The “imdb.csv” file is a dataset that showcases the data contained on the IMDB database regarding the most popular films and series (shows). This dataset was chosen to be the topic that we would cover because it contains a suitable number of entries (located in the description of the dataset), the content covered in the dataset makes it easier for us to perform EDA, and most importantly the topic was one that everyone in the group was comfortable with working on/analyzing.

The dataset’s description provided below briefly goes over what data has been collected from IMDB’s database. This dataset was retrieved from [here](https://www.kaggle.com/mazenramadan/imdb-most-popular-films-and-series) with full credits going to the author who gathered the data to make this dataset usable for others. The source of the data is from the [IMDB](https://www.imdb.com/) database.