Marvel vs DC

Business Intelligence – Project

Data Set Info:

This data set focuses on all known ‘live-action’ box-offices movies. It contains a list of movies from Marvel and DC produced between the years from 1944 to 2020. You can find more information on the data set on Kaggle and download the data [here.](https://www.kaggle.com/datasets/jcraggy/marvel-vs-dc-imdb-rotten-tomatoes?resource=download) The data set contains 19 columns and nearly 90 rows.

Data Exploration:

* Will any data cleaning be necessary before uploading the data to Tableau/Power BI?
* Are any of the columns superfluous?
* For whom was this data set created? What is its intended use?

We can use Power Query in Power BI to make any data cleansing, however, opening the data excel file and viewing the data, it seems like we only need to remove some columns for our specific purpose. Apparently, this data set was scraped from the web and was validated for being good content. The intended use was to get a full list of movies from Marvel/DC where the films were produced between 1944 and 2020. It provides information about rating of the film, director, important cast, and description. There is a variety of tasks that can be done with this data set, for example, create a recommendation engine for super hero movies or compare films between the two companies.

Planning:

Our goal for this project is to create a dashboard to visually compare the Marvel and DC superhero films. We can answer the following questions: which brand has the higher average IMBD rating, which brand has been more profitable over the years, and how has the average rating of each brand changed over time? A person who might be interested in this dashboard might be a company stakeholder to see how their customers are view their films over time. Has their average rating decreased or increased over time or stayed stagnate. If the average ratings have decreased over time? We might want to answer why and investigate further as a data analyst since our stakeholders will be expecting this.