STATISTICAL DATA ANALYSIS

Course code	*FUN107*
Course title	*Statistical data analysis*
Type of course	*Compulsory*
Stage of study	*Undergraduate*
Year of study	*2nd*
Semester	*Autumn*
ECTS	*6: 26 hours of theory, 12 hours of practice (open labs and consultations), 12 hours of examinations, 94 hours of self-study*
Coordinating lecturer	*Alan Freeman*
Studies form	*Full-time*
Prerequisites	-
Language of instruction	*English*

Course description

This is an introductory course of statistical data analysis, designed for the students of the Economics and Politics program. 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 laboratory sessions 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. Statistical methods and SPSS® software will be applied to the analysis of concrete economic and political data.

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, lab, and self-study	Midterm exam and Assignment 1
CLO3. To understand the basic concepts of probability	Lecture and self-study	Midterm exam and Assignment 2
CLO4. To understand the basic principles of inferential statistics	Lecture, lab, and self-study	Final exam
CLO5. To be able to utilize the correct statistical test based on sample, and hypothesis	Lecture, lab, and self-study	Final exam and Assignments 2, 3
CLO6. To understand the difference between parametric and nonparametric tests	Lecture and self-study	Final exam
CLO7. To be able to apply basic descriptive statistics to an available data base	Lecture and lab	Assignment 1
CLO8. To be able to apply the appropriate basic inferential statistics to the decision making process	Lecture and lab	Final exam and all Assignments
CLO9. To be able to make generalizations about a population based on a sample from that population	Lecture, lab, and self-study	Midterm exam, Final exam and all Assignments
CLO10. To be able to apply statistical techniques to evaluate basic business hypothesis	Lecture and lab	Midterm exam, Final exam and all Assignments

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

Week	Topic	Contact Hours		Readings
		Lecture	Computer class	
1	**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. **2. 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	3 [open lab]	Gravetter (1 & 2)
	3. Central Tendency. Measures of central tendency: mean, median, mode, central tendency and the shape of the distribution. **4. Variability.** Measures of variability: range and interquartile range, standard deviation, variance (population / sample)	2		Gravetter (3 & 4)
	5. 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		Gravetter (5)
	Consultation	0		
2	**Assignment 1**	0	2	
	6. 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	3 [open lab]	Gravetter (6 & 7)
	7. 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		Gravetter (8)
	8. 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	2		Gravetter (9)
	Consultation	0		
3	**Midterm exam**	2	0	
	9. 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	3 [open lab]	Gravetter (10)
	10. 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		Gravetter (11)
	11. Estimation. Overview of estimation, estimation with the z statistic, estimation with the t statistic	2		Gravetter (12)
	12. Introduction to Analysis of Variance. Analysis of variance (ANOVA) is a hypotesis-testing procedure that is used to evaluate mean differences between two or more populations	2		Gravetter (13)
	Consultation	0		
4	**Assignment 2**	0	2	

	13. Correlation. Overview of correlation, the Pearson correlation, understanding and interpreting the Pearson correlation, hypothesis tests with correlation, the Spearman correlation	2	3 [open lab]	Gravetter (16)
	14. Introduction to Regression. Introduction to linear regression, testing the significance of the regression equation, analysis of regression	2		Gravetter (17)
	15. 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		Gravetter (18)
	Consultation	0		
5	**Assignment 3**	0	2	
TBA	**Homework Evaluation**	0	2	
	Total (contact hours):	**28**	**20**	
One week after last lecture	**Final exam**	2	0	

Course assignments and assessment of achievements

Type of assignment	Due	Topics	Total hours of self study	Evaluation, %
Assignment 1: data, descriptive statistics, plots	Monday of 2nd week	1 – 5	9	5
Assignment 2: one sample and two sample t tests	Monday of 4th week	8 – 12	9	5
Assignment 3: correlation and regression analysis	Monday following last lecture	13 – 14	9	5
Homework	TBA	1 – 15	9	5
Midterm exam	First meeting of 3rd week	1 – 7	29	40
Final exam	One week after last lecture	8 – 15	29	40
		TOTAL:	**94**	**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. Practical **assignments** will count for the **15%** of the final evaluation. There will be 3 assignments, **each worth 5%**. Students are expected to attend exercises (which will be organized in an open lab form), where they will be instructed on practical work of data analysis with SPSS. Results should be saved in electronic form, showed and properly explained to receive an evaluation.
2. **Homework** will count for the **5%** of the final evaluation. Homework will be assigned daily covering the material presented in lecture. One time during the course homework will be collected and graded. The day the homework will be collected will not be announced but rather collected at the beginning of the class. Thus, the student must always be prepared as he/she will not know the day the homework will be collected.
3. Two-hour written **midterm exam** will count for the **40%** of the final evaluation and will include open and multiple-choice questions on the topics discussed during the lecturers of the first half of the course. Midterm exam will be held the first lecture day of the 3rd week of the course. Only non-text calculators and appropriate tables will be allowed.
4. Two-hour written **final exam** will count for the **40%** of the final evaluation and will include open and multiple-choice questions on the topics discussed during the lecturers of the second half of the course. Final exam will be held one week after the final lecture of the 4th week of the course. Only non-text calculators and appropriate tables will be allowed.

Students must score for all tasks of the course at the specified time (see *Weekly course content*, *Course assignments and assessment of achievements*). Postponing of *practical assignments* and *homework* is impossible and explicit retake of the *midterm exam* will not be allowed. Meanwhile, in case of negative final evaluation, **retake** is possible, but topics will cover the material of the whole course and will comprise **80%** of the final grade, midterm exam and final exam results will be annulled; only non-text calculators and appropriate tables 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. (2009). Statistics for the Behavioral Sciences (8th Edition). Toronto: Thompson.
2. SPSS Tutorial v.15. 2007.
3. Powerpoint Presentations

Optional:

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