UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistics for The Social Sciences

2627-1-E2006P006

Learning area

2: Psychosocial aspects underlying communication.

Learning objectives

Knowledge and understanding

The course focuses on the formation of data-driven opinions and decisions. In particular, it examines the epistemological processes through which knowledge is generated from data. To this end, the course first emphasizes statistical literacy in a context of rapidly increasing information production, where communicative overload requires individuals to navigate the "asymmetry of information" and distinguish between data that are scientifically authoritative and data that are not.
The course familiarizes students with ISTAT, the Italian National Institute of Statistics, as the official statistical authority and a public good serving citizens. It then broadens its scope to include other official and alternative data sources, as well as Big Data. The development and use of metadata become crucial tools for interpreting data, together with systems for assessing and validating data quality. These aspects are examined within the broader framework of supranational statistical systems such as Eurostat, the OECD, and the World Health Organization (WHO).
From the perspective of statistical methodology, the course starts from the main constructs used in the study of behavior, which represent an interdisciplinary thread running through the degree program. Beginning with the conceptualization of these constructs, the course illustrates how they are operationalized for empirical research through the construction of statistical variables and related measures.
The course then introduces data analysis within the two main branches of statistics: descriptive statistics and statistical inference. Particular emphasis is placed on the dual role of descriptive statistics, both in its own right as a means of understanding a population or group, and as an essential prerequisite for the correct application of inferential methods.

Applying knowledge and understanding

Concepts, definitions, and disciplinary tools will be systematically presented from an interdisciplinary perspective, fostering the ability to integrate the statistical procedures learned in the course with information sources of different kinds.
The competencies developed in the course have a dual applied focus.
First, becoming familiar with data collection processes and the associated measurement systems used in official statistical production enables students to interpret data correctly and avoid misleading approaches. This is achieved through the ability to identify and investigate the key contextual and methodological elements that have generated the data, as well as through the practice of accurate data reading, appropriate data processing, and critical interpretation of results.
Second, through the methodological foundations of the discipline, students will learn how to organize collected data based on their understanding of measurement scales. The practical application of these scales will enable them to derive the appropriate descriptive measures for both univariate and bivariate analysis. Descriptive summaries will also serve as a guide to the inferential possibilities available.
Finally, the use of open-source statistical software will strengthen students' proficiency in applying statistical methods and in understanding the meaning and interpretation of different statistical measures.

Making judgements

The ability to exercise independent judgment and critical thinking in understanding the vast amount of data that permeates both scientific literature and information more generally is an essential driver for navigating contemporary communication, which relies on multiple forms of language.
Developing sensitivity to the processes through which data are generated enables students to interpret data correctly, place them within their original contexts, and understand the conditions and limitations of the measurement procedures that define their content. During the course, students will engage in the direct examination of quantitative information drawn from different sources, exercising critical thinking to avoid superficial interpretations and refining strategies and approaches aimed at preventing oversimplification.
This perspective will then be extended to alternative data sources and, more broadly, to public communication, with particular attention to decoding the strategies used in the presentation of information.
Independent judgment and critical thinking will be explicitly fostered through data processing activities, the construction of indices and indicators, and the selection of the most appropriate statistical measures according to the characteristics of the phenomenon under investigation.
Statistical modeling will be introduced through bivariate analysis across different measurement scales, with attention to identifying misleading interpretations such as spurious correlation and Simpson’s paradox.
These competencies will be developed both through classroom and computer-lab activities and through the assessment methods adopted in the course.

Communication skills

Statistical literacy is a crucial component of the skill set of graduates in Communication Sciences, both for a full understanding of the scientific literature—which relies heavily on empirical validation—and for the ability to communicate complex content through appropriately diverse and integrated forms of expression.
Students will develop the ability to present summarized data in both numerical and graphical forms, while simultaneously explaining their empirical meaning, definitions, and metadata, thereby clarifying their scope and preventing naïve interpretative distortions.
They will also become proficient in communicating statistical phenomena through graphs, tables, and verbal descriptions, adapting the format to the intended purpose and audience.
Classroom activities and individual exercises, which will be discussed and reviewed during dedicated sessions, will help students consolidate these communication skills. These competencies will also form part of the assessment process in the course examinations.

Learning skills

The teaching structure of the course is aligned with the overall perspective of the degree program, which seeks to

promote a style of learning that integrates different disciplines within a multidimensional educational framework. This approach is aimed at enhancing the knowledge and skills needed to address the complexity of contemporary communication.
The continuation of academic training is oriented toward information systems for communication and information management, as well as evidence-based humanities disciplines.
The development of students' autonomy and the consolidation of their learning abilities will be supported through both individual and group activities and will subsequently be assessed through the final examinations

Contents

The course provides the theoretical and analytical tools necessary for accessing, processing, and interpreting quantitative and qualitative–quantitative information. Methods of data generation are examined with particular reference to official national statistics within the Eurostat network and are subsequently extended to alternative data sources and Big Data.
The operationalization of concepts and phenomena into variables, together with the identification of the corresponding measurement methods, is illustrated through surveys and studies on multidisciplinary topics, with a particular emphasis on the interdisciplinary nature of the degree program.
The course promotes the autonomous and critical acquisition of appropriate interpretative frameworks and guidelines for understanding information and statistical-computational reasoning, both in learning basic quantitative techniques and in accessing and interpreting the results of surveys and their dissemination through different media.
The distinction between descriptive statistics and inferential statistics includes an introduction to the fundamentals of statistical inference based on experimental data. In addition, the course addresses more recent analytical techniques, including the analysis of textual data and social networks.

Detailed program

- Statistical terminology.
- Types of data.
- Classification of statistical sources.
- Measurement scales.
- Data collection.
- Official statistical sources. Types of surveys.
- Data quality: concepts and definitions.
- Sampling strategies. Some sampling issues in social research.
- Basic aspects of questionnaire and aits dministration.
- Notes on scaling techniques.
- Data visualization.
- Distribution measures. Position and analytical means. Variability.
- The shape of the distribution. Asymmetry and kurtosis.
- Comparisons between quantities. Ratios and indexes. Composite indicators.
- Elements of probability. Normal distribution as a significant distribution.
- Hypothesis testing for the equality of means.
- Bivariate analysis: contingency, cogradation and correlation.
- Hypothesis testing in bivariate analysis.
- Introduction to text analytics and social networks.

Prerequisites

Mathematics: Algebraic skills commonly acquired in lower secondary education (middle school level) are required.
Computer Science: Students are expected to possess the competencies covered in the Computer Science course (a compulsory course offered during the first semester).
A review of the fundamental mathematical and computing knowledge will be provided, making it possible to identify and implement remedial activities where necessary.

Teaching methods

The course is delivered in a blended-learning format through different teaching methods and is worth 8 ECTS credits (60 hours):

- 5 two-hour classes delivered in person in mixed mode, combining lectures for the presentation of introductory topics with interactive activities aimed at checking students' understanding of the material;
- 6 two-hour online lectures, delivered in a traditional lecture format, focusing on the more complex concepts of the discipline;
- 10 two-hour online interactive sessions, involving practical applications and the analysis of official statistical websites and data sources;
- 9 two-hour laboratory sessions with computer-based applications, organized for three separate groups of students and conducted in person in a computer laboratory using an interactive teaching approach.

The course is structured around theoretical lectures on statistical methodology, applied practical exercises, and online activities.
The statistical methodology component is organized into two thematic modules.
Module 1: Data Sources, Official Statistics, and Data Quality
The first module addresses methodological issues related to available data, beginning with official statistics. Students explore the websites of ISTAT within the broader Eurostat network, covering a variety of disciplinary domains. Particular attention is given to the associated online documentation, especially metadata and data quality standards.
This module is delivered through online activities, including:
live-streamed lectures providing the methodological framework;
hands-on activities, supported by the instructor and tutor, focused on exploring official statistical websites and databases.
Module 2: Statistical Theory and Applied Analysis
The second module focuses on theoretical knowledge, emphasizing the understanding of concepts and the coherence of analytical reasoning, with particular attention to the specific features of psycho-social research.
Practical sessions in computer laboratories systematically translate methodological concepts into their applied counterparts, enabling students to contextualize the information acquired through the use of statistical software, preferably open-source analytical tools.
Computer-lab sessions are conducted in person, as are the theoretical lectures at the beginning and end of the course, whereas the remaining lectures are delivered remotely in a live lecture format.
Online Activities and Group Work
Part of the online component consists of group activities, exercises, and freely accessible simulations made available through the learning platform. These activities are designed to integrate the two modules by placing the statistical measures learned at the methodological level within the broader context of statistical information encountered in everyday life.
From a pedagogical perspective, the development of this competence encourages the discussion and critical evaluation of numerical information presented in the media, including comparisons with official statistics.
These discussions take place in groups organized according to students' academic interests and thematic preferences.

Assessment methods

The final assessment is entirely written and consists of two components: a project work and a computer-based examination on Moodle. No mid-term assessments (in itinere) are scheduled.

Project Work
The project work consists of a critical analysis of a phenomenon or topic that can be investigated using statistical data produced by official institutions.
Students are required to identify and use official statistical sources, examining their data collection methods, metadata, and the definitions adopted to delimit and describe the phenomenon under investigation. The information obtained must then be critically compared with the way the same topic is represented in the media and in public communication.
The project assesses students' ability to:

- understand and apply the fundamental principles of descriptive statistics;
- correctly interpret quantitative data and information in light of the data collection methods employed;
- recognize the strengths and limitations of information sources;
- identify the most effective ways of presenting and communicating data;
- develop a critical understanding of statistical information disseminated through the media.
  In addition to the correct application of methodological principles, students are expected to provide an original interpretation of the phenomenon under study and a critical evaluation of the use of official statistics in public and journalistic communication.
  The project work contributes 35% of the final grade.

Computer-Based Moodle Examination
The computer-based examination, conducted through the Moodle platform, consists of the statistical analysis of a dataset simulating a real-world survey. The examination assesses students' ability to apply the main tools of descriptive statistics and introductory concepts of inferential statistics using statistical software.
The examination includes:

- multiple-choice questions;
- True/False questions;
- closed-ended exercises;
- production and interpretation of graphs;
- analysis and interpretation of statistical results.
  The final section includes methodological multiple-choice questions aimed at assessing students' ability to identify the most appropriate statistical method for a given problem.
  The examination is designed to assess students' ability to:
- recognize the characteristics of variables and statistical measures;
- select the most appropriate analytical procedures;
- correctly interpret statistical results;
- use statistical and computational tools to solve applied problems.
  The computer-based Moodle examination contributes 65% of the final grade.

Textbooks and Reading Materials

- Iezzi, Domenica Fioredistella, (2024) *Dai dati alla conoscenza. Statistica per le decisioni*. Roma: Carocci.
- Websites and additional learning material as indicated by instructors.

Although this course is held in Italian, for Erasmus students, course material is available also in English, and

students can sit the exam in English if they wish. They can adopt the text: Howitt, D. Cramer, D. Introduction to statistics in psychology, 5th ed. Harlow : Pearson, 2011, available at the University Library.

Sustainable Development Goals

GENDER EQUALITY | PEACE, JUSTICE AND STRONG INSTITUTIONS