TOP 1000 HIGHEST GROSSING HOLLYWOOD MOVIES (1972-2021)

PROJECT 3 - DATA VISUALIZATIONS

GROUP MEMBERS

Alfredo Garcia
Connor Grant
Neel Patel
Yousuf Amin Alfatwa

Northwestern University
Data Science Visualization Boot Camp (2022)

DATA SOURCE

KAGGLE

https://www.kaggle.com/datasets/sanjeetsinghnaik/top-1000-highest-grossing-movies

DATA

Analyzed Dataset which contained data for movies that includes Movie Title, Year it was made, Domestic & Foregin Sales, Genre, Movie Run Time.

POSTGRES

postgresql://postgres:postgres@localhost:5432/hollywood_movies

ETL

Python Flask-powered API used with PostgreSQL (pgAdmin4).

VISUALIZATIONS

Used plot.ly to create visualizations. Webpage created using Javascript, HTML and CSS.

PROJECT OBJECTIVE

- **Utilizing data from IMDB regarding the highest grossing Hollywood films, our team sought to analyze and visually explain the financial results of these films, as well as seek out how successful films were in relation to their average run-time.**

PROCESS

1. Data in csv file.
2. Format/Clean the data.
3. Create Schema and upload the csv file to PostgreSQL (pgAdmin 4).
4. Use SQLAlchemy to create API Flask (as a reference point).
5. Road Block: Create Primary Key in pgAdmin 4.
6. Create JSON file based on the original data.
7. Use JSON file to make interactive dashboard.
8. Use the same JSON data to create graphs.
   a. Used Plot.ly

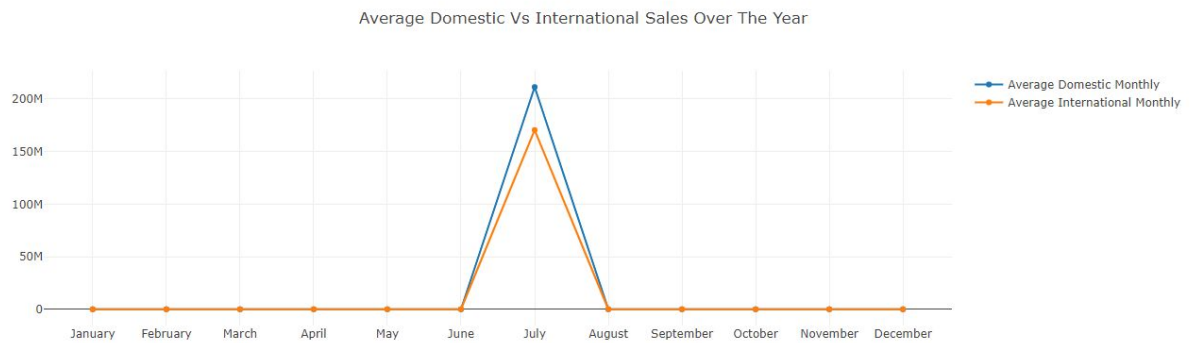| index [PK] integer	title character varying (255)	movie_info character varying	distributor character varying (255)	month integer	day integer	year integer	domestic_sales integer	international_sales integer	world_sales bigint	genre character varying (255)	movie_runtime character varying (255)	license character varying (255)
1	0 Star Wars: Episode VII - The Force Awakens (2015)	As a new threat to the galaxy rises, Rey, a desert scaveng...	Walt Disney Studios Motion Pictures	12	16	2015	936662225	1132859475	2069521700	[Action', 'Adventure', 'Sci-Fi']	2 hr 18 min	PG-13
2	1 Avengers: Endgame (2019)	After the devastating events of Avengers: Infinity War, the ...	Walt Disney Studios Motion Pictures	4	24	2019	858373000	1939128328	2797501328	[Action', 'Adventure', 'Drama', 'Sci-Fi']	3 hr 1 min	PG-13
3	2 Avatar (2009)	A paraplegic Marine dispatched to the moon Pandora on ...	Twentieth Century Fox	12	16	2009	760507625	2086738578	2847246203	[Action', 'Adventure', 'Fantasy', 'Sci-Fi']	2 hr 42 min	PG-13
4	3 Titanic (1997)	A seventeen-year-old aristocrat falls in love with a kind bu...	Paramount Pictures	12	19	1997	659363944	1542283320	2201647264	[Drama', 'Romance']	3 hr 14 min	PG-13
5	4 Jurassic World (2015)	A new theme park, built on the original site of Jurassic Pa...	Universal Pictures	6	10	2015	652385625	1018130819	1670516444	[Action', 'Adventure', 'Sci-Fi']	2 hr 4 min	PG-13
6	5 The Avengers (2012)	Earth's mightiest heroes must come together and learn to...	Walt Disney Studios Motion Pictures	4	25	2012	623357910	895457605	1518815515	[Action', 'Adventure', 'Sci-Fi']	2 hr 23 min	PG-13
7	6 Star Wars: Episode VIII - The Last Jedi (2017)	The Star Wars saga continues as new heroes and galacti...	Walt Disney Studios Motion Pictures	12	13	2017	620181382	712517448	1332698830	[Action', 'Adventure', 'Fantasy', 'Sci-Fi']	2 hr 32 min	PG-13
8	7 The Lion King (2019)	After the murder of his father, a young lion prince flees hi...	Walt Disney Studios Motion Pictures	7	11	2019	543638043	1119261396	1662899439	[Adventure', 'Animation', 'Drama', 'Family', 'Musi...	1 hr 58 min	PG
9	8 The Dark Knight (2008)	When the menace known as the Joker wreaks havoc and ...	Warner Bros.	7	16	2008	534858444	471115201	1005973645	[Action', 'Crime', 'Drama', 'Thriller']	2 hr 32 min	PG-13
10	9 Rogue One: A Star Wars Story (2016)	In a time of conflict, a group of unlikely heroes band toget...	Walt Disney Studios Motion Pictures	12	14	2016	532177324	523880396	1056057720	[Action', 'Adventure', 'Sci-Fi']	2 hr 13 min	PG-13

DASHBOARD / GRAPHS

Top 1000 Highest Grossing Hollywood Films

Home

Release Year:
1985

Average Domestic Vs International Sales Over The Year

Average Domestic Vs International Sales Over The Year

- Average Domestic Monthly
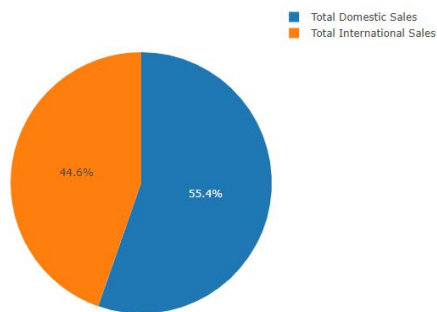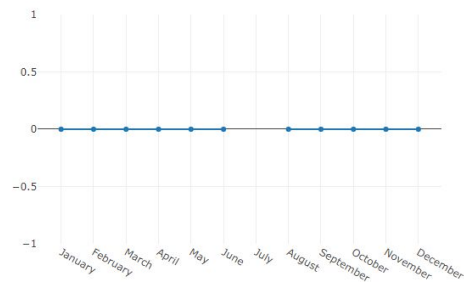- Average International Monthly

Total Domestic Vs International Sales in The Year

Total Domestic Vs International Sales in The Year

- Total Domestic Sales
- Total International Sales

44.6%
55.4%

Average Movie Runtime Over The Year

Average Movie Runtime Over The Year

SUMMARY

- **Dataset contains information about the top 1000 highest-grossing Hollywood movies.**

- **Struggles:**
  - Don't use javascript libraries that we are not familiar with.

- **Further Updates:**
  - Create another webpage that has a drop down menu which lists all the movie. Based on that selection, the movie poster is displayed along with a trailer of the movie.
  - Fix average run time graph.
  - Use Chart.js library.

- **Lesson Learned:**
  - Pick a dataset that has a good geoJSON file so we can have graphical visualizations.
  - Our dataset has a list within the JSON file and many loops are required extract the data into an interactive dashboard.
  - Extract and transform the data in Python by creating a function and then using JS to call that function for dashboard.
  - Use Bootstrap to make better website and visualizations.
  -

THANK YOU

- **Link to Github for more information:**
  - https://github.com/nealp1910/NU_Bootcamp_2022_Project_3_Data_Visualizataion

- **We really appreciate the help we got from Mo and all the TA's in this project!**