

Project: Investigate a Dataset (The Movies Database)

Table of Contents

- [Introduction](#)
- [Data Wrangling](#)
- [Exploratory Data Analysis](#)
- [Conclusions](#)

Introduction

This dataset contains information about 10,000 movies collected from The Movie Database, including user ratings and revenue. In this report, I will be exploring the following questions: (1) which year has the highest patronage of movies. (2) what kinds of properties are associated with movies that have high revenues. (3) which genres are most popular from year to year. (4) do customers have preference for production from a particular company. (5) what factors or variables significantly impact the popularity of the movies. (6) which year did the film industry make the highest profit

Data Set Up

```
In [1]: import pandas as pd
import numpy as np
% matplotlib inline
```

Data Wrangling