# **Business data science**

**Faculty**

Sergio Venturini

**Obiettivi formativi**

The growing amount of data available today represents one of the greatest revolutions in recent human history. An essential part of this revolution is the transformation of these enormous amounts of data into valuable insights, which are used by decision-makers around the world to identify the best strategy to undertake.

This course essentially has two objectives: on the one hand to provide the student with the analytical and statistical tools most frequently used today in digital and traditional marketing. On the other hand, the course aims to encourage the formation of the ability to model the world around us, which represents a fundamental skill for the analysis of economic and social phenomena. The presentation of the various topics is illustrated through the practical discussion of numerous cases related to different business problems, with a particular focus in marketing. In parallel with the methodological presentation, the course also aims to show the use of R, one of the software currently most used for the development of data science applications in business.

**Risultati di apprendimento attesi**

KNOWLEDGE AND UNDERSTANDING

At the end of the course, the student will be able to:

* Identify the correct methodology to be used to solve the problem under study.
* Carry out the analyzes presented in the course with the R software.
* Recognize and correctly interpret the results of statistical analyzes applied to business problems.

APPLYING KNOWLEDGE AND UNDERSTANDING

At the end of the course, the student will be able to:

* Properly synthesize a set of data using visualizations and summary indexes.
* Compare alternative methods of analysis and identify the best method for the problem under consideration.
* Create analytical reports for various business applications.
* Correctly interpret the results of elaborate statistical analyses produced independently or by third parties.
* Use the R software for the application of the methods presented.

MAKING JUDGEMENTS

* Learning the statistical concepts that are fundamental for working autonomously in searching, selecting and elaborating corporate data and in the development of analytical reports.

COMMUNICATION SKILLS

* Learning the terminology and advanced statistical methods essential to implement and communicate appropriately the results of the analyzes conducted in different business contexts.

**Programma del corso**

The course is composed of the following topics:

1. Descriptive analytics review.
* Data synthesis using graphical displays.
* Data synthesis through the use of numerical indices.
1. Review of inferential statistics.
* Sample variability. The concept of sampling distribution.
* Point estimate and confidence intervals for some notable cases.
* Introduction to hypothesis testing and presentation of some remarkable cases. P-value of a test.
1. Reduction of the data dimensionality: the principal component analysis.
2. Main methods of predictive analytics.
* The linear regression model. Examples (prediction of purchasing behaviors).
* The logistic regression model for binary responses. Examples (credit scoring and churn models).

**Modalità di insegnamento**

* Standard frontal lectures

**Modalità di verifica dell'apprendimento**

The exam is written and allows you to obtain a maximum score of 31/30. The exam text will contain both practical exercises and theoretical questions (definitions) in two possible formats: 1) multiple choice questions and 2) questions with open answers.

The exam will be considered passed if the final grade is greater than or equal to 18/30. A final score of 31/30 is equivalent to a grade of 30 with honors.

The exam will cover all the material presented during the course.

Students learn their grade through the institutional channels. After 5 days from the communication, the exam grade is considered accepted and will be registered automatically. Otherwise, within the same deadline the student must explicitly express their intention to refuse the grade through the procedures established by the University.

For each test, together with the grade obtained, a single date will be communicated for the paper show.

The student has the right to take the exam in the way described above for a **maximum of three sessions** in an academic year. For the next attempts (i.e. from the fourth onwards), the exam must be carried out **exclusively in oral form**.

The examination methods detailed above aim to verify:

* The ability to identify the correct methodology to solve a given problem.
* The ability to compute specific statistical indicators with the software.
* The ability to interpret the analysis output obtained with the software.

**Testi consigliati e bibliografia**

The material for the course will be provided by the instructor through the BlackBoard platform.

For further details, we recommend referring to the following textbooks:

* Chapman, C., McDonnell Feit, E., *R for Marketing Research and Analytics*. 2nd edition. Springer-Verlag, 2019.
* James, G., Witten, D., Hastie, T., Tibshirani, R., *An Introduction to Statistical Learning*. 2nd edition. Springer-Verlag, 2021.
* Kuhn, M., Johnson, K. *Applied Predictive Modeling*. Springer-Verlag, 2013.

**Note**

1. The rules discussed above apply indistinctly to all students.
2. During the course it is presumed that students are familiar with basic mathematical tools (powers, logarithms, concept of function).
3. Although not necessary, the use of a standard hand calculator is allowed during the exams.
4. Due to the current COVID-19 pandemics, the actual assessment methods may change.

**Orario e luogo di ricevimento degli studenti**

Time and place for the office hours will be made available on the teacher’s personal page, available on the website <http://www.unicatt.it/>.