UČNI NAČRT PREDMETA / COURSE SYLLABUS (leto / year 2017/18)			
Predmet:	Analiza podatkov s programom R		
Course title:	Data analysis with program R		

Študijski program in stopnja Study programme and level	Študijska smer Study field	Letnik Academic year	Semester Semester
Univerzitetni študijski program Finančna matematika	ni smeri	2	drugi
First cycle academic study programme Financial Mathematics	none	2	second

Vrsta predmeta / Course type	obvezni / compulsory

Univerzitetna koda predmeta / University course code: M0356

Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30		30			90	5

Nosilec predmeta / Lecturer: prof. dr. Andrej Bauer

Jeziki / Languages:	Predavanja / Lectures:	slovenski / Slovene
	Vaje / Tutorial:	slovenski / Slovene

Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:	Prerequisites:
Vpis v letnik študija. Opravljen predmet Uvod v programiranje.	Enrolment in the programme. Completed course Introduction to programming.
Vsebina:	**Content (Syllabus outline):**

Uvod. R kot računalo.	Introduction. R as a calculator.
Razpredelnice, enote, spremenljivke. Merske lestvice. Priprava podatkov. NA. Excel, CSV. Branje in shranjevanje.	Spreadsheets, units, variables. Measurement scales. Data preparation. NA. Excel, CSV. Reading and storing.
Številski podatki. Vektorji. Povzetek. Histogram, škatlasti prikaz, gostota porazdelitve.	Numerical data. Vectors. Summary. Histogram, boxplot, distribution density.
Urejenostni in imenski podatki. Zapis v R-u. Povzetek. Stolpčni in krožni prikaz.	Ordinal and nominal data. Representation in R. Summary. Bar chart and pie chart.
Prikazi podatkov na zemljevidu. Barve.	Data presentation on maps. Colors.
Razvrščanje v skupine. Različnosti. Metoda združevanj in metoda voditeljev.	Clustering. Measures of similarity. Agglomerative method and leaders algorithm.
Povezanosti med spremenljivkami. pairs, QQplot. Metoda najmanjših kvadratov. Regresijska premica.	Associations between variables. pairs, QQplot. Least squares method. Regression line.
Nelinearne povezanosti. Modeli. Glajenje in prileganje.	Non-linear associations. Models. Smoothing and fitting.
Delo z besedili. Unicode. Regularni izrazi. Zipfov zakon.	Working with texts. Unicode. Regular expressions. Zipf's law.
Podatki z Interneta. Branje s spletnih strani. XML. Sprehajanje po straneh na spletu.	Data from internet. Reading from web pages. XML. Crawling web pages.
Osnove dela s časovnimi vrstami.	Basic operations on time series.
Prikazi podatkov s knjižnico ggplot2.	Visualizations using the ggplot2 library
Osnove metode Monte Carlo.	Basics of Monte Carlo method

Temeljni literatura in viri / Readings:

M.J. Crawley: The R Book. Wiley, 2007.

J. Maindonald, J. Braun: Data Analysis and Graphics Using R, Cambridge Univ. Press, Cambridge, 2003.

J.M. Chambers: Programming with R Software for Data Analysis. Springer, 2008.

P. Murrell: R Graphics, Chapman #ali#, Hall/CRC, Boca Raton, 2005.

C.P. Robert, G. Casella: Introducing Monte Carlo Methods with R. Springer 2010.

spletna stran http://www.r-project.org

Cilji in kompetence:

Študent spozna programski jezik R skupaj s pripadajočim okoljem in ob tem osnove slikovnega prikaza ter statističnih obdelav podatkov.

Objectives and competences:

Students learn programming language R with the corresponding environment. Using the language they learn basics of statistical data analysis and visualization.

Predvideni študijski rezultati:

Znanje in razumevanje: Študent dobro spozna programski paket R ustvarjen za statistično analizo in prikaze podatkov. Ob tem nadgradi poznavanje osnovnih programerskih prijemov in spozna nekatere posebnosti jezika R.

Uporaba: Priprava lastnih knjižnic, izdelava grafikonov, preprosta analiza podatkov.

Refleksija: Pomen sodobne informacijske tehnologije za analizo večjih količin podatkov, pomen slikovnih prikazov pri pregledovanju podatkov in pri predstavitvi rezultatov.

Prenosljive spretnosti – niso vezane le na en predmet: Delo z računalnikom, algoritmični način razmišljanja.

Intended learning outcomes:

Knowledge and understanding: Student learns programming package R designed primarily for statistical data analysis and visualization. Student upgrades her/his knowledge of basic programming techniques and learns some special features of language R.

Application: Builiding of user's libraries, preparation od charts, simple data analysis.

Reflection: The importance of modern information technology in analysis of large amounts of data, the importance of visualization in data exploration and presentation of results.

Transferable skills: Working with a computer, algorithmic way of thinking.

Metode poučevanja in učenja:

Learning and teaching methods:

Predavanja, vaje, domače naloge, konzultacije	Lectures, exercises, homework, consultations

Načini ocenjevanja: Delež (v %) / Weight (in %)   **Assessment:**

| Način (pisni izpit, ustno izpraševanje, naloge, projekt):

domače naloge, zaključni projekt

izpit iz teorije

ocene: 1-5 (negativno), 6-10 (pozitivno) (po Statutu UL) | 50%
50% | Type (examination, oral, coursework, project):

homeworks, final project

theoretical exam

grading: 1-5 (fail), 6-10 (pass) (according to the Statute of UL) |

Reference nosilca / Lecturer's references:

Andrej Bauer:

BAUER, Andrej, STONE, Christopher A. RZ: a tool for bringing constructive and computable mathematics closer to programming practice. Journal of logic and computation, ISSN 0955-792X, 2009, vol. 19, no. 1, str. 17-43. [COBISS.SI-ID 15325785]

BAUER, Andrej, TAYLOR, Paul. The Dedekind reals in abstract Stone duality. Mathematical structures in computer science, ISSN 0960-1295, 2009, vol. 19, iss. 4, str. 757-838. [COBISS.SI-ID 15322201]

BAUER, Andrej, BIRKEDAL, Lars. Continuous functionals of dependent types and equilogical spaces. V: CLOTE, Peter G. (ur.). Computer science logic : 14th international workshop, CSL 2000, annual conference of the EACSL, Fischbachau, Germany, August 21-26, 2000 : proceedings, (Lecture notes in computer science, ISSN 0302-9743, 1862). Berlin [etc.]: Springer, 2000, vol. 1862, str. 202-216. [COBISS.SI-ID 10606681]