, 1997). For example, when interpreting the data, one usually compares different data sets presented in graphs and tables to make predictions about an unknown case, to generalize to a population, or to discern a trend.

There appears to be little research on learners' comprehension of tables, despite the pervasive use of data tables in statistical data analysis and textbooks of statistics. The limited existing literature in statistics education addresses table learning in children (Brizuela and Lara-Roth, 2002; Ben-Zvi and Sharett-Amir, 2005; Marti, 2009; Brizuela and Alvarado, 2010; Gabucio et al., 2010; Marti et al., 2010). Brizuela and Lara-Roth (2002) showed that 7-year-old students who had not received direct instruction in the use and configuration of tables, could use information from a table to work on a problem. The tables used by students in the research study of Brizuela and Lara-Roth's were produced without imposing any specific structure on the primary students. Estrella and Mena (2014) investigated primary-school children's comprehension of statistical frequency tables, when the students producedmore tables while trying to analyse some data. They identified different levels of conceptualization of tables in these students, such as text lists with and without counting, tables with icons with and without counting, tables with text with and without counting, and tables with text without individual counts but with marginal totals. These primary students' conceptualization of tables allowed Estrella and Mena's (2014) to explore how students register data in a table, count in a table, list of elements belonging to a class, using partitioning, equivalence relations, and counting that allow for ordering data to obtain information in order to place data in rows, columns, cells, and to use written language to label headings.

Kemp and Kissane (2010) described a five-step framework to help both teachers and students in primary, secondary, and tertiary mathematics education, to interpret data in the form of tables or graphs. The framework for interpreting tables and graphs provides a progression from simple numerical reading of a table to more complex interpretations of tables and graphs required for a better understanding of data in their context. Another classification of patterns of reading and comprehending graphs, and hence of interpreting of graphs, was developed by Curcio (1987) who assessed fourth and seventh grade students' interpretations of school graphs. From the analysis of the students' responses, he developed a framework for assessing and building learners' graphical comprehension that has three levels: reading the data, reading within the data, and reading beyond the data. These three types of related components of graph comprehension comprise the classical Curcio schema:

1. "Reading the data," which involves "Lifting" information to answer explicit questions for which the

obvious answer is right there in the graph (e.g., What is the least preferable means of transportation of students who travelled to school on Mondays?).

2. "Reading between the data," which involves finding relationships in the data presented in a graph by making comparisons (e.g., Is the number of students who travel to school by car on Mondays the same as the number of students who travel to school by bus?).

3. "Reading beyond the data," which involves extrapolating, predicting, or inferring from the representation to answer implicit questions (e.g., If we ask the students of ten schools, about how they get to school on Monday, how many students who travel by bus might they expect to find?).

This framework has helped statistics educators in building instructional strategies for facilitating student understanding of different graphical representations.

Although Curcio's framework (1987) has undoubtedly made a very important contribution for understanding the processes involved in the interpretation of graphical representations, it has been criticized in recent years for limiting its investigation to the kinds of graphs used in school contexts (e.g., simplistic tables of limited purpose in real life, cf. Monteiro & Ainley, 2007), and hence for restricting the range of situations to which the interpretation of data tables and graphs is applied (Sharma, 2013). According to Sharma (2013), Curcio did not investigate how students evaluated and critically commented on information displayed in tables and graphs. Additionally, the questions of Curcio's research did not provide students with any opportunities to explain their choices, hence, to gain an insight into students' thinking.

A more elaborated framework given by Friel, Curcio, and Bright (2001), developed the original Curcio framework, splitting each stage into two parts, to more precisely describe the behaviours associated with graph comprehension:

1) recognizing the components of graphs, the interrelationships of these components, and the impact of these components on the graphical presentation of information (Reading the data)

2) speaking the language of particular graphs, when reasoning about the information displayed in graphs (Reading the data)

3) understanding the relationships among graphs (Reading within the data),

4) making sense of a graph (Reading within the data)

5) interpreting information in a graph (Reading beyond the data)

6) recognizing if the graph is appropriate (Reading beyond the data)

Monteiro and Ainley (2007) argue that familiarity with the above components is not sufficient to ensure understanding of specific graphs. They claim that that context may be the key factor in understanding the comprehension of graphs. According to Monteiro and Ainley, data displays used for analytical purposes are predominantly tools for detection of important or unusual features in the data. On the other hand, graphs used for communication are defined as pictures intended to convey information about numbers and relationships among numbers. The authors add that the use of the kinds of "school graphs" which were used within Friel et al.'s study (for example displaying information about "the number of letters in students' names" or "how many raisins are in various boxes") have limited purpose, in terms of analysing or communicating information which relates to interesting problems. Moreover, Monteiro and Ainley state that in the specific examples used by Friel et al., the term looking beyond the data does not imply a need to look critically at the data and ask worrying questions (see Gal, 2004). Indeed we might look beyond the data (extrapolating, predicting, or inferring from the representation) without being prompted to question the main idea presented in the data display.

Similarly, Monteiro and Ainley (2007) have criticized the Friel, Curcio, and Bright (2001) framework, arguing that learners' familiarity with its components is not sufficient to ensure understanding of data displays, and that context might be the most important factor in graph comprehension. Recognizing the important role of context in statistical analysis, Shaughnessy (2007), added a fourth level beyond Curcio's three levels of graph comprehension: "reading behind the data or graph," which emphasizes the need for interpreting data displays based on the context and situation underlying the graph being constructed. But Shaughnessy's work, too, was done with simple data tables like those found in text books; in this study we look to see if similar patterns of reasoning hold with more complex data tables.

3. Methodology

This study uses qualitative analysis to examine students' reasoning about large-scale data based on experimental data.

3.1 Sample

The research study involves two schools in Cyprus. Eight groups, each of four year-9 students, were drawn from six different classes in the same school (N = 32) to participate in a workshop that examined the components of population change in the EU and member countries. The mathematics teachers of each class selected students who came from the same class, so interpersonal relationships had been already established prior to the research. The teachers were asked to choose articulate students who would have no difficulty in setting up a friendly group. The groups were selected by the teacher so as to include, in their assessment, two girls and two boys from a "middle" attainment in mathematics. The researcher spent one 100-minute double period with each group. Scripts from each group and rough working sheets from each group were collected. Written reflections of each group of students were included in the data. The researcher also made field notes during and immediately after students' engagement with the workshop.

3.2 Instrument

A statistics-learning situation was implemented with paper and pencil during a workshop designed to provide opportunities for students to engage in investigating real data published by EUROSTAT, the statistical office of the European Union situated in Luxemburg.

The students were provided with the workshop sheet (Appendix A) and they were asked to examine five data tables taken from the EUROSTAT website representing number of live births (Appendix B), crude rates of population change (Appendix C), immigration rates (Appendix D), emigration rates (Appendix E), and population by citizenship-Foreigners (Appendix F). The students did not have direct access to computers or mobile devices, therefore I presented the five tables to the students in a paper format.

The workshop called for participants to examine the components of population change in Europe and candidate countries, namely natural increase of population, net overseas migration for Europe and their country, and total population growth.

The workshop included reflective questions designed to provoke students to pause and reflect on the data-tables, seeking interesting aspects of the graphs such as possible reasons for the higher rate of natural increase in some of these countries, or variations in growth rates among different countries, and to discuss their observations. The questions provided opportunities to query students as to the reason underlying their reasoning, and thus, gain an insight into students' way of thinking. In particular, the participants in the workshop were asked to examine the data-tables and complete the following tasks in the order presented:

1) Compare the indicators of population change in Cyprus: number of live births,[1] crude rate of population change,[2] immigration,[3] emigration,[4] and population by citizenship–Foreigners.[5]
2) Discuss and explain their observations regarding growth in Cyprus from 2001-2012, including possible social, historical, environmental, economic, and political factors that might have caused this change. Identify and justify the dominant factors.
3) Identify the European countries that have a net loss of migrants, explain why these countries may be experiencing that loss, and identify and reason about the dominant social, historical, environmental, economic, and political factors that might have influenced the change of the population.

In this paper, we report on students' written answers to question three, supplemented by observations of the students working in groups and discussions with the students to clarify their reasoning. During the interviews, students presented and discussed the conclusions they had drawn regarding the tasks. In this way we obtained further clarification of the nature and type of reasoning they used and the difficulties that emerged during the reflective activity. As mentioned earlier, the purpose of this paper is to elaborate more precisely the nature of students' reasoning about large-scale data presented in tables, to enable us an in-depth analysis of key aspects of

---

[1] Eurostat defines the number of live births as the number of births of children that showed any sign of life (total births minus stillbirths).
[2] Eurostat defines the crude rate of population change as the ratio of the population change during the year to the average population in that year. The value is expressed per 1000 inhabitants. Population change is the difference between the population sizes on 1 January of two consecutive years.
[3] Eurostat defines immigration the total number of long-term immigrants into the reporting country during the reference year.
[4] Eurostat defines emigration the total number of long-term emigrants from the reporting country during the reference years.
[5] Eurostat defines population by citizenship-Foreigners as the total number of foreigners residing in the country, including citizens of other EU Members States and non-EU citizens, usually resident in the reporting country.

participants' reasoning.

3.3 Analysis

Participants' reflective responses to questions, in conjunction with the working sheets from each group and researcher's field notes, were analysed at the macro level to identify episodes of students' reasoning while examining the data tables.

Each episode was coded based on common elements of participants' reasoning then subjected to microanalysis to see if there were shared characteristics of the reasoning. Finally, the analysis identified typical instances of students' reflective activities in the workshop engaged in by students that capture the category of students' reasoning and the competencies that underpin such a reasoning category.

3.4 Results

The results are presented according to the four organizational categories defined by Shaughnessy, which guided this study. In each category is presented an episode from the data that is representative of the category.

<u>First category</u>: Reading data

This first category is concerned with how students engage with tables and how they recognise the components of the table (e.g., the raw data) and the interrelations among these components and then use this information to answer explicit questions. After studying the data-tables of Emigration versus time, and Number of live births versus time, students commented:

Group 3: We observe that from 2001 to 2012, emigration in Italy has increased from 56077 to 106216, while in Cyprus immigration has increased from 13909 to 18105. In Italy, it has doubled (from 56077 to 106216), but in Cyprus the increase was less than 25%. The number of live births in Italy has remained the same from 2001 to 2012, while in Cyprus, it has increased by approximately 2000. The crude rate of population change in Cyprus is 11.4% in 2001, 11.5% in 2002, 12.8% in 2003, 14.0% in 2004, 14.8% in 2005, 18.5% in 2006, and it increases substantially to 24% in 2007. The change in the crude rate in 2008 is 26.2%, 17.5% in 2009, and then it decreases to 24.8% in 2010, increases to 26.2% in 2011 and then it dramatically decreased to 4.5% in 2012.

These students appeared to be chiefly confined to the reading of data in order to report the variations of change seen in rows of the tables of Emigration versus time, and Number of live births versus time. They did not seem to understand the deep structure of the data in their totality, through making comparisons among the countries that students have chosen.

<u>Second category</u>: Reading within data

This type of interaction occurs when students are interpolating and finding relationships in the data while reasoning about the information displayed in tables. A second group of students read the data displayed in columns of the tables (vertical reading), comparing the variations in population change among different countries from 2001-2012:

Group 2: In 2001, Liechtenstein had the highest crude rate of change (19.9%), then Ireland followed with 17.3%, Turkey (13.8%), Spain (13.7%), and Cyprus (11.4%). Luxemburg also had the same crude rate (11.4%) as Cyprus (11.4%). When we looked at the data for 2012, the crude range of population change of Liechtenstein was decreased to 9.9% in 2012, while Ireland's crude rate decreased to 1.8%.

This group of students observed the data in columns (vertical reading) to identify the country that had the highest numerical value of a data point for a certain year. They usually combined the vertical reading with reading the data in rows (horizontal reading) in order to make comparisons of data between different countries (vertical reading) and within a country (horizontal reading) for different years. Although, they made correct comparisons within the data-tables, students did not make any reference to the contextual factors that impacted on the reported population change.

<u>Third category</u>: Reading beyond the data

This category is concerned with extrapolating, predicting, or inferring from the data table to answer implicit questions. Some groups of students attended to different variables of the data and seemed able to integrate the information provided by those variables:

Group 4: In general, almost all the countries of Europe have been affected by the economic crisis. When we look at the table of the crude rate of population change, we observe that there is a decrease of the crude rate in almost all of the European countries. However, the crude rate of some countries was decreased a lot. For example, we observe that after 2008 the crude rate of population in Greece decreased from 3.4 (in 2007) to -5.5 (in 2012). Similarly, Portugal's crude rate decreased from 2.0 (in 2008) to -5.2 (in 2012). UK's crude rate decreased from 8.1 (in 2008) to 6.3 (in 2012). However, in Romania, we observe an increase in the crude rate from -23.7 (in 2007) to -1.9 (in 2012)

In this category of reasoning, students seemed to pay attention to the entire distribution of data and then they focused particularly on individual cases that exhibit distinctive variability in the measurement, providing appropriate qualitative inferences about the possible meaning of the data within their context. They acknowledged, however, the students acknolwedged that the many factors impacting populations change meant that they were not fully able to explain the observed changes in the the variables presented in the data table. Other students engaged critically in a familiar context when they observed the data-tables of immigration:

Group 1: The countries where the immigration is decreasing during the last few years are Esthonia (from 3709 in 2011 to 2629 in 2012), Italy (from 558019 in 2007 to 350772 in 2012), Cyprus (20206 in 2010 to 17476 in 2012) and Greece (from 119079 in 2010 to 110139 in 2012). We cannot claim that all these countries experience financial crisis. For example, we observe a pattern of immigration in Estonia; immigration goes up one year and then down the next year. Similarly, in Cyprus. On the contrary, in Italy immigration decreases steadily from 2007 to 2012, and similarly in Greece it decreases from 2010 to 2012. So, the immigration (the people who go to work in a foreign country). . . The countries that have not been affected by the economic crisis and where the immigration is increasing after 2010 (including 2010) are Belgium, France, and Austria because they have money to pay people, so people immigrate to these countries.

These students seemed to appreciate variation and to qualitatively interpret the existence of variation in context. They demonstrated awareness of relevant features of the table, however these features are predominantly based on both the data and the context. When using this type of reasoning, the students appeared to be able to focus on the data interpretation and they exclusively based their answers on the different variables in the data tables.

Fourth category: (reading behind the data)

In this type of reasoning, the students seemed to move beyond the data, and attempted to give an answer that drew upon prior knowledge about issues directly related to the data presented in the tables. In such situations, students' reasoning related prior knowledge to components of open data tables, which allowed more complex inferences:

Group 5: Knowing that the economic crisis has been very intense for the following countries: Greece, Cyprus, Ireland, Italy, Portugal, Romania and Spain, when we observe the data-table of emigration versus time, we understand that Spain was the first of these countries that began to suffer from the economic crisis, in 2003; when its citizens began to leave the country in an attempt to find work in other countries. Afterwards, the data tell us that economic crisis has affected Portugal in 2004, Ireland in 2006, and Italy in 2008. Later on, in 2009, financial crisis influenced Greece. In 2010 Cyprus joined the group of the EU countries affected by the economic crisis. We can understand when one country has been affected by the EU economic crisis from the increase we observe in the Emigration data table, since the citizens of a country leave their country during economic crisis with the intent to settle permanently in another country. . However, this trend is not observed in Romania, because the emigration there is decreasing instead. We can observe a data table of another indicator for Romania to be able to tell when Romania was affected by crisis. The data table of immigration vs. time for Romania shows that the number of Immigrants decreased from 2008 to 2009, increased slightly in 2010, decreased slightly in 2011 and increased substantially in 2012. The data do not provide us with adequate evidence to deduce the effect of the financial crisis on Romania. We need to look at the data table of another indicator. The data from the table of the number of live births shows an increase from 2001 to 2012, so it is not clear when Romania was influenced by a rise in poverty. We should look at other data-tables that can provide us with appropriate evidence to be able to draw any reliable conclusions. For example, we need to look at the table for the crude rate of population change for Romania. We look at it and we observe that the crude rate of population is increasing from 2007 (-23.7) to 2012 (-1.9). We do not have enough evidence to deduce whether Romania was one of the five countries of the EU that was most affected by the economic crisis.

Students in this group attempted to give an answer that integrates prior knowledge of issues directly related to the data-tables. They seemed to understand the purpose of the data, and of the inferences made. These students used the relevant features of the data and background contextual knowledge, and utilized different tables of data of the same variable to appropriately answer a question. They acknowledged that the quantitative data included in a single table might not show a particular trend in the data, thus examination of other variables is required to get a more complete picture of the situation at hand.

4. Conclusions

The open data movement has provided unprecedented access to authentic, large-scale data sets on a wide range of socially important topics. Competent use of large-scale data predominantly requires comprehension of tables and other visual representations of statistical data, since these are routinely used in daily life and in the workplace to communicate information. Thus, statistics instruction at the school level should give more emphasis to enhancing students' comprehension and interpretation of large-scale data displayed in tables and graphs.

In this paper, I investigated the emerging reasoning about data-tables of a group of year 9 secondary school students (15 year olds) in Cyprus. The findings not only provide empirical confirmation of the four-part framework of Shaughnessy in this research on large-scale data-table comprehension, they help establish a theoretical framework that can address different levels of large-scale data-table comprehension. This extension of the framework described by Shaughnessy to a novel context gives support to the emergence of four levels of large-scale data-table comprehension and it shows that when drawing conclusions about samples of large-scale data, at level 3 and level 4 of the framework, comprehension, comparison, and interpretation of different variable of the multivariate data sets is central.

Level 1, an elementary reading (reading data) is characterised by simply reading the data either horizontally or vertically, following the rows and columns of a two-dimensional table to answer specific questions for which the obvious answer is in the data-table, without making any judgements with regard to comparing any variations in growth rates among different countries.

Level 2, an intermediate reading (reading within data), is focused on making comparisons of data between different countries and within a country for different years. Students at this level attend to one or more relevant aspects of the data but have difficulty in integrating those aspects into their context.

Level 3, an overall reading (reading beyond the data) is characterised by interpreting the numerical values of the data, and attempting to contextualise the data by providing qualitative interpretations of what might have impacted the variation in data values. Additionally, students reading beyond the data begin to gain an awareness of how a few of the possible social, historical, environmental, economic and political factors might have caused similarities and/or differences in the data. The students of this group appeared to be aware that many complicated questions about data might be answered by examining data tables of different variables.

Level 4, an advanced reading (reading behind the data) is characterised by attempting to give an answer that takes into account prior knowledge about a question that is directly related to the data-tables. In such situations, students' reasoning related to comprehension of the components of open data tables is characterised by inference from the data to develop answers to questions (e.g., we are aware of Europe's economic crisis, but we do not know the number of the countries that have been very badly affected by the economic crisis. Can you tell from the data-tables of the given indicators of population change—number of live births, crude rate of population change, immigration, emigration, population by citizenship-foreigners—which are these countries?).

Concurring with the findings of previous studies (Sharma, 2013), findings from the current study indicate that students' reasoning about large-scale data changes over time due to natural developmental process from reading data to focusing on interpreting data with respect to the data's context.

Furthermore, Sharma (2013) argues that a "number of research studies from different theoretical perspectives seem to show that students are particularly weak in drawing inferences and predicting from tables and graphs (e.g., Bright and Friel, 1998; Curcio, 1987; Estepa, Bataneo, and Sanchez, 1999; Pereira-Mendoza and Mellor, 1991, Sharma, 1997)" (p. 52). This could be the result of the instructional neglect of concepts related to the interpretation of tabular representations in context. Student encounters with data-tables in the mathematics classroom are restricted to some "school tables", which do not support developing understanding of complex and challenging tabular representations of authentic data such as those presented to students in the current study. The analysis of the results of this study does not suggest shortcomings of the participants in any meaningful way; it shows that some students reason in simpler patterns than others, but not in any way that we can generalize about overall performance of students. The study isn't set up to assess ability; it is set up to characterize patterns of reasoning.

Although the study has provided some valuable insights into students' conceptions of data tables, very little is still known about this important aspect of statistical reasoning. More research needs to be carried out to investigate and support comprehension of tables by students of different age groups and educational and cultural backgrounds. As the research literature tells us very little about how comprehension of data-tables develops, a possible direction for future research is to find ways to scaffold students' learning in terms of reading and understanding tables, and connecting them with other numerical and graphical representations of data.

Another possible research direction is to study how contextual knowledge affects comprehension of data-tables, and to find ways to help students relate information displayed in a table to the context of the situation. This is essential since, as shown in the current study, students' comprehension of tables and other data representations is reliant not only upon their understanding of the features of the visual display under study, but also on their prior knowledge of the context from which the presented data is drawn, as well as on their ability to utilize this contextual knowledge to make sense of the situation displayed in the table or chart.

4.1 Limitations of the Study

This study discussed in this paper involved relatively a small sample of students. I have reported only few groups of students' reasoning, the clearest illustrations of the emerging ideas. Even had it been possible to analyse all data and the examples I presented were representative of the sample that I had drawn, the findings must be regarded as tentative because this was a small sample and we cannot generalise. Fortunately this has opened up opportunities for future research at a macro-level on students' reasoning about large-scale data displayed in tables.

Future research on students' emerging reasoning about large-scale data should begin with this study as a cornerstone. Implications for research and teaching are outlined below.

4.2 Implications for Teaching

The data revolution provides challenges and opportunities for statistics educators to educate an entire population and create instructional materials for curricula that devote particular attention to engaging students with a broader variety of novel techniques that encourage the comprehension and interpretation of large-scale data displayed in tables and graphs. The comprehension of large-scale open data sets that are two-dimensional tables (both rows and columns) can be achieved when answer questions dealing with several variables. The exploration and interpretation of large-scale multivariate data sets (Ridgway, Nicholson, & McCusker, 2013) is very challenging. A wide range of visualisation tools (e.g., Gapminder) may help students to simplify multidimensional datasets, thus interpret complex data sets.

Prodromou (2013) argues that what is to be communicated to the student is not just the technique of partitioning the complex data set as a building block of process, but also the value of the final partitioning of the dimensions of a data set, when identified and explicitly labeled. Partitioning leads to a focus on a part or segment of the data. When this segment is rendered relatively homogeneous with respect to some features of the complete data, its internal complexity is reduced. Thus the selected data segment's own particular internal data patterns are more likely to emerge in any data summarising activity.

For statistics educators who teach big data, traditional methods and techniques for analysis of data cannot be applied, and novel methods must be developed through collaboration of statistics educators with computer scientists.

Fundamental ideas, such as data quality, the principles of measurement, and drawing inferences in the face of uncertainty, request particular attention. In addition, the habit of thinking from samples to inferences about populations was a function of a "small data" environment. At their core, big data and open data are about making generalisations and predictions, similar to making statistical inferences. For example, large-scale data may be used to predict consumers' future purchases based on their interactions at different sites on the internet or even the performance of a stock market.

To provide effective instruction, teachers need to increase their knowledge of the three emerging trends that have impacted this revolution in our increasingly data-driven society and of how to teach these new trends.

Because of the recent emphasis on large-scale data and data analysis, these concerns have only recently become an important necessity in the secondary school mathematics curriculum. Consequently, teachers may not have had adequate opportunities to learn about large-scale data. More visualisation tools need to be developed to fill this gap. But beyond the materials, thought should be given to how professional-development experiences can be structured so that teachers learn not only how to better interpret large-scale data displayed in tables and graphs, but also how to help students develop similar skills. In order to take into account the full complexity of data, we have to change the way we think about controlling and handling data. This view calls for another change to the constructs of statistical literacy (Gal, 2002) and the introduction of new constructs and principles needed for the revolution of data.

In order to immerse students to this new culture of data, it seems important to give students many opportunities to construct their own meaningful data visualizations that highlight emerging important aspects of data and promote their reasoning about covariation between multiple variables while using the cycle of inquiry and visual analysis (Prodromou, 2014). In particular, I think it will be helpful to encourage students to revisit their specific kinds of inferences while inventing and revising their visual representations of data. In this way they will be able to attend to the changing role of variables from data visualisation to data visualisation.

4.3 Implications for Future Research

The revolution of large-scale data challenges people to become better informed about the ways in which they can harness vast bodies of data rather than small datasets, and simultaneously harness the technology. This attention to graphical developments increases the need to research the psychological aspects of data visualisation. These understandings will provide us with feedback about how students reason when using graphical displays, what aspects of formal inference are needed given current visualisation tools, and which methods foster students' ability to understand conventional formal conceptions and characteristics of large-scale data sets. Ideally, there should be further progress in the formal theory of data visualisation. Nevertheless, current growth of the field already leads to the challenges of integrating data visualisation in statistics education for students so that they are enabled to become competent citizens in the large-scale, big data era.

One crucial issue related to this process that was outside the scope of this study was the question of how people can use visualization to recognize biased or otherwise distorted data. One major concern for our society is the potential misuse of what might be called "big data," which, in contrast to the open data provided by governments and researchers, is proprietary and used for the profit of large corporations. The scope of the study only allowed for consideration of the students' experiences with large-scale open data, and not with big data. Further study could build on the foundation provided here to examine students' interactions with big data. In pursuing that research, there is an important role for visualization technologies that were not incorporated into this study. The reasoning about big data used by expers is different from common reasoning, because of the inherent complexity of data, and supporting dynamic visualisations of data are required. As Prodromou (2014) showed, 14- to 16-year-old students interpreted representations of multivariate data generated by a dynamic visualisation tool while they constructed their own meaningful data visualizations that highlighted emerging important aspects of data. Such a use of visualisation tools promoted students' articulations of the diverse inferences from data visualisations and reasoning about covariation between multiple variables while using the cycle of inquiry and visual analysis. In that study, students revisited their specific inferences while using complex data visualisation tools, inventing and revising their visual representations of data. Once they obtained some necessary insight, they readily made an informed decision.

References

Batanero, C, Estepa, A,. Godino, J. D., & Green, D. (1996). Intuitive strategies and preconceptions about association in contingency tables. *Journal for Research in Mathematics Education, 27*(2), 151-169. http://dx.doi.org/10.2307/749598

Ben-Zvi, D., & Sharett-Amir, Y. (2005). How do Primary School Students Begin to Reason about Distributions? In K. Makar (Ed.), Reasoning about distribution: A collection of current research studies. *Proceedings of the Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, University of Auckland, New Zealand.

Bright, G.W., & Friel, S. N. (1998). Graphical representations: Helping students interpret data. In S. P. Lajoie (Ed.), *Reflections on Statistics: Learning, Teaching, and Assessment in Grades K-1 2* (pp. 63-88). Mahwah, NJ: Lawrence Erlbaum Associates.

Brizuela, B. M., & Alvarado, M. (2010). First graders' work on additive problems with the use of different notational tools. Revista IRICE, 21, 37-43.

Brizuela, B. M., & Lara-Roth, S. (2002). Additive relations and function tables. *Journal of Mathematical Behavior, 20*(3), 309-319. http://dx.doi.org/10.1016/S0732-3123(02)00076-7

Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18*, 382–393. http://dx.doi.org/10.2307/749086

Estepa, A., Bataneo, C., & Sanchez, F. T. (1999). Students' intuitive strategies in judging association when comparing two samples. *Hiroshima Journal of Mathematics Education, 7*, 17-30.

Estrella, S., & Mena, A., (2014). Vergnaud's theory applied to basic school students' statistical representations. In K. Makar, B. de Sousa, & R. Gould (Eds.), Sustainability in statistics education. *Proceedings of the Ninth International Conference on Teaching Statistics (ICOTS9, July, 2014)*, Flagstaff, Arizona, USA. Voorburg, The Netherlands: International Statistical Institute.

Friel, S. N., Curcio, F., & Bright, G. W. (2001). Making sense of graphs: critical factors influencing comprehension and instructional implications. *Journal for Research in mathematics Education, 32* (2), 124-158. http://dx.doi.org/10.2307/749671

Gabucio, F., Martí, E., Enfedaque, J., Gilabert, S., & Konstantinidou, A. (2010). Niveles decomprensión de las tablas en alumnos de primaria y secundaria. *Cultura y Educación, 22* (2), 183-197. http://dx.doi.org/10.1174/113564010791304528

Gal, I. (2004). Statistical literacy: Meanings, components, responsibilities. In J. Garfield & D. Ben-Zvi (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 47-78). Dordrecht, The Netherlands: Kluwer. http://dx.doi.org/10.1007/1-4020-2278-6_3

Gal, I. (2002). Adult statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*(1), 1-25. http://dx.doi.org/10.1111/j.1751-5823.2002.tb00336.x

Kemp, M., & Kissane, B. (2010). A five step framework for interpreting tables and graphs in their contexts. In C. Reading (Ed.), Data and context in statistics education: Towards an evidence-based society. *Proceedings of the Eighth International Conference on Teaching Statistics* (ICOTS8, July, 2010), Ljubljana, Slovenia. Voorburg, The Netherlands: International Statistical Institute. Online: www.stat.auckland.ac.nz/~iase/publications.php

Martí, E. (2009). Tables as cognitive tools in primary education. *Representational systems and practices as learning tools* (pp. 133-148). Rotterdam: Sense Publishers.

Martí, E., García-Mila, M., Gabucio, F., & Konstantinidou, K. (2010). The construction of a double-entry table: a study of primary and secondary school students' difficulties. *European Journal of Psychology of Education, 26* (2), 215-234. http://dx.doi.org/10.1007/s10212-010-0046-1

Monteiro, C. & Ainley, J. (2007). Investigating the Interpretation of Media Graphs among Student Teachers. *International Electronic Journal of Mathematics Education, 2* (3), 188-207.

Pereira-Mendoza, L. & Mellor, J. (1991). Students' concepts of bar graphs–Some preliminary findings. In D. Vere-Jones (Ed.), *Proceedings of the Third International Conference on Teaching Statistics. Vol. 1. School and General Issues* (pp. 150-157). Voorburg, The Netherlands: International Statistical Institute.

Prodromou, T. (2014). Drawing inference from data visualisations. *International Journal of Secondary Education, 2* (4), 66-72. http://dx.doi.org/10.11648/j.ijsedu.20140204.12

Prodromou, T. (2013). Data Visualisation and Statistics from the Future. *Proceedings of the 59th ISI World Statistics Congress* (pp. 150-157), Hong Kong, China: International Statistical Institute (ISI). Online: http://2013.isiproceedings.org/Files/IPS049-P3-S.pdf

Ridgway, J., Nicholson, J., & McCusker, (2013). 'Open data' and the semantic web require a rethink on statistics education. *Technology Innovations in Statistics Education, 7* (2). Online: http://escholarship.org/uc/item/6gm8p12m

Sharma, S., (2013). Assessıng Students' Understandıng of Tables and Graphs: Implıcatıons for Teachıng and Research. *International Journal of Educational Research and Technology, 4*(4), 61-69.

Sharma, S. (1997). Statistical ideas of high school students: Some findings from Fiji. Unpublished doctoral thesis. Waikato University, Hamilton, New Zealand.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester, Jr., (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 957-1009). Reston: The National Council of Teachers of Mathematics.

Zawojewski, J. S., & Shaughnessy, J. M. (2000). Mean and median: Are they really so easy? *Mathematics Teaching in the Middle School, 5* (7), 436–440.

Appendix A

Workshop

1.1. Overview

This activity examines the components of population change for EU and candidate countries. It uses data published by EUROSTAT

http://epp.eurostat.ec.europa.eu/portal/page/portal/about_eurostat/introduction

Eurostat is the statistical office of the European Union situated in Luxembourg. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions.

The change of population for the period 2001 **-**2012 is examined, graphed and mapped. The opportunity exists to discuss variations in growth rates among different countries.

1.1.1 How many people live in EU and candidate countries? How many people live in your country?

1.1.2 As a group complete the table below:

	Size of the population of EU and candidate countries
	Size of your country's population
	Size of your country's population
	Size of your country's population

Eurostat is the statistical office of the European Union situated in Luxembourg. Its task is to provide the European Union with statistics at European level that enable comparisons between countries and regions.

Discuss the concept of population growth and explain the meaning of the following terms:

- Number of live births.
- Immigration
- Emigration
- Population by citizenship - Foreigners

1.1.3 Representing the Growth of Population in your country

1.1.4 Complete Table

TABLE 1. POPULATION change in your country from 2001-2013

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
Number of live births.												
Immigration												
Emigration												
Population by citizenship – Foreigners												

Using the data in Table 1:

Create one or more graphs that compare the following indicators of population change in your country: number of live births, Immigration, Emigration, population by citizenship-foreigners.

After studying the graph(s) and the table, write at least 10 lines explaining your observations regarding growth change in your country from 2001-2012, discussing possible social, historical, environmental, economic and political factors that might have caused this change. Identify and justify the dominant factors.

Share the data and your graphs with students from another country and compare your data and graphs, explaining (by writing at least 10 lines) your observations regarding similarities and/or differences in growth change between your countries from 2001-2012. Discuss and compare the possible social, historical, environmental, economic and political factors that might have caused these similarities and/or differences. Identify and justify the dominant factors.

1.2. EXPLAINING IMMIGRATION TRENDS

Study the following table (Immigration Table):

http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&language=en&pcode=tps00176&plugin=1

Identify: (a) the EU and candidate countries with a considerable decrease in the total number of long-term immigrants into the country;

(b) the EU and candidate countries with a considerable increase in the total number of long-term immigrants into the country;

1.2.1 Provide possible explanations as to why the countries you identified in part (a) are experiencing a loss of migrants, whereas the countries identified in part (b) are experiencing high increases in immigration.

APPENDIX B

Eurostat – Tables, Graphs and Maps Interactive (TGM) Table printer Prieview Page 1of 1

Number of live births

geo time	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
EU (28 countries)	5062948	5033705	5080465	5157173	5176850	5264100	5323425	5469434	5412572	5411129	5266162	5231177
EU (27 countries)	5021955	4993611	5040797	5116866	5134358	5222654	5281515	5425681	5367995	5367768	5224965	5189406
Euro area (18 countries)	3320208	3314946	3332210	3370485	3360486	3400181	3417552	3492593	3435284	3442269	3363927	3319557
Euro area (17 countries)	3300482	3294819	3311059	3349934	3338607	3377310	3393594	3468196	3413240	3422488	3345102	3299660
Belgium	115372	113096	114005	117295	119622	122529	124095	127205	127198	130100	128705	128051
Bulgaria	68180	66499	67359	69886	71075	73978	75349	77712	80956	75513	70846	69121
Czech Republic	90715	92786	93685	97664	102211	105831	114632	119570	118348	117153	108673	108576
Denmark	65458	64149	64682	64609	64282	64984	64082	65038	62818	63411	58998	57916
Germany	734475	719250	706721	705622	685795	672724	684862	682514	665126	677947	662685	673544
Estonia	12632	13001	13036	13992	14350	14877	15775	16028	15763	15825	14679	14056
Ireland	57854	60503	61529	61972	61372	65425	71389	75173	75554	75174	74033	72225
Greece	102282	103569	104420	105655	107545	112042	111926	118302	117933	114766	106428	100371
Spain	405313	417688	440531	453172	464811	481295	491138	518503	493717	485252	470553	453348
France	804052	793606	793893	800240	807787	830288	819605	829311	825564	833654	824263	821844
Croatia	40993	40094	39668	40307	42492	41446	41910	43753	44577	43361	41197	41771
Italy	535282	538198	544063	562599	554022	560010	563933	576659	568857	561944	546585	534186
Cyprus	8167	7883	8088	8309	8243	8731	8575	9205	9608	9801	9622	(10161)
Latvia	19726	20127	21151	20551	21879	22871	23958	24397	22044	19781	18825	19897
Lithuania	31185	29541	29977	29769	29510	29606	30020	31536	32165	30676	30268	30459
Luxembourg	5459	5345	5303	5452	5371	5514	5477	5596	5638	5874	5639	6026
Hungary	97047	96804	94647	95137	97496	99871	97613	99149	96442	90335	88049	90269
Malta	3957	3918	4050	3887	3858	3779	3765	4013	4029	3898	4165	4130
Netherlands	202603	202083	200297	194007	187910	185057	181336	184634	184915	184397	180060	175959
Austria	75458	78399	76944	78968	78190	77914	76250	77752	76344	78742	78109	78952
Poland	368205	353765	351072	356131	364383	374244	387873	414499	417589	413300	388416	386257
Portugal	112774	114383	112515	109298	109399	105449	102492	104594	99491	101381	96855	89841
Romania	220368	210529	212459	216261	221020	219483	214728	221900	222388	212199	196242	201104
Slovenia	17477	17501	17321	17961	18157	18932	19823	21817	21856	22343	21947	21938
Slovakia	51136	50841	51713	53747	54430	53904	54424	57360	61217	60410	60813	55535
Finland	56189	55555	56630	57758	57745	58840	58729	59530	60430	60980	59961	59493
Sweden	91466	95815	99157	100928	101346	105913	107421	109301	111801	115641	111770	113177
United Kingdom	669123	668777	695549	715996	722549	748563	772245	794383	790204	807271	807776	812970
Iceland	4091	4049	4143	4234	4280	4415	4560	4835	5026	4907	4492	4533
Liechtenstein	401	395	347	372	381	361	351	350	406	329	395	357
Norway	56696	55434	56458	56951	56756	58545	58459	60497	61807	61442	60220	60255
Switzerland	72295	72372	71848	73082	72903	73371	74494	76691	78286	80290	80808	82164
Montenegro	8839	8499	8344	7849	7352	7531	7834	8258	8642	7418	7215	7459
Former Yugoslav Republic of Macedonia, the	27010	27761	27011	23361	22482	22585	22688	22945	23684	24296	22770	23568
Serbia	78435	78101	79025	78186	72180	70997	68102	69083	70299	68304	65598	67257
Turkey	1362000	1362000	1361000	1360000	1361000	1362000	1266503	1262333	1263289	1238970	1241412	1279864
Albania	54283	45515	47012	43022	39612	34229	33163	36251	:	:	32628	:
Bosnia and Herzegovina	37717	36485	35234	34167	34627	34033	33835	34176	34550	33528	31875	32072 (p)
Kosovo (under United Nations Security Council Resolution 1244/99)	:	36136	31994	35063	37218	34187	33112	34399	34458	33751	27626	27743

:=not available p=provisional

Source of Data Eurostat
Last update: 24.04.2014
Date of extraction: 25 Апр 2014 15:05:53 CEST
Hyperlink to the table: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tps00111
General Disclaimer of the EC website: http://ec.europa.eu/geninfo/legal_notices_en.htm
Short Description: Live births are the births of children that showed any sign of life (total births minus stillbirths).
Code: tps00111

APPENDIX C

Eurostat – Tables, Graphs and Maps Interactive (TGM) Table printer Prieview Page 1of 1

Crude rate of population change

Per 1 000 inhabitants

geo \ time	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
EU (28 countries)	1.9 (b)	3.6	3.9	4.2	3.7	3.6	4.1 (b)	3.5	2.4	2.5	2.2 (b)	2.2 (b)
EU (27 countries)	1.9 (b)	3.6	3.9	4.2	3.8	3.6	4.1 (b)	3.6	2.4	2.5	2.2 (b)	2.3 (b)
Euro area (18 countries)	4.9 (b)	5.6	5.5	5.8	5.0	4.8	6.0	4.2	2.4	2.6	2.3 (b)	2.3 (b)
Euro area (17 countries)	5.1 (b)	5.7	5.7	5.9	5.1	4.9	6.1	4.3	2.6	2.8	2.5 (b)	2.4 (b)
Belgium	4.5	4.5	3.9	4.7	6.3	6.9	7.7	8.0	8.0	14.7	8.5 (b)	6.0
Bulgaria	-32.2	-5.8	-5.7	-5.2	-5.5	-5.1	-7.2 (b)	-6.8	-6.1	-7.1	-5.7	-5.8
Czech Republic	-3.0	-0.8	0.3	0.3	2.4	3.0	8.7	7.9	3.5	2.4	1.8	1.0
Denmark	3.6	2.8	2.6	2.5	3.0	3.6	5.3	6.5	4.2	4.7	3.6	4.0
Germany	2.2	1.2	-0.1	-0.4	-0.8	-1.5	-1.2	-2.6	-2.4	-0.6	1.1	2.4
Estonia	-6.6	-6.0	-6.5	-5.4	-6.0	-5.8	-3.3	-2.0	-1.8	-2.7	-3.3	-3.8
Ireland	17.3	16.4	16.2	20.3	23.2	30.9	26.7	14.2	6.2	4.7	2.6	1.8
Greece	3.1	2.8	3.5	3.3	3.5	2.8	3.4	0.8	-0.6	-5.4	0.0	-5.5
Spain	13.7	19.1	17.1	17.4	16.3	17.4	19.6	12.4	5.3	3.9	3.2	-1.9
France	7.3	7.1	6.9	7.7	7.3	6.5	5.7	5.3	4.8	4.9	4.7	4.4 (b)
Croatia	2.3 (b)	0.0	0.1	1.2	0.4	0.2	-0.4	-0.5	-1.6	-3.0	-3.2	-3.2
Italy	0.5	2.5	6.4	6.6	3.3	2.7	7.3	5.9	3.2	2.9	0.5	4.9
Cyprus	11.4	11.5	12.8	14.0	14.8	18.5	24.0	26.2	27.5	24.8	26.2	4.5
Latvia	-13.9	-9.3	-10.0	-11.8	-9.8	-8.6	-7.7	-13.3	-19.8	-21.9	-14.5	-10.3
Lithuania	-9.3	-6.7	-9.5	-12.9	-19.7	-12.2	-11.6	-9.0	-13.2	-28.9	-16.2	-10.6
Luxembourg	11.4	9.5	14.7	13.7	16.9	15.0	15.9	19.9	17.2	19.3	24.7	23.0 (b)
Hungary	-2.5	-3.2	-2.5	-1.9	-2.1	-1.0	-2.1	-1.4	-1.7	-2.9	-2.8	-2.3 (b)
Malta	8.2 (b)	6.7	6.5	7.0	5.8	1.5	5.4	7.6	7.5	2.3	6.1	9.1
Netherlands	7.4	5.4	4.0	2.9	1.8	1.5	2.9	4.9	5.4	4.9	4.5	2.9
Austria	5.3	4.5	5.2	7.2	6.4	3.5	4.3	4.4	2.4	3.5	0.5	5.2
Poland	-0.3	-0.6	-0.7	-0.4	-0.4	-0.8	-0.3	0.5	0.8	0.9	0.2 (b)	-0.1
Portugal	6.2	4.8	2.7	2.1	1.6	2.0	2.0	0.9	1.0	-0.1	-2.9	-5.2
Romania	-27.0	-9.5	-4.9	-6.5	-5.9	-6.0	-23.7	-9.5	-7.1	-4.7	-5.1	-1.9
Slovenia	2.0	0.5	0.7	0.6	2.9	3.5	7.7	10.9 (b)	7.2	1.6	2.6	1.6
Slovakia	0.0	-0.8	-0.6	0.2	0.0	0.0	0.5	1.2	1.5	0.4	2.2	1.2
Finland	2.7	2.2	2.6	3.2	3.6	4.1	4.4	4.9	4.7	4.4	4.8	4.7
Sweden	3.0	3.5	3.9	4.0	4.0	7.2	7.6	8.0	9.1	8.0	7.1	7.7
United Kingdom	4.1	4.4	4.9	6.5	7.3	7.4	8.1	7.6	7.5	8.2	7.5	6.3
Iceland	11.3	6.6	7.2	10.3	21.3	25.6	25.0	12.3	-5.5	2.6	3.5	7.1
Liechtenstein	19.9	10.0	12.6	8.9	8.8	7.5	5.3	6.6	8.5	7.1	9.0	9.9
Norway	4.6	6.2	5.5	6.3	7.3	8.8	11.9	13.0	12.2	12.7	13.2	13.0
Switzerland	7.1	8.0	6.9	6.9	5.9	6.6	11.2	14.2	10.8	10.0	10.7 (b)	10.6
Montenegro	3.7	3.6	3.2	2.8	1.0	2.1	4.2	4.2	4.4	2.9 (b)	2.2	2.5
Former Yugoslav Republic of Macedonia, the	3.7	-7.4	3.1	2.6	1.6	1.7	1.6	1.7	2.0	2.2	1.2	1.2
Serbia	-0.3	-1.5	-2.8	-1.9	-4.1	-3.8	-4.4	-4.2	-3.9	-4.2	-4.8 (b)	-4.9
Turkey	13.8	13.4	13.1	12.9	12.6	12.4	12.8	13.1	14.5	15.9	13.5	12.0
Albania	6.8	6.0	5.4	4.9	4.5	1.1	5.5	4.6	:	:	:	:
Bosnia and Herzegovina	6.2	4.5	1.8	1.3	0.0	0.4	0.0	0.0	0.0	-0.3	-1.0	-0.9 (p)
Kosovo (under United Nations Security Council Resolution 1244/99)	:	:	15.5	12.3	28.5	12.6	12.4	12.7	12.5	:	2.5 (b)	9.4

:=not available b=break in time series p=provisional

Source of Data Eurostat
Last update: 24.04.2014
Date of extraction: 25 Апр 2014 15:03:41 CEST
Hyperlink to the table: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tps00006
General Disclaimer of the EC website: http://ec.europa.eu/geninfo/legal_notices_en.htm
Short Description: The crude rate of population change is the ratio of the population change during the year to the average population in that year. The value is expressed per 1 000 inhabitants. Population change is the difference between the population sizes on 1 January of two consecutive years.
Code: tps00006

APPENDIX D

Eurostat – Tables, Graphs and Maps Interactive (TGM) Table printer Prieview Page 1of 1

Immigration

Persons

geo \ time	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
EU (28 countries)	:	:	:	:	:	:	:	:	:	:	:	:
EU (27 countries)	:	:	:	:	:	:	:	:	1731100	1811300	1750600	1665900
Euro area (18 countries)	:	:	:	:	:	:	:	:	:	:	:	:
Euro area (17 countries)	:	:	:	:	:	:	:	:	:	:	:	:
Belgium	110410	113857	112060	117236	132810	137699	146409	164152	:	135281	144698	147387
Bulgaria	:	:	:	:	:	:	1561	1236	:	:	:	14103
Czech Republic	12918	44679	60015	53453	60294	68183	104445	108267	75620	48317	27114	34337
Denmark	55984	52778	49754	49860	52458	56750	64656	57357	51800	52236	52833	54409
Germany	879217	842543	768975	780175	707352	661855	680766	682146	346216	404055	489422	:
Estonia	241	575	967	1097	1436	2234	3741	3671	3884	2810	3709	2639
Ireland	64925	61725	58875	78075	102000	139434	122415	82592	50604	52339	53224	54439
Greece	:	:	:	:	:	:	:	:	:	119070	110823	110139
Spain	414772	483260	672266	684561	719284	840844	958266	599075	392962	360705	371331	304053
France	:	:	:	:	:	301544	293980	296608	296970	307111	319816	327431
Croatia	24415	20365	18455	18383	:	14978	14622	14541	:	:	8534	8959
Italy	208252	222801	470491	444566	325673	297640	558019	534712	442940	458856	385793	350772
Cyprus	17485	14370	16779	22003	24419	15545	19017	14095	11675	20206	23037	17476
Latvia	5376	6642	4063	4844	6691	8212	7517	4678	3731	4011	10234	13303
Lithuania	4694	5110	4728	5553	6789	7745	8609	9297	6487	5213	15685	19843
Luxembourg	12135	12101	13158	12872	14397	14352	16675	17758	15751	16962	20268	20478
Hungary	22079	19855	21327	24298	27820	25732	24361	37652	27894	25519	28018	33702
Malta	472	533	:	:	187	1829	6730	6043	6161	4275	5465	7111
Netherlands	133404	121250	104514	94019	92297	101150	116819	143516	122917	126776	130118	124566
Austria	89928	108125	111869	122547	114465	98535	72862	73772	69295	70978	82230	91557
Poland	6625	6587	7048	9495	9364	10802	14995	47880	189166	155131	157059	217546
Portugal	74800	79300	72400	57920	49200	38800	46300	29718	32307	27575	19667	14606
Romania	:	:	:	:	:	:	:	138929	135844	149885	147685	167266
Slovenia	7803	9134	9279	10171	15041	20016	29193	30693	30296	15416	14083	15022
Slovakia	2023	2312	6551	10390	9410	12611	16265	17820	15643	13770	4829	5419
Finland	18955	18113	17838	20333	21355	22451	26029	29114	26699	25636	29481	31278
Sweden	60795	64087	63795	62028	65229	95750	99485	101171	102280	98801	96467	103059
United Kingdom	372206	385901	431487	518097	496470	529008	526714	590242	566514	590950	566044	498040
Iceland	5002	4215	3704	5350	7773	9832	12546	10288	3921	3948	4073	4960
Liechtenstein		:	:	:	:	:	:	578	584	591	650	671
Norway	34263	40122	35957	36482	40148	45776	61774	58123	55953	69214	70337	69908
Switzerland	122494	126080	119783	120188	118270	127586	165634	184297	160623	161778	148799	149051
Montenegro	:	:	:	:	:	:	:	:	:	:	:	:
Former Yugoslav Republic of Macedonia, the	458	2259	:	1718	2671	2077	1320	1052	:	:	1464	1715
Serbia	:	:	:	:	:	:	:	:	:	:	:	:
Turkey	:	:	:	:	:	:	:	:	22470	29905	27273	:
Bosnia and Herzegovina	:	:	:	:	:	:	5620	:	:	:	:	:

:=not available b=break in time series d=definition differs, see metadata p=provisional e=estimated

Source of Data Eurostat
Last update: 24.04.2014
Date of extraction: 25 Апр 2014 14:30:24 CEST
Hyperlink to the table: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tps00176
General Disclaimer of the EC website: http://ec.europa.eu/geninfo/legal_notices_en.htm
Short Description: Total number of long-term immigrants into the reporting country during the reference year
Code: tps00176

APPENDIX E

Eurostat – Tables, Graphs and Maps Interactive (TGM) Table printer Prieview Page 1of 1

Emigration

Persons

geo \ time	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012
EU (28 countries)	:	:	:	:	:	:	:	:	:	:	:	:
EU (27 countries)	:	:	:	:	:	:	:	:	1149500 (bdp)	1167800 (bdp)	1253000 (bdp)	1302000 (bdp)
Euro area (18 countries)	:	:	:	:	:	:	:	:	:	:	:	:
Euro area (17 countries)	:	:	:	:	:	:	:	:	:	:	:	:
Belgium	75261	75960	79399	83895	86899	88163	91052	100275	:	66013 (b)	67475	74720
Bulgaria	:	:	:	:	:	:	2958	2112	:	:	:	16615 (p)
Czech Republic	21469	32389	34226	34818	24065	33463	20500	51478 (b)	61782	61069	55910	46106
Denmark	43980	43481	43466	45017	45869	46786	41566	38356 (b)	39899	41456	41593	43663
Germany	606494	623255	626330	697632	628399	639064	636854	737889	286582 (b)	252456	249045	:
Estonia	2175	2038	3073	2927	4610	5527	4384	4406 (b)	4658	5294	6214	6321
Ireland	25750	28375	27200	28675	34350	44409 (b)	48040	65934	69672	78099	87053	89436
Greece	:	:	:	:	:	:	:	:	:	119985	125984	154435
Spain	:	36605	64298 (b)	55092	68011	142296 (b)	227065	288432 (b)	380121	403377	409034	446606
France	:	:	:	:	:	189403 (b)	220354	239796	264631	269531 (b)	280556 (p)	288331 (p)
Croatia	7488	11767	6534	6812	:	7692	9002	7488	:	:	12699	12877
Italy	56077	49383	62970	64849	65029	75230	65196	80947 (p)	80597 (p)	78771	82461	106216
Cyprus	13909	7485	4437	6279	10003	6874	11389	10500	9829	4293	4895	18105
Latvia	24539	15837	15647	20167	17643	17019	15463	27045	38208	39651	30311 (b)	25163
Lithuania	27841 (b)	16719	26283	37691	57885	32390	30383	25750	38500	83157	53863	41100
Luxembourg	8824	9452	7746	8480	8287	9001	10674	10058	9168	9302	9264	10442
Hungary	2591	3126	3122	3820	3658	4314 (p)	4500	9591 (b)	10483	13365 (b)	15100	22880
Malta	:	96	:	:	:	1908 (e)	5029 (e)	3719 (b)	3868	4201	3806	4005
Netherlands	63318	66728	68885	75049	83399	91028	91287	90067	92825 (b)	95970	104201	110431
Austria	72654	74831 (b)	71996	71721	70133	74432	49898 (b)	51563	53244	51651	51197	51812
Poland	23368	24532	20813	18877	22242	46936	35480	74338 (b)	229320 (b)	218126	265798	275603
Portugal	9800 (p)	9300 (p)	8900 (p)	10680 (p)	10800 (p)	12700 (p)	26800 (p)	20357 (b)	16899	23760	43998	51958
Romania	:	:	:	:	:	:	:	302796 (b)	246626	197985	195551	170186
Slovenia	4811	7269	5867	8269	8605	13749	14943	12109 (b)	18788	15937	12024	14378
Slovakia	1011	1411	4777 (b)	6525	2784	3084	3570	4857	4753	4447	1863 (b)	2003
Finland	13153	12891	12083	13656	12369	12107	12443	13657	12151	11905	12660	13845
Sweden	32141	33009	35023	36586	38118	44908	45418	45294	39240	48853	51179	51747
United Kingdom	251369	305931	313960	310389	328408	369470	317587	427207 (b)	368177	339306	350703	321217
Iceland	4034	4490	3837	4820	3913	4577	7414	9144	6874 (b)	5459	4812	4758
Liechtenstein	:	:	:	:	:	:	:	:	:	428	467	439
Norway	26309	22948	24672	23271	21709	22053	22122	12976 (b)	17072	25835	20349	22693
Switzerland	82235	78425	76756	79726	82090	88218	90175	86130	86036	96839	96494 (b)	103881
Montenegro	:	:	:	:	:	:	:	:	:	:	:	:
Former Yugoslav Republic of Macedonia, the	312	:	:	669	1300	1108	240	751	:	:	1290	1415
Serbia	:	:	:	:	:	:	:	:	:	:	:	:
Turkey	:	:	:	:	:	:	:	:	:	:	:	:

:=not available b=break in time series d=definition differs, see metadata p=provisional e=estimated

Source of Data Eurostat

Last update: 24.04.2014

Date of extraction: 25 Απρ 2014 14:31:06 CEST

Hyperlink to the table: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tps00177

General Disclaimer of the EC website: http://ec.europa.eu/geninfo/legal_notices_en.htm

Short Description: Total number of long-term emigrants from the reporting country during the reference year

Code: tps00177

APPENDIX F

Eurostat – Tables, Graphs and Maps Interactive (TGM) Table printer Prieview Page 1of 1

Population by citizenship - Foreigners

Persons

geo \ time	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
EU (28 countries)	:	:	:	:	:	:	:	:	:	:	:	:
EU (27 countries)	:	:	:	:	:	:	:	:	:	20502067 (bdp)	20679035 (bdp)	20370366 (bdp)
Euro area (18 countries)	:	:	:	:	:	:	:	:	:	:	:	:
Euro area (17 countries)	:	:	:	:	:	:	:	:	:	:	:	:
Belgium	846734	850077	860287	870862	900473 (e)	932161	971448	1009055 (p)	1052844 (p)	1162608 (b)	1224904	1253902
Bulgaria	:	:	:	:	26000 (e)	25500 (e)	36265	37132	38002	38815	39432	45201
Czech Republic	163805	179154	195394	193480	258360	296236	347649	407541	424419	416737	422966	422280
Denmark	266729	265424	271211	267604	270051	278096	298450	320033	329797	345884	358714	374569
Germany	7318263	7347951	7341820	7287980	7289149	7255949	7255395	7185921	7130919	7198946	7409754	7696413
Estonia	:	:	:	:	242000 (e)	236400 (e)	229300 (e)	214437	212659	208038 (b)	206558	197141
Ireland	272643	329743	348797	386388	429281	498706 (b)	559021	579770	570190	560478	548915	543636
Greece	:	:	891197 (e)	:	884000 (e)	887600 (e)	906400 (e)	929530	954784	956007	975374	862381
Spain	1560724	2189213	2771962	3371394	4002509	4606474	5262095	5386661	5402578	5312444	5236030	5072680
France	:	3263186	:	3623063	3510000 (e)	3685544 (e)	3709814 (e)	3750406	3824590	3875096	3943700	4089051
Croatia	:	:	:	:	:	36200	37100 (e)	:	:	:	:	27854
Italy	:	1549373	1990159	2402157	2670514	2938922	3432651	3891295	4235059	4570317	4825573	4387721
Cyprus	66100	72500	83500	98100	110200 (e)	118100	125300 (e)	124649	163102	167783	172427	170076
Latvia	556801	534534	514966	487212	456758 (e)	432951	404876	382704	362378	342799	324288	315414
Lithuania	32665	33609	33708	33672	32685	32240	31998	30907	27318	24031	22865	22224
Luxembourg	166700	170700	177600	183600	191400	198213	205889	214848	215699	220705	229870	238844
Hungary	116429 (p)	115888	130109	143774	156160	167873	176580	186365	200005	209202	143125	141122
Malta	9564	10358	11000	11999	12000 (e)	13877	15460	16791	18952	19139	20302	22466
Netherlands	690393	699954	702185	699351	691357	681932	688375	637136	652188	673235	697741	714552
Austria	730261	746753	754216	774401	796666	804779	824974	852604	876068	905435	945176	997038
Poland	41375	41650	41950	42763	49499 (e)	54883	57842	48167	45464	47261 (p)	57450 (p)	58859
Portugal	224932	238746	:	:	276000 (e)	434887	446333	440277	454191	445262	436822	417042
Romania	25645	:	25645	25929	25993	26069	26100	31354	:	:	36536 (p)	70666
Slovenia	45273	44693	45294	44285	48968	53555	68621	70554 (b)	82176	82746	85555	91385
Slovakia	:	29854	29855 (b)	22251	25563	32130	40904	52545	62882	67976	70727 (b)	72925
Finland	98577	103682	107003	108346	113852	121739	132708	142288	154623	166627	181697	194250
Sweden	475986	474099	476076	481141	479899	491996	524488	547664	590475	622275	646095	659374
United Kingdom	:	2760031	2941400	3066055	3425000 (e)	3659900 (e)	4020800 (e)	4184106	4362174	4486644 (p)	4802331	4929710
Iceland	9850	10221	10180	10636	13778	:	23420	24379	21701	21143	20957	21446
Liechtenstein	:	11566	11786	11852	:	:	:	11770	11886	12004	12144	12337
Norway	185863	197668	204731	213303	222277	238305	266260	302908	331618	368475	409193	457396
Switzerland	1447553	1476966	1500907	1524663	1541912	1554527	1602093	1669715	1714004	1765750 (b)	1815063	1869070
Montenegro	:	:	:	:	:	:	:	:	:	44324	:	:
Former Yugoslav Republic of Macedonia, the	:	:	:	:	:	:	:	:	:	:	:	:
Serbia	:	:	:	:	:	:	:	:	:	:	:	:
Turkey	:	:	:	:	292000	:	98064	103753	167344	175384	235067	272842

:=not available b=break in time series d=definition differs, see metadata p=provisional e=estimated

Source of Data Eurostat

Last update: 24.04.2014

Date of extraction: 25 Απρ 2014 15:08:16 CEST

Hyperlink to the table: http://epp.eurostat.ec.europa.eu/tgm/table.do?tab=table&init=1&plugin=1&language=en&pcode=tps00157

General Disclaimer of the EC website: http://ec.europa.eu/geninfo/legal_notices_en.htm

Short Description: Total number of foreigners including citizens of other EU Member States and non-EU citizens, usually resident in the reporting country. January, 1

Code: tps00157