



Универзитет у Београду
Математички факултет

СТАТИСТИЧКИ СОФТВЕР 1 (3В)

17. август 2025. године

1. База података `iris` садржи флоралне мере за три врсте цвета (`setosa`, `versicolor`, `virginica`).
 - (а) [1 поен] Користећи функцију `summarise()` израчунати за сваку меру (`Sepal.Length`, `Sepal.Width`, `Petal.Length`, `Petal.Width`) укупан распон и укупну дисперзију, без обзира на врсту.
 - (б) [2 поен] Креирати нову променљиву `Petal.Ratio` као однос `Petal.Length / Petal.Width`, а затим израчунати просечну вредност `Petal.Ratio` по врсти цвета и рангирати врсте од најмање до највеће вредности овог односа.
 - (в) [1 поен] Користећи функције `group_by()` и `summarise()` одредити за сваку врсту просечну `Sepal.Length`, медијану `Sepal.Width`, и максималну `Petal.Length`.
2. У бази `murders` (паке `dslabs`) налазе се подаци о броју убиства и стопи злочина по савезним државама САД, а у бази `presidential_election_2016.csv` подаци о резултатима председничких избора 2016. године по државама.
 - (а) [3 поен] Из базе `presidential_election_2016` издвојити резултате за Доналда Трампа и Хилари Клинтон, па их функцијом `spread()` трансформисати у широку табелу са новом колоном `trump_margin` која представља разлику у процентима гласова.
 - (б) [1 поен] Проверити да ли су називи држава (`state`) потпуно усклађени, а затим користећи функцију `full_join()` спојити табелу из дела под (а) и базу `murders` по колони `state` у нову табелу (`analyse`).
 - (в) [2 поен] У табели `analyse`, креирати следеће нове променљиве:
 - `murder_rate` - стопа убиства на 100.000 становника;
 - `murder_level` - категоричка променљива са вредностима:
 - `low`: ако је стопа убиства (`murder_rate`) мања од 2;
 - `midium`: ако је стопа убиства (`murder_rate`) 2 и 5;
 - `high`: ако је стопа убиства (`murder_rate`) већа 5.
 - (г) [1 поен] Израчунати просечан број изборних гласова (`electoral_votes`) по нивоу убиства (`murder_level`)
 - (д) [2 поен] Израчунати средњу вредност разлике у процентима између Трампа и Клинтонове по регионима (`region`), и визуализовати је као тракасти дијаграм (паке `ggplot2`), обојен по `murder_level`.
 - (ђ) [2 поен] Приказати `scatter plot` (паке `ggplot2`) за `murder_rate` према `trump_margin`, и приказати тренд помоћу `geom_smooth()`.