Week 5 - Visualizations Activity

Raquel Alexandria

2023-06-22

```
# Install and load the gridExtra package
if (!require(gridExtra)) {
  install.packages("gridExtra")
}
```

```
## Loading required package: gridExtra
```

```
library(gridExtra)

# Reading and combining the data from two CSV files. Mainly using readr and dplyr here
library(readr)
library(dplyr)
```

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

```
## The following object is masked from 'package:gridExtra':
##
##     combine
```

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

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

```
library(ggplot2)
library(grid)

file1 <- "https://raw.githubusercontent.com/fivethirtyeight/data/master/comic-characters/marvel-wikia-da
file2 <- "https://raw.githubusercontent.com/fivethirtyeight/data/master/comic-characters/dc-wikia-data.

data1 <- read_csv(file1)
```

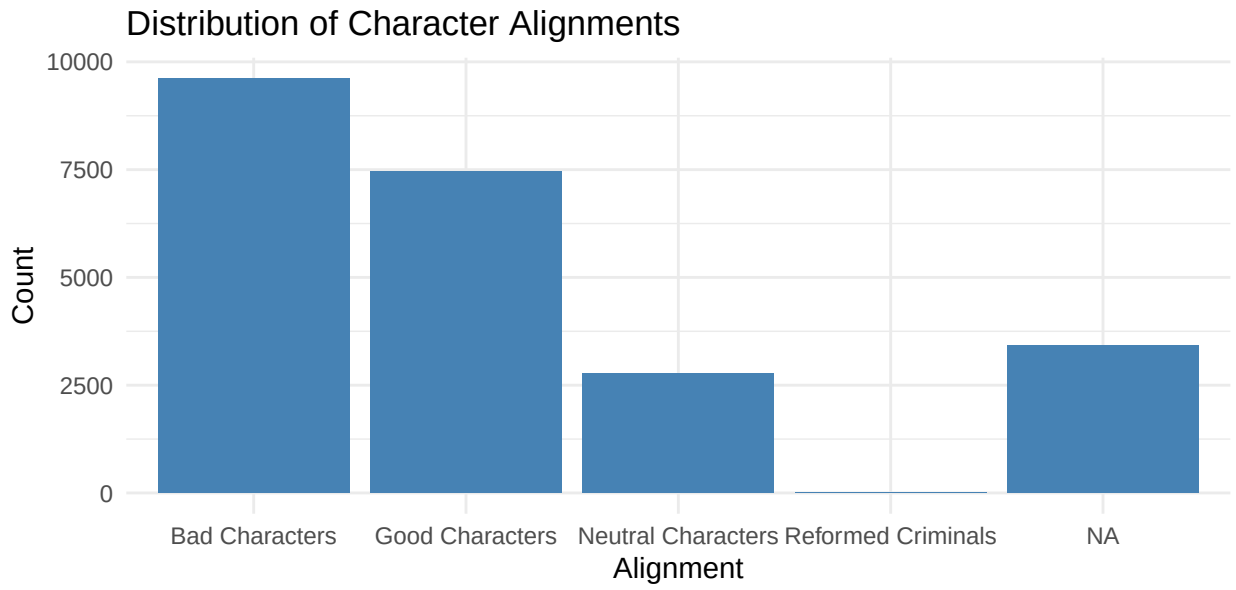
```
## Rows: 16376 Columns: 13
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (10): name, urlslug, ID, ALIGN, EYE, HAIR, SEX, GSM, ALIVE, FIRST APPEAR...
## dbl  (3): page_id, APPEARANCES, Year
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
data2 <- read_csv(file2)
```

```
## Rows: 6896 Columns: 13
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (10): name, urlslug, ID, ALIGN, EYE, HAIR, SEX, GSM, ALIVE, FIRST APPEAR...
## dbl  (3): page_id, APPEARANCES, YEAR
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

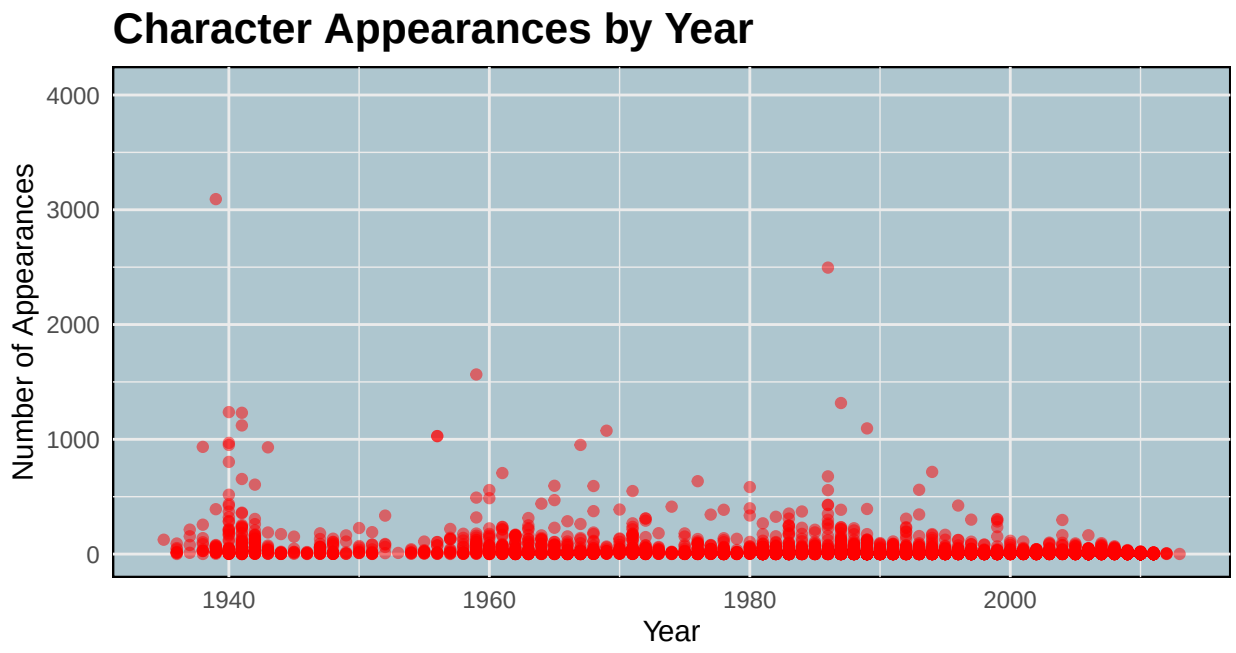
```
combined_data <- bind_rows(data1, data2)

# Creating the bar chart
alignment_chart <- ggplot(combined_data, aes(x = ALIGN)) +
  geom_bar(fill = "steelblue") +
  labs(title = "Distribution of Character Alignments",
       x = "Alignment",
       y = "Count") +
  theme_minimal()

# Adding a description before the bar chart
description_bar <- "The bar chart above illustrates the
distribution of character alignments in the comic dataset.
Each bar represents a different alignment category,
and the height of the bar indicates the count of characters
belonging to that alignment category."

# Creating the textbox for the description
textbox_bar <- grobTree(rectGrob(gp = gpar(col = "black", fill = "white", lwd = 1)),
                        textGrob(description_bar, x = 0.05, y = 0.05, hjust = 0, vjust = 0, gp = gpar(f

# Displaying the description and the bar chart
grid.arrange(alignment_chart, textbox_bar, nrow = 2, heights = c(3, 1.3))
```



The bar chart above illustrates the distribution of character alignments in the comic dataset. Each bar represents a different alignment category, and the height of the bar indicates the count of characters belonging to that alignment category.

```
# Selecting columns for the scatterplot
scatterplot_data <- combined_data %>%
  select(APPEARANCES, YEAR)

# Creating a scatterplot
scatterplot <- ggplot(scatterplot_data, aes(x = YEAR, y = APPEARANCES)) +
  geom_point(color = "red", alpha = 0.5) +
  labs(title = "Character Appearances by Year",
       x = "Year",
       y = "Number of Appearances") +
  theme_minimal() +
  theme(plot.title = element_text(size = 16, face = "bold"),
        panel.background = element_rect(fill = "#AEC6CF"))

# Adding a description before the scatterplot
description_scatter <- "This scatterplot illustrates the relationship
between the year and the number of character appearances.
Each red dot represents a character, with its position on the plot
indicating the year and the number of appearances for that character."

# Creating the textbox for the description
textbox_scatter <- grobTree(rectGrob(gp = gpar(col = "black", fill = "white", lwd = 1)),
                            textGrob(description_scatter, x = 0.05, y = 0.05, hjust = 0, vjust = 0, gp =

# Displaying the description and scatterplot
grid.arrange(scatterplot, textbox_scatter, nrow = 2, heights = c(3, 0.9))
```

```
## Warning: Removed 16791 rows containing missing values (`geom_point()`).
```



This scatterplot illustrates the relationship between the year and the number of character appearances. Each red dot represents a character, with its position on the plot indicating the year and the number of appearances for that character.