

ProCivicStat © - Teacher's Manual, 5.409

## Gender equity and pay gap

Anna Trostianitser | Iddo Gal

[Anna.Trostianitser@gmail.com](mailto:Anna.Trostianitser@gmail.com) | [iddo@research.haifa.ac.il](mailto:iddo@research.haifa.ac.il)

Haifa University, Israel

Pedro Campos | Sónia Teixeira | Paula Lopes

[pcampos@fep.up.pt](mailto:pcampos@fep.up.pt) | [soniaacteixeira@gmail.com](mailto:soniaacteixeira@gmail.com) | [rosyfogo09@gmail.com](mailto:rosyfogo09@gmail.com)

Porto University, Portugal

<b>What?</b>	Teaching Material for learning statistics about society: <b><i>Possible reasons and consequences for gender equity and pay gap.</i></b>
<b>Why?</b>	Learning statistics with real data and motivating topics of high relevance for informed citizenship and civic engagement.
<b>Statistics topics</b>	Students will be able to investigate basic notion of statistics: different types of variables, measures of location and dispersion, descriptive statistics.
<b>Level</b>	College
<b>Prerequisites</b>	Knowledge of descriptive statistics, Plot of mean, average calculation of a quantitative variable in function to a group (qualitative variable)
<b>Digital tools</b>	R or R-Studio and R-Commander
<b>Resources needed</b>	R, R-Studio or R-Commander needs to be installed on PCs.
<b>Lesson time</b>	4-8 hours
<b>Further remark</b>	Suitable for individual/group work during class or homework assignment. Intro to the software should be provided previously.

### What is Civic Statistics?

To be fully engaged, citizens need to be aware of and understand statistics regarding past trends, present situations, and possible future changes in demographics, employment, wages, migration, health, crime, poverty, access to services, energy, education, human rights, and other domains. Statistics and data about these and related topics are collectively called here civic statistics. Information about civic statistics is provided by official statistics agencies and



other public and private/non-profit statistics providers, and some of it is mediated to the public via print and visual media. Understanding of civic statistics is required for participation in democratic societies, but involves data that often are open, large-scale, official, multivariate in nature, and/or presented dynamically. Such statistics are usually not at the core of regular statistics instruction. These materials fill an important gap, not addressed in general education at school or university level.

ProCivicStat is a cooperative project of six partners in five countries, funded through the Erasmus+ program of the European Union. ProCivicStat developed new methods for statistics instruction for high schools and universities that will contribute to young people's ability to understand quantitative evidence about key social phenomena that permeate civic life. ProCivicStat offers a platform for continuing cooperation, exchange of ideas, exploration and dissemination of theoretical concepts and concrete teaching materials for promoting civic engagement via exploration of evidence and understanding of statistics about society.

For more information and material, visit the ProCivicStat website: <http://www.procivicstat.org>

## Introduction:

This document is intended for college mathematics teachers.

After reading this manual, you will have a deeper understanding on the following topics:

- ✓ The importance of using real data at class
- ✓ The topics that can be covered when using this lesson plan
- ✓ Solutions of questions presented in this lesson plan
- ✓ Programs that can be used
- ✓ Advantages / Disadvantages of a particular program
- ✓ Manual for acquiring basic knowledge about preferred software and additional information.

### 1. Why should the teacher use real data in class?

By using real databases on a particular subject, we will have a better understanding of the reality that we live in. And when we use real database we will have the understanding if the results achieved do or do not make sense.

### 2. What topics will your students learn if the teacher uses this lesson plan?

- Basic notion of statistics
- Graphs and measures more suitable for each type of variables
- Descriptive statistics
- Basic notion of statistics
- Type of variables
- Measures of location and dispersion
- Descriptive statistics

### 3. What program should teachers use for analysis?

It is recommended to use the R commander and R or RStudio program to maintain this lesson plan. These programs are totally free and are used worldwide to make statistical analysis from the most basic to the most complex ones. R Commander is a R package that facilitates analysis primarily for people who have no previous knowledge of programming in R.



Link for more information on R Commander: <http://www.rcommander.com/>

Link to install the R: <https://www.r-project.org/>

Link to install the R-Studio: <https://www.rstudio.com/>

#### 4. Advantages vs. Disadvantages for using R Commander:

There are plenty of advantages for using R Commander for this lesson plan:

- R Commander is a free R package that can be installed anywhere in the world
- The R Commander has a friendly interface (identical to SPSS), therefore, you do not have to know how to program with R Commander
- R Commander accepts files from other programs like XLS, CSV, Access, Txt etc.
- It is possible to change the nature of the variables and to recode them
- It is possible to perform basic statistical analysis as well as a more complex one (such as, Chi-square association test, linear regression model, generalized linear regression model, principal component analysis, etc.)

With use of the video manual, in less than an hour, you are ready to work on R Commander.

#### Disadvantages:

The tests of normality that are implemented in R Commander allows to test only the normality for a quantitative variable. Thus, it does not allow to test the normality of a quantitative variable in function of groups (for example, to test the normality of the final grade by sex of students). But we can write the code in the R script of the R Commander, which allows this test to be performed.

#### 5. Where to start?

Proceed to the following pages to learn more about lesson plan as well as the use of R Commander programs for this matter.

**Appendix 1:** How to use R Commander.

**Appendix 2:** Worksheet with answers and explanations on every question how to do it in program and what to highlight.

#### Appendix 1: How to use R Commander.

**This section will give you the basic knowledge on R Commander.**

Proceed to the following link to learn more about R Commander:

<http://www.rcommander.com/>

In addition, see the "Introduction Chapter" in the following link: <https://cran.r-project.org/doc/contrib/Karp-Rcommander-intro.pdf>

#### Appendix 2: Worksheet with answers.

This part of the document will provide you with worksheet with answer and explanations on every question about program use and what to highlight.

### On Students Worksheets:

Please use the program R Commander to analyze the data and answer the following questions.

**Note:** Before importing this data set to the R Commander, first of all, exclude the 9999999 from the IncTo variable that correspond to the NIU (not in universe).

**Suggestion:** Use the Excel override function.

After importing this data set in the R commander, all variables in study appear in R Commander as numeric, but pay attention, that out of all 5 variables in this study, only one variable is numeric (**IncTot**). For this reason, the first task to be accomplished before continuing with the analysis itself, is to transform these variables into categorical.

- a) Transforms the categorical variables that are currently sorted as numeric in the program, into categorical ones. (See description of the variables in Annex I)

**Solution:** Go to **Data** and choose the **Manage variables in active data set** option and choose the option to **Convert numeric variables to factors** and choose the variables that we want to transform into factors or categories by pressing the control key and choose the option of **Use numbers** and click ok and choose the option to write over the existing.

### Questions:

1. For starters, you would like to have a first impression of the situation in which the Brazilian population is in. Calculate the mean, median, minimum, maximum, standard deviation, 1st and 3rd quartiles for the variable IncTot. Interpret these values in statistical terms and summarize it in your answer.

**Solution:** To answer this question go to **Statistics** and choose the **Summaries** option, then choose the **Numerical Summaries** option and select the variable **IncTot** and click ok.

**Observation:** The total of the variable **IncTot** is 274544 because in this variable were excluded in total 65645 that correspond to the 9999999 NIU (not in universe).

2. The economics section of the newspaper that you read recently published the following statement: “The highest average yield for women and men were recorded for age group of [40.44] years”. Do you agree with this statement? Justify your answer.

**Solution:** Make **Plot of means**. Go to **Graphs** and choose the **Plot of means** option and choose the **variables Age2 and Sex to Factors** (pressing the control key) and **IncTot to Response variable**.



3. It is common-knowledge that USA have a higher percentage of people in relation to Brazil. Analyze the data in this dataset and check if it's true.

**Solution:** To answer this question go to *Statistics* and choose the *Summaries* option, then choose the *Frequency distributions* option and select the variable *Country* and click *ok*.

4. Another common-knowledge is that the highest yield requires a more demanding position. Analyze the real data and check whether it is true that Legislators, senior officials and managers are those who have a higher average yield. Justify your answer.

**Solution:** To answer this question go to *Statistics* and choose the *Summaries* option, then choose the *Table of statistic* option and choose as to *Factors* the variable *OccISCO* and to *Response variable* *IncTot* and in *Statistic* choose *average* and click *ok*.

Alternatively, you can do *Plot of means*. Go to *Graphs* and choose the *Plot of means* option and choose *OccISCO* to *Factors* and *IncTot* to *Response variable*.

5. A colleague of yours from different channel, claims that higher average yields have been recorded for mining industries. Do you agree with this opinion? Please explain your answer.

**Solution:** The procedure is the same as in question 4 with the only difference of using variable *IndGen* instead of the *OccISCO* variable.

6. Analyze the average income per age group. **Note:** Begin to analyze the results of the age group 4, since it makes no sense to see the age group's income before the age group 4 (age group 0 to 14 years) (see the Annex I.). Does this result make any sense to you? Justify your answer.

**Solution:** The procedure is the same as in question 4 with the only difference of using variable *Age2* instead of the *OccISCO* variable.

- 6a) What might be the consequences for these results? Propose at least 2 consequences that government will need to deal with and how.

7. Calculate the average total income for the 2 countries in study. What can you conclude from this analysis? What might be the reasons for this result? Propose at least 2 reasons.

**Solution:** To answer this question go to *Statistics* and choose the *Summaries* option, then choose the *Table of statistic* option and choose as to *Factors* the variable *Country* and to *Response variable* *IncTot* and in *Statistic* choose *average* and click *ok*.

Alternatively, you can do *Plot of means*. Go to *Graphs* and choose the *Plot of means* option and choose *Country* to *Factors* and *IncTot* to *Response variable*.

8. The media presents the opinion that men have an average yield higher than that of women. Does this dataset support this point of view or are these reports biased? Propose at least two reasons that will explain this result.

**Solution:** The procedure is the same as in question 4 with the only difference of using variable *Sex* instead of the *OccISCO* variable.

9. From all the categories of the occupation variable (**OccISCO**), men always present a higher average yield than women, except for category 10 (Armed Forces). Does this result make any sense to you? Explain your answer.

**Note:** For the interpretation of the results discards the Categories 98 and 99 that correspond unknown and NIU (not in universe) respectively.

**Solution:** Make *Plot of means*. Go to *Graphs* and choose the *Plot of means* option and choose the variables *OccISCO* and *Sex* to *Factors* (pressing the control key) and *IncTot* to *Response variable*.

- 9a) Why do you think that gap happened? Propose at least 2 possible explanations.

10. The local media suggests that men have a higher average income than women in virtually all industrial sectors. Verify if the data supports this statement. Present at least two reasons for your results.

***Solution:** Make Plot of means. Go to Graphs and choose the Plot of means option and choose the variables IndGen and Sex to Factors (pressing the control key) and IncTot to Response variable.*

### Annex I- Description the variables- Gender equity

				Description
<b>Country</b>	Country (Brazil and USA)	Qualitative	Nominal	076-Brazil 840-USA
<b>Year</b>	Year of analysis	Qualitative	Ordinal	2000 2010
<b>Sample</b>	IPUMS sample identifier (id)	Qualitative	Nominal	<b>Sample</b> identifies the IPUMS sample from which the case is drawn. Each sample receives a unique 9-digit code. The code is structured as follows: - The first 3 digits are the ISO/UN codes used in COUNTRY -The next 4 digits are the year of the census/survey -The final 2 digits identify the sample within the year. For the last two digits, censuses or large census-like surveys have a value "0" (e.g, 01) in the second-to-last digit, household surveys have a value of "2" (e.g., 21), and employment surveys have a value of "4" (e.g., 41)
<b>Serial</b>	Household serial number (id)	Qualitative	Nominal	<b>Serial</b> is an identifying number unique to each household in a given sample
<b>PerNum</b>	Person number	Qualitative	Nominal	<b>PerNum</b> numbers all persons within each household consecutively (starting with "1" for the first person record of each household)
<b>PerWt</b>	Person weight	Qualitative	Nominal	<b>PerWt</b> indicates the number of persons in the actual population represented by the person in the sample

<b>Age2</b>	Age, grouped in intervals of 5 in 4 (17 groups)	Qualitative	Ordinal	<ul style="list-style-type: none"> <li>1-Age [0,4]</li> <li>2- Age [5,9]</li> <li>3-Age [10,14]</li> <li>4-Age [15,19]</li> <li>8-Age [20,24]</li> <li>9-Age [25,29]</li> <li>10-Age [30,34]</li> <li>11-Age [35,39]</li> <li>12-Age [40,44]</li> <li>13-Age [45,49]</li> <li>14-Age [50,54]</li> <li>15-Age [55,59]</li> <li>16-Age [60,64]</li> <li>17-Age [65,69]</li> <li>18-Age [70,74]</li> <li>19-Age [75,79]</li> <li>20-Age &gt;=80</li> </ul>
<b>Sex</b>	Sex	Qualitative	Nominal	<ul style="list-style-type: none"> <li>1-Male</li> <li>2-Female</li> </ul>
<b>EmpStat</b>	Activity status (employment status) [general version]	Qualitative	Nominal	<ul style="list-style-type: none"> <li>0-NIU (not in universe)</li> <li>1-Employed</li> <li>2-Unemployed</li> <li>3-Inactive</li> </ul>
<b>EmpStatD</b>	Activity status (employment status) [detailed version]	Qualitative	Nominal	<ul style="list-style-type: none"> <li>0-NIU (not in universe)</li> <li>110-At work</li> <li>120-Have job, not at work in reference period</li> <li>131-Armed forces, at work</li> <li>132-Armed forces, not at work in reference period</li> <li>200-UNEMPLOYED, not specified</li> <li>300-INACTIVE (not in labor force)</li> </ul>
<b>OccISCO</b>	Occupation, ISCO general	Qualitative	Nominal	<ul style="list-style-type: none"> <li>1-Legislators, senior officials and managers</li> <li>2-Professionals</li> <li>3-Technicians and associate professionals</li> <li>4-Clerks</li> <li>5-Service workers and shop and market sales</li> <li>6-Skilled agricultural and fishery workers</li> <li>7-Crafts and related trades workers</li> <li>8-Plant and machine operators and assemblers</li> <li>9-Elementary occupations</li> <li>10-Armed forces</li> <li>11- Other occupations, unspecified or n.e.c.</li> <li>98-Unknown</li> <li>99-NIU (not in universe)</li> </ul>





## Promoting Civic Engagement via Exploration of Evidence: Challenges for Statistics Education

Co-funded by the  
Erasmus+ Programme  
of the European Union



<b>IndGen</b>	Industry, general recode	Qualitative	Nominal	<p>0- NIU (not in universe)</p> <p>10-Agriculture, fishing, and forestry</p> <p>20-Mining</p> <p>30-Manufacturing</p> <p>40-Electricity, gas and water</p> <p>50-Construction</p> <p>60- Wholesale and retail trade</p> <p>70-Hotels and restaurants</p> <p>80-Transportation, storage and communications</p> <p>90-Financial services and insurance</p> <p>100-Public administration and defense</p> <p>111-Real estate and business services</p> <p>112-Education</p> <p>113-Health and social work</p> <p>114-Other services</p> <p>120-Private household services</p> <p>130-Other industry, n.e.c.</p> <p>999- Unknown</p>
<b>ClassWk</b>	Status in employment (class of worker) [general version]	Qualitative	Nominal	<p>0-NIU (not in universe)</p> <p>1- Self-employed</p> <p>2-Wage/salary worker</p> <p>3- Unpaid worker</p>
<b>ClassWkD</b>	Status in employment (class of worker) [detailed version]	Qualitative	Nominal	<p>0-NIU (not in universe)</p> <p>101-Self-employed, unincorporated</p> <p>102-Self-employed, incorporated</p> <p>110-Employer</p> <p>120-Working on own account</p> <p>123-Subsistence worker, own consumption</p> <p>200-WAGE/SALARY WORKER</p> <p>209-Employee without legal contract</p> <p>210-Wage/salary worker, private employer</p> <p>213-Wage/salary worker, non-profit, NGO</p> <p>220-Wage/salary worker, government</p> <p>221-Federal, government employee</p> <p>222-State government employee</p> <p>223-Local government employee</p> <p>230-Domestic worker (work for private household)</p> <p>310-Unpaid family worker</p> <p>340-Apprentice or trainee</p>
<b>IncTot</b>	Total income	Quantitative	Discrete	<p><b>IncTot</b> is a 7-digit numeric variable that reports the person's total personal income from all sources in the previous month or year. Observation: 9999999 = NIU (not in universe)</p>

ProCivicStat © - Student's Worksheet, 5.409

## Gender equity and pay gap

Anna Trostianitser | Iddo Gal

[Anna.Trostianitser@gmail.com](mailto:Anna.Trostianitser@gmail.com) | [iddo@research.haifa.ac.il](mailto:iddo@research.haifa.ac.il)

Haifa University, Israel

Pedro Campos | Sónia Teixeira | Paula Lopes

[pcampos@fep.up.pt](mailto:pcampos@fep.up.pt) | [soniaacteixeira@gmail.com](mailto:soniaacteixeira@gmail.com) | [rosyfogo09@gmail.com](mailto:rosyfogo09@gmail.com)

Porto University, Portugal



Source: <https://www.pexels.com/photo/close-up-of-coins-on-table-332304/>



### Introduction:

Suppose you're an American journalist of well-known TV channel, that broadcasts worldwide. The topic of your current scoop is gender pay gap and equity between sexes. Your boss suggested that it would be more interesting to compare different cultures, therefore, you were sent to Brazil. You prepared a list of topics and questions that you would like to focus on. But upon your arrival to Brazil, you understood that in order to show the situation in unbiased way, you have to understand the root of the situation and it means – to use basic statistics to analyze the data.

### Background:

In the high and middle-income countries that make up the OECD, the median wage of a woman working full-time is 85% than of a man. Despite a prevailing opinion, this is not because employers pay a woman less than they would have paid a man in her place. Data from 25 countries collected by Korn Ferry, a consultancy agency, show that women earn 98% as much as men who do the same job for the same employer. The real reason is twofold. Women outnumber men in positions with lower salaries and little chance of promotion. And men and women are segregated between occupations and industries; those where women predominate - pay less.

Women, besides earning less, still suffer more from unemployment. The survey shows a male unemployment rate of 7.7% while female unemployment rate is as high as 11.6%. According to Cristiane Soares, women have a lower job placement than men, because they have always been linked to the care of the home and children. When the market is hot, it seeks out these women as more labor, but when there is a crisis, the impact is greater for them.

## Variables and further information:

This data set represents the Gender equity / pay gap for Brazil and USA for the years 2000 and 2010. With this data set we can analyze what can influence the income of men and women.

This database contains 15 variables and 340189 records, as shown in the table below:

Country	Country (Brazil and USA)	Qualita-	Nominal
Year	Year	Qualita-	Ordinal
Sample	IPUMS sample identifier (id)	Qualita-	Nominal
Serial	Household serial number (id)	Qualita-	Nominal
PerNum	Person number	Qualita-	Nominal
PerWt	Person weight	Qualita-	Nominal
Age2	Age, grouped into intervals	Qualita-	Ordinal
Sex	Sex	Qualita-	Nominal
EmpStat	Activity status (employment status) [general ver-	Qualita-	Nominal
EmpStat	Activity status (employment status) [detailed	Qualita-	Nominal
OccISCO	Occupation, ISCO general	Qualita-	Nominal
IndGen	Industry, general recode	Qualita-	Nominal
ClassWk	Status in employment (class of worker) [general	Qualita-	Nominal
Class-	Status in employment (class of worker) [detailed	Qualita-	Nominal
IncTot	Total income	Quantita-	Discrete

However, for our analysis of the 15 variables we will use only the following 7:

**Country, Year, Age2, Sex, OccISCO, IndGen and IncTot.**

**Note:** For more information on the categories of variables see Annex I

For more information about this database see the link below:

**IPUMS:** <https://international.ipums.org/international-action/variables/group>

Please use the program R Commander to analyze the data and answer the following questions.

**Note:** Before importing this data set to the R Commander, first of all, exclude the 9999999 from the IncTo variable that correspond to the NIU (not in universe).

**Suggestion:** Use the Excel override function.

After importing this data set in the R commander, all variables in study appear in R Commander as numeric, but pay attention, that out of all 5 variables in this study, only one variable is numeric (**IncTot**). For this reason, the first task to be accomplished before continuing with the analysis itself, is to transform these variables into categorical.

- a) Transforms the categorical variables that are currently sorted as numeric in the program, into categorical ones. (See description of the variables in Annex I)

***Solution:** Go to **Data** and choose the **Manage variables in active data set** option and choose the option to **Convert numeric variables to factors** and choose the variables that we want to transform into factors or categories by pressing the control key and choose the option of **Use numbers** and click ok and choose the option to write over the existing.*

### Questions:

1. For starters, you would like to have a first impression of the situation in which the Brazilian population is in. Calculate the mean, median, minimum, maximum, standard deviation, 1st and 3rd quartiles for the variable IncTot. Interpret these values in statistical terms and summarize it in your answer.

2. The economics section of the newspaper that you read recently published the following statement: “The highest average yield for women and men were recorded for age group of [40.44] years”. Do you agree with this statement? Justify your answer.

3. It is common-knowledge that USA have a higher percentage of people in relation to Brazil. Analyze the data in this dataset and check if it's true.



4. Another common-knowledge is that the highest yield requires a more demanding position. Analyze the real data and check whether it is true that Legislators, senior officials and managers are those who have a higher average yield. Justify your answer.

5. A colleague of yours from different channel, claims that higher average yields have been recorded for mining industries. Do you agree with this opinion? Please explain your answer.

6. Analyze the average income per age group. Note: Begin to analyze the results of the age group 4, since it makes no sense to see the age group's income before the age group 4 (age group 0 to 14 years) (see the Annex I.). Does this result make any sense to you? Justify your answer.

- 6a) What might be the consequences for these results? Propose at least 2 consequences that government will need to deal with and how.



7. Calculate the average total income for the 2 countries in study. What can you conclude from this analysis? What might be the reasons for this result? Propose at least 2 reasons.

8. The media presents the opinion that men have an average yield higher than that of women. Does this dataset support this point of view or are these reports biased? Propose at least two reasons that will explain this result.

9. From all the categories of the occupation variable (**OccISCO**), men always present a higher average yield than women, except for category 10 (Armed Forces). Does this result make any sense to you? Explain your answer.

**Note:** For the interpretation of the results discards the Categories 98 and 99 that correspond unknown and NIU (not in universe) respectively.

- 9a) Why do you think that gap happened? Propose at least 2 possible explanations.

10. The local media suggests that men have a higher average income than women in virtually all industrial sectors. Verify if the data supports this statement. Present at least two reasons for your results.

### ***Annex I- Description the variables- Gender equity***

Description				
<b>Country</b>	Country (Brazil and USA)	Qualitative	Nominal	076-Brazil 840-USA
<b>Year</b>	Year of analysis	Qualitative	Ordinal	2000 2010
<b>Sample</b>	IPUMS sample identifier (id)	Qualitative	Nominal	<b>Sample</b> identifies the IPUMS sample from which the case is drawn. Each sample receives a unique 9-digit code. The code is structured as follows: - The first 3 digits are the ISO/UN codes used in COUNTRY -The next 4 digits are the year of the census/survey -The final 2 digits identify the sample within the year. For the last two digits, censuses or large census-like surveys have a value "0" (e.g, 01) in the second-to-last digit, household surveys have a value of "2" (e.g., 21), and employment surveys have a value of "4" (e.g., 41)
<b>Serial</b>	Household serial number (id)	Qualitative	Nominal	<b>Serial</b> is an identifying number unique to each household in a given sample
<b>PerNum</b>	Person number	Qualitative	Nominal	<b>PerNum</b> numbers all persons within each household consecutively (starting with "1" for the first person record of each household)
<b>PerWt</b>	Person weight	Qualitative	Nominal	<b>PerWt</b> indicates the number of persons in the actual population represented by the person in the sample



<b>Age2</b>	Age, grouped in intervals of 5 in 4 (17 groups)	Qualitative	Ordinal	<ul style="list-style-type: none"> <li>1-Age [0,4]</li> <li>2- Age [5,9]</li> <li>3-Age [10,14]</li> <li>4-Age [15,19]</li> <li>8-Age [20,24]</li> <li>9-Age [25,29]</li> <li>10-Age [30,34]</li> <li>11-Age [35,39]</li> <li>12-Age [40,44]</li> <li>13-Age [45,49]</li> <li>14-Age [50,54]</li> <li>15-Age [55,59]</li> <li>16-Age [60,64]</li> <li>17-Age [65,69]</li> <li>18-Age [70,74]</li> <li>19-Age [75,79]</li> <li>20-Age &gt;=80</li> </ul>
<b>Sex</b>	Sex	Qualitative	Nominal	<ul style="list-style-type: none"> <li>1-Male</li> <li>2-Female</li> </ul>
<b>EmpStat</b>	Activity status (employment status) [general version]	Qualitative	Nominal	<ul style="list-style-type: none"> <li>0-NIU (not in universe)</li> <li>1-Employed</li> <li>2-Unemployed</li> <li>3-Inactive</li> </ul>
<b>EmpStatD</b>	Activity status (employment status) [detailed version]	Qualitative	Nominal	<ul style="list-style-type: none"> <li>0-NIU (not in universe)</li> <li>110-At work</li> <li>120-Have job, not at work in reference period</li> <li>131-Armed forces, at work</li> <li>132-Armed forces, not at work in reference period</li> <li>200-UNEMPLOYED, not specified</li> <li>300-INACTIVE (not in labor force)</li> </ul>
<b>OccISCO</b>	Occupation, ISCO general	Qualitative	Nominal	<ul style="list-style-type: none"> <li>1-Legislators, senior officials and managers</li> <li>2-Professionals</li> <li>3-Technicians and associate professionals</li> <li>4-Clerks</li> <li>5-Service workers and shop and market sales</li> <li>6-Skilled agricultural and fishery workers</li> <li>7-Crafts and related trades workers</li> <li>8-Plant and machine operators and assemblers</li> <li>9-Elementary occupations</li> <li>10-Armed forces</li> <li>11- Other occupations, unspecified or n.e.c.</li> <li>98-Unknown</li> <li>99-NIU (not in universe)</li> </ul>



**Promoting Civic Engagement via Exploration of Evidence:  
Challenges for Statistics Education**

Co-funded by the  
Erasmus+ Programme  
of the European Union



<b>IndGen</b>	Industry, general recode	Qualitative	Nominal	0- NIU (not in universe) 10-Agriculture, fishing, and forestry 20-Mining 30-Manufacturing 40-Electricity, gas and water 50-Construction 60- Wholesale and retail trade 70-Hotels and restaurants 80-Transportation, storage and communications 90-Financial services and insurance 100-Public administration and defense 111-Real estate and business services 112-Education 113-Health and social work 114-Other services 120-Private household services 130-Other industry, n.e.c. 999- Unknown
<b>ClassWk</b>	Status in employment (class of worker) [general version]	Qualitative	Nominal	0-NIU (not in universe) 1- Self-employed 2-Wage/salary worker 3- Unpaid worker
<b>ClassWkD</b>	Status in employment (class of worker) [detailed version]	Qualitative	Nominal	0-NIU (not in universe) 101-Self-employed, unincorporated 102-Self-employed, incorporated 110-Employer 120-Working on own account 123-Subsistence worker, own consumption 200-WAGE/SALARY WORKER 209-Employee without legal contract 210-Wage/salary worker, private employer 213-Wage/salary worker, non-profit, NGO 220-Wage/salary worker, government 221-Federal, government employee 222-State government employee 223-Local government employee 230-Domestic worker (work for private household) 310-Unpaid family worker 340-Apprentice or trainee
<b>IncTot</b>	Total income	Quantitative	Discrete	<b>IncTot</b> is a 7-digit numeric variable that reports the person's total personal income from all sources in the previous month or year. Observation: 9999999 = NIU (not in universe)