



PYTHON AND PANDAS CHALLENGE

EXERCISE OVERVIEW In the Python course, you will be teaching the Pandas data analysis package. We want to give you the opportunity to show off your Pandas skills in this code challenge. You will be building a Python script with a few requirements for this instructor challenge.

DELIVERABLE For this instructor challenge, you will be creating a Pandas script that performs an analysis of IMDB movie data. We normally teach using IPython notebooks, but feel free to submit this challenge in the format you think best.

REQUIREMENTS There are a few parts to this challenge, the first is to perform some analysis on the top 100 IMDB movies. We have attached a CSV for you to work with called `imdb_100.csv`.

First, load the data into Pandas, then perform some analysis on this file using.

- How many rows and columns are in the dataset?
- What are the column names?
- How many unique genres are there?
- How many movies are there per genre?

You will be using the [OMBD API](#) for the next part of the challenge. We will be answering the question: How do Rotten Tomato scores compare to IMDB ratings? Please make an API key on their website using the instructions found there. Then, use the Python requests library to pull from their site.

Your script should pull the Rotten Tomato score for each of the movies in the top 100 IMDB movies and join it onto your Pandas dataframe. You may need to clean up the data!

Now answer the following questions:

- What are the top five R-Rated movies?
- What is the average Rotten Tomato score for the top 100 IMDB films?
- What is the Five Number Summary like for top rated films as per IMDB? Is it skewed?
- Create a column that is the ratio between Rotten Tomato rating vs IMDB rating. What film has the highest IMDB : Rotten Tomato ratio? The lowest?
- Create your own question... then answer it!

For each of the following prompts, create a plot to visualize the answer. Consider what plot is most appropriate to explore the given prompt.

- What is the relationship between IMDB ratings and Rotten Tomato ratings?
- What is the relationship between IMDB rating and movie duration?
- How many movies are there in each category?
- What does the distribution of Rotten Tomato ratings look like?

Your script should be clear and understandable, since your student facing code should be beginner friendly! Include comments briefly explaining your thought process and your code.

Then, write up a quick summary of your findings in a way that would be educational and accessible for students.