

SEB: Available datasets and problems for the Final Assignment

Ferrara University - A.Y. 2019/2020

Prof. Bonnini and Mini

As we told you several times during classes, the SEB final exam is composed by 2 parts: the written test and the final assignment.

THIS DOCUMENT ADDRESSES FINAL ASSIGNMENT RULES AND GIVE YOU A LIST OF AVAILABLE PROBLEMS

The final assignment is a team work (you can create groups, 3 students each) through which you will demonstrate to handle, understand and explain one of the statistical techniques we discuss during our course. To run the statistical analysis you can:

- choose one of the following Available Problems (based on one of the datasets professors provide);
- or you can construct/adopt your own dataset. In this case you are demanded to show and explain your dataset and the related research question to prof. Bonnini and Mini before to run the analysis. Only if you obtain a positive feedback, you are allowed to carry on the assignment based on your own data.

The statistical analysis will be explained in a written research paper and through an oral presentation.

Once you completed your analysis and the written research paper, you must send your work (pdf written paper, ppt presentation and txt R commands) to professors, few days before the oral presentation.

The paper should follow the structure we present during lecture (an example is available on our course website).

The presentation will take around 15' and each group member should explain part of the analysis; at the end of the presentation, professors can ask you some questions about the done work.

Thanks to the Assignment (paper and presentation) you'll gain from 0 to 3 points. The points are given to each team member: obviously, considering different performances, the given points can be dissimilar from one to another.

If you have already obtained points for your Final Assignment (for example, last year you have already submitted and presented your assignment to professors who communicated to you the related gained points), the points you gained are still official.

IN THE FOLLOWING PARTS OF DIS DOCUMENTS YOU'LL FIND:

1. List of Available Datasets 2020
2. List of Available Problems 2020

Once your group select a problem, please write up to professors to obtain the related dataset indicating the problem and dataset (i.e DT3_PR9), thus you can start your statistical analysis.

1. LIST OF AVAILABLE DATASET

Database	Brief description	Statistical units	Variables	Database	Brief description	Statistical units	Variables
DT_1	European countries described by real GDPpc and number of firms adopting a certain typology of technology	European countries (22)	Number of firms adopting various technologies	DT_3	Italian Main Cities described by population, houses and tourism's topics. Data from Eurostat.	Italian Cities (92)	All variables refer to 2015: total population (total ,by gender and age - very young and old people), number of museum visitors, number of nights spent in tourists' accommodations, number of beds in tourism accommodations, persons with different levels of education, current situation of houses and households related,
DT_2	European countries described by number of firms adopting technology in logistics, shipment, organizational processes and products	European countries (14)	Number of firms adopting various technologies	DT_4	Italian Counties described by some economics, innovation, society and energy's indicators.	Italian Counties (110)	All variables are based on 2015 and they describes counties through indicators of energy production, firms' performances, innovation, social care and status.

2.LIST OF AVAILABLE PROBLEMS

<p>DT1_1 - Research question: Does the wealth of nations depend on typology of R&D adopted by their firms? A case study based on European countries. Using DT1 we construct a Multiple linear regression analysis to investigate the causal relationship between GDPpc (a national wealth proxy) and the numbers of firms adopting Cloud Computing, Big Data, Integration with partners, Internal integration, 3D printing and robotics, In-House R&D and External R&D.</p>	<p>DT1_2 : Research question: Does the wealth of nations depend on product-process innovation adopted by their firms? A case study based on European countries. Using DT1 we construct a Multiple Linear Regression Analysis to investigate the causal relationship between GDPpc (a national wealth proxy) and the numbers of firms adopting innovative process or product, innovative organization or MKTing, new products for world, new products for the firm, new product the market.</p>	<p>DT1_3 : Research question: What are the determinants of innovation on the supply side in Europe? A case study base on 22 countries. Using DT1 we perform a Principal Component Analysis or factor analysis. (nb: we are not interested in GDP)</p>	<p>DT1_4 : Research questions: Can we profile countries using their firm's innovation strategies? Grouping innovation systems across Europe. Using DT1 we perform a Cluster Analysis to identify groups of countries based on innovation strategies adopted by their firms. (nb: we are not observing GDP per capita; however, it can be used in the final part of the analysis to compare countries).</p>	<p>DT1_5 : Evaluating firms' innovativeness in European courtiers using Composite Indicators. An empirical comparison of different procedures. Using all the composites indicators explained in our SEB course, we want to compare and explain different results. At the end we chose the best indicator for our purpose.</p>
--	---	--	---	--

<p>DT2_1: What are the key drivers of innovation in logistics? A supply side case study based on 14 European countries. Using DT2 we construct a Multiple Linear Regression Analysis to investigate the causal relationship between the numbers of firms introducing innovation in logistics and the numbers of firms introducing new technology in shipment, organizational processes and products</p>	<p>DT2_2 : Research question: What are the determinants of innovation on the supply side in Europe? A case study base on 14 countries. Using DT2 we perform a Principal Component Analysis or Factor Analysis to understand the organizational, logistics and products innovativeness of 14 European countries.</p>	<p>DT2_3 : Research questions: Can we profile countries using their firm's innovativeness? Using DT2 we perform a Cluster Analysis to identify groups of countries based on innovation strategies adopted by their firms.</p>	<p>DT2_4: Evaluating supply side innovativeness in European Countries using Composite Indicators. An empirical comparison of different procedures. Using all the composites indicators explained in our SEB course, we want to compare and explain different results. At the end we chose the best indicator for our purpose.</p>
--	--	--	--

<p>DT3_PR1: Studying key drivers of tourism in Italian Cities. An applied study based on 92 cases. Using DT_3 we construct a Multiple Linear Regression Analysis to investigate the causal relationship between the total nights spent in tourism accommodation and the available infrastructure (number of beds in tourism accommodations), the number of museum visitors, the level of population's education and the number of citizens very young or very old.</p>	<p>DT3_PR2: Key determinants of wellbeing defined as population living in their own dwelling. A case study based on 92 Italian Cities. Using DT_3 we construct a Multiple Linear Regression Analysis to investigate the causal relationship between the number of people living in their dwelling and variables of educational level, demographic characteristics and employment status.</p>	<p>DT3_PR3: Can we characterize Italian Cities using their tourism profile? Using DT3 we perform a Cluster Analysis to identify groups of Italian cities based on characteristics related to tourism and city attraction</p>	<p>DT3_PR4: Evaluating tourism and population's education in Italian Cities using Composite Indicators. An empirical comparison of different procedures. Using all the composites indicators explained in our SEB course, we want to compare and explain different results. At the end we chose the best indicator for our purpose.</p>
---	---	---	--

<p>DT3-PR5: Can we characterize Italian Cities using their households and dwelling profile? Using DT3 we perform a Cluster Analysis to identify groups of Italian cities based on characteristics related to households and dwelling profiles</p>	<p>DT3_PR6: Evaluating households and dwelling status of Italian Cities using Composite Indicators. An empirical comparison of different procedures. Using all the composites indicators explained in our SEB course, we want to compare and explain different results. At the end we chose the best indicator for our purpose.</p>	<p>DT3_PR7: What are the key latent variables able to describe our cities? A case study based on 92 Italian Cities measured by Eurostat. Using DT3 we perform a Principal Component Analysis or Factor Analysis.</p>
--	--	---

<p>DT4_PR1: Studying key drivers of Export in sector with a global dynamic demand. An applied study based on 110 Italian Counties. Using DT_4 we construct a Multiple Linear Regression Analysis to investigate the causal relationship between the total Export in globally dynamic sector and variables which can affect this performance.</p>	<p>DT4_PR2: The role of social, economic and innovation status of territories to explain the employment rate. A case study based on 110 Italian Provinces. Using DT_4 we construct a Multiple Linear Regression Analysis to investigate the causal relationship between the employment rate and local variables which can explain it.</p>	<p>DT4_PR3: Can we Characterize Italian Provinces using their economic profile? Using DT4 we perform a Cluster Analysis to identify groups of Italian provinces based on their economic characteristics</p>
---	--	--

<p>DT4_PR4: Can we Characterize Italian Provinces using their social, innovation and energy's data? Using DT4 we perform a Cluster Analysis to identify groups of Italian provinces based on their innovativeness, energy profile and social characteristics.</p>	<p>DT4_PR5: Territorial competitiveness in Italy: a case study based on Factor Analysis. Using DT4 we perform a Factor Analysis to individuate the factors able to explain the Territorial competitiveness across Provinces.</p>	<p>DT4_PR6: Territorial competitiveness in Italy: a case study based on Principal Component Analysis. Using DT4 we perform a PCA to individuate the main latent variables (components) able to explain the Territorial competitiveness across Counties.</p>	<p>DT4_PR7: Territorial Competitiveness indicator a case study based on 110 Italian provinces. Using all the composites indicators explained in our SEB course, we want to compare and explain different results. At the end we chose the best indicator for our purpose.</p>
--	---	--	--