

Loading required packages

```
library(rstudioapi)
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.4
## v forcats    1.0.0      v stringr    1.5.1
## v ggplot2    3.4.4      v tibble     3.2.1
## v lubridate  1.9.3      v tidyr      1.3.0
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(RColorBrewer)
```

Loading the data

```
current_path <- rstudioapi::getActiveDocumentContext()$path
current_path <- path.expand(current_path)
current_path <- dirname(current_path)
to_csv <- file.path(current_path, "Data ", "Survey_AI.csv")
to_csv
```

```
## [1] "/Users/komputer/Desktop/RR_git1/Project/RR_Final_Projekt/Data /Survey_AI.csv"
```

```
data <- read.csv(to_csv)
```

Checking the structure of the data

```
str(data)
```

```
## 'data.frame':   91 obs. of  35 variables:
## $ ID                                     : int  1 2 3 4 5 6 7 8 9 10 ...
## $ Q1.AI_knowledge                       : int  8 7 5 5 4 5 7 6 6 4 ...
## $ Q2.AI_sources                         : chr  "Internet;Books/Scientific papers (physical/online form
## $ Q2.1.Internet                        : int  1 1 1 1 1 1 0 1 1 1 ...
## $ Q2.2.Books.Papers                    : int  1 0 1 0 0 0 1 0 1 0 ...
## $ Q2.3.Social_media                     : int  0 1 0 1 0 0 1 0 0 0 ...
## $ Q2.4.Discussions                      : int  0 0 0 0 0 1 1 0 1 0 ...
## $ Q2.5.NotInformed                     : int  0 0 0 0 0 0 0 0 0 0 ...
## $ Q3.1.AI_dehumanization                : int  1 2 2 4 1 3 1 1 2 4 ...
## $ Q3.2.Job_replacement                  : int  2 3 1 4 2 4 2 1 4 4 ...
## $ Q3.3.Problem_solving                  : int  5 4 4 5 5 4 4 5 5 4 ...
## $ Q3.4.AI_rulling_society               : int  1 1 1 3 1 3 1 1 2 3 ...
## $ Q4.1.AI_costly                        : int  4 3 3 4 3 4 5 3 5 4 ...
## $ Q4.2.Economic_crisis                  : int  2 3 1 3 1 3 4 2 2 3 ...
## $ Q4.3.Economic_growth                  : int  4 4 3 3 4 3 2 3 3 3 ...
## $ Q4.4.Job_loss                         : int  2 3 2 4 2 4 4 2 4 4 ...
## $ Q5.Feelings                          : int  1 1 1 1 1 2 1 1 1 1 ...
## $ Q6.Domains                           : chr  "Education;Medicine;Marketing" "Medicine;Agriculture;C
## $ Q6.1.Education                       : int  1 0 1 1 1 1 1 1 0 1 ...
## $ Q6.2.Medicine                        : int  1 1 0 1 1 0 1 1 0 1 ...
## $ Q6.3.Agriculture                     : int  0 1 0 0 1 1 0 0 1 1 ...
## $ Q6.4.Constructions                    : int  0 1 0 0 1 0 1 0 0 1 ...
## $ Q6.5.Marketing                       : int  1 0 1 0 0 0 0 0 1 0 ...
```

```
## $ Q6.6.Administration      : int  0 0 1 0 0 1 0 0 1 0 ...
## $ Q6.7.Art                 : int  0 0 0 0 0 0 0 0 0 0 ...
## $ Q7.Utility_grade         : int  9 6 6 9 8 6 10 8 8 7 ...
## $ Q8.Advantage_teaching    : int  3 2 3 1 3 1 1 3 3 1 ...
## $ Q9.Advantage_learning    : int  1 2 3 2 2 2 2 3 2 1 ...
## $ Q10.Advantage_evaluation : int  2 1 3 2 3 3 1 3 2 3 ...
## $ Q11.Disadvantage_educational_process: int  3 2 4 3 4 1 3 3 1 1 ...
## $ Q12.Gender               : int  1 2 2 1 1 1 1 2 1 1 ...
## $ Q13.Year_of_study        : int  2 2 2 2 2 2 2 2 2 2 ...
## $ Q14.Major                : int  2 2 2 2 2 2 2 2 2 3 ...
## $ Q15.Passed_exams         : int  1 1 0 1 1 1 0 0 0 1 ...
## $ Q16.GPA                  : num  9.2 7.7 7.2 8.2 7.7 7.7 7.2 6.7 7.7 8.2 ...
```

Deleting character variables (they're only as comments for the variables for Q2.1-Q2.5 and Q6.1-Q6.7)

```
data <- subset(data, select = -c(Q2.AI_sources, Q6.Domains, ID))
```

Changing all variables to numeric

```
data <- mutate_all(data, as.numeric)
```

Checking for missing values

```
print(missing_values <- colSums(is.na(data)))
```

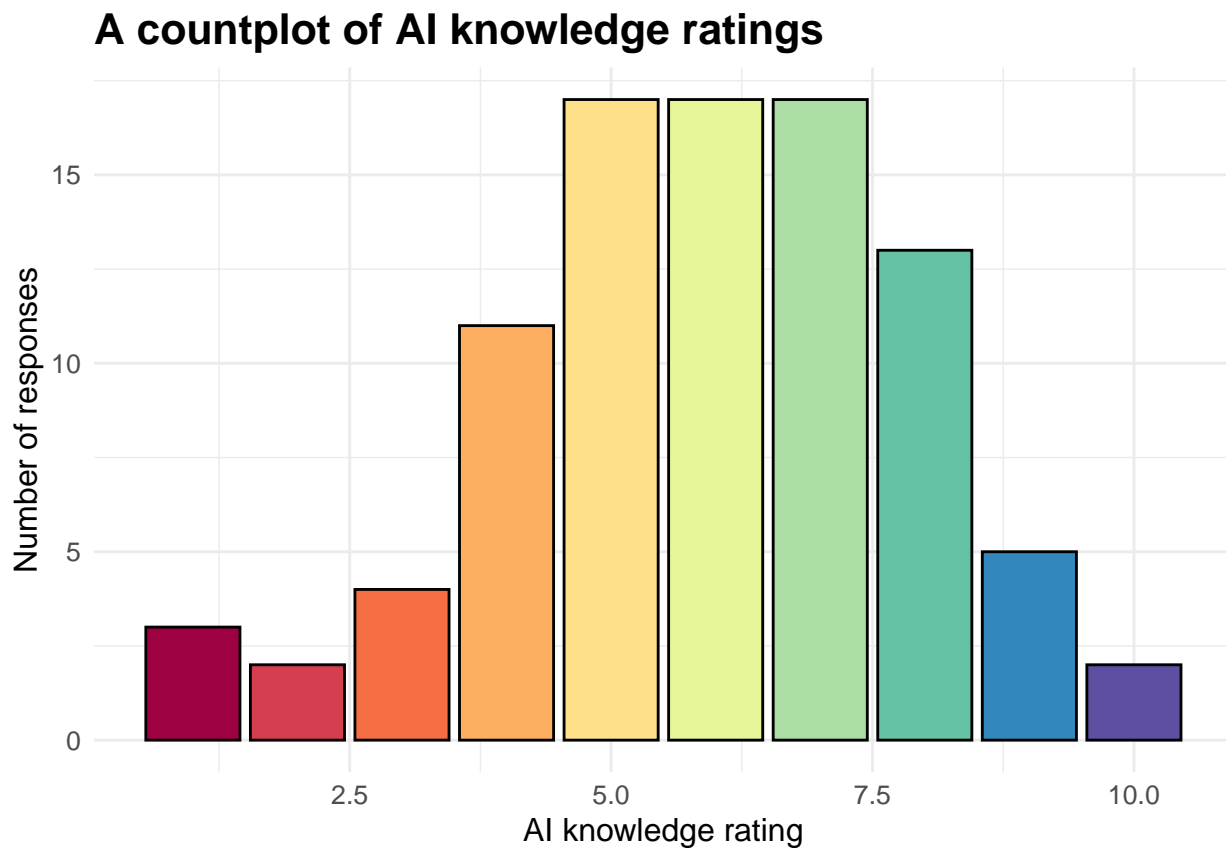
```
##           Q1.AI_knowledge           Q2.1.Internet
##           0                      0
##           Q2.2.Books.Papers       Q2.3.Social_media
##           0                      0
##           Q2.4.Discussions         Q2.5.NotInformed
##           0                      0
##           Q3.1.AI_dehumanization   Q3.2.Job_replacement
##           0                      0
##           Q3.3.Problem_solving     Q3.4.AI_rulling_society
##           0                      0
##           Q4.1.AI_costly           Q4.2.Economic_crisis
##           0                      0
##           Q4.3.Economic_growth     Q4.4.Job_loss
##           0                      0
##           Q5.Feelings              Q6.1.Education
##           0                      0
##           Q6.2.Medicine            Q6.3.Agriculture
##           0                      0
##           Q6.4.Constructions       Q6.5.Marketing
##           0                      0
##           Q6.6.Administration      Q6.7.Art
##           0                      0
##           Q7.Utility_grade         Q8.Advantage_teaching
##           0                      0
##           Q9.Advantage_learning    Q10.Advantage_evaluation
##           0                      0
##           Q11.Disadvantage_educational_process Q12.Gender
##           0                      0
##           Q13.Year_of_study        Q14.Major
##           0                      0
##           Q15.Passed_exams         Q16.GPA
##           0                      0
```

No missing values found, so we can proceed with the barplot of students' knowledge

```
custom_colors <- brewer.pal(10, "Spectral")

fig1 <- ggplot(data, aes(x = Q1.AI_knowledge)) +
  geom_bar(aes(fill = factor(Q1.AI_knowledge)), color = "black") +
  scale_fill_manual(values = custom_colors) +
  xlab("AI knowledge rating") +
  ylab("Number of responses") +
  ggtitle("A countplot of AI knowledge ratings") +
  theme_minimal() +
  theme(plot.title = element_text(size = 16, face = "bold"),
        axis.title = element_text(size = 12),
        axis.text = element_text(size = 10),
        legend.title = element_blank(),
        legend.position = "none")

print(fig1)
```



Creating data frame with sources of students' knowledge with their counts and producing an appropriate plot

```
source_names <- c("Internet", "Books/Papers", "Social Media", "Discussions", "Not Informed")
source_counts <- sapply(data[, c("Q2.1.Internet", "Q2.2.Books.Papers", "Q2.3.Social_media", "Q2.4.Discussions", "Q2.5.Not_informed")], FUN = function(x) sum(x == "Yes"))

source_data <- data.frame(Source = source_names, Count = source_counts)

source_data$Proportion <- source_data$Count / sum(source_data$Count)
```

```

custom_colors <- brewer.pal(5, "Spectral")
fig2 <- ggplot(source_data, aes(x = reorder(Source, -Proportion), y = Proportion, fill = Source)) +
  geom_bar(stat = "identity", color = "black") +
  labs(x = "Source", y = "Proportion", fill = "Source") +
  ggtitle("Proportion of Students Using Each Source") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_text(size = 16, face = "bold"),
        axis.title = element_text(size = 12),
        axis.text = element_text(size = 10),
        legend.title = element_text(size = 12),
        legend.position = "right") +
  scale_fill_manual(values = custom_colors)

print(fig2)

```

