Exam Results Analysis

Liena Kāpiņa

12/08/2025

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

[Add description of data!!]

Cleaning the data

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

```
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.

```
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.

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

Descriptive 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 are secondary school, distance learning secondary school, state gymnasium, professional education and University of Latvia. Note that in University of Latvia only people already with a secondary school diploma take exams.

```
library(pander)
```

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

```
proportions <- table (data$Urbanizacija, data$Tips)
rownames(proportions) <- c("Riga", "State city", "Regional city", "Countryside")
colnames(proportions) <- c("Secondary", "Distance", "Gymnasium", "Pofessional ed", "UL")
pander(proportions, caption = "Number of Exam Takers")
```

Table 1: Number of Exam Takers Let's calculate the mean scores of the first part of the exam. The maximum number of points in this part was 75.

	Secondary	Distance	Gymnasium	Pofessional ed	UL
Riga	3020	752	988	1638	55
State city	1808	38	1202	1573	0
Regional city	1931	171	920	456	0
Countryside	758	788	0	509	0

```
split <- split(data, data$Urbanizacija)
riga <- split$`1`
city <- split$`2`
town <- split$`3`
country <- split$`4`
means <- aggregate(riga$`1 dala`, list(riga$Tips), FUN=mean)
colnames(means)[colnames(means) == 'Group.1'] <- 'Type'
colnames(means)[colnames(means) == 'x'] <- 'Riga'
col <- aggregate(city$`1 dala`, list(city$Tips), FUN=mean)
col[nrow(col) + 1,] <- c(9, NA)
means$'City' <- col$'x'
col <- aggregate(town$`1 dala`, list(town$Tips), FUN=mean)
col[nrow(col) + 1,] <- c(9, NA)
means$'Town' <- col$'x'
col <- aggregate(country$`1 dala`, list(country$Tips), FUN=mean)
col[nrow(col) + 1,] <- c(9, NA)
col <- rbind(col[1:2,], c(6, NA), col[-(1:2),])
means$'Country' <- col$'x'
means
```

```
##   Type     Riga     City     Town  Country
## 1    3 40.15348 33.73424 33.97048 32.15369
## 2    4 15.44548 11.78947 18.54971 18.21066
## 3    6 59.25455 47.08153 42.61576       NA
## 4    7 23.17491 21.22441 16.69518 16.62181
## 5    9 18.16364       NA       NA       NA
```

Then do the same for the second part of the exam. The maximum number of points in this part was 25.

```
means
```

```
##   Type      Riga      City     Town  Country
## 1    3  9.094371  7.065542 6.521233 6.255277
## 2    4  2.215426  1.197368 2.748538 2.826777
## 3    6 17.904352 11.167221 9.310870       NA
## 4    7  3.630342  3.159886 2.246711 2.261297
## 5    9  2.990909        NA       NA       NA
```

Let's look at the histograms of first part results from secondary schools.