**DAUGAVPILS UNIVERSITĀTES**

**STUDIJU KURSA APRAKSTS**

|  |  |
| --- | --- |
| Studiju kursa nosaukums | **Datu analīzes rīki** |
| Studiju kursa kods (DUIS) |  |
| Zinātnes nozare |  |
| Kursa līmenis |  |
| Kredītpunkti | 4 |
| ECTS kredītpunkti | 6 |
| Kopējais kontaktstundu skaits | 64 |
| Lekciju stundu skaits | - |
| Semināru stundu skaits | - |
| Praktisko darbu stundu skaits | 64 |
| Laboratorijas darbu stundu skaits | - |
| Studējošā patstāvīgā darba stundu skaits | 96 |
|  | |
| Kursa autors(-i) | |
| Dr.sc.comp., doc. Vija Vagale | |
| Kursa docētājs(-i) | |
| Dr.sc.comp., doc. Vija Vagale Mg.sc.comp., lekt. Olga Perevalova Mg.sc.comp., lekt. Andrejs Radionovs | |
| Priekšzināšanas | |
|  | |
| Studiju kursa anotācija | |
| KURSA ANOTĀCIJA:  Kursa ietvaros tiek sniegts ieskats mūsdienās izmantojamajos liela apjoma datu un statistikas metožu apstrādes rīkos. Īpaša uzmanība tiek veltīta programmēšanas un datu vizualizācijas rīkiem. Kursā tiek apskatīta datu ieguve, modelēšana, apstrāde, analīze un vizualizācija.  KURSA MĒRĶIS ir sniegt zināšanas par biznesa procesu modelēšanu izmantojot programmēšanas un datu vizualizācijas rīkus.    KURSA UZDEVUMI:   1. Sniegt izpratni par datu analīzes nozīmīgumu; 2. Sniegt izpratni par datu analīzē izmantojamajiem programmēšanas un datu vizualizācijas rīkiem; 3. Sniegt zināšanas un praktiskas iemaņas datu ieguvē, transformācijā, modelēšanā, un vizualizācijā. | |
| Studiju kursa kalendārais plāns | |
| Studiju kursa struktūra: praktiskie darbi (P) – 64 st., studējošo patstāvīgais darbs (Pd) – 96 st.   1. Programmēšanas rīki datu apstrādei. P2, Pd2 2. Datu kopu ieguve un izveide. P10, Pd10 3. Datu vizualizēšana. P20, Pd20 4. Datu analīze un statistikas paņēmienu izmantošana R programmēšanas valodā. P16, Pd32 5. Datu analīze un statistikas paņēmienu izmantošana Python programmēšanas valodā. P16, Pd32 | |
| Studiju rezultāti | |
| ZINĀŠANAS:   1. Zina datu organizēšanas pamatprincipus; 2. Iegūta izpratne par datu analīzē izmantojamajiem programmēšanas un datu vizualizācijas rīkiem un to izmantošanas iespējām.   PRASMES:   1. Spēj izveidot un iegūt no tīmekļa analizējamās datu kopas; 2. Spēj transformēt iegūtās datu kopas; 3. Spēj vizualizēt datus un veidot atskaites; 4. Spēj pielietot vienkāršu statistiku datu analīzē.   KOMPETENCES:   1. Spēj izvēlēties piemērotu risinājumu datu analīzei, veikt datu analīzi un izdarīt patstāvīgus secinājumus. | |
| Studējošo patstāvīgo darbu organizācijas un uzdevumu raksturojums | |
| Studējošo darbs tiek organizēts individuāli.  Patstāvīgie uzdevumi:   1. Literatūras un interneta avotu studēšana atbilstoši kursa tematikai; 2. Praktisko starppārbaudījumu pildīšana.   Starppārbaudījumi:   1. Datu skrāpēšana, apstrāde, analīze un vizualizācija ar MS Power BI. 2. Datu ieguve, apstrāde, analīze un vizualizācija R valodā. 3. Datu ieguve, apstrāde, analīze un vizualizācija Python valodā. | |
| Prasības kredītpunktu iegūšanai | |
| Diferencēta ieskaite.   1. Nodarbību apmeklējums vismaz 70% no visa nodarbību skaita (veido 20% no kopējās atzīmes). 2. Praktiskie starppārbaudījumi (veido 80% no kopējās atzīmes).   STUDIJU REZULTĀTU VĒRTĒŠANA   |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | | Pārbaudījumu veidi |  | Studiju rezultāti | | | | | | | 1. | 2. | 3. | 4. | 5. | 6. | 7. | | 1. starppārbaudījums | + | + | + | + | + | + | + | | 2. starppārbaudījums | + | + | + | + | + | + | + | | 3. starppārbaudījums | + | + | + | + | + | + | + |   Studiju kursa apguve tā noslēgumā tiek vērtēta 10 ballu skalā saskaņā ar Latvijas Republikas normatīvajiem aktiem un atbilstoši "Nolikumam par studijām Daugavpils Universitātē" (apstiprināts DU Senāta sēdē 17.12. 2018., protokols Nr.15) vadoties pēc šādiem kritērijiem: iegūto zināšanu apjoms un kvalitāte, iegūtās prasmes un kompetences atbilstoši plānotajiem studiju rezultātiem.  Studējošo zināšanas, prasmes un kompetence tiek izvērtēta starpparbaudījumu rezultātu novērtēšanā. | |
| Kursa saturs | |
| Kursa struktūra: praktiskie darbi (P) - 64 st.  Praktisko darbu tēmas:   1. Ievads datu apstrādē izmantojamajos rīkos. 2. Tīmekļa datu ieguve (datu skrāpēšana) izmantojot JavaScript bibliotēkas. 3. Iegūto datu tīrīšana. 4. Ievads MS Power BI. 5. Tīmekļa datu ieguve ar MS Power BI. 6. Iegūto datu transformācija. Datu importēšana un eksportēšana. 7. Vienkārša datu analīze un vizualizācija. 8. Datu modelēšana. Attiecību starp tabulām veidošana. 9. Formulu veidošana ar DAX palīdzību. 10. Atskaišu veidošana un noformēšana. 11. Datu apvienošanas un grupēšanas automatizācija. 12. Datu vizualizācijas automatizēšana. 13. Datu noformēšana pēc nosacījuma. 14. Datu analīze. 15. Tendences un prognozes. 16. Atskaišu publicēšana MS Power BI. 17. Ievads R programmēšanas valodā. Datu tipi. 18. Datu vizualizācijas pakotnes (ggplot2, ggmap, shiny). 19. Datu eksportēšana un importēšana. Darbs ar failiem. 20. Datu manipulācijas un analīzes pakotnes (dplyr, stringr, lubridate, DataExplorer). 21. Statistikas pielietošana R valodā. 22. Statistikas metodes R valodā. 23. Mašīnmācīšanās pakotnes. 24. Dziļās mācīšanās pakotnes. 25. Python programmēšanas valoda. Datu tipi. 26. Datu vizualizācija ar Pyplot bibliotēkas palīdzību. 27. Darbs ar failiem. 28. Datu imports un analīze ar Panda bibliotēkas palīdzību. 29. Datu apstrāde ar NumPy bibliotēkas palīdzību. 30. Datizrace ar Orange Data Mining bibliotēkas palīdzību. 31. Datu apstrāde ar SciPy bibliotēkas palīdzību. 32. Mašīnmācīšanās bibliotēka Scikit-learn. | |
| Obligāti izmantojamie informācijas avoti | |
| 1. Amos David, Bader Dan, Jablonski Joanna, Heisler Fletcher, (2021). Puthon Basics: A Practical Introduction to Lython 3. Real Python, 635 pp. ISBN: 978-1775093329. 2. Douglas Alex, Roos Deon, Mancini Francesca, Couto Ana, Lusseau David (2023). An Introduction to R, https://intro2r.com 3. Dr. Troy A. Adair Jr., (2020). Introduction to R for Business (Intro to Data Science for Business). DataDicts, LLC; 2nd edition, 437 pp. ISBN: 978-1734113723 4. Jeremey Arnold, (2022). Learning Microsoft Power BI: Transforming Data into Insights. O'Reilly Media, 307 pp. ISBN: 978-1098112844 5. O'Connor Errin, (2019). Microsoft Power BI Dashboards Step by Step. Microsoft Press, 304 pp. ISBN: 978-1509308033 6. Philip Robbins, (2023). Python Programming for Beginners. Independently published, 114pp. ISBN: 979-8376161821. | |
| Papildus informācijas avoti | |
| 1. De Vries Andrie, (2015). R For Dummies. For Dummies; 2nd edition, 432 pp. ISBN: 978-1119055808 2. Field Andy, Miles Jeremy, Field Zoe, (2012). Discovering Statistics Using R. SAGE Publications Ltd, 957 pp. ISBN: 978-1446200469 3. Matloff Norman, (2011). The Art of R Programming: A Tour of Statistical Software Design. No Starch Press, 404 pp. ISBN: 978-1593273842 4. McGrath Mike, (2018). R for Data Analysis in easy steps - R Programming essentials. In Easy Steps Limited, 192 pp. ISBN: 978-1840787955 5. Nelli Fabio, (2015). Python Data Analytics (Data Analysis and Science Using Pandas, matplotlib, and the Python Programming Language). Apress, 337 pp. ISBN: 978-1-4842-0959-2. 6. Tilman, Davies (2016).The Book of R: A First Course in Programming and Statistics. No Starch Press, 832 pp. ISBN: 978-1593276515. 7. Wickham Garrett, Grolemund Garrett, (2017). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O′Reilly, 522 pp. ISBN: 978-1491910399. | |
| Periodika un citi informācijas avoti | |
| 1. Exercises in R, https://alexd106.github.io/intro2R/exercises.html 2. Get webpage data by providing examples, https://learn.microsoft.com/en-us/power-bi/connect-data/desktop-connect-to-web-by-example 3. Introduction to R, https://www.datacamp.com/courses/free-introduction-to-r 4. Free Visuals Elements, https://github.com/deldersveld/power-bi-sunset-visuals/tree/master/Custom-visuals 5. R Built-in Data Sets, http://www.sthda.com/english/wiki/r-built-in-data-sets#list-of-pre-loaded-data 6. R for Data Science: A Complete Guide for Beginners, https://www.knowledgehut.com/blog/data-science/r-for-data-science 7. R Plot Function, https://www.datamentor.io/r-programming/plot-function/ 8. R Programming for Data Science, https://www.geeksforgeeks.org/r-programming-for-data-science/ 9. R Scripts, http://mercury.webster.edu/aleshunas/R\_learning\_infrastructure/R%20scripts.html 10. The Comprehensive R Archive Network, https://cran.r-project.org/ 11. What are custom visuals in Power BI and where can I get them, https://learn.microsoft.com/en-us/power-bi/developer/visuals/power-bi-custom-visuals | |
|  | |
| Piezīmes | |
| Akadēmiskā maģistra studiju programmas “Datorzinātnes” studiju kurss.  Kurss tiek docēts latviešu valodā. | |