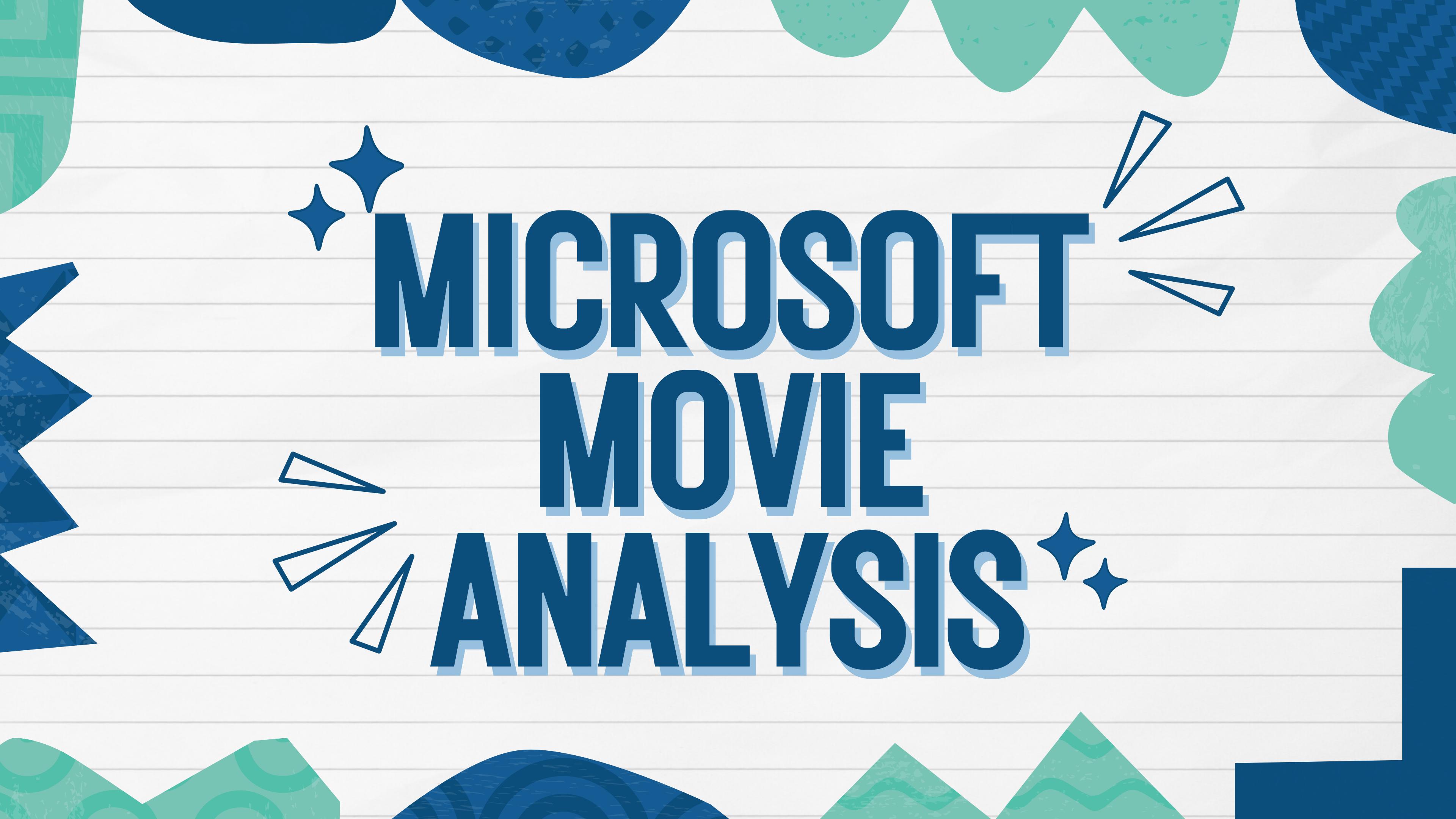


# MICROSOFT MOVIE ANALYSIS



# BUSINESS PROBLEM

Microsoft, an American multinational technology corporation, has taken an interest in all the big companies creating products in the niche of **creating original video content** and wants to venture in.

They have created a movie studio but have inadequate knowledge and expertise to create movies.



# CONTEXT: BACKGROUND AND JUSTIFICATION

This project explores **trends in movies** currently performing well in the market.

Valuable information obtained will be used to provide actionable insights that can help Microsoft decide on the type of films to create.

# METHODOLOGY: PROCEDURES AND TOOLS

The procedures followed are:

1. Data extraction
2. Data Exploration
3. Data preparation
4. Data modelling
5. Evaluation
6. Recommendation

Tools used include:

1. Python language
2. Pandas package
3. Matplotlib - plotting
4. Seaborn- plotting



# TIMELINE: PHASES

**01**

## Data Extraction

Use relevant movie datasets for the analysis

**02**

## Data exploration

The movie datasets are explored to find the number of records, information about the movies available and possibly missing and/or duplicate data

**03**

## Data Preparation

The datasets are “cleaned” by removing duplicates if any, handling missing values and removing or adding new columns to the datasets.

**04**

## Data Modelling

Statistical operations are carried out on the datasets and visualizations are created to gain insight on the data

**05**

## Evaluation

The analysis is reviewed for accuracy or any changes prompting an iteration after which a final draft is made

**06**

## Recommendations

Microsoft is given recommendations on how to go about creating video content based on insights from the analysis

# BOX OFFICE MOJO (BOM)

Information available to us from BOM include the **title**, **domestic and foreign gross earnings**, **studio** and **year of release**

		title	studio	domestic_gross	foreign_gross	year
2649		Until Forever (Por Siempre)	CZ	105000.0	NaN	2016
1350		Disconnect	LD	1400000.0	NaN	2013
213		Kites	Relbig.	1600000.0	1300000	2010
1329		Ram-Leela	Eros	2700000.0	NaN	2013
459		Shaolin	Vari.	66100.0	33600000	2011

# INTERNET MOVIE DATABASE (IMDB)

The information available to us from IMDB includes two dataset. One contains **information on the movie** and the other has info on **ratings** the **id(tconst)** of the movie, **primary title**, **the original name of movie**, **year of release**, **runtime** and **genres**

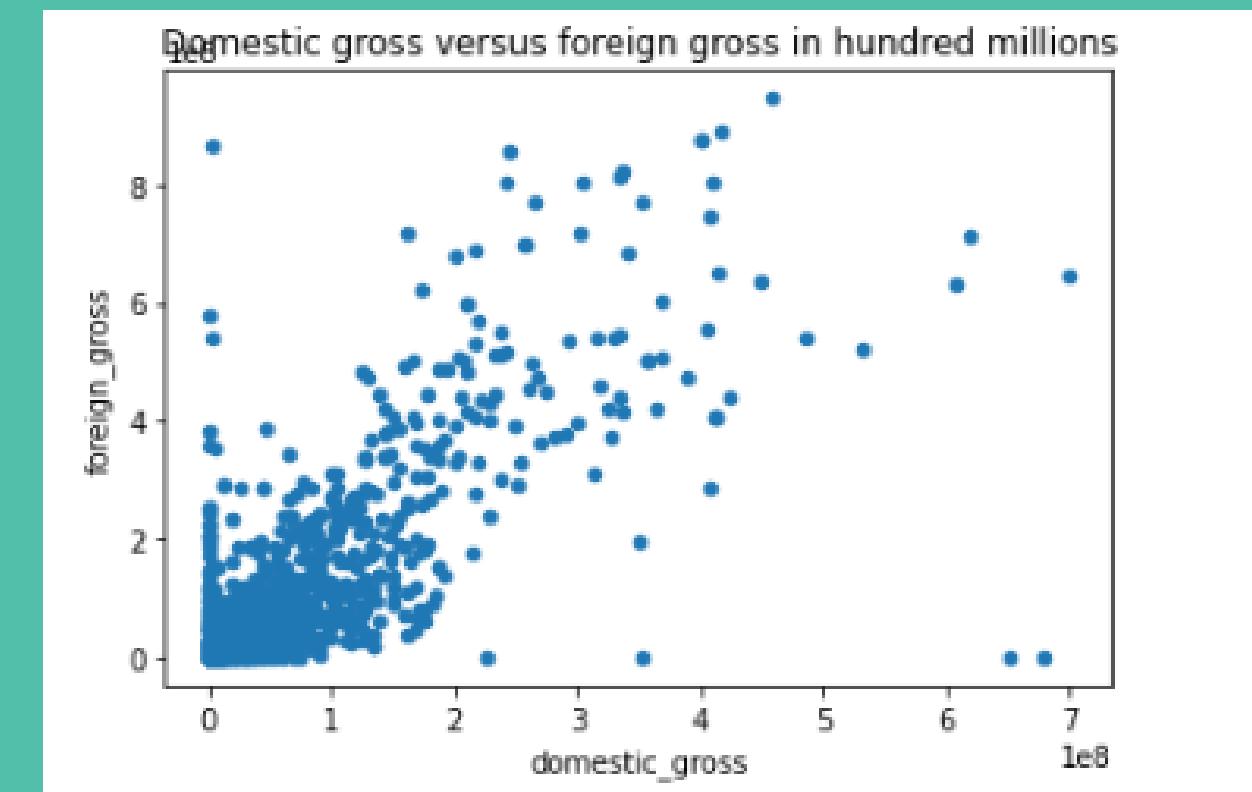
	tconst	primary_title	original_title	start_year	runtime_minutes	genres
0	tt0063540	Sunghursh	Sunghursh	2013	175.0	Action,Crime,Drama
1	tt0066787	One Day Before the Rainy Season	Ashad Ka Ek Din	2019	114.0	Biography,Drama
2	tt0069049	The Other Side of the Wind	The Other Side of the Wind	2018	122.0	Drama
3	tt0069204	Sabse Bada Sukh	Sabse Bada Sukh	2018	NaN	Comedy,Drama
4	tt0100275	The Wandering Soap Opera	La Telenovela Errante	2017	80.0	Comedy,Drama,Fantasy

# IMDB (CONT.)

The other has info on **ratings** the **id(tconst)** of the movie, **average rating** and **number of votes (numvotes)**

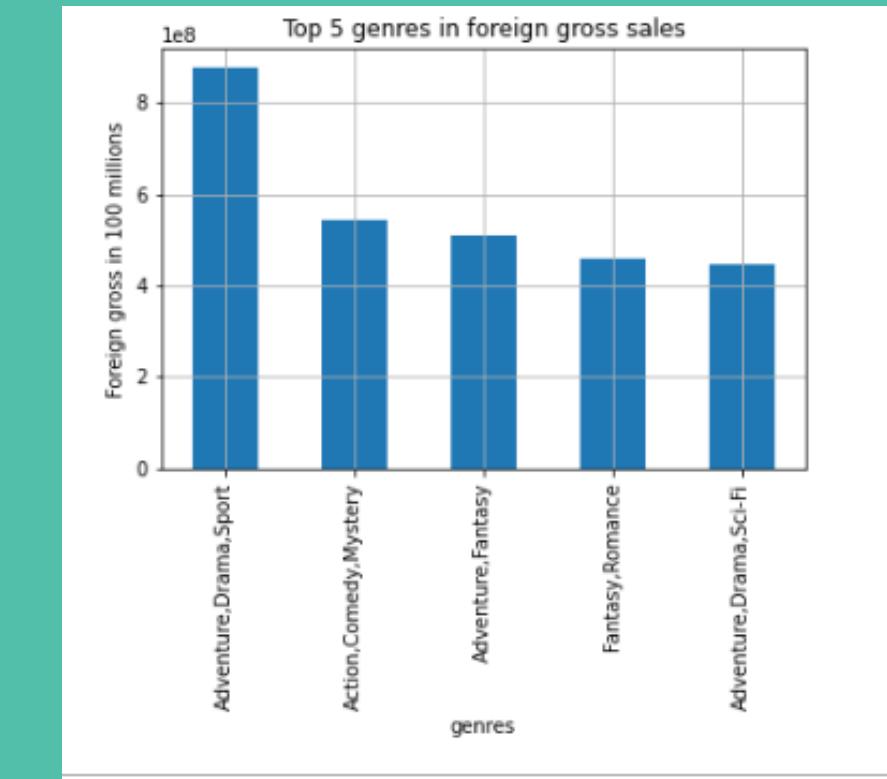
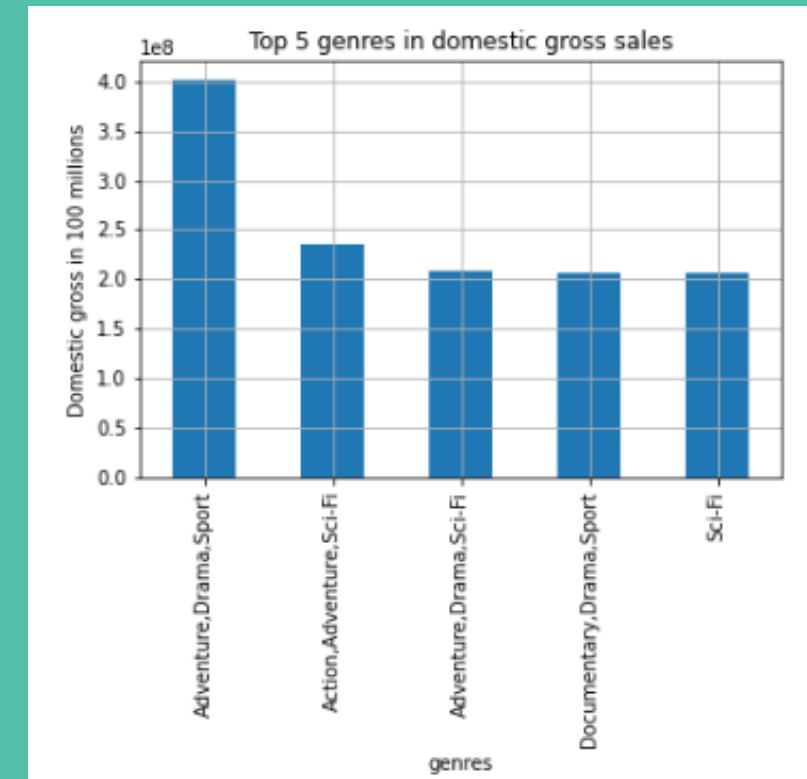
	tconst	averagerating	numvotes
0	tt10356526	8.3	31
1	tt10384606	8.9	559
2	tt1042974	6.4	20
3	tt1043726	4.2	50352
4	tt1060240	6.5	21

# DOMESTIC VS FOREIGN GROSS



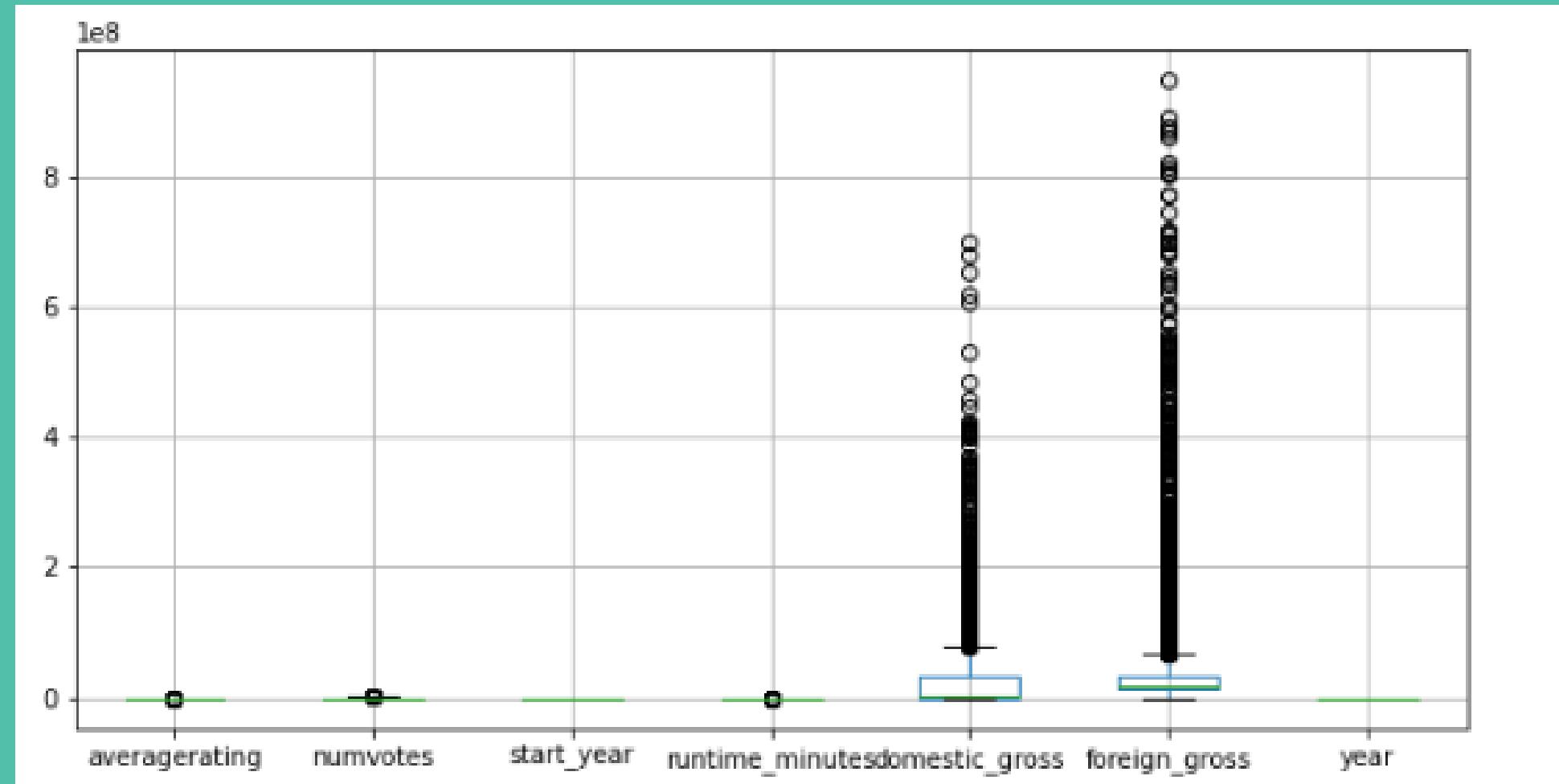
Domestic and foreign gross sales have a **positive correlation** meaning an increase in gross sales can lead to an increase in foreign sales.

# TOP 5 GROSS INCOME / GENRES



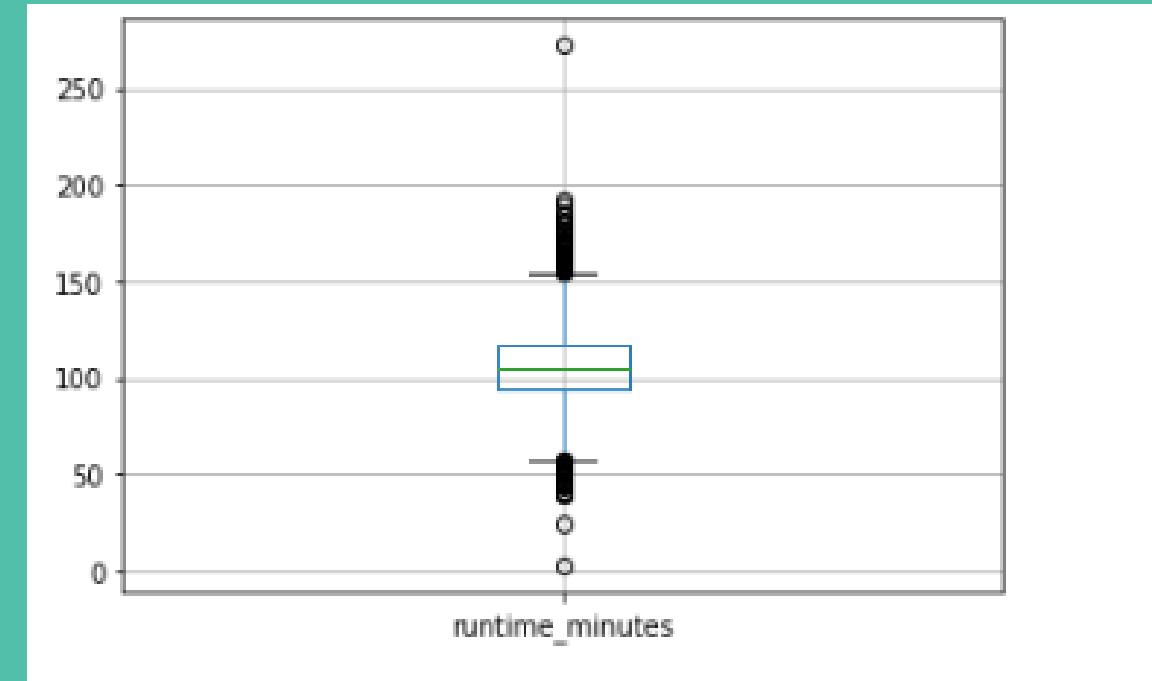
Action, Adventure, Drama, Sci-Fi and Sports are the top 5 genres in domestic sales while Mystery and fantasy add to more gross sales in foreign markets

# OVERALL BOXPLOT



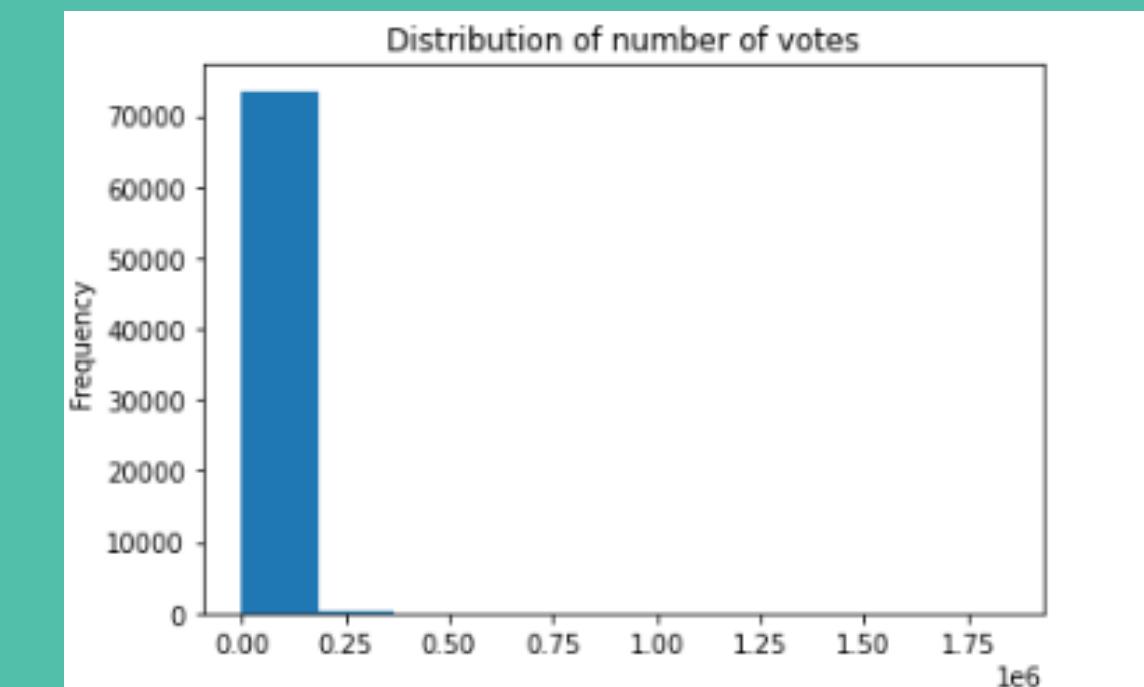
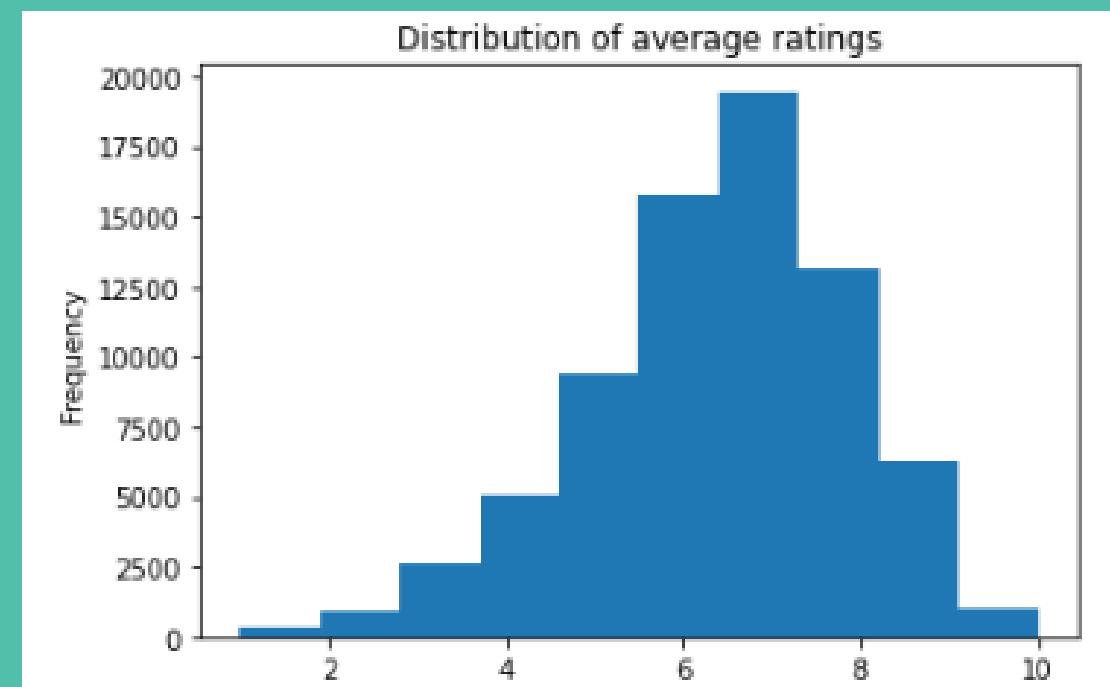
The boxplot shows that domestic and foreign gross sales are typically within **\$100million dollars** with some going as much as over **\$800million dollars**

# MOVIE RUNTIME



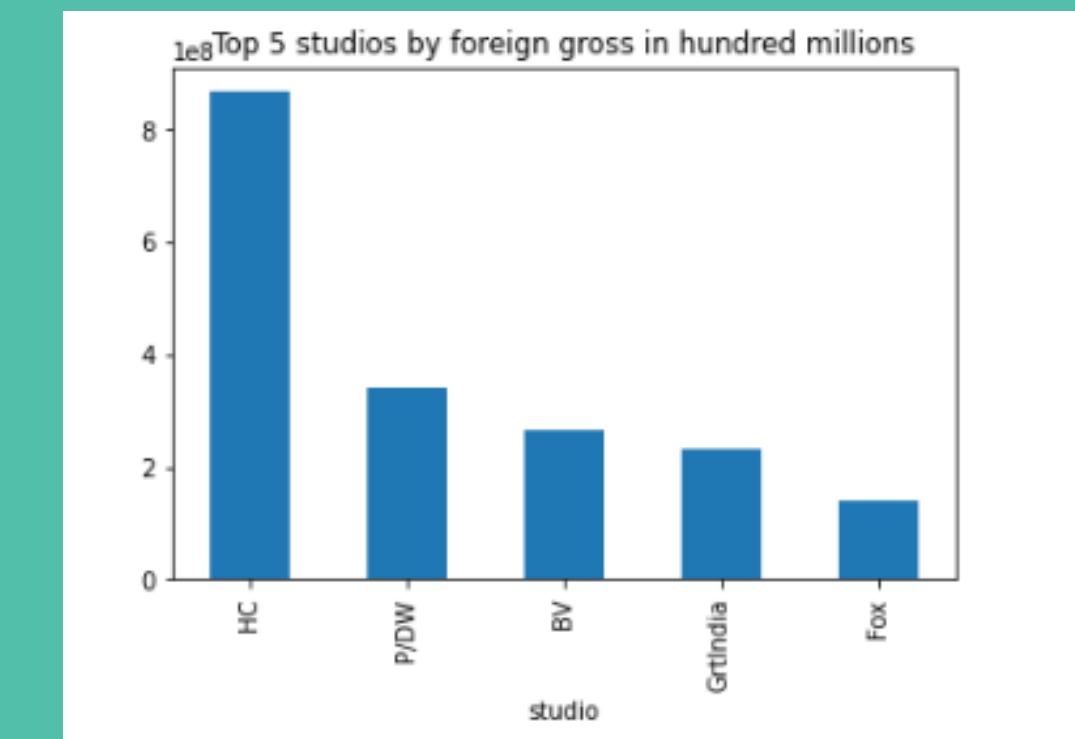
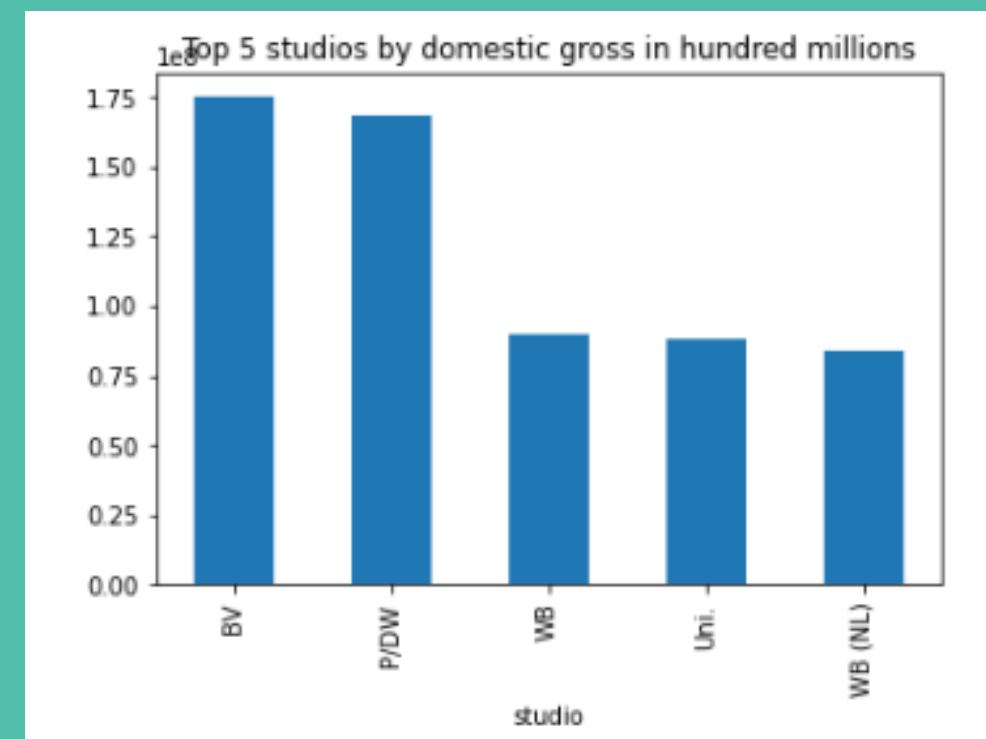
The boxplot shows that movies typically have a runtime of around 100 minutes.

# AVG. RATINGS AND NUMBER OF VOTES



The average rating per movie is ~7/10 with total number of votes at over 70,000

# STUDIOS



Buena Vista and **Dreamworks** studios have the highest total domestic gross earnings in the top 5 of **~\$1750million dollars**, whereas **HC** and **Dreamworks** have higher total foreign sales of over **\$8000million** and **~\$3500million** respectively

# RECOMMENDATION S & CONCLUSION



1. Microsoft should consider creating movies in the `Action`, `Drama`, `Adventure` and `Sci-Fi` as those receive high gross return in both domestic and foreign markets.
2. Microsoft should invest in high budget movies because movies that performed well and had high returns were from studios such as `Buena Vista (BV)`, `Warner Bros. (WB)` and `Dreamworks` who are well known for creating best selling movies.
3. Microsoft should create movies with an average runtime of around `100 minutes` as this is the most common watch time as per majority of other movies.

**THANK  
YOU VERY  
MUCH!**