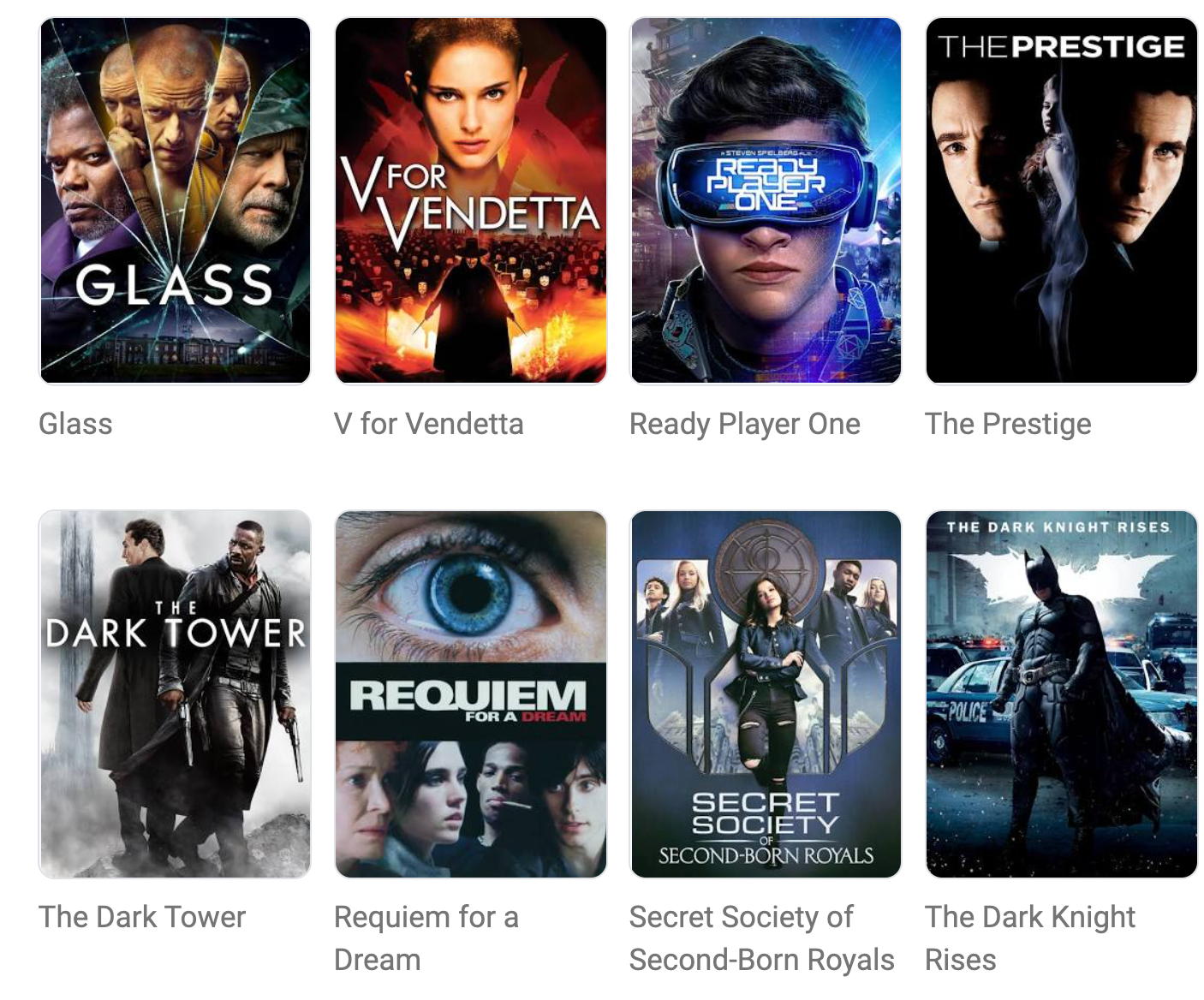
**Day 77 Goals: what you will make by the end of the day**

In this lesson, we're going to be looking at movie budget and revenue data. This dataset is perfect for trying out some new tools like scikit-learn to run a linear regression and seaborn, a popular data visualisation library built on top of Matplotlib.



The question we want to answer today is: Do higher film budgets lead to more revenue in the box office? In other words, should a movie studio spend more on a film to make more?

**Today you'll learn:**

* How to use a popular data visualisation library called Seaborn
* How to run and interpret a linear regression with scikit-learn
* How to plot a regression a scatter plot to visualise relationships in the data
* How to add a third dimension to a scatter plot to create a bubble chart
* How to cleverly use floor division // to convert your data



**Download and add the Notebook to Google Drive**

As usual, download the .zip file from this lesson and extract it. Add the .ipynb file into your Google Drive and open it as a Google Colaboratory notebook.

**Add the Data to the Notebook**

The .zip file also includes a .csv file called cost\_revenue\_dirty. This is the data for the project. Add this file to your notebook.

