Interna IT izobraževanja KIS

Uporaba statističnega programa R

Jana Obšteter

Kazalo vsebine

[Modul A: Osnove R-a, R-studia in pisanja kode 2](#_Toc46984003)

[1. Spoznavanje z R-studiem 2](#_Toc46984004)

[3. Osnove organizacije dela z datotekami 2](#_Toc46984005)

[Modul B: Osnovna obdelava podatkov ter pisanje kode v R 3](#_Toc46984006)

[Modul C: Naprednejša statistika z R in pisanje funkcij 4](#_Toc46984007)

# Modul A: Osnove R-a, R-studia in pisanja kode

## 1. Spoznavanje z R-studiem

* struktura R studia
* pisanje programske kode in izvrševanje ukazov
* nameščanje paketov in nalaganje knjižnic
* spletni viri:
  + programska koda
  + R pomoč in zgledi kode

## 2. Osnove organizacije dela z datotekami

* kako nastaviti delovni imenik
* branje / pisanje različnih tipov datotek
* lastnosti prebranih podatkov
* primer branja podatkov
* primer ustvarjanja tabele

## 3. Spoznavanje s kodo

* podatkovni tipi
* podatkovne strukture
* osnovne računske operacije
* osnovne “vgrajene” funkcije in pomoč
* dobre prakse pisane kode
* samodejno dopolnjevanje
* najpogostejše napake

## 4. Pomoč za R

* stack overflow

# Modul B: Osnovna obdelava podatkov ter pisanje kode v R

## 1. Povzemanje podatkov

* Opisna statistika (srednje vrednosti, razpršenost podatkov; numerične / kategorične spremenljivke)
* funkcije v base / dplyr paketih
* apply družina funkcij

## 2. Manipulacija s podatki

* preurejanje tabel
* združevanje tabel
* knjižnice reshape, tidyr, tidyverse

## 3. Osnovna statistika – preverjanje hipotez

* Kritične vrednosti
* t-test
* Chi2
* ANOVA
* osnovni post-hoc testi

## 4. Vizualizacija podatkov z lattice / ggplot2

* Različne vrste grafikonov
  + histogram
  + gostota podatkov
  + stolpčni grafikoni
  + razstreseni grafikoni
  + črtni grafikon
  + škatlasti grafikon
* shranjevanje grafov

## 5. Shranjevanje delovnega prostora v R-u

# Modul C: Naprednejša statistika z R in pisanje funkcij

## 1. Linearni modeli

* preverjanje predpostavk modela
* priprava modela
* koeficienti in rezultati modela
* preverjanje ustreznosti modela
* napovedovanje
* post-hoc
* večrazsežna (“multiple”) linearna regresija

## 2. Logistična regresija

## 3. Pisanje zank in funkcij

## 4. Druge možnosti R-a

* shiny R
* knit
* genomika