

Lab 1 Practice

Your Name Here

2025-09-05

Part 1: Set up your work environment

1. Create a folder for this class.
2. Create a project for this class in that folder.
3. Put this RMarkdown file into that folder.

Part 2: Equations practice

Type the equation for the mean in LaTeX code so that it prints on a new line.

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

Figure 1: Equation for Mean

YOUR EQUATION HERE

Part 3: Knitting

1. Try to knit this document as is.
2. Now create a new R chunk and edit the chunk options so that it is included in the knitted PDF output. Try knitting again to make sure it worked.

Part 4: Write a function

Write a function that tells you whether an input is between 6 and 36. (hint - the `ifelse()` command might be helpful here – look it up in the console to see how it works)

```
# write function here
```

Uncomment these lines to test your function

```
#your_function_name(4)
#your_function_name(30)
```

Part 5: Practicing tidyverse commands

1. Load in the `tidyverse` package.
2. Load in a dataset of LGBTQ movies. This data is from [TidyTuesdays](#), which is a cool source to poke around if you're looking for interesting (and already cleaned – huge plus) datasets!

```
# loading in data
lgbtq_movies <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/main/data/

## Rows: 7165 Columns: 12
## -- Column specification -----
## Delimiter: ","
## chr  (5): title, original_title, original_language, overview, genre_ids
## dbl  (4): id, popularity, vote_average, vote_count
## lgl  (2): adult, video
## date (1): release_date
##
## 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.
```

3. How many rows and columns are in this dataset?
4. Which movies have higher ratings on average – those released in the 20th century or those released in the 21st century? The ratings for each movie can be found in the `vote_average` column.
5. Create a dataset of the top 10 most popular (`popularity`) Spanish language (`es`) LGBTQ movies.
6. Create a variable that contains the average rating for a movie only if it has more than 100 votes. Otherwise this variable should be `NA`.

Part 6: Summary statistics

1. Create a dataset of summary statistics by language that contains:
 - (i) The earliest release date of an LGBTQ movie in that language
 - (ii) The latest release date of an LGBTQ movie in that language
 - (iii) The average rating of LGBTQ movies in that language