

UČNI NAČRT PREDMETA / COURSE SYLLABUS						
Predmet:		Izbrane teme iz analize podatkov				
Course title:		Topics in data analysis				
Študijski program in stopnja Study programme and level		Študijska smer Study field		Letnik Academic year	Semester Semester	
Univerzitetni študijski program Matematika		ni smeri		3	drugi	
First cycle academic study programme Mathematics		none		3	second	
Vrsta predmeta / Course type				izbirni		
Univerzitetna koda predmeta / University course code:				M0260		
Predavanja Lectures	Seminar Seminar	Vaje Tutorial	Klinične vaje work	Druge oblike študija	Samost. delo Individ. work	ECTS
30		30			90	5
Nosilec predmeta / Lecturer:		doc. Alen Orbanič, prof. Alexander Keith Simpson				
Jeziki / Languages:		Predavanja / Lectures: slovenski/Slovene				
		Vaje / Tutorial: slovenski/Slovene				
Pogoji za vključitev v delo oz. za opravljanje študijskih obveznosti:				Prerequisites:		
Opravljen predmet Uvod v programiranje.				Completed course Introduction to programming.		
Vsebina:				Content (Syllabus outline):		
Multivariatne metode za analizo večrazsežnih podatkov: glavne komponente, večrazsežno lestvičenje, razvrščanje v skupine.				Multivariate methods for analysis of multidimensional data: principal components, multidimensional scaling, clustering. Modern approaches to data visualization: e.g.		

<p>Sodobni pristopi k prikazom podatkov: npr. ggplot2, Many Eyes, D3.js.</p> <p>Osnove analize časovnih in prostorskih podatkov. Analiza časovnih vrst: osnovne značilnosti, glajenje, ARIMA.</p> <p>Osnove analize omrežij: osnovne značilnosti, prikazi omrežij, porazdelitve stopenj in brezlestvična omrežja, pomembni deli omrežij.</p> <p>Velika podatkovja: simbolna analiza podatkov, vzporedno in porazdeljeno računanje (Hadoop).</p> <p>Teme bodo večinoma podane kot črne škatle - opisano bo KAJ s posameznimi metodami dosežemo in pri KATERIH pogojih, ne pa tudi podrobno KAKO je to narejeno. Za tiste, ki bi jih to zanimalo, bodo podane reference, kjer lahko spoznajo tudi podrobnosti o ozadju.</p>	<p>ggplot2, Many Eyes, D3.js</p> <p>Basics of analysis of temporal and spatial data. Time series analysis: basic properties, smoothing, ARIMA.</p> <p>Basic network analysis: basic properties, network visualization, distribution of vertex degrees, scale-free networks, important parts of networks.</p> <p>Big data: symbolic data analysis, parallel and distributed computation (Hadoop).</p> <p>The topics will be mostly presented in a 'black box' manner - it will be described WHAT can be achieved using particular methods subject to WHICH conditions, but it will not be explained in detail HOW the methods work. For students interested in details and background of methods, references will be given.</p>
---	--

Temeljni literatura in viri / Readings:

B. Everitt, T. Hothorn: An Introduction to Applied Multivariate Analysis with R. Springer, 2011.

W.N. Venables, B.D. Ripley: Modern Applied Statistics with S (fourth edition). Springer, 2002.

L. Wilkinson. The Grammar of Graphics (second edition). Springer, 2005.

H. Wickham: ggplot2 - Elegant Graphics for Data Analysis. Springer, 2009.

R.H. Shumway, D.S. Stoffer: Time Series Analysis and Its Applications With R Examples (third edition). Springer, 2011.

R.S. Bivand, E.J. Pebesma, V. Gómez-Rubio: Applied Spatial Data Analysis with R. Springer, 2008.

W. de Nooy, A. Mrvar, V. Batagelj: Exploratory Social Network Analysis with Pajek, (revised and expanded second edition). Cambridge University Press, 2012.

E. Diday, M. Noirhomme-Fraiture (eds.): Symbolic Data Analysis and the SODAS Software. Wiley, 2008.

Q.E. McCallum, S. Weston: Parallel R. O'Reilly, 2012.

spletna stran: <http://d3js.org/>

spletna stran: <http://www-958.ibm.com/software/analytics/manyeyes/>

Cilji in kompetence:

Študent preko praktičnega dela na raznolikih izbranih podatkovnih zbirkah spozna osnove analize in slikovnih prikazov različnih tipov podatkov. Praktično spozna in preizkusi relativno širok spekter analitičnih in vizualizacijskih metod a brez poglobljenega razumevanja ozadja.

Objectives and competences:

Through a practical work on various selected data sets students learn basics on data analysis and data visualization. They experience basic use of a relatively wide range of different analytic and visualization methods, but without in-depth understanding of the methods.

Predvideni študijski rezultati:

Znanje in razumevanje: Študent se spozna z različnimi tipi podatkovij ter širokim spektrom osnovnih pristopov in metod za njihovo analizo in grafično predstavitev. Ob tem nadgradi poznavanje podatkovno-analitičnih in deloma tudi programerskih prijemov.

Uporaba: Izvedba osnovnih analiz na različnih podatkovjih. Priprava lastnih enostavnih analitičnih metod. Izdelava prikazov podatkov.

Refleksija: Pomen sodobne informacijske tehnologije za analizo večjih količin podatkov, pomen prikazov podatkov.

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

Intended learning outcomes:

Knowledge and understanding: Student familiarises him/herself with various types of data sets and wide range of basic approaches and methods for their data analysis and visualization. He also upgrades data-analytic and in part also programming skills.

Application: Carrying out basic data analysis on various types of data sets. Building of analytic methods. Preparation of charts and graphical presentations of data.

Reflection: The importance of modern information technology to analyze large amounts of data, the importance of data visualization.

Transferable skills: Working with a computer, data-analytic and algorithmic way of thinking.

Metode poučevanja in učenja:**Learning and teaching methods:**

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

Delež (v %) /

Načini ocenjevanja:

Weight (in %)

Assessment:

Način (pisni izpit, ustno izpraševanje, naloge, projekt): domače naloge, končni projekt	50%	Type (examination, oral, coursework, project): homeworks, final project
ocene: 1-5 (negativno), 6-10 (pozitivno) (po Statutu UL)	50%	grading: 1-5 (fail), 6-10 (pass) (according to the Statute of UL)

Reference nosilca / Lecturer's references:

Alen Orbanić:

- ŠIROK, Brane, BIZJAN, Benjamin, ORBANIĆ, Alen, BAJCAR, Tom. Mineral wool melt fiberization on a spinner wheel. Transactions of the Institution of Chemical Engineers. Part A, Chemical engineering research and design, ISSN 0263-8762, 2014, vol. 92, issue 1, str. 80-90, ilustr [COBISS.SI-ID 13057819]
- BIZJAN, Benjamin, ORBANIĆ, Alen, ŠIROK, Brane, KOVAČ, Boštjan, BAJCAR, Tom, KAVKLER, Iztok. A computer-aided visualization method for flow analysis. Flow measurement and instrumentation, ISSN 0955-5986. [Print ed.], Aug. 2014, vol. 38, str. 1-8, ilustr [COBISS.SI-ID 13484571]
- ORBANIĆ, Alen. Tools for networks. V: ALHAJJ, Reda (ur.), ROKNE, Jon (ur.). Encyclopedia of social network analysis and mining. New York: Springer, cop. 2014, str. 2166-2175, ilustr [COBISS.SI-ID 17145433]

Alexander Keith Simpson:

- EGGER, Jeff, MØGELBERG, Rasmus Ejlers, SIMPSON, Alex. The enriched effect calculus: syntax and semantics. Journal of logic and computation, ISSN 0955-792X, 2014, vol. 24, iss. 3, str. 615-654 [COBISS.SI-ID 17090137]
- EGGER, Jeff, MØGELBERG, Rasmus Ejlers, SIMPSON, Alex. Linear-use CPS translations in the enriched effect calculus. Logical methods in computer science, ISSN 1860-5974, 2012, vol. 8, iss. 4, paper 2 (str. 1-27) [COBISS.SI-ID 17090905]

