

Мастер академске студије
Информациони инжењеринг

Статистика у информационом инжењерингу

Основно програмирање помоћу језика R

(материјали за вежбе)

1. Употреба конзоле
2. Помоћне команде
3. Задаци за основне типове објеката
4. Задаци за посебне сложене типове објеката
5. Извори и литература

Активирање конзоле

начини за активирање конзоле

активирање конзоле директно (*R.exe*)

кроз терминал

активирање графичке апликације за конзолу (*Rgui.exe*)

кроз терминал или прегледач датотека

приступање програмима за *R* кроз терминал

потребно да одговарајући инсталациони директоријум буде доступан

потребно подесити системску променљиву *PATH*

на *Windows* оперативним системима потребно *PATH* проширити путањом

C:\Program Files\R\R-4.0.0\bin\i386 (пример за 32-битни *Windows*)

или

C:\Program Files\R\R-4.0.0\bin\x64 (пример за 64-битни *Windows*)

Корисне „пречице” у конзоли

преглед извршених команди

тастери ▲ и ▼

допуњавање команде приликом куцања

тастер **Tab**

уклањање приказаног текста

комбинација тастера **Ctrl + L**

прекид задавања или извршавања команде

тастер **Esc**

1. Употреба конзоле
- 2. Помоћне команде**
3. Задаци за основне типове објеката
4. Задаци за посебне сложене типове објеката
5. Извори и литература

Помоћне команде

преглед основних података о текућој сесији

команда **sessionInfo()**

отварање уграђене документације за језик R

команда **help.start()**

преглед претходних команди

команда **history()**

гашење сесије

команда **q()**

Помоћне команде

преглед објеката у окружењу

команда **ls()**

уклањање објеката

команда **rm(...)**

уклањање скоро свих објеката помоћу позива **rm(list=ls())**

Помоћне команде

текстуални приказ вредности

команда **print(...)**

команда **cat(...)**

команда **paste(...)**

мерење времена протеклог у извршавању команде

команда **system.time(...)**

Помоћне команде

постављање глобалних подешавања

команда **options(...)**

очитавања путање до радног директоријума

команда **getwd()**

подешавање путање до радног директоријума

команда **setwd("putanja-do-direktorijuma")**

покретање скрипт датотеке

команда **source("putanja-do-skripta")**

1. Употреба конзоле
2. Помоћне команде
- 3. Задаци за основне типове објеката**
4. Задаци за посебне сложене типове објеката
5. Извори и литература

Задаци за основне типове објеката

Задатак А1

Написати наредбе које решавају систем две линеарне једначине с две непознате!

две једначине су представљене помоћу два вектора, нпр.

```
j1 <- c(1, -2, 0) # x - 2y = 0
```

```
j2 <- c(1, 1, 3) # x + y = 3
```

решење је у облику вектора састављеног од два елемента, нпр.

```
[1] 2 1
```

дозвољена је употреба уграђених оператора

Задачи за основне типове објеката

Задатак А2

У оквиру конзоле реализовати сортирање вектора!

Задатак А3

Саставити скрипт којим се сортира вектор!

Задатак А4

Саставити скрипт којим се генерише вектор Фибоначијевих бројева са n елемената!

Задаци за основне типове објеката

Задатак А5

Написати функцију за сортирање вектора!

Задатак А6

Написати функцију за генерисање вектора Фибоначијевих бројева са n елемената!

Задатак А7

Написати функцију за сортирање вектора по алгоритму *Quicksort*!

Задаци за основне типове објеката

Задатак A8

Написати функцију за одређивање медијане нумеричког вектора!

Задатак A9

Написати функцију за одређивање модуса вектора!

Задатак A10

Написати функцију за пребрајање модуса вектора!

Корисне функције

формирање помоћних вектора

функција **c(...)**

оператор **:**

функција **seq(...)**

функција **sample(...)**

функција **rnorm(...)**

функција **rbinom(...)**

функција **runif(...)**

провера облика и типа за објекат

функција **is.numeric(...)**

функција **is.vector(...)**

функција **is.list(...)**

Задаци за основне типове објеката

Корисне функције

формирање грешке

функција **stop(...)**

функција **stopifnot(...)**

уклањање поновљених вредности

функција **unique(...)**

извршавање задате функције над појединачним елементима

функција **lapply(...)**

функција **sapply(...)**

1. Употреба конзоле
2. Помоћне команде
3. Задаци за основне типове објеката
- 4. Задаци за посебне сложене типове објеката**
5. Извори и литература

Задаци за посебне сложене типове објеката

Задатак Б1

Формирати скуп података *medalje!*

скуп података **medalje**

бројеви освојених медаља на ОИ 2016 по врстама медаља и државама
подаци агенције Ројтерс (од 21. 8. 2016)

<https://www.reuters.com/article/oly-medals/olympics-complete-medals-table-at-end-of-day-fifteen-idUSISS188500>

(приступљено 28. 2. 2024)

drzava	zlato	srebro	bronza
CHN	26	18	26
GBR	27	23	17
GER	17	10	15
RUS	19	18	19
USA	46	37	38

Задаци за посебне сложене типове објеката

Задатак Б2

Проширити скуп података *medalje* колоном у којој су забележени укупни бројеви освојених медаља!

Задатак Б3

Проширити скуп података *medalje* колоном у којој су забележени рангови држава по укупном броју медаља!

Задаци за посебне сложене типове објеката

Задатак Б4

Написати функцију за сортирање скупа података *medalje* по називу државе у обрнутом редоследу!

Задатак Б5

Написати функцију за сортирање скупа података *medalje* по бројевима освојених медаља!

користити метод табеле медаља за ОИ

сортирање прво по броју златних медаља

у случају истих бројева златних медаља, сортирање по броју сребрних медаља

у случају истих бројева сребрних медаља, сортирање по броју бронзаних медаља

у случају истих бројева бронзаних медаља, навођење по ознаци државе

Задатак Б6

Написати функцију за сортирање скупа података с произвољним бројем колона по узору на метод табеле медаља за ОИ!

Задаци за посебне сложене типове објеката

скуп података **lh**

регистар хуманих лекова у Републици Србији

датотека *lekovi_humani.csv* и датотека *lekovi_humani.xls*

подаци Агенције за лекове и медицинска средства Србије

Регистар лекова за употребу у хуманој медицини (формат CSV, од 19. 4. 2023)

https://www.alims.gov.rs/lekovi/lekovi_humani.csv

(преузето 28. 2. 2024)

Регистар лекова за употребу у хуманој медицини (формат XLS, од 19. 4. 2023)

https://www.alims.gov.rs/lekovi/lekovi_humani.xls

(преузето 28. 2. 2024)

Портал отворених података Републике Србије – ЛЕКОВИ ЗА УПОТРЕБУ У ХУМАНОЈ МЕДИЦИНИ

<https://data.gov.rs/sr/datasets/lekovi-za-upotrebu-u-khumanoj-meditcini/>

за све очекиване обраде дат одговарајући текстуални опис обраде

Задаци за посебне сложене типове објеката

Задатак В1

Учитати скуп података *lh* из *CSV* датотеке!

Задатак В2

Учитати скуп података *lh* из *XLS* датотеке!

Задатак В3

Приликом читавања скупа података *lh*, празне ћелије и ћелије које садрже искључиво косу црту третирају се као ћелије с недостајућом вредношћу!

Задатак В4

Приликом читавања скупа података *lh*, учитати вредности колоне *JKL* као знаковне вредности!

Задаци за посебне сложене типове објеката

Задатак В5

Уредити називе колона у скупу података *lh*!

Задатак В6

Уредити типове колона у скупу података *lh*!

Задатак В7

Проверити да ли у скупу података *lh* постоје лекови с истом вредношћу *JKL*!

Ако постоје такви лекови, приказати њихове називе!

Приказати назив лека чија се вредност *JKL* највише пута помиње!

Задатак В8

Пребројати лекове у скупу података *lh* по свакој појединачној врсти лека!

Задаци за посебне сложене типове објеката

скуп података **bk**

преглед података о запосленима у Белој кући (САД)

датотека *2013_Report_to_Congress_on_White_House_Staff.csv*

Извештај Конгресу о особљу Беле куће из 2013. (формат *CSV*, од 20. 2. 2014)

<https://open.obamawhitehouse.archives.gov/dataset/2013-report-congress-white-house-staff>

https://open.obamawhitehouse.archives.gov/sites/default/files/2013_Report_to_Congress_on_White_House_Staff.csv

(преузето 28. 2. 2024)

open.obamawhitehouse.gov

<https://open.obamawhitehouse.archives.gov/>

Задаци за посебне сложене типове објеката

Задатак Г1

Учитати скуп података **bk** из *CSV* датотеке!

водити рачуна о типовима података
посебно за обележје о имену (обележје *Name*)

Задатак Г2

Израчунати основне дескриптивне статистике за обележје о плати из скупа података **bk**!

Задатак Г3

Нацртати хистограм за обележје о плати из скупа података **bk**!

водити рачуна о дужинама X и Y осе
ширину појединачног интервала на X оси подесити на 10000

Задатак Г4

Нацртати графикон функције густине вероватноће за обележје о плати из скупа података **bk**!

Задаци за посебне сложене типове објеката

Задатак Г5

Нацртати бројчани графикон за обележје о плати из скупа података *bk*!

Задатак Г6

Испитати да ли вредности обележја о плати из скупа података *bk* прате нормалну расподелу!

применом одговарајућег статистичког теста

применом одговарајућег графикона

Задатак Г7

Формирати у скупу података *bk* ново обележје за позицију радника *Position Cat* које ће имати само две дозвољене вредности

Analyst која одговара вредностима из обележја *Position Title* са садржаним појмом *ANALYST*

NonAnalyst која одговара свим осталим вредностима из обележја *Position Title*

Задаци за посебне сложене типове објеката

Задатак Г8

Испитати да ли у скупу података *bk* постоје значајне разлике у плати између групе радника са статусом *Employee* и групе радника са статусом *Detailee*! Користити ниво значајности $\alpha = 0,05$!

применити одговарајући параметарски тест

проверити да ли су задовољени предуслови за адекватну примену теста

применити одговарајући непараметарски тест

проверити да ли су задовољени предуслови за адекватну примену теста

нацртати помоћни кутијаста графикон

Задатак Г9

У скуп података *bk* додати обележје *Salary Cat* с вредностима

Low која одговара износу плате мањем од 100000 долара

High која одговара износу плате који је једнак или већи од 100000 долара

Задаци за посебне сложене типове објеката

Задатак Г10

Испитати да ли за скуп података *bk* може да се одбаци претпоставка о независности обележја *Salary Cat* и *Position Cat*! Користити ниво значајности $\alpha = 0,05$!

применити одговарајући статистички тест

проверити да ли су задовољени предуслови за адекватну примену теста

Задатак Г11

У скуп података *bk* додати обележје *Position Cat New* с вредностима

Assistant за случај када назив радног места садржи реч *ASSISTANT*

Director за случај када назив радног места садржи реч *DIRECTOR*

Other за све преостале случајеве

Задаци за посебне сложене типове објеката

Задатак Г12

Над скупом података *bk* спровести једнофакторску анализу варијансе над обележјима *Salary* (зависна променљива) и *Position Cat New* (независна променљива)! Протумачити добијене резултате!

проверити да ли су задовољени предуслови за спровођење анализе

Задаци за посебне сложене типове објеката

скуп података **zajednica**

подаци из процене заједнице у вези са сигурношћу и гостољубивошћу у граду Болдер (Колорадо, САД)

датотека *Community_Assessment_Survey_-621747622075900632.csv*

отворени скуп података из анкете за процену заједнице у 2017. (формат CSV, од 27. 8. 2020)

https://open-data.bouldercolorado.gov/datasets/7fb80cf7a352462a9dfe3f2110c2e8f4_0/about

(преузето 17. 3. 2024)

https://open-data.bouldercolorado.gov/datasets/7fb80cf7a352462a9dfe3f2110c2e8f4_0/explore

отворени подаци града Болдер (Колорадо, САД)

<https://open-data.bouldercolorado.gov/>

лиценца и могућности коришћења

CC0 1.0

<https://creativecommons.org/publicdomain/zero/1.0>

Задаци за посебне сложене типове објеката

Задатак Д

Спровести анализу скупа података *zajednica* и о томе саставити извештај користећи формат *R Markdown!*

описати коришћени скуп података

припремити податке

учитати скуп података из *CSV* датотеке и при томе водити рачуна о типовима података

спровести препознавање садржаја датумских обележја

увести поредак између категорија за обележја код којих је то смислено

приказати податке о анкети која је описана у скупу података

приказати период у у којем је изведена анкета

приказати број учесника у анкети

приказати податке о учесницима у анкети

нацртати одговарајући графикон о старости учесника

нацртати одговарајући графикон о образовању учесника

нацртати одговарајући графикон о вероисповести учесника

нацртати одговарајући графикон о приходима учесника

...

Задаци за посебне сложене типове објеката

Задатак Д

Спровести анализу скупа података *zajednica* и о томе саставити извештај користећи формат *R Markdown!*

...

приказати расподелу одговора учесникâ на питања у вези са сигурношћу

нацртати одговарајући графикон о степену осећања сигурности

нацртати одговарајући графикон о доживљеном осећању несигурности

анализирати одговоре учесникâ на питање у вези с прихваћеношћу од стране људи

нацртати одговарајући графикон о осећању прихваћености од стране људи

осмислити погодну бројчану представу одговора датих на питање у вези с прихваћеношћу од стране људи

израчунати аритметичку средину и стандардну девијацију за одговоре у бројчаној представи

није потребно обухватити недостајуће или неодређене одговоре

...

Задаци за посебне сложене типове објеката

Задатак Д

Спровести анализу скупа података *zajednica* и о томе саставити извештај користећи формат *R Markdown*!

...

испитати независност обележја о осећању прихваћености од стране људи и обележја о томе да ли учесник студира на Универзитету Колорада у Болдеру

случајеви у којима није познато осећање прихваћености од стране људи или за које се то осећање не може поуздано утврдити могу бити изузети из испитивања
случајеви у којима није познато да ли учесник студира на Универзитету Колорада у Болдеру могу бити изузети из испитивања

осећања прихваћености од стране људи разврстати у три категорије (позитивну, неутралну и негативну)

за потребе испитивања независности два обележја користити одговарајући статистички тест и приказати резултате

проверити да ли су задовољени евентуални предуслови за адекватну примену теста

годашње најомене

у извештају треба приказати изворни код и резултате

у извештају уз приказ резултата треба да буду дата тумачења и да буде изнета одговарајућа дискусија

Задаци за посебне сложене типове објеката

скуп података **prisustvo**

присуство ђака настави у школама града Њујорк (Њујорк, САД)

датотека *2018-2019_Daily_Attendance_20240325.csv*

подаци Управе за образовање Града Њујорка

отворени скуп података о присуству ђака настави на дневном нивоу у школама у граду Њујорк у 2018/2019 (формат CSV, од 11. 9. 2019)

<https://data.cityofnewyork.us/Education/2018-2019-Daily-Attendance/x3bb-kg5j>
(преузето 25. 3. 2024)

отворени подаци града Њујорк (Њујорк, САД)

<https://opendata.cityofnewyork.us/>

Задаци за посебне сложене типове објеката

скуп података **ssko1e**

евиденција средњих школа града Њујорк (Њујорк, САД)

датотека *2018_DOE_High_School_Directory_20240325.csv*

подаци Управе за образовање Града Њујорка

отворени скуп података о средњим школама у граду Њујорк за 2018.
(формат CSV, од 23. 10. 2017)

<https://data.cityofnewyork.us/Education/2018-DOE-High-School-Directory/vw9i-7mzq>

(преузето 25. 3. 2024)

отворени подаци града Њујорк (Њујорк, САД)

<https://opendata.cityofnewyork.us/>

Задаци за посебне сложене типове објеката

Задатак Ђ1

Учитати скуп података *prisustvo* из *CSV* датотеке!

водити рачуна о називима обележја

водити рачуна о типовима података

посебно за обележје о датуму (обележје *Date*)

испитати да ли у подацима постоје недостајуће вредности

Задатак Ђ2

На основу скупа података *prisustvo* формирати регресиони модел у којем је зависно обележје оно о броју одсутних ученика а независно обележје оно о броју уписаних ученика!

испитати појединости о формираном регресионом моделу

нацртати одговарајући графикон на којем ће бити приказани коришћени подаци и представљен формиран регресиони модел

проверити да ли су задовољени предуслови за адекватну примену регресионе анализе

Задатак ЂЗ

На основу скупа података *prisustvo* формирати регресиони модел у којем је зависно обележје оно о броју одсутних ученика а независна обележја она о броју уписаних ученика и томе да ли је у току зима!

формирати ново нумеричко обележје о томе да ли је у току зима

проверити каква је расподела вредности за ново обележје

испитати појединости о формираном регресионом моделу

проверити да ли су задовољени предуслови за адекватну примену регресионе анализе

Задатак Ђ4

На основу скупа података *prisustvo* формирати регресиони модел у којем је зависно обележје оно о броју одсутних ученика а независна обележја она о броју уписаних ученика, томе да ли је у току зима и томе који је дан у седмици!

формирати ново нумеричко обележје о томе који је дан у седмици

проверити каква је расподела вредности за ново обележје

испитати појединости о формираном регресионом моделу

проверити да ли су задовољени предуслови за адекватну примену регресионе анализе

Задаци за посебне сложене типове објеката

Задатак Ђ5

Формирати нови скуп података који ће садржати податке из скупа података *prisustvo* проширене подацима из скупа података *sskole* тако да у новом скупу података буду обухваћени подаци само о средњим школама!

учитати скуп података *sskole* из *CSV* датотеке

задржати само податке за обележја о ознаци школе (обележје *dbn*) и успешности у завршетку школе (обележје *graduation_rate*)

водити рачуна о називима обележја

уклонити појаве у којима постоје недостајуће вредности

спојити податке из скупа података *prisustvo* и податке из скупа података *sskole* у један скуп података

Задаци за посебне сложене типове објеката

Задатак Ђ6

На основу новог скупа података формирати регресиони модел у којем је зависно обележје оно о броју одсутних ученика а независна обележја она о броју уписаних ученика, томе да ли је у току зима и томе који је дан у седмици!

испитати појединости о формираном регресионом моделу
проверити да ли су задовољени предуслови за адекватну примену регресионе анализе

Задатак Ђ7

На основу новог скупа података формирати регресиони модел у којем је зависно обележје оно о броју одсутних ученика а независна обележја она о броју уписаних ученика, томе да ли је у току зима, томе који је дан у седмици и успешности у завршетку школе!

испитати појединости о формираном регресионом моделу
проверити да ли су задовољени предуслови за адекватну примену регресионе анализе

Задаци за посебне сложене типове објеката

скуп података **parking.crvena**

црвена зона паркирања у Зрењанину

датотека *crvena-zona-2019-08-10-2019.csv*

подаци Градске управе Зрењанин

Паркинзи на територији града Зрењанина – црвена зона паркирања (формат CSV, од 15. 1. 2022)

<https://data.gov.rs/s/resources/parking-1/20220115-025247/crvena-zona-2019-08-10-2019.csv>

(преузето 25. 3. 2024)

Портал отворених података Републике Србије – Паркинг – ГУ Зрењанин

<https://data.gov.rs/sr/datasets/parking-1/>

за све очекиване обраде дат одговарајући текстуални опис обраде

Задаци за посебне сложене типове објеката

скуп података **parking.zuta**

жута зона паркирања у Зрењанину

датотека *zuta-zona-2019-08-10-2019.csv*

подаци Градске управе Зрењанин

Паркинзи на територији града Зрењанина – жута зона паркирања (формат CSV, од 15. 1. 2022)

<https://data.gov.rs/s/resources/parking-1/20220115-025400/zuta-zona-2019-08-10-2019.csv>

(преузето 25. 3. 2024)

Портал отворених података Републике Србије – Паркинг – ГУ Зрењанин

<https://data.gov.rs/sr/datasets/parking-1/>

за све очекиване обраде дат одговарајући текстуални опис обраде

Задаци за посебне сложене типове објеката

скуп података **parking.zelena**

зелена зона паркирања у Зрењанину

датотека *zelena-zona-2019-08-10-2019.csv*

подаци Градске управе Зрењанин

Паркинзи на територији града Зрењанина – зелена зона паркирања (формат CSV, од 15. 1. 2022)

<https://data.gov.rs/s/resources/parking-1/20220115-025343/zelena-zona-2019-08-10-2019.csv>

(преузето 25. 3. 2024)

Портал отворених података Републике Србије – Паркинг – ГУ Зрењанин

<https://data.gov.rs/sr/datasets/parking-1/>

за све очекиване обраде дат одговарајући текстуални опис обраде

Задаци за посебне сложене типове објеката

Задатак Е

Спровести анализу скупова података *parking.crvena*, *parking.zuta* и *parking.zelena* ослањајући се првенствено на употребу пакета из колекције *tidyverse*!

учитати појединачне скупове података

учитати скупове података *parking.crvena*, *parking.zuta* и *parking.zelena* из одговарајућих **CSV** датотека

утврдити димензије учитаних скупова података

објединити скупове података

повезати скупове података *parking.crvena*, *parking.zuta* и *parking.zelena* у обједињени скуп података

приказати део обједињеног скупа података и утврдити димензије обједињеног скупа података

прочистити обједињени скуп података уклањањем евентуалних појава у којима недостају вредности за бројеве места на паркингу

ако у обједињеном скупу података постоје појаве код којих недостају вредности за обележја о бројевима места на паркингу, онда такве појаве треба уклонити
утврдити димензије обједињеног скупа података након евентуалног прочишћења

...

Задаци за посебне сложене типове објеката

Задатак Е

Спровести анализу скупова података *parking.crvena*, *parking.zuta* и *parking.zelena* ослањајући се првенствено на употребу пакета из колекције *tidyverse*!

...

на основу обједињеног скупа података приказати податке о паркинзима у црвеној зони паркирања

за сваки паркинг у црвеној зони паркирања приказати назив зоне, назив улице, вредности за бројеве места на паркингу и радно време

појаве приликом приказа сортирати по броју места на паркингу и то у обрнутом поретку

у обједињени скуп података додати бројчано обележје о зони паркирања

свакој зони паркирања одговара другачија бројчана ознака

црвеној зони паркирања може одговарати бројчана ознака 1

жутој зони паркирања може одговарати бројчана ознака 2

зеленој зони паркирања може одговарати бројчана ознака 3

бројчано обележје о зони паркирања поставити да стоји непосредно иза текстуалног обележја о зони паркирања

појаве сортирати по бројчаној ознаци зоне паркирања и то у растућем поретку

...

Задаци за посебне сложене типове објеката

Задатак Е

Спровести анализу скупова података *parking.crvena*, *parking.zuta* и *parking.zelena* ослањајући се првенствено на употребу пакета из колекције *tidyverse*!

...

у обједињеном скупу података обележје у вези с наплатом једног сата паркирања преименовати тако да му нови назив буде *Sat*

у обједињени скуп података додати засебна обележја за телефонски број преко којег је могуће уплатити један сат паркирања и за износ који се наплаћује за један сат паркирања

на основу обједињеног скупа података за сваку зону паркирања приказати број паркинга у тој зони и просечан број места на паркингу у тој зони

на основу обједињеног скупа података нацртати стубичасти графикон на којем ће бити приказани бројеви паркинга по зонама паркирања и то редом за црвену, жуту и зелену зону паркирања

на основу обједињеног скупа података нацртати хистограм за број места на паркингу

Задаци за посебне сложене типове објеката

скуп података **dobra**

културна, архитектонска и градитељска добра у Шапцу

датотека *sabac-nepokretnakulturnadobra-poli.geojson*

подаци Градске управе Шабац

Непокретна културна добра и архитектонско и градитељско наслеђе града Шапца (формат *GeoJSON*, креирање од 1. 2. 2022. а последња промена од 2. 2. 2022)

<https://data.gov.rs/s/resources/nepokretna-kulturna-dobra-i-arkhitektonsko-i-graditeljsko-nasledje-grada-shaptsa/20220202-180827/sabac-nepokretnakulturnadobra-poli.geojson>

(преузето 25. 3. 2024)

Портал отворених података Републике Србије – Непокретна културна добра и архитектонско и градитељско наслеђе града Шапца – НКД Шабац

<https://data.gov.rs/sr/datasets/nepokretna-kulturna-dobra-i-arkhitektonsko-i-graditeljsko-nasledje-grada-shaptsa/>

за све очекиване обраде дат одговарајући текстуални опис обраде

Задаци за посебне сложене типове објеката

Задатак Ж

Спровести анализу скупа података *dobra* ослањајући се првенствено на употребу пакета из колекције *tidyverse*!

учитати скуп података

учитати скуп података *dobra* из одговарајуће *GeoJSON* датотеке и при томе трансформисати податке тако да буду у облику таблице (енгл. *tibble*) али задржавајући само вредности основних обележја без координата

приказати део скупа података и утврдити димензије скупа података

у скуп података додати обележје о веку изградње добра

вредности новог обележја одредити на основу вредности обележја о години изградње

обележје о веку изградње поставити да стоји непосредно иза обележја о години изградње

у скуп података додати обележје о старости добра

вредности новог обележја одредити на основу вредности обележја о години изградње

обележје о старости поставити да стоји непосредно иза обележја о добу изградње

...

Задаци за посебне сложене типове објеката

Задатак Ж

Спровести анализу скупа података *dobra* ослањајући се првенствено на употребу пакета из колекције *tidyverse*!

...

нацртати наслагани стубичасти графикон за број добара по нивоима значаја

за сваки ниво значаја добра нацртати по један матични стуб

висина матичног стуба одговара броју добара за придружени ниво значаја
сваки матични стуб треба да буде састављен од вертикално наслаганих
компонентних стубова који одговарају нивоима категорије добра

висина компонентног стуба одговара броју добара за придружене нивое значаја
и категорије

компонентне стубове обојити у зависности од придруженог нивоа категорије

нацртати груписани стубичасти графикон за број добара по нивоима значаја

за сваки ниво значаја добра нацртати по једну групу стубова

стубови унутар групе одговарају различитим нивоима категорије добра
висина стуба унутар групе одговара броју добара за придружене нивое значаја
и категорије

стубове унутар групе обојити у зависности од придруженог нивоа категорије

...

Задаци за посебне сложене типове објеката

Задатак Ж

Спровести анализу скупа података *dobra* ослањајући се првенствено на употребу пакета из колекције *tidyverse*!

...

нацртати кутијаста графикон за старост добра

нацртати упоредо кутијасте графиконе за старост добра по различитим нивоима значаја добра

нацртати виолински графикон за старост добра

на графикону нацртати ознаке за први, други и трећи квартил

нацртати графикон расејања за добра

X оса одговара години изградње добра

Y оса одговара површини добра

боја симбола одговара категорији добра

облик симбола одговара значају добра

ознака уз симбол одговара идентификатору добра

избећи преклапање ознака симбола на графикону

Задаци за посебне сложене типове објеката

Корисне функције – основни пакети

преглед основних података о скупу података

функција **summary(...)**

функција **str(...)**

одређивање и уређивање поретка унутар скупа података

функција **rank(...)**

функција **order(...)**

учитавање података из *CSV* датотека

функција **read.csv(...)**

функција **read.csv2(...)**

прилагођење знаковних вредности синтакси за називе у језику *R*

функција **make.names(...)**

Задаци за посебне сложене типове објеката

Корисне функције – основни пакети

формирање табеле контингенције на основу задатих фактора

функција **table(...)**

функција **xtabs(...)**

примена регуларних израза у претраживању и измени знаковних вредности

функција **grep(...)**

функција **grep1(...)**

функција **gsub(...)**

рад над датумским вредностима

функција **as.Date(...)**

функција **weekdays(...)**

спајање два скупа података

функција **merge(...)**

Задаци за посебне сложене типове објеката

Корисне функције – пакети колекције *tidyverse*

учитавање података из *XLS* и *XLSX* датотека

функција **`readxl::read_excel(...)`**

учитавање података из *JSON* датотека

функција **`jsonlite::fromJSON(...)`**

помоћни преглед података

функција **`pillar::glimpse(...)`**

рад над недостајућим подацима

функција **`tidyr::drop_na(...)`**

функција **`tidyr::complete(...)`**

векторско извођење гранања

функција **`dplyr::case_when(...)`**

Задаци за посебне сложене типове објеката

Корисне функције – пакети колекције *tidyverse*

рад над знаковним вредностима

функција **stringr::str_trim(...)**

функција **stringr::str_sub(...)**

функција **stringr::str_locate(...)**

рад над датумским и временским вредностима

функција **lubridate::now(...)**

функција **lubridate::year(...)**

рад над групама података

функција **dplyr::ungroup(...)**

визуализација података

функција **ggplot2::geom_violin(...)**

1. Употреба конзоле
2. Помоћне команде
3. Задаци за основне типове објеката
4. Задаци за посебне сложене типове објеката
- 5. Извори и литература**

Основни извори и литература

- ◆ R Project. An introduction to R – Notes on R: A programming environment for data analysis and graphics version 4.5.1 (2025-06-13). Internet: <https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf>
- ◆ R Project. R language definition – Version 4.5.1 (2025-06-13) DRAFT. Internet: <https://cran.r-project.org/doc/manuals/r-release/R-lang.pdf>
- ◆ R Project. R: A language and environment for statistical computing – Reference index – The R Core Team – Version 4.5.1 (2025-06-13). Internet: <https://cran.r-project.org/doc/manuals/r-release/fullrefman.pdf>
- ◆ Adler J. R in a nutshell: A desktop quick reference. 2nd edition. O'Reilly; 2012.
- ◆ RStudio. R Markdown. Internet: <https://rmarkdown.rstudio.com/>

Основни извори и литература

- ◆ Tidyverse. Tidyverse. Internet: <https://www.tidyverse.org/>
- ◆ Posit Software. Posit cheatsheets. Internet: <https://posit.co/resources/cheatsheets/>
- ◆ R Project. CRAN: Package tidyverse. Internet: <https://cran.r-project.org/web/packages/tidyverse/index.html>

Извори података

- ◆ скуп података **medalje**
 - ◆ бројеви освојених медаља на ОИ 2016 по врстама медаља и државама
 - ◆ подаци агенције Ројтерс (од 21. 8. 2016)
 - ◆ <https://www.reuters.com/article/oly-medals/olympics-complete-medals-table-at-end-of-day-fifteen-idUSISS188500>
(приступљено 28. 2. 2024)

Извори података

- ◆ скуп података **lh**
 - ◆ датотека *lekovi_humani.csv* и датотека *lekovi_humani.xls*
 - ◆ подаци Агенције за лекове и медицинска средства Србије
 - ◆ Регистар лекова за употребу у хуманој медицини (формат CSV, од 19. 4. 2023)
 - ◆ https://www.alims.gov.rs/lekovi/lekovi_humani.csv (преузето 28. 2. 2024)
 - ◆ Регистар лекова за употребу у хуманој медицини (формат XLS, од 19. 4. 2023)
 - ◆ https://www.alims.gov.rs/lekovi/lekovi_humani.xls (преузето 28. 2. 2024)
 - ◆ Портал отворених података Републике Србије – ЛЕКОВИ ЗА УПОТРЕБУ У ХУМАНОЈ МЕДИЦИНИ
 - ◆ <https://data.gov.rs/sr/datasets/lekovi-za-upotrebu-u-khumanoj-meditsini/>
 - ◆ за све очекиване обраде дат одговарајући текстуални опис обраде

Извори података

- ◆ скуп података **bk**
 - ◆ датотека *2013_Report_to_Congress_on_White_House_Staff.csv*
 - ◆ преглед података о запосленима у Белој кући (САД)
 - ◆ Извештај Конгресу о особљу Беле куће из 2013. (формат CSV, од 20. 2. 2014)
 - ◆ <https://open.obamawhitehouse.archives.gov/dataset/2013-report-congress-white-house-staff>
 - ◆ https://open.obamawhitehouse.archives.gov/sites/default/files/2013_Report_to_Congress_on_White_House_Staff.csv (преузето 28. 2. 2024)
 - ◆ *open.obamawhitehouse.gov*
 - ◆ <https://open.obamawhitehouse.archives.gov/>

Извори података

- ◆ скуп података **zajednica**
 - ◆ датотека *Community_Assessment_Survey_-621747622075900632.csv*
 - ◆ подаци из процене заједнице у вези са сигурношћу и гостољубивошћу у граду Болдер (Колорадо, САД)
 - ◆ отворени скуп података из анкете за процену заједнице у 2017. (формат CSV, од 27. 8. 2020)
 - ◆ https://open-data.bouldercolorado.gov/datasets/7fb80cf7a352462a9dfe3f2110c2e8f4_0/about
(преузето 17. 3. 2024)
 - ◆ https://open-data.bouldercolorado.gov/datasets/7fb80cf7a352462a9dfe3f2110c2e8f4_0/explore
 - ◆ отворени подаци града Болдер (Колорадо, САД)
 - ◆ <https://open-data.bouldercolorado.gov/>
 - ◆ лиценца и могућности коришћења
 - ◆ *CC0 1.0*
 - ◆ <https://creativecommons.org/publicdomain/zero/1.0>

Извори података

- ◆ скуп података **prisustvo**
 - ◆ датотека *2018-2019_Daily_Attendance_20240325.csv*
 - ◆ подаци Управе за образовање Града Њујорка
 - ◆ присуство ђака настави у школама града Њујорк (Њујорк, САД)
 - ◆ отворени скуп података о присуству ђака настави на дневном нивоу у школама у граду Њујорк у 2018/2019 (формат CSV, од 11. 9. 2019)
 - ◆ <https://data.cityofnewyork.us/Education/2018-2019-Daily-Attendance/x3bb-kg5j> (преузето 25. 3. 2024)
 - ◆ отворени подаци града Њујорк (Њујорк, САД)
 - ◆ <https://opendata.cityofnewyork.us/>

Извори података

- ◆ скуп података **skskole**
 - ◆ датотека *2018_DOE_High_School_Directory_20240325.csv*
 - ◆ подаци Управе за образовање Града Њујорка
 - ◆ евиденција средњих школа града Њујорк (Њујорк, САД)
 - ◆ отворени скуп података о средњим школама у граду Њујорк за 2018. (формат *CSV*, од 23. 10. 2017)
 - ◆ <https://data.cityofnewyork.us/Education/2018-DOE-High-School-Directory/vw9i-7mzq>
(преузето 25. 3. 2024)
 - ◆ отворени подаци града Њујорк (Њујорк, САД)
 - ◆ <https://opendata.cityofnewyork.us/>

Извори података

- ◆ скуп података **parking.crvena**
 - ◆ датотека *crvena-zona-2019-08-10-2019.csv*
 - ◆ подаци Градске управе Зрењанин
 - ◆ Паркинзи на територији града Зрењанина – црвена зона паркирања (формат CSV, од 15. 1. 2022)
 - ◆ <https://data.gov.rs/s/resources/parking-1/20220115-025247/crvena-zona-2019-08-10-2019.csv>
(преузето 25. 3. 2024)
 - ◆ Портал отворених података Републике Србије – Паркинг – ГУ Зрењанин
 - ◆ <https://data.gov.rs/sr/datasets/parking-1/>
 - ◆ за све очекиване обраде дат одговарајући текстуални опис обраде

Извори података

- ◆ скуп података **parking.zuta**
 - ◆ датотека *zuta-zona-2019-08-10-2019.csv*
 - ◆ подаци Градске управе Зрењанин
 - ◆ Паркинзи на територији града Зрењанина – жута зона паркирања (формат CSV, од 15. 1. 2022)
 - ◆ <https://data.gov.rs/s/resources/parking-1/20220115-025400/zuta-zona-2019-08-10-2019.csv>
(преузето 25. 3. 2024)
 - ◆ Портал отворених података Републике Србије – Паркинг – ГУ Зрењанин
 - ◆ <https://data.gov.rs/sr/datasets/parking-1/>
 - ◆ за све очекиване обраде дат одговарајући текстуални опис обраде

Извори података

- ◆ скуп података **parking.zelena**
 - ◆ датотека *zelena-zona-2019-08-10-2019.csv*
 - ◆ подаци Градске управе Зрењанин
 - ◆ Паркинзи на територији града Зрењанина – зелена зона паркирања (формат CSV, од 15. 1. 2022)
 - ◆ <https://data.gov.rs/s/resources/parking-1/20220115-025343/zelena-zona-2019-08-10-2019.csv>
(преузето 25. 3. 2024)
 - ◆ Портал отворених података Републике Србије – Паркинг – ГУ Зрењанин
 - ◆ <https://data.gov.rs/sr/datasets/parking-1/>
 - ◆ за све очекиване обраде дат одговарајући текстуални опис обраде

Извори података

- ◆ скуп података **dobra**
 - ◆ датотека *sabac-nepokretnakulturnadobra-poli.geojson*
 - ◆ подаци Градске управе Шабац
 - ◆ Непокретна културна добра и архитектонско и градитељско наслеђе града Шапца (формат *GeoJSON*, креирање од 1. 2. 2022. а последња промена од 2. 2. 2022)
 - ◆ <https://data.gov.rs/s/resources/nepokretna-kulturna-dobra-i-arkhitektonsko-i-graditeljsko-nasledje-grada-shaptsa/20220202-180827/sabac-nepokretnakulturnadobra-poli.geojson>
(преузето 25. 3. 2024)
 - ◆ Портал отворених података Републике Србије – Непокретна културна добра и архитектонско и градитељско наслеђе града Шапца – НКД Шабац
 - ◆ <https://data.gov.rs/sr/datasets/nepokretna-kulturna-dobra-i-arkhitektonsko-i-graditeljsko-nasledje-grada-shaptsa/>
 - ◆ за све очекиване обраде дат одговарајући текстуални опис обраде

Мастер академске студије
Информациони инжењеринг

Статистика у информационом инжењерингу

Основно програмирање помоћу језика R

(материјали за вежбе)