COURSE OUTLINE

(1) GENERAL

SCHOOL	SCHOOL OF INFORMATION SCIENCES & TECHNOLOGY		
ACADEMIC UNIT	DEPARTMENT OF STATISTICS		
LEVEL OF STUDIES	1st Cycle (UNDERGRADUATE)		
COURSE CODE	**6127**	**SEMESTER**	**7th**
COURSE TITLE	**Methods of Statistical and Machine Learning**		

INDEPENDENT TEACHING ACTIVITIES	**WEEKLY TEACHING HOURS**	**CREDITS**
Lectures	4	8
Workshops		
Labs	2	

COURSE TYPE	Elective
PREREQUISITE COURSES:	Understanding subjects related to Statistical Inference, Distribution Theory and Linear Algebra will be useful.
LANGUAGE OF INSTRUCTION and EXAMINATIONS:	GREEK
IS THE COURSE OFFERED TO ERASMUS STUDENTS	YES
COURSE WEBSITE (URL)	https://www.dept.aueb.gr/en/stat/content/methods-statistical-and-machine-learning-8-ects

(2) LEARNING OUTCOMES

Learning outcomes
Upon completion of the course the students will be able to: apply contemporary statistical methods using the R software to analyze large volumes of data, chart and understand relationships in the data, find groups of observations, create classification rules, apply methods and work with large data sets. At the end of the course, the student will be able to construct graphs and understand relationships between data, identify observation clusters in the data, be able to build classification rules.
General Competences
• Search, analysis and synthesis of data and information, using the necessary technologies • Adaptation to new situations • Decision-making • Autonomous work • Teamwork

- Working in an interdisciplinary environment
- Generation of new research ideas
- Promotion of free, creative and inductive thinking

(3) SYLLABUS

Distinguishing statistical learning methods as supervised and unsupervised and determining the type of statistical problems they treat, the concept of distance in Statistics, Clustering (K-means, Hierarchical clustering, Model-based clustering), Classification (LDA, QDA, K-nearest neighbors, decision trees, random forests, naïve Bayes, Support Vector Machines, ANN Fisher's discriminant analysis and other methods). Resampling methods (cross-validation, bootstrap), linear model selection and regularization (subset selection, shrinkage, dimension reduction), multinomial regression, , step functions, regression splines, tree methods, support vector machines, neural networks.

(4) TEACHING and LEARNING METHODS - EVALUATION

<table>
<tr><td>DELIVERY
Face-to-face, Distance learning, etc.</td><td colspan="3">Face-to-face</td></tr>
<tr><td>USE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY</td><td colspan="3"><table>
<tr><td>Use in Teaching</td><td>Yes</td><td>During the lectures, applications of the methods using computer programs are presented.</td></tr>
<tr><td>In Labs</td><td>Yes</td><td>All labs are performed on PC’s and computational software.</td></tr>
<tr><td>In communicating with the students</td><td>Yes</td><td>Virtual meetings through Microsoft Teams and email</td></tr>
</table></td></tr>
<tr><td>TEACHING METHODS</td><td colspan="3"><table>
<tr><th>Activity</th><th>Semester workload</th></tr>
<tr><td>Lectures</td><td>52</td></tr>
<tr><td>Lab Exercise</td><td>26</td></tr>
<tr><td>Studying and Analyzing Bibliography</td><td>50</td></tr>
<tr><td>Tutorial</td><td>22</td></tr>
<tr><td>Assignment</td><td>50</td></tr>
<tr><td>Course Total</td><td>200</td></tr>
</table></td></tr>
<tr><td>STUDENT PERFORMANCE EVALUATION
.</td><td colspan="3">Written examination at the end of the semester: 80%
Project: 20%

Information is available at eclass</td></tr>
</table>

(5) ATTACHED BIBLIOGRAPHY

- Bartholomew D.J., Steele F., Moustaki I., Galbraithe J.I., Ανάλυση Πολυμεταβλητών Τεχνικών στις Κοινωνικές Επιστήμες, Εκδόσεις Κλειδάριθμος ΕΠΕ, 2011.
- Ιωαννίδης Δ., Αθανασιάδης Ι., Στατιστική και Μηχανική Μάθηση με την R, Εκδόσεις Τζιόλα, 2017.
- Rajaraman A., Ullman D.J., Εξόρυξη από Μεγάλα Σύνολα Δεδομένων, Εκδόσεις Νέων Τεχνολογιών, 2014.
- Sidney B., Everitt, Casella G., Fienberg, S., Ingram O., An R and S-PLUS Companion to Multivariate Analysis, Springer-Verlag London Limited, 2005.
- Hastie, Tibshirani and Friedman (2009) Elements of Statistical Learning, 2nd edition Springer
- James, Witten, Hastie and Tibshirani (2011) Introduction to Statistical Learning with applications in R, Springer
- B. S. Everitt, S. Landau, M. Leese, and D. Stahl (2011) Cluster Analysis, Fifth Edition, Wiley