

Data Analytics - Career Accelerator

# Spreadsheets Gym Session



# Spreadsheets Gym Session

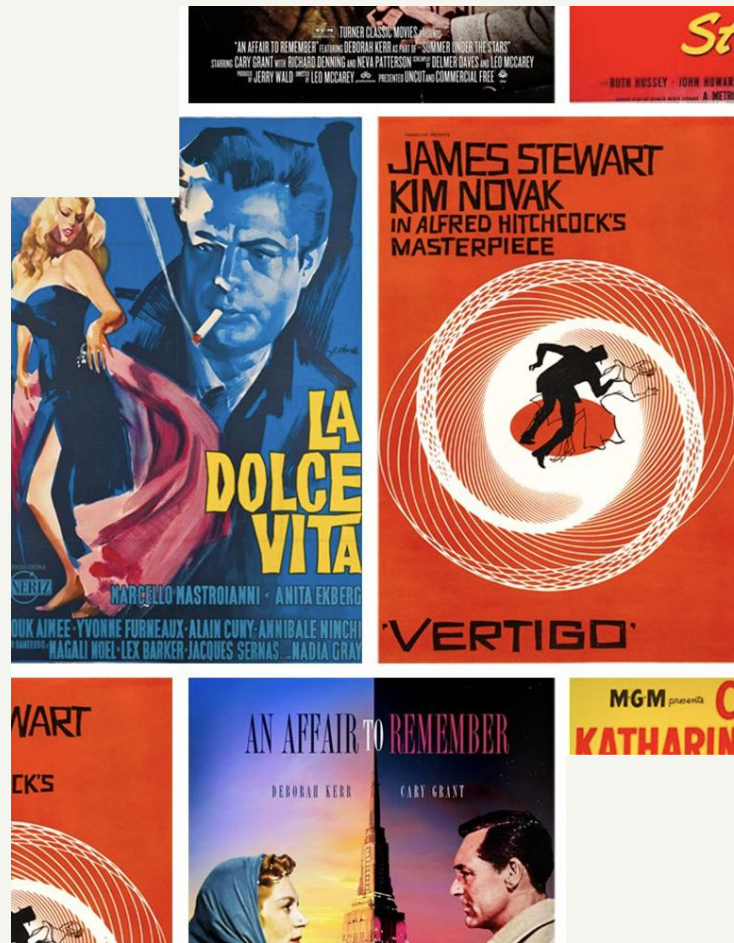
**While we're waiting to begin,  
please make a copy of the Google  
Sheet in the Zoom chat.**



# Films Dataset

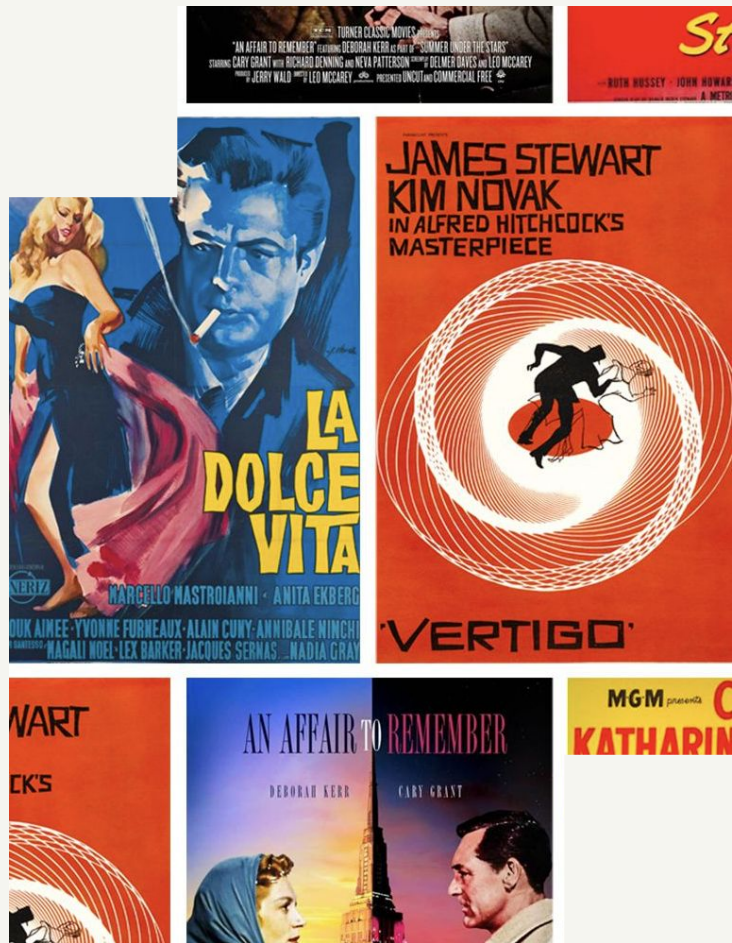
- Make a copy of the [spreadsheet](#) (link in the Zoom chat)
- 1600+ movies with properties such as length, main actor and actress, director and popularity.
- Take a minute to get familiar with the data

Source: [Project Datasets](#)



# Dataset Columns

- **Year:** film release year
- **Length:** run time
- **Title:** movie title
- **Subject:** film genre
- **Actor/Actress:** lead actor/actress in film
- **Director:** film director
- **Popularity:** popularity rating



# Question 1



# Question 1

How many films are there per year?



## Question 2

## Question 2

Plot the count of films per year in a bar chart





# Question 3

## Question 3

If you look at the values of the subject column, you'll discover two subjects that seem very similar but differ by a letter. Create a new column called genre to clean up this value.



# Question 4

## Question 4

Create a pie chart representing the number of films per genre.  
Use the new genre column created in the previous exercise.



# Question 5

## Question 5

Plot a line chart of the number of films released per year broken down by Genre





# Question 6

## Question 6

Create a box plot chart based on movie popularity for directors with at least 10 movies



# Thank You

And Keep Practicing!