STATISTICAL DATA ANALYSIS

Course code	*FUN107*
Course title	*Statistical Data Analysis*
Type of course	*Compulsory*
Stage of study	*Undergraduate*
Year of study	*2nd*
Semester	*Fall*
ECTS	*6: 24 hours of theory, 24 hours of practice, 112 hours of self- study*
Coordinating lecturer	*Marius Kušlys, markus@ism.lt*
Studies form	*Full-time*
Prerequisites	-
Language of instruction	*English*

Course description

This is an introductory course of statistical data analysis. The material covered in the course includes methods of data collection, analysis, presentation of results as well as the use of statistical data analysis software, SPSS®. The course encompasses methods of descriptive statistics, statistical estimates, data comparison hypothesis testing, correlation as well as regression analysis. During the seminars in computer class students will learn to use the statistical data analysis software, SPSS®, types of data, user tools, as well as procedures for data analysis and presentation.

Course aim

The goal of the course is to provide students with the theoretical knowledge and practical skills necessary for the analysis of economic and political data. At the end of the course the students should be able to identify and apply the key methods of data analysis, carry out the analysis using specialized software, and to interpret the results.

Course learning outcomes (CLO)	Study methods	Assessment methods
CLO1. To understand the basic terminology used in statistics	Lecture and self-study	Midterm exam and Final exam
CLO2. To understand the basic principles of descriptive statistics	Lecture, seminar in computer class, and self-study	Midterm exam and Defence of homework 1
CLO3. To understand the basic concepts of probability	Lecture and self-study	Midterm exam
CLO4. To understand the basic principles of inferential statistics	Lecture, seminar in computer class, and self-study	Final exam
CLO5. To be able to utilize the correct statistical test based on sample, and hypothesis	Lecture, seminar in computer class, and self-study	Final exam and Defences of homework 2, 3
CLO6. To understand the difference between parametric and nonparametric tests	Lecture, seminar in computer class, and self-study	Final exam, Defence of homework 3
CLO7. To be able to apply basic descriptive statistics to an available data base	Lecture and seminar in computer class	Defence of homework 1
CLO8. To be able to apply the appropriate basic inferential statistics to the decision making process	Lecture and seminar in computer class	Final exam and Defences of homework 2, 3
CLO9. To be able to make generalizations about a population based on a sample from that population	Lecture, seminar in computer class, and self-study	Midterm exam, Final exam and Defences of homework
CLO10. To be able to apply statistical techniques to evaluate basic business hypothesis	Lecture and seminar in computer class	Midterm exam, Final exam and Defences of homework

Quality assurance issues

The lecturer will strive to ensure a variety of teaching methods as well as modes of self-assessment. The feedback from students will always be highly valued and appreciated.

Cheating prevention

The teaching and testing methods are chosen taking into account the purpose of the minimization of cheating opportunities. The course is based and promotes the value of integrity. Lack of academic integrity (erg. plagiarism, copying another person's work, the use of unauthorized aids on examinations, cheating, facilitating acts of academic dishonesty by others) will not be tolerated. Consequences for violations range from zero grade given for the assignments over failure of the course up to disciplinary measures for severe cases.

Course content

No.	Topic	Contact Hours		Readings
		Lecture	Computer class	
1.	**Introduction.** Study object of statistics. Data collection. Statistical observation. Population and sample. Data structure, research methods and statistics, variables and measurement, statistical notation, ways of obtaining a sample. **Frequency Distributions.** Frequency distributions, frequency distribution tables, frequency distribution graphs, the shape of frequency distributions, percentiles, percentile ranks and interpolation, stem and leaf displays, boxplots.	2	2	Gravetter (1 & 2)
2.	**Central Tendency.** Measures of central tendency: mean, median, mode, central tendency and the shape of the distribution. **Variability.** Measures of variability: range and interquartile range, standard deviation, variance (population / sample)	2	2	Gravetter (3 & 4)
3.	**Introduction to z- Scores.** Concept and use of the z-score: z-scores and the location in a distribution, using z-scores to standardize a distribution, other standardized distributions based on z-scores, computing z-scores for a sample	2	2	Gravetter (5)
4.	**Overview of Probability.** Brief overview of counting technics in probability, the probability and normal distribution, probabilities and proportions for scores from a normal distribution, probability and the binomial distribution. This will include continuous random variables, the normal distribution, the mean, dispersion and standard deviation of a continuous random variable, the binomial distribution, the exponential distribution.	2	0	Gravetter (6 & 7)
	Defence of homework 1	0	2	
5.	**Introduction to Hypothesis Testing.** The logic of hypothesis testing, uncertainty and errors in hypothesis testing, directional hypothesis tests, the general elements of hypothecs testing	2	0	Gravetter (8)
	Midterm exam	2	0	
6.	**Introduction to the t Statistic.** The t statistic- an alternative to z, hypothesis tests with the t statistic, measuring effect of size for the t statistic, directional test for the t statistic **The t Test for Two Independent Samples.** Intro to the t statistic for independent measures research design, the assumptions underlying the independent measure t formula	2	2	Gravetter (9 & 10)
7.	**The t Test for Two Related Samples.** Intro to the t statistic related measure design, hypothesis tests and effect size for repeated measures design, uses and assumptions for related measures t tests	2	2	Gravetter (11)
8.	**Introduction to Analysis of Variance.** Introduction to Analysis of variance (ANOVA) for independent-measures design, post hoc tests.	2	2	Gravetter (12)
	Defence of homework 2	0	2	
9.	**Correlation.** Overview of correlation, the Pearson correlation, understanding and interpreting the Pearson correlation, hypothesis tests with correlation, the Spearman correlation	2	2	Gravetter (15)
10.	**Introduction to Regression.** Introduction to linear regression, testing the significance of the regression equation, analysis of regression	2	2	Gravetter (16)
11.	**The Chi-Square Statistic: Tests for Goodness of Fit and Independence.** Parametric vs nonparametric tests, the chi-square test for goodness of fit / for independence, assumptions and restrictions for chi-square tests, special applications of chi-square test	2	2	Gravetter (17)
	Defence of homework 3	0	2	
	Total:	24	24	
	Final exam	2	0	

Course assignments and assessment of achievements

Type of assignment	Topics	Total hours of self study	Evaluation, %
Defence of homework 1	1 – 3	10	10
Defence of homework 2	6 – 8	10	10
Defence of homework 3	9 – 11	10	10
Midterm exam	1 – 5	41	35
Final exam	6 – 11	41	35
	TOTAL:	**112**	**100**

The overall assessment of the course (total maximum of 100% is possible) will be composed from evaluations of tasks, which are described as follows:

1. The **defences of homework** will count for the 30% of the final evaluation. There will be 3 defences, each worth 10%. During the defence students are expected to perform statistical calculations with SPSS and to answer the questions given by the lecturer on the corresponding topic
2. The **midterm exam** will count for the 35% of the final evaluation and will include problems and multiple-choice questions on the topics discussed during the lecturers of the first half of the course. Only non-text calculators and provided appropriate tables and formulas will be allowed.
3. The **final exam** will count for the 35% of the final evaluation and will include problems and multiple-choice questions on the topics discussed during the lecturers of the second half of the course. Only non-text calculators and provided appropriate tables and formulas will be allowed.

Students must score for all tasks of the course at the specified time. Postponing of *the defence of homework* is impossible and explicit retake of the *midterm exam* will not be allowed. Meanwhile, in case of failing final evaluation, **retake** is possible, but topics will cover the material of the whole course and will comprise 70% of the final grade; midterm exam and final exam results will be annulled; only non-text calculators and provided appropriate tables and formulas will be allowed. Precision of composite evaluations is left intact (up to 2 decimal places) until the end of the course and only the final evaluation will be subject to rounding.

Literature

Obligatory:
1. Gravetter F. J., Wallnau L. B. (2013). Statistics for the Behavioral Sciences (9th Edition). Toronto: Thompson.
2. SPSS Tutorial v.25.

Optional:
3. Elliot A. C., Woodward W. A. (2007). Statistical Analysis Quick Reference Guidebook: With SPSS Examples.
4. Weiss N. A. (2008). Elementary Statistics (7th Edition). Boston: Pearson Education.
5. Lind D. A., Marchal W. G., Wathen S. A. (2010). Basic Statistics for Business and Economics. New York: McGraw.