



ONLINE Ljubljana Doctoral Summer School 2023

10 – 14 July 2023

9:00 – 13:00 (CEST, Ljubljana)

Course title:

Data Production and Analysis for Decision Support: Methods and Tools

ECTS credits: 4

Lecturer: GONZALEZ-FELIU Jesus, Excelia Business School, France

Contact: gonzalezfeliuj@excelia-group.com

Aims of the course:

The purpose of this intensive course is to present the basics of data production and analysis, i.e. the main concepts and methods for the production, processing and analysis of data (as an input and an output of a model) to understand and describe phenomena and subsequently forecasting or estimation procedures. This course does not focus on the mathematical foundations of statistical methods or models, but on a scientific methodological approach for the global process of data production and analysis (defining objectives, identifying the sample/population to enquire, collecting or estimating data, processing those data for the given objectives, analysing them and interpreting the results, to finally relating the given outputs to the searched reality).

Course syllabus:

In this course, students will address the following content:

1. The different ways of the scientific path: deduction, induction and abduction.
2. Basics of data production and collection for both quantitative and qualitative data.
3. The different stages of modelling and estimation (from objective definition to model validation, including data retrieval)
4. How to choose the most suitable method and how to define the variables for data collection and estimation.
5. Main elements of data processing and analysis: most common assumptions and validity of uses.
6. Small set statistics: which methods and estimates can be used with very small sets of data.
7. Hypothesis testing and choice of the most suitable tests.
8. Interpretation, replicability and research and practical considerations

Tentative schedule:

Module 1: Research, estimation and data production



1. An introduction to scientific path and its three ways of knowledge-seeking: deduction, induction and abduction.
2. Data production and collection: definitions and main methods (for both quantitative and qualitative data, based on observations, declarations, measurements or estimations, among others).
3. Population and sample, sampling and adjusting: main considerations when dealing with sample-based data production.

Reading assignments: Lecture notes and basics given before the course (Gonzalez-Feliu, 2023).

Work to be done before class: Read the notes.

Module 2: The data production process

1. The cycle of modelling and the role of data production and collection.
2. Objectives, variables and choice of the suitable data production method.
3. The basics of sampling: types of samples, choice of the sample size, reliability considerations.
4. Collecting data: practical considerations in operating observations and surveys
5. Data completion, modelling and simulation: main considerations for producing data by estimations.
6. Adjusting and processing collected or produced data.

Reading assignments: Book chapter about data production (Gonzalez-Feliu, 2019), book chapter on questionnaire design and sampling (Parfitt, 2013).

Work to be done before class: Read the assignments.

Module 3: Processing and analysing data

1. Characteristics of dispersion, asymmetry and flattening: how to choose the suitable indicators and interpret them.
2. Considerations on using normality, statistical distributions and probability theories in analysing and interpreting data.
3. An introduction to small set statistics: methods, conditions of applications and validity
4. An introduction to hypothesis testing on quantitative and qualitative data: main types of tests and guide for choosing the most relevant one for the chosen objectives.
5. Case study 1: Mobility plan development.

Reading assignments: Two blog issues on small set statistics (Sauro, 2013; Lewis and Sauro, 2022).

Work to be done before class: Read the assignments.



Module 4: Interpretation, research and practical considerations

1. From data to theories: how concluding and discussing obtained data in a scientific way.
2. Main research considerations of data-driven and data production approaches.
3. Replicability and transferability of data production and analysis methods.
4. Case study 2: Opinion poll on reaction to crises.

Reading assignments: Two research papers on the impacts on collection/production choices in data quality and interpretation keys (Sanchez, 1992; Gonzalez-Feliu, 2019)

Work to be done before class: Read the assignments, read case study 1.

Module 5: Exercises, cases and final examination

1. Exchanges on Case study 1
2. Exchanges on Case study 2
3. Practical considerations of data production and analysis methods
4. Presentation of students' works (for grades)

Reading assignments: None.

Work to be done before class: Finish the two case studies.

List of readings:

Gonzalez-Feliu, J. (2019). Data Production for Urban Goods Transport Planning and Management: How to Manage "No Data" Situations in the Big Data Era. In *Logistics and Transport Modeling in Urban Goods Movement* (pp. 29-54). Hershey: IGI Global.

Gonzalez-Feliu, J. (2023). Introduction to data production, processing and analysis. Preliminary and basic notions. Notes specially prepared for this course.

Gonzalez-Feliu, J., & Sánchez-Díaz, I. (2019). The influence of aggregation level and category construction on estimation quality for freight trip generation models. *Transportation Research Part E: Logistics and Transportation Review*, 121, 134-148.

Lewis, J., & Sauro, J. (2022). Five Styles of Statistical Rhetoric. Retrieved from MeasuringU webpage, <https://measuringu.com/five-styles-of-statistical-rhetoric/>, on Jan 17th 2023.

Parfitt, J. (2013). Questionnaire design and sampling. In *Methods in human geography* (pp. 78-109). Routledge.

Sanchez, M. E. (1992). Effects of questionnaire design on the quality of survey data. *Public opinion quarterly*, 56(2), 206-217.

Sauro, J. (2013). Best Practices for Using Statistics on Small Sample Sizes. Retrieved from MeasuringU webpage, <https://measuringu.com/small-n/>, on Jan 17th 2023.

**Teaching methods:**

- Lectures
- Case studies
- Work in mini-groups
- Exercises
- Inversed class/reading discussion

Course credit:

Students needing course credit for their PhD studies will have to successfully pass an oral presentation on the last day of the course and submit a small report of a case study.

Lecturer's biographical note:

Jesus Gonzalez-Feliu is Full Professor in Supply Chain Management at Exceia Business School since 2020. Previously, he was associate professor at Mines Saint-Etienne and research engineer at CNRS. He has a PhD. in computer and systems sciences from Politecnico di Torino, Italy and a Habilitation to Direct Researches from Université Paris Est, France. Before, he was associate professor at Ecole des Mines de Saint-Etienne and Research Engineer in Data Production at the National Center of Scientific Research, France, among others. His main teaching and research subjects are the design of logistics and transport systems, the estimation of urban freight and shopping demand, the processes of decision and management driven by data and information, the evaluation of supply chains including sustainability, resilience and maturity, and collaborative decision making. He is author or editor of five books, co-author of more than 50 peer-reviewed journal papers, guest editor of ten special issues in international journals and has coordinated various courses on the field in international weeks in different countries, at Bachelor, Master or Doctoral level.