Data Processing – Final Project

Date of entry: 17.05.2025

What I've worked on: creating a git account, setting up the repository and getting started on the coding

What problems I encountered: grouping by genre, because each cell has several genres but all as one string

What I learned: how to separate each string into genres, making plots with seaborn

Which resources did I use: seaborn documentation, Stack Overflow

Date of entry: 20.05.2025

What I've worked on: Graphs for TV show genres, writing a project description

What problems I encountered: setting the xtick labels on the graph so one can read them

What I learned: Customizing graphs & axis

Which resources did I use: geeksforgeeks website

Date of entry: 22.05.2025

What I've worked on: pie chart for types, top ten countries on Netflix, started looking into making a colour palette from Netflix colours for all graphs

What problems I encountered: having in the pie chart the percentage and count of the types

What I learned: adding information in pie charts, learned more about seaborn barplots

Which resources did I use: geeksforgeeks, matplotlib pie chart documentation, seaborn documentation, ww3 schools

Date of entry: 23.05.2025

What I've worked on: creating a colour palette using image colour picker from a picture of the Netflix intro screen, started implementing the palette for the graphs (top 10 countries graph), making a nested pie chart, graphs for the duration of movies

What problems I encountered: making only one bar of a barplot a different colour, making the nested pie chart (determining the inner and outer sizes of the pie charts), Setting the colours for the pie chart

What I learned: making a nested pie chart, learned a lot about how to style and customize graphs (colours, grid, axes)

Which resources did I use: https://imagecolorpicker.com/, https://www.statology.org/seaborn-barplot-color/, https://www.w3schools.com/python/matplotlib_pie_charts.asp , https://how.dev/answers/how-to-make-a-nested-pie-chart-in-python,

Date of entry: 25.04.2025

What I've worked on: graph for duration of TV shows, determining age groups based on the rating, donut chart for age groups

What problems I encountered: 3 rows of shows had the duration in the rating column, fixed that and replaces the values, redid the graph for the duration of movies again since 3 values were missing

What I learned: inserting a white circle in a pie chart to make a donut

Which resources did I use: IMBd to add the missing ratings, seaborn documentation, https://en.wikipedia.org/wiki/TV_Parental_Guidelines (to make age groups from ratings), matplotlib documentation

Date of entry: 26.05.2025

What I've worked on: making a wordcloud, presentation, style (colours, text, ect) of a couple of graphs, writing a read me

What problems I encountered: When saving graphs some of the axis labels were cut of

What I learned: making a wordcloud, adding arguments to savefig for saving graphs(specifying higher resolution (dpi), making sure all labels are included (bbox_inches))

Which resources did I use: https://medium.com/@m3redithw/wordclouds-with-python-c287887acc8b (for making the wordcloud), matplotlib documentation, canva for making the presentation

Date of entry: 27.05.2025

What I've worked on: going through the presentation, reading through the read me, description and process book, pushing all the files to git