

Course title/Course Neptun ID: Multivariate Statistics/ GAZDT506N	Credit value: 6
Professor in charge (name, position, academic degree): György Kövér, associate professor, PhD	
Professor(s) involved in teaching, if any (name, position, academic degree):	
Course category: compulsory / <u>elective</u>	
Course type: <u>lecture</u> /seminar/practical course/consultation, and total number of course sessions in the given semester: 21/ 0/0/0 course sessions	
Methods, (<i>specific</i>) tools, and characteristics applied to convey knowledge: e.g.: case reports, case study analyses, project work, other... The course is based on case studies with carefully selected datasets.	
Method of assessing knowledge (exam / practical course grade / other): exam	
Grading method(s) applied for assessing knowledge: Homework and oral presentation	
The course's placement in the curriculum (in which semester the course is taken): first semester	
Compulsory studies prior to the course (if any):	
Course description: a brief yet informative description of the course's objectives and the material to be covered	
This course covers some of the most useful tools of the multivariate statistics. The course focuses on concepts, statistical modeling and practical applications. The students finishing this course will be able to select the suitable statistical method, use the covered software packages, visualize their data and know how to publish their results. This course is based on the usage of the R and the STATA software packages.	
Course topics/semester syllabus: - Brief coverage of the preliminary studies - Visualization of the multivariate dataset. Histograms, scatter plots, rotating 3D plots - The multivariate normal distribution, the confidence ellipse. - Correlation, covariance and partial correlations - Multiple linear regression. The assumptions of regression. Regression statistics. Regression coefficients, confidence intervals, hypothesis tests. Diagnostic tests. Non-linear fit. Model selection methods. Detecting multicollinearity, heteroscedasticity and autocorrelation. - Logistic regression. The odds and odds ratios. The logistic model. Dependent variable. Factors and covariates. Wald statistics. Logistics coefficients and interpretation. Goodness of fit. - Principal component analysis. The main objective is data reduction. Eigenvalues, eigenvectors. The interpretation of the principal components. Performing other statistical methods based on the principal components.	

- Factor analysis. The factor model. The number of the required factors. The communalities. . The interpretation of the factors. Factor rotation. Factor extraction by principal component, maximum likelihood and principal factor methods. Model assumptions.
- Cluster analysis. Objectives. Measures of distance. Binary variables. Agglomerative and divisive clustering. Cluster description. K-means clustering. Ward method
- Scales. Internal consistency. Cronbah's alpha. Measures of agreement. Measures of Discrimination.
- The confirmatory Factor Analysis (CFA).
- The Structural Equations Modeling (SEM) The fitting of the model. The measurement and structural model. Validity.

A list of the 2-5 most relevant compulsory literature (notes, textbooks) with bibliographic information (author, title, publication details, page numbers, ISBN):

A list of the 2-5 most relevant recommended literature (notes, textbooks) with bibliographic information (author, title, publication details, page numbers, ISBN):

Jeffrey M. Wooldridge - Introductory Econometrics: A Modern Approach. A the second or later edition.

Naresh K. Malhotra: Marketing research. The sixth or later edition..

Professional competencies, elements of competence, and skills to which the course typically and substantially contributes (knowledge, ability, attitude, autonomy, and responsibility):

Knowledge

- Knows some of the most useful methods of the multivariate statistics.
- Able to use the R and the STATA software packages.

Skills

- Capable of selecting and applying the proper statistical method
- Capable of evaluating the computational results.
- Capable of creating the required plots to visualize the results.

Attitudes

- Open to understanding and applying further statistical methods.
- Committed to publishing the interpretation and explanation of computational results in an accessible manner

Autonomy and Responsibility

- Independently initiates the uncovering of relationships within the databases.
- Capable of engaging in a professional dialogue on equal terms with representatives of related disciplines and the practical sphere.
- Adheres to ethical research guidelines.