



LOVELY
PROFESSIONAL
UNIVERSITY

Report on “Top 1000 Highest Grossing Movies Of All Time”

Name: Sai Shankar Gattu

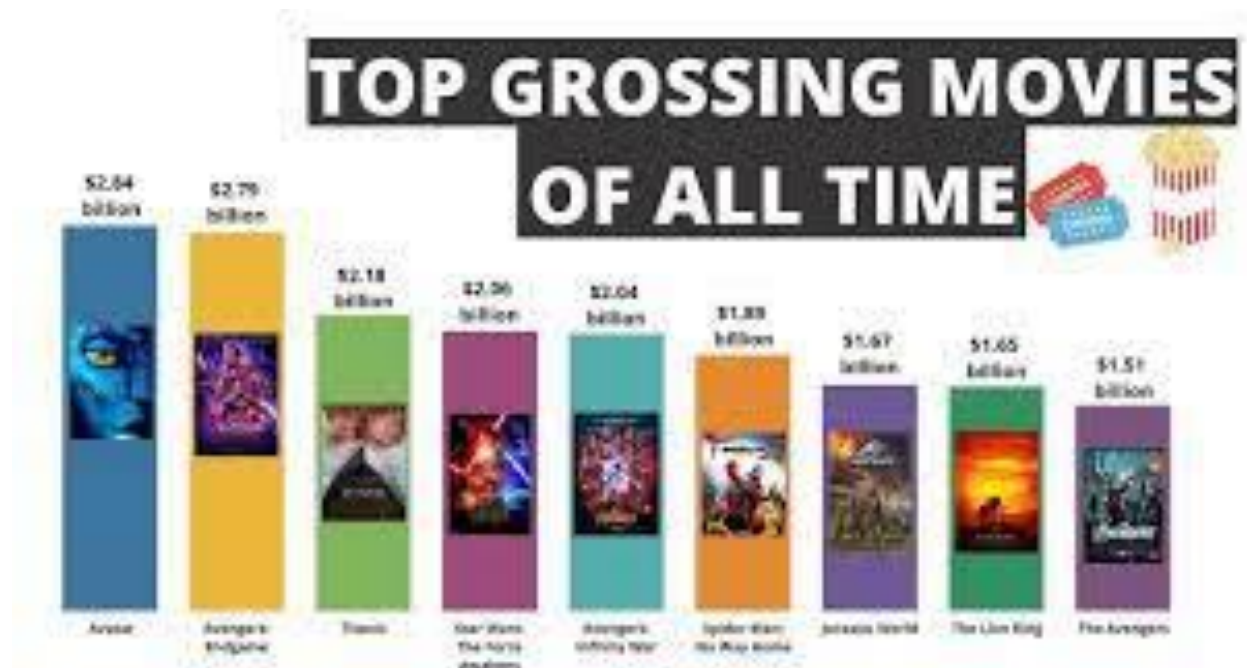
Registration no: 12016277

Section: K20RU

Roll no: RK20RUA32

Course: INT353

Submitted to: Abhijeet Dutta



Overview

This Dataset displays only the top 1000 highest grossing feature films of all time as of September 5, 2022.

Dataset overview

In this notebook we will explore this dataset, a detailed sample of top 1000 highest grossing movies of all time. We will probe the basic dataset attributes and hopefully uncover some interesting effects from the data!

Columns	Description
Movie Title	The name of the movie
Year of Release	The year the movie was released
Genre	Categories where the movie belongs.

Movie Rating	Ratings given by IMDb registered users (on a scale of 1 to 10)
Duration	Movie running time in minutes.
Gross	Gross earnings in U.S. dollars.
Worldwide LT Gross	Worldwide Lifetime Gross (International + Domestic totals).

Metascore	Weighted average of many reviews coming from reputed critics (on a scale of 0 to 100)
Votes	Number of votes cast by IMDb registered users
Logline	A one or two sentence summary of the film.

Insights and Approach on the Dataset

- ✚ First of all, Import the Dataset into Jupiter Notebook to work on the dataset.
- ✚ Analyse the columns and drop any unnecessary columns that doesn't provide any valid insights.
- ✚ Then check for any null values in the data, if any column contains lot of null values, then drop the column.
- ✚ Replace the null values using filtering data techniques. Standardize the values.
- ✚ After cleaning the dataset, I work on columns that gives me most insights.
- ✚ I will work on some attributes so that I get which movie is released in which year and where it is released.
- ✚ Working on duration versus total gross

✚ Working on metascore versus total gross

✚ Working on votes versus total gross