

Lab 4 - Data Analysis

Dr Austin R Brown

INSTRUCTIONS: For this assignment, please thoroughly read the below instructions and complete the tasks as outlined in each coding language. Once you have completed the assignment, please download your .sas, .R, and .ipynb files and submit them to the appropriate assignment folder in D2L.

1. Using the Movie Dataset 2015.csv file, please:
 - a. Upload the file into RStudio Cloud, SAS Studio, and Google Colab.
 - b. Read the file into R, SAS, and Python.
 - c. Using the `runtime` variable, create a histogram in R, SAS, and Python.
 - d. Using the `budget` variable, calculate the mean, median, and standard deviation in R, SAS, and Python.
 - e. Create a new dataset in each language that only includes `genres` which are equal to `Action`, `Adventure`, or `Science Fiction`. (HINT: In R and Python the `|` operator is used for OR, in SAS the keyword `OR` is used).
 - f. Using the new dataset created in part e, tabulate the frequency of genres in R, SAS, and Python.
 - g. Using the new dataset created in part e, create a bar chart in R, SAS, and Python that displays the frequency of genres.