

# BAS4SC Survey Analysis

Dejan

2024-12-07

## Introduction

This document analyzes data from the BAS4SC survey. The analysis covers responses from students, teachers, and companies across several countries, including Serbia, Croatia, Slovenia, and Poland. Each section of the analysis includes data processing, summarization, and the creation of joint tables.

## Loading Libraries

We start by loading the required libraries for data manipulation and reading Excel files.

```
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

## Summarization Function

We define a function `sumarizator` to summarize the survey responses. This function processes a specific set of columns from the dataset, calculates the frequency of each response level, and returns a matrix.

```
sumarizator <- function(dt) {
  dt <- dt[, 6:106]
  tt <- t(dt)
  mat <- matrix(0, ncol = 5, nrow = dim(dt)[2])

  nivoi <- levels(as.factor(tt))

  colnames(mat) <- nivoi
  rownames(mat) <- colnames(dt)
```

```

for (i in 1:nrow(tt)) {
  for (level in nivoi) {
    mat[i, level] <- summary(as.factor(tt[i, ]))[level]
  }
}

mat[is.na(mat)] <- 0
return(mat)
}

```

## Students Data Analysis

In this section, we process survey responses from students across different countries. The results are saved as separate CSV files and combined into a joint table.

### Serbia

```

# Load data
data_tab <- read_excel("./SERBIA_Students, postgraduates - BAS4SC (Odpowiedzi).xlsx")

# Process data
nivoi <- levels(as.factor(t(data_tab[, 6:106])))
mat <- sumarizator(data_tab)

# Reorder columns
col_order <- c("Not important", "Important", nivoi[3], nivoi[5], nivoi[4])
srb <- mat[, col_order]

# View and save
View(srb)
# write.csv(srb, "srb.csv")

```

### Croatia

```

data_tab <- read_excel("./CROATIA_Students, postgraduates - BAS4SC (Odpowiedzi).xlsx")
nivoi <- levels(as.factor(t(data_tab[, 6:106])))
mat <- sumarizator(data_tab)
col_order <- c(nivoi[1], nivoi[3], nivoi[2], nivoi[5:4])
cro <- mat[, col_order]
# write.csv(cro, "cro.csv")

```

### Slovenia

```

data_tab <- read_excel("./SLOVENIA_Students, postgraduates - BAS4SC (Odpowiedzi).xlsx")
nivoi <- levels(as.factor(t(data_tab[, 6:106])))
mat <- sumarizator(data_tab)

```

```
col_order <- c(nivoi[2], nivoi[3], nivoi[1], nivoi[4:5])
slo <- mat[, col_order]
# write.csv(slo, "slo.csv")
```

## Poland

```
data_tab <- read_excel("./PL_Students, postgraduates - BAS4SC (Odpowiedzi).xlsx")
nivoi <- levels(as.factor(t(data_tab[, 6:106])))
mat <- sumarizator(data_tab)
col_order <- c(nivoi[4], nivoi[5], nivoi[3], nivoi[2:1])
pl <- mat[, col_order]
# write.csv(pl, "pl.csv")
```

## Joint Table

```
joint_table <- srb + cro + slo + pl
View(joint_table)
# write.csv(joint_table, "jt.csv")
```

## Teachers Data Analysis

We follow the same procedure to analyze teacher survey responses. The `sumarizator` function is adjusted to handle the column range for teachers.

```
sumarizator2 <- function(dt) {
  dt <- dt[, 9:109]
  tt <- t(dt)
  mat <- matrix(0, ncol = 5, nrow = dim(dt)[2])

  nivoi <- levels(as.factor(tt))

  colnames(mat) <- nivoi
  rownames(mat) <- colnames(dt)

  for (i in 1:nrow(tt)) {
    for (level in nivoi) {
      mat[i, level] <- summary(as.factor(tt[i, ]))[level]
    }
  }

  mat[is.na(mat)] <- 0
  return(mat)
}

# Process and combine teacher data here (similar to student section)
```

## Companies Data Analysis

Similar steps are taken to process company survey responses. The data is loaded, summarized, and combined into a joint table.

```
# Home Work: Define the summarization function and repeat the process for companies
```

## Companies Data Analysis

In this section, we process survey responses from companies across different countries. The results are saved as separate CSV files and combined into a joint table.

```
sumarizator3 <- function(dt) {  
  # Select relevant columns (from 10 to 110)  
  dt <- dt[, 10:110]  
  
  # Transpose the data  
  tt <- t(dt)  
  
  # Initialize matrix with appropriate dimensions (rows = columns of dt, columns = number of unique lev  
  mat <- matrix(0, nrow = nrow(tt), ncol = length(levels(as.factor(tt))))  
  
  # Get the factor levels  
  nivoi <- levels(as.factor(tt))  
  
  # Set column names of matrix to the factor levels  
  colnames(mat) <- nivoi  
  rownames(mat) <- colnames(dt)  
  
  # Loop through each row (question) and count occurrences of each level  
  for (i in 1:nrow(tt)) {  
    level_counts <- summary(as.factor(tt[i, ]))  
  
    # Loop over each factor level and populate the matrix  
    for (level in nivoi) {  
      mat[i, level] <- level_counts[level]  
    }  
  }  
  
  # Replace NA values with 0  
  mat[is.na(mat)] <- 0  
  return(mat)  
}
```

## Serbia

```
# Load data for companies in Serbia  
data_tab <- read_excel("./SERBIA_Companies - BAS4SC (Odpowiedzi).xlsx")  
  
# Process data using the summarizer function  
nivoi <- levels(as.factor(t(data_tab[, 10:110])))
```

```
mat <- sumarizator3(data_tab)

# Reorder columns for clarity
col_order <- c(nivoi[5], nivoi[2], nivoi[3], nivoi[1], nivoi[4])
srb3 <- mat[, col_order]

# View and save results
View(srb3)
# write.csv(srb3, "srb3.csv")
```

## Croatia

```
# Load data for companies in Croatia
data_tab <- read_excel("./CROATIA_Companies - BAS4SC (Odgovori).xlsx")

# Process data using the summarizer function
nivoi <- levels(as.factor(t(data_tab[, 10:110])))
mat <- sumarizator3(data_tab)

# Reorder columns for clarity
col_order <- c(nivoi[4], nivoi[1], nivoi[2], nivoi[3], nivoi[5])
cro3 <- mat[, col_order]

# View and save results
View(cro3)
# write.csv(cro3, "cro3.csv")
```

## Slovenia

```
# Load data for companies in Slovenia
data_tab <- read_excel("./SLOVENIA_Companies - BAS4SC (Odpowiedzi).xlsx")

# Process data using the summarizer function
nivoi <- levels(as.factor(t(data_tab[, 10:110])))
mat <- sumarizator3(data_tab)

# Reorder columns for clarity
col_order <- c(nivoi[3], nivoi[1], nivoi[2], nivoi[5:4])
slo3 <- mat[, col_order]

# View and save results
View(slo3)
# write.csv(slo3, "slo3.csv")
```

## Poland

```
# Load data for companies in Poland
data_tab <- read_excel("./PL_Companies - BAS4SC (Odpowiedzi).xlsx")
```

```

# Process data using the summarizer function
nivoi <- levels(as.factor(t(data_tab[, 10:110])))
mat <- sumarizator3(data_tab)

# Reorder columns for clarity
col_order <- c(nivoi[3], nivoi[2], nivoi[1], nivoi[5:4])
pl3 <- mat[, col_order]

# View and save results
View(pl3)
# write.csv(pl3, "pl3.csv")

```

## Joint Table

```

# Combine the data from all countries into a joint table
joint_table_comp <- srb3 + cro3 + slo3 + pl3
View(joint_table_comp)

# Save the joint table

```

## Bonus

### Weighted Analysis and Classification

Finally, we perform a weighted analysis of the survey data and classify subjects into categories based on evaluation scores.

```

overall <- read_excel("./bonus/excel/overall2.xlsx")

# Calculate total evaluations
overall$total_evaluations <- (overall$Students + overall$Teachers + overall$Companies) / 3

# Sort and classify into categories
overall_sorted <- overall[order(overall$total_evaluations, decreasing = TRUE), ]
n_subjects <- nrow(overall_sorted)
n_A <- ceiling(n_subjects * 0.3)
total_evaluations_A <- overall_sorted$total_evaluations[n_A]

overall_sorted$category <- ifelse(overall_sorted$total_evaluations >= total_evaluations_A, "A",
                                ifelse(overall_sorted$total_evaluations > total_evaluations_A * 0.7, "B", "C"))

# Save results
write.csv(overall_sorted, "overall_sorted.csv")
View(overall_sorted)

```

## Conclusion

This analysis provides insights into the BAS4SC survey data across students, teachers, and companies. The results are saved as CSV files for further exploration and reporting.