

- 1) Load in the data set (Netflix)
- 2) Check for missing values
- 3) Check for an overview of the data
- 4) Drop the few null rows in Date added, rating and duration columns respectively
- 5) Fill in the null rows in director, cast and country column with ("None Specified", "None listed" and "Non Specified ") respectively
- 6) Recheck for missing values
- 7) Check the value count for Duration column
- 8) Check the value count for type column
- 9) Using pie chart in matplotlib visualize the type column (Movie and TV Show)
- 10) Subset the data to only include Movies in the type Column
- 11) Get the counts of directors in the dataset
- 12) Visualize the top movies directors
- 13) Check for outliers in the duration column
- 14) What are the top countries by movies counts
- 15) Visualize the top countries by movie count
- 16) Select movies type that are been released in the last 10 year (where years ≥ 2012)
- 17) Using bar plot, show the count of movies and series produced in the last 10 years (where years ≥ 2012)
- 18) Using a count plot show the total movies and TV show been released in the last 10 years
- 19) Check for unique rating
- 20) Explore the data