



Co-funded by the  
Erasmus+ Programme  
of the European Union



## Course Package “Data Analysis”

<b>Work Package</b>	WP3: Development of Course Materials for the Reformed MA Programmes, Deliverable 3.1
<b>Author(s)</b>	Armen Ktoyan
<b>E-mail Address</b>	arm.ktoyan@gmail.com
<b>Institution</b>	Armenian State University of Economics



**Erasmus+ Capacity Building in Higher Education Project:**

**“Reforming Master Programmes in Finance in Armenia and Moldova” (REFINE)**

[www.reforming-finance.eu](http://www.reforming-finance.eu) | Project No. 585784-EPP-1-2017-1-AT-EPPKA2-CBHE-JP

## Document History

Version	Date	Author(s)	Description
1	01.12.2019	Armen Ktoyan	Draft 1
2	29.01.2020	Armen Ktoyan, Marek Gróf	Final Package

### Disclaimer:

*This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.*

## Content

1. General information about the course.....	4
2. Learning outcomes of the course .....	4
3. Syllabus of the course .....	5
4. Teaching methodology of the course.....	7
5. Labour market relevance of the course .....	7
6. Assessment and grading .....	8
7. References.....	8
8. Course assignments.....	9
8.1 Assignment 1 .....	9
8.2 Assignment 2 .....	9
Annex: Presentation slides .....	10

## 1. General information about the course

Explanation: Please fill in the table below.

<b>Title of the course (as specified in the reformed curriculum)</b>	Data Analysis
<b>Name of the teacher</b>	Armen Ktoyan
<b>Novelty of the course (please select as appropriate)</b>	This course is an updated and revised version of a course which already existed in the curriculum
<b>Year of the course in the curriculum</b>	2nd year
<b>Semester of the course in the curriculum</b>	3rd
<b>Language of instruction of the course</b>	Armenian
<b>Number of ECTS credits</b>	5

## 2. Learning outcomes of the course

Explanation: Please specify the learning outcomes of the course.

- Implement R in data analysis tasks
- implement T-tests for comparison of means and interpret the results
- analyse within-group and between-group Variations
- conduct ANOVA (using software) and interpret the results
- Implement factor analysis in solution of financial issues,
- understand the purpose and procedures of clustering
- assess the quality of classification of financial objects
- conduct agglomeration clusterization (using software) and interpret the results
- conduct K-means method (using software) and interpret the results,
- analyze cluster profiles,
- conduct logistic regression in R,
- implement ROC-analysis, built ROC-curve and interpret the results,
- Implement logistic regression results in financial risks assessment.

### 3. Syllabus of the course

Explanation: Please provide a detailed syllabus of the course (broken down in weeks) – maximum 2 pages

14 weeks	Content (Lectures)	Seminars, Laboratory Works
<b>Week 1</b>	Course Outline, Main Problems and Issues of Data Analysis <ul style="list-style-type: none"> <li>• Key Terms of Data Analysis</li> <li>• Basic Stages of Data Analysis</li> <li>• Application of Data Analysis in Finances</li> <li>• Basics of R</li> <li>• Installation of R</li> <li>• Data Types, Data Structures in R</li> <li>• Quantitative and Qualitative Data</li> </ul>	
<b>Week 2</b>	Functions and Simulations in R <ul style="list-style-type: none"> <li>• Functions in R</li> <li>• Vectors, Vector Operations</li> <li>• Functions for Statistical Distributions</li> <li>• Simulation Programming in R</li> </ul>	Seminar: Interviews and discussions about the basics and main issues of data analysis. Laboratory work on basic points of R (brief, private, 5-minute interview of student).
<b>Week 3</b>	T – tests and Data Visualization <ul style="list-style-type: none"> <li>• T-tests for Comparison of Means</li> <li>• Methods for Multiple Comparisons</li> <li>• Conducting T-tests in R</li> <li>• Data Visualization in R</li> <li>• Data Importing/Exporting in R</li> </ul>	
<b>Week 4</b>	Dispersion Analysis <ul style="list-style-type: none"> <li>• The Essence of Dispersion Analysis</li> <li>• Dispersion Analysis Application Conditions</li> <li>• One-factor ANOVA</li> <li>• Within-group and Between-group Variations</li> <li>• Conducting ANOVA in R</li> </ul>	
<b>Week 5</b>	Multi-factor Analysis of Variance <ul style="list-style-type: none"> <li>• Two-factor ANOVA</li> <li>• Multi-factor ANOVA</li> <li>• Conducting Multi-factor ANOVA in R</li> </ul>	Laboratory Work: Conducting ANOVA in R, based on financial data-sets.
<b>Week 6</b>	Factor analysis <ul style="list-style-type: none"> <li>• The Purpose and Issues of Factor Analysis</li> <li>• Factor Structure</li> <li>• Components of Dispersion in Factor Analysis</li> <li>• Principal Component Method</li> </ul>	
<b>Week 7</b>	Implementation of Factor Analysis <ul style="list-style-type: none"> <li>• Comparison of the Results of the Component and Factor Analysis.</li> <li>• Commentary on the Results of Factor Analysis</li> <li>• Implementation of Factor Analysis in Finances</li> </ul>	Laboratory Work: Conducting factor analysis in R, based on financial data-sets. interviewing students on the results of the analysis.

	<ul style="list-style-type: none"> <li>Conducting Factor Analysis in R</li> </ul> Task Assignment 1	
<b>Week 8</b>	Presentation of Assignment 1: Homework	
<b>Week 9</b>	Cluster Analysis: Basics and Logics <ul style="list-style-type: none"> <li>The problem of Classification of Financial Objects.</li> <li>Cluster Analysis Methods.</li> <li>Distance Between Financial Objects</li> <li>Quality of Classification</li> <li>Procedures of Clustering</li> </ul>	
<b>Week 10</b>	Cluster Analysis: Implication in Finances <ul style="list-style-type: none"> <li>Agglomeration Clusterization</li> <li>K-means method</li> <li>Two-Stages Clusterization</li> <li>Cluster Profiles and their Analysis.</li> <li>Conducting Cluster Analysis in R</li> </ul>	Laboratory Work: Conducting cluster analysis in R, based on financial data-sets. interviewing students on the results of the analysis.
<b>Week 11</b>	Discriminant Analysis: Basics and Logics <ul style="list-style-type: none"> <li>Theoretical Backgrounds of Discriminant Analysis</li> <li>Areas and Methods of Application of Discriminant Analysis</li> <li>General Problems for Discriminant Analysis</li> </ul>	
<b>Week 12</b>	Discriminant Analysis: General Algorithm and Implementation <ul style="list-style-type: none"> <li>Methods for Evaluating Indicator Information</li> <li>Stepping Algorithm</li> <li>Conducting Discriminant Analysis in R</li> </ul>	Laboratory Work: Conducting discriminant analysis in R, based on financial data-sets. interviewing students on the results of the analysis.
<b>Week 13</b>	Logistic Regression <ul style="list-style-type: none"> <li>Binary Logistic Regression Model</li> <li>Implementation of Logistic Regression in Financial Risks Assessment</li> <li>Conducting Logistic Regression in R</li> <li>ROC Analysis</li> </ul>	Laboratory Work: Conducting logistic regression in R, based on financial data-sets. Interviewing students on the results of the analysis.
<b>Week 14</b>	Assignment 2 -presentation	

## 4. Teaching methodology of the course

Explanation: Please explain the teaching methodology and pedagogical approaches of the course – maximum ½ page

### **The following pedagogical approaches are used:**

- Student-Centered Approach to Learning.
- High Tech Approach to Learning.

### **The following methods and forms of study are used in the course:**

- lectures (2 hours a week).
- seminars and laboratory works (2 hours per two weeks: interviews, practical exercises, conducting analysis on statistical package (R, in some cases - SPSS), based on real data sets, discussions of results, assignments and case presentation).
- group work.
- self-study.
- current control and grading include: participation in interviews, laboratory work, group work results and case presentations.
- intermediate control includes mid-term exam (individual analysis of a financial data, using a software).
- final exam (a short research, based on implementation of data analysis methods).

## 5. Labour market relevance of the course

Explanation: Please explain the labour market relevance of the course (linked to findings of WP1) – maximum ½ page

To determine the current demands and expectations of employers from the MA students and graduates in the context of data analysis some surveys have been implemented within some group employers from RA market. The employer's list covered the financial institutions, non-financial corporations and state governmental bodies.

The employers have outlined the following skills and competences, which have relations to data analysis and that are of great importance.

- ✓ select and apply the necessary tools for data analysis and research
- ✓ be able to learn new methods of data analysis on their own,
- ✓ analyse financial processes as a system, describe the links between parts of a system,
- ✓ predict trends of development of processes,
- ✓ interpret results in a meaningful way,
- ✓ choose and use modern software packages for data analysis,
- ✓ be able to search and find the necessary information according to the analytical issue, critically analyse different sources of data,
- ✓ identify the basics and the reasons of specific situations,
- ✓ apply mathematical, statistical, financial analysis methods and models, identify their strengths and weaknesses, and propose ways to improve them.

## 6. Assessment and grading

Explanation: Please explain the form of assessment of the course – maximum ½ page

Assessment		
	Performance	Grade ratio
Class attendance and participation Interviews (brief, private, 10-minute interview of student), discussions, activity of laboratory works.	40-100	20%
Group works (slide presentations)	40-100	20%
Case study	40-100	10%
Mid-term exam	40-100	10%
Final exam (short research)	40-100	40%

Performance and grade			
Percentage	Criteria	Grading numbers	Grade
0-39%		D	insufficient
40-70%	basic criteria met	C- C C+	Sufficient
71-85%	average performance with some errors	B- B B+	Good
86-95%	above average performance with minor errors	A- A	Very good
96-100%	outstanding performance	A+	Outstanding

## 7. References

Explanation: Please provide the main references and recommended reading for the course – maximum 1 page

1. Matloff N. The Art of R Programming. A Tour of Statistical Software Design, No Starch Press, Inc. San Francisco  
(<http://diytranscriptomics.com/Reading/files/The%20Art%20of%20R%20Programming.pdf>)
2. Johnson, R.A. & D. W. Wichern (2012). *Applied multivariate statistical analysis* Sixth Edition, PHI.
3. Hair, J. F. Jr., Black, W. C., Babin, B. J., Anderson, R. E. (2010). *Multivariate data analysis* (7th ed). UpperSaddle River, NJ: Prentice Hall.

4. Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed). Boston, MA: Pearson.
5. *Методы статистического моделирования для экономистов : учебное пособие* / Г.А. Соколов, Н.А. Чистякова, Рос. экон. ун-т им. Г.В. Плеханова. – М.: Изд-во РЭУ им. Г.В. Плеханова, 2011.

## 8. Course assignments

Explanation: Please provide two assignments for the course (e.g. group work, project, essay, case study, homework).

### 8.1 Assignment 1

*Homework: Analysing Data Sets*

- compose a data base from a set of financial-economic indicators of a sample of companies or objects, on Excel basic,
- implement ANOVA and factor analysis,
- comment the results,
- assess the trends and main features of development of a sample.

### 8.2 Assignment 2

*Group work: Analysing Data Sets*

The students will be divided into groups (4-6 students in each group). They will be asked to implement analysis on a big data base and according to it results answer the following questions.

- what kind of latent variables can be produced from initial indicators,
- which is the optimal way of classifying the objects,
- how can be classified the new objects,
- how can be analysed the cluster profiles,
- are their significant differences between means of the groups,
- what kind of financial and managerial decisions can be formed based on the results of data analysis.

## **Annex: Presentation slides**

Explanation: Please provide presentation slides for your course (this can be done in a separate document, e.g. Power Point (Minimum: 25 slides))

The slides are attached in ppt format.