

Please download the data set `R01_assignment_dataset.csv` from the GitHub repository. Load it as a data frame `df`:

```
> df = read.csv("R01_assignment_dataset.csv")
```

Task 1.1 (Exploring the data set).

- (a) What data does the data set contain?
- (b) How many variables (columns) and how many observations (rows) does the data set contain?
- (c) Use the `table()` function to find out how many authors are male and female.
- (d) Is the column `pronoun` simply the sum of the columns `i`, `we`, `you`, `SheHe`, `they`, `ImpersonalPronouns`?

Hint: You might tackle this question with simple array calculations on the columns.

Task 1.2 (Investigating trends in the data).

- (a) Take a closer look at the data to answer the following question:
Do male and female authors differ with respect to the average number of words they write per sentence?
 - (i) Use `ggplot2` to create a barplot with `geom_bar` that shows the average number of words per sentence separately for the male and female authors.
 - (ii) Use `ggplot2` to create a density plot with `geom_density` that shows the distribution of number of words per sentence separately for the male and female authors.
- (b) Choose another *continuous* variable in the data set. Visualize its distribution separately for male and female authors.

Please solve the assignment in the `.Rmd` format and export it as `R01-<LastName><FirstName>.PDF`