Exam Results Analysis

Liena Kāpiņa

12/08/2025

Data obtained from https://www.viaa.gov.lv/lv/valsts-parbaudes-darbu-statistika

Cleaning the data

Renaming data so important column names match between the two types of exams and important fields do not contain special characters.

```r
library(readxl)
ol <- read_excel("visi_dati_2024_0810.xlsx", sheet = "MATOL")
al <- read_excel("visi_dati_2024_0810.xlsx", sheet = "MATAL")
colnames(al)[5] <- '1 dala'
colnames(al)[7] <- '2 dala'
colnames(al)[9] <- '3 dala'
colnames(al)[11] <- '4 dala'
colnames(ol)[5] <- '1 dala'
colnames(ol)[7] <- '2 dala'
colnames(al)[colnames(al) == 'Urbanizācija'] <- 'Urbanizacija'
colnames(ol)[colnames(ol) == 'Urbanizācija'] <- 'Urbanizacija'
```

Remove unnecessary columns

```r
ol$N.p.k. <- NULL
ol$Norise <- NULL
ol$`Procenti Zināšanas. izpratne un prasmes` <- NULL
ol$`Procenti Kompleksu problēmu risināšana` <- NULL
ol$`Procenti kopā` <- NULL
al$N.p.k.<- NULL
al$Norise <- NULL
al$`Procenti Zināšanas. izpratne un prasmes (optimālā līmeņa saturs)` <- NULL
al$`Procenti Kompleksu problēmu risināšana (optimālā līmeņa saturs)` <- NULL
al$`Procenti Zināšanas. izpratne un prasmes (augstākā līmeņa saturs)` <- NULL
al$`Procenti Kompleksu problēmu risināšana (augstākā līmeņa saturs)` <- NULL
al$`Procenti kopā` <- NULL
```

Add missing columns so data frames for MATOL and MATAL can be merged into one.

```r
ol$`3 dala` <- as.double(NA)
ol$`4 dala` <- as.double(NA)
data <- rbind(ol, al)
```

Then split data into multiple frames based on urbanization and type of school. Data from persons taking the exam at a university will be omitted later because of irrelavance to the research question.

```r
split <- split(data, data$Urbanizacija)
riga <- split$`1`
city <- split$`2`
town <- split$`3`
country <- split$`4`

split <- split(riga, riga$Tips)
riga_vsk <- split$`3`
riga_talm <- split$`4`
riga_gymn <- split$`6`
riga_prof <- split$`7`
LU <- split$`9`

split <- split(city, city$Tips)
city_vsk <- split$`3`
city_talm <- split$`4`
city_gymn <- split$`6`
city_prof <- split$`7`

split <- split(town, town$Tips)
town_vsk <- split$`3`
town_talm <- split$`4`
town_gymn <- split$`6`
town_prof <- split$`7`

split <- split(country, country$Tips)
country_vsk <- split$`3`
country_talm <- split$`4`
country_gymn <- split$`6`
country_prof <- split$`7`
```

Data Analysis

Let's compare the number of people taking the exam in each type of school. Rows are Riga, then state significance cities, regional significance cities and finally countryside. Columns 3, 4, 6, 7 and 9 are secondary school, distance learning secondary school, state gymnasium, professional education and University of Latvia. Note that in University of Latvia only people with a secondary school education are taking exams.

```r
library(pander)
```

```
## Warning: package 'pander' was built under R version 4.4.3
```

```r
proportions <- table (data$Urbanizacija, data$Tips)
pander(proportions, round = 3, caption = "Number of Exam Takers")
```

Table 1: Number of Exam Takers

3	4	6	7	9
3020	752	988	1638	55
1808	38	1202	1573	0
1931	171	920	456	0
758	788	0	509	0