Academic subject: Statistics and Data Science for Marketing

Degree Class: LM-40 - Matematica		**Degree Course:** Mathematics	**Academic Year:** 2019/2020
		Kind of class: optional	**Year:** / **Period:** 2
			ECTS: 7 divided into **ECTS lessons:** 6.5 **ECTS exe/lab/tutor:**

Time management, hours, in–class study hours, out–of–class study hours
lesson: 52 exe/lab/tutor: 8 in–class study: 60 out–of–class study: 90

Language: Italian	**Compulsory Attendance:** no		
Subject Teacher: Marcello De Giosa	**Tel:** +39 080 5442707 **e–mail:** marcello.degiosa@uniba.it	**Office:** Department of Mathematics Room 12, Floor 4	**Office days and hours:** Monday 11 a.m. – 1 a.m. or by appointment.

Prerequisites:
Probability Calculus and Statistics, as usually offered during a degree of L-35 class.

Educational objectives:
Acquiring statistical concepts, techniques and R software skills for typical applications of Data Science and Statistics in modern Marketing problems.

Expected learning outcomes (according to Dublin Descriptors)

Knowledge and understanding:
Students should acquire the modern fundamental concepts and techniques of Data Science and Statistics for Marketing. Students should acquire the advanced R software skills required in typical modern Marketing problems.

Applying knowledge and understanding:
The acquired theoretical knowledge and R software skills are useful in facing typical practical problems of modern Marketing.

Making judgements:
Ability to deal with practical Marketing and business problems and to approach them with the right statistical tecniques. Problem solving skills should be supported by the capacity in implementing Data Science and Analytics protocols consistent with the theoretical approach.

Communication:
Acquiring the Data Science and Statistical language and formalism needed to comprehend and translate practical problems in data analytics problems, to solve theme with the right protocols and software capabilities, to explain and present the acquired knowledge and results to a wide audience.

Lifelong learning skills:
Students should acquire suitable learning methods and the abilities needed to interpret and solve also practical problems not directly considered in the course.

Course program

1. Overview of the R language and Packages Used During the Course.

2. Advanced Linear Modeling Topics.
Multiple regression. Logistic Regression. Hierarchical Linear Models

3. Reducing Data Complexity and Dimension.
Principal Component Analysis (PCA). Perceptual Maps. Exploratory Factor Analysis (EFA).
Rotations. Mathematical Insights on PCA and EFA.

4. Confirmatory Factor Analysis and Structural Equation Modeling.
Scale Assessment: CFA. General Models: Structural Equation Models.

5. Segmentation. Clustering and Classification.
Hierarchical Clustering. Mean-Based Clustering. Model Based Clustering. Classi_cation.
Mathematical Insights on Clustering and Classification.

6. Association Rules and Market Basket Analysis.
Metrics. Finding and Visualizing Rules. Exploring Segments. Mathematical Insights on Associations and Market Basket Analysis.

7. Choice Modeling
Choice-Based Conjoint Analysis Surveys. Simulating Choice Data. Fitting a Choice Model. Adding Consumer Heterogeneity to Choice Models. Hierarchical Bayes Choice Models. Design of Choice-Based Conjoint Surveys.

8. Social Network Analytics.
Directed vs. Undirected Networks. Visualizing and Analyzing Networks. Social Data Metrics and Taxonomy. Using Network Metrics in Prediction and Classification. Collecting Social Network Data with R.

9. Text Mining.
The Tabular Representation of Text: Term-Document Matrix and "Bag-of-Words". Bag-of-Words vs. Meaning Extraction at Document Level. Preprocessing the Text. Implementing Data Mining Methods.

Teaching methods:
Lectures and exercise sessions with the software R.

Auxiliary teaching:
Notes and slides will be given during the course.

Assessment methods:
Software R session and oral exam.

Bibliography:
1) G. Shmueli, P.C. Bruce, I. Yahav, N.R. Patel, K.C. Lichtendahl (2018) – Data Mining for Business Analytics. Concepts, Techniques, and Applications in R. - Wiley.
2) C. Chapman, E.McDonnel Feit (2019) - R for Marketing and Analytics. - Springer.

Further reading
3) W.W. Piegorsch (2015) - Statistical Data Analytics. Foundations for Data Mining, Informatics, and Knowledge Discovery. - Wiley
4) T. W. Miller (2015) - Marketing Data Science - Pearson.
5) M. Sarstedt, E. Mooi (2019) - A Concise Guide to Market Research - Springer.