

Lab 3 - Data Transformations

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 2023.csv file, please:
 - a. Upload the file to RStudio Cloud, SAS Studio, and Google Colab.
 - b. Read the file into R, SAS, and Python.
 - c. Create a new dataset in each language that includes only the following variables:
 - i. title
 - ii. release_date
 - iii. runtime
 - iv. popularity
 - v. genres
 - d. Using the dataset created in part c, subset the data to include only movies with a runtime greater than 120 minutes using each language.
 - e. Using the dataset from part b, create a new variable in each language that calculates the net profit for each movie. Net profit is calculated as revenue minus budget.