

MICROSOFT MOVIE ANALYSIS



DAPHINE LUCAS

01
Business understanding



02
Data description and analysis



03
Data visualization



04
Insights



BUSINESS UNDERSTANDING

Business problem

Microsoft aims to produce profitable movies and seeks to understand which genres are most successful in achieving this objective. To address this inquiry, an analysis of domestic and foreign sales data, as well as the production budget relative to domestic and worldwide gross, was conducted. This analysis aimed to identify the genres that yield the highest financial returns. Additionally, the average rating and number of votes for each genre were examined to gauge the correlation between popularity and financial success.

Data sets used:

bom.movie_gross

imdb.title.basics

imdb.title.ratings

tn.movie_budgets





Analysis questions

1

What are the most voted movies and highly rated movie types?



2

What movie types have the highest average domestic gross and foreign gross income?



3

What specific metrics are considered when defining the success of a film genre at the box office? (domestic gross sale, foreign gross sales, worldwide gross sales, production budget, average rating, number of votes, genre specific metrics, rankings)



Analysis questions

4

How do production budgets correlate with gross revenue for different genres, and what implications does this have for investment decisions?



5

What factors must be considered when determining the ideal combination of genres for Microsoft's film production initiatives?

METHODS

1

The analytical process commenced with a thorough examination of each dataset to ascertain column names and identify any null values present.

2

Subsequently, I merged two datasets, `df_movie_gross` and `title_basics`, utilizing the 'title' column as a unique identifier, resulting in the creation of a new dataframe.

3

I integrated the `title_rating` dataset with the joined dataframe, using the `tconst` as the unique identifier, thereby generating a comprehensive combined dataset.

4

I then concatenated the merged data with `movie_budget`. Upon reviewing the information contained within `complete_df`, I proceeded with the analysis

DATA DESCRIPTION

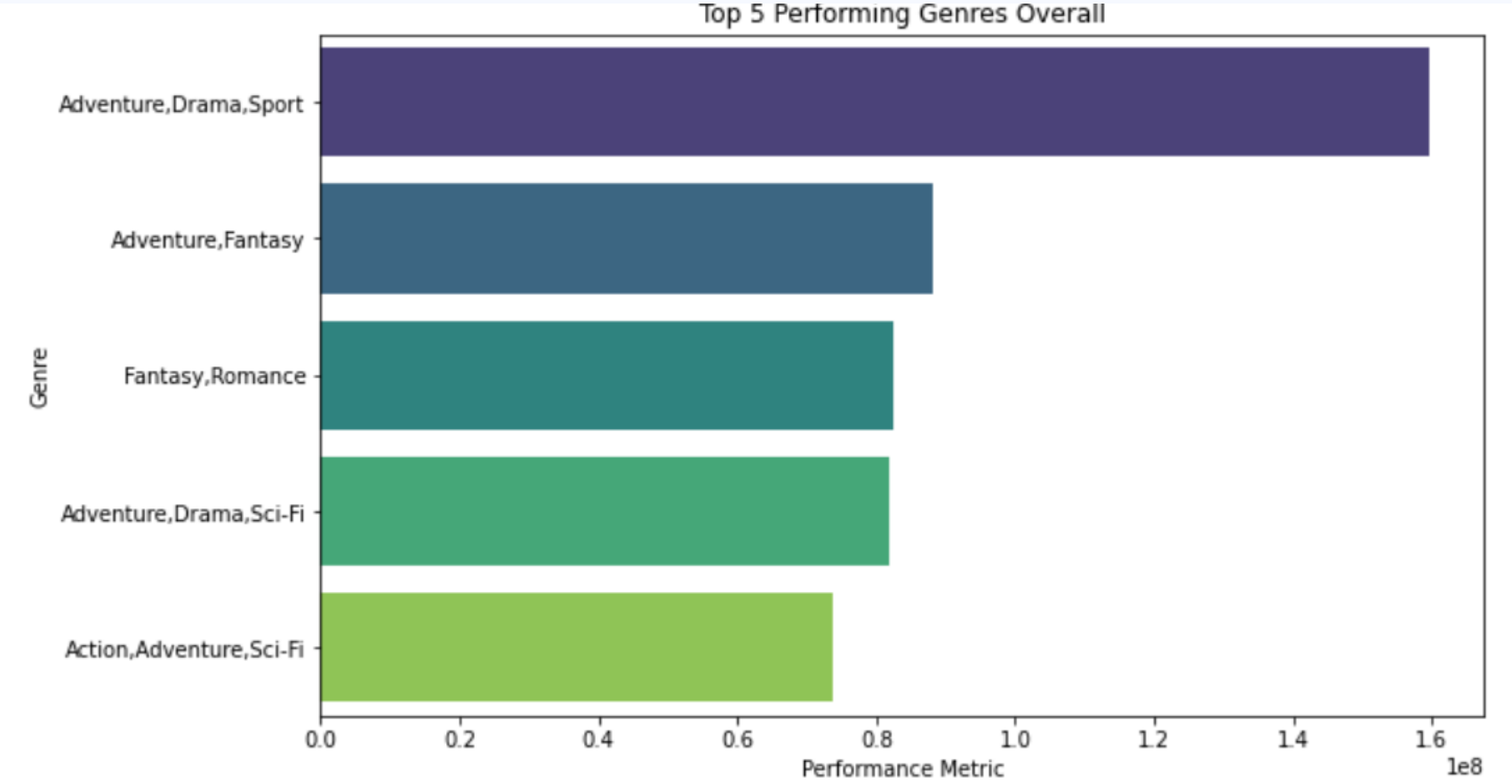
#	Column	Non-Null Count		Dtype
--	-----	----	-----	-----
0	title	2597	non-null	object
1	studio	2595	non-null	object
2	domestic_gross	2576	non-null	float64
3	foreign_gross	1608	non-null	object
4	year	2597	non-null	int64
5	tconst	2597	non-null	object
6	original_title	2597	non-null	object
7	start_year	2597	non-null	float64
8	runtime_minutes	2560	non-null	float64
9	genres	2593	non-null	object
10	averagerating	2597	non-null	float64
11	numvotes	2597	non-null	int64

	domestic_gross	year	start_year	runtime_minutes	averagerating	numvotes
count	2.576000e+03	2597.000000	2597.000000	2560.000000	2597.000000	2.597000e+03
mean	3.292019e+07	2013.699653	2013.243743	106.886719	6.452214	6.902101e+04
std	7.008290e+07	2.272781	2.209441	19.854740	0.987042	1.328612e+05
min	1.000000e+02	2010.000000	2010.000000	3.000000	1.600000	5.000000e+00
25%	1.510000e+05	2012.000000	2011.000000	94.000000	5.900000	3.231000e+03
50%	2.600000e+06	2014.000000	2013.000000	104.000000	6.500000	1.705500e+04
75%	3.540000e+07	2016.000000	2015.000000	118.000000	7.100000	7.432900e+04
max	7.001000e+08	2018.000000	2019.000000	272.000000	9.200000	1.841066e+06

DATA INFORMATION

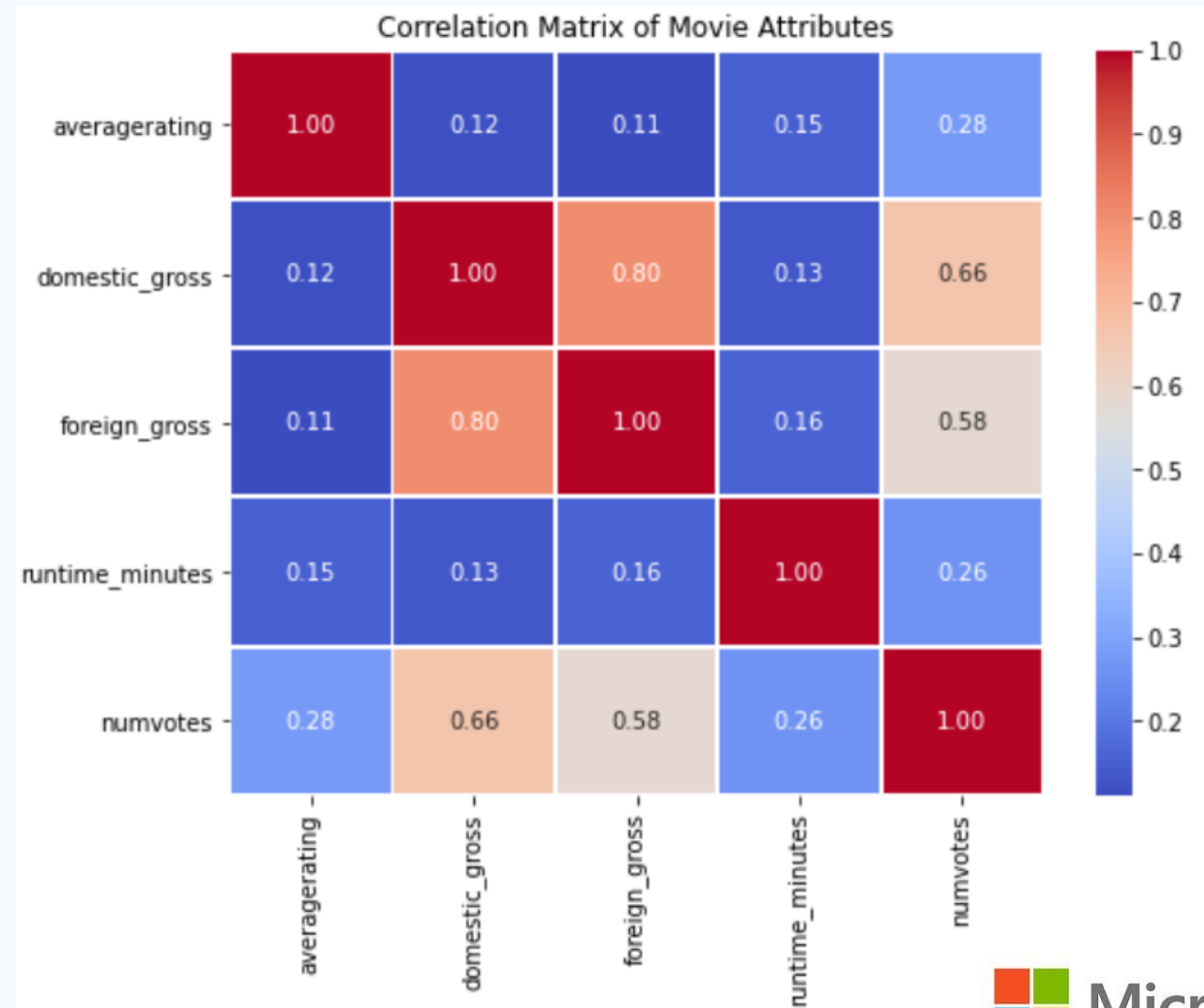
DATA DESCRIPTION

GENRE ANALYSIS



CORRELATION ANALYSIS

CORRELATION BETWEEN
AVERAGERATING,
DOMESTIC_GROSS,
FOREIGN_GROSS,
RUNTIME_MINUTES,
NUMVOTES





DATA ANALYSIS SUMMARY

Number of genres was 322

Highest average rating genre: Adventure

Lowest average rating genre: Comedy, Thriller

Genre with highest domestic gross: Adventure, Drama, Sport

Genre with lowest domestic gross: Comedy,

Genre with highest foreign gross: Adventure, Drama, Sport

Genre with lowest foreign gross: Biography, Documentary, Thriller

Genre with longest runtime: Drama, History,

Genre with shortest runtime: Action, Sport

Genre with most numvotes: Adventure, Drama, Sci-Fi

Genre with least numvotes: Documentary, Drama, Romance

Best Performing Genre: Adventure, Drama, Sport



Insights

Microsoft should put its focus on adventure, drama and sport genres.

Recommendations

Implement strategies to continually monitor and adjust to evolving audience preferences and market trends in the film industry



THANK YOU!

EMAIL: DAPHINELUCAS235@GMAIL.COM

GITHUB: DAPHINELUCAS

LINKEDIN: [HTTPS://WWW.LINKEDIN.COM/IN/DAPHINE-LUCAS-338136213?UTM_SOURCE=SHARE&UTM_CAMPAIGN=SHARE_VIA&UTM_CONTENT=PROFILE&UTM_MEDIUM=IOS_APP](https://www.linkedin.com/in/daphine-lucas-338136213?utm_source=share&utm_campaign=share_via&utm_content=profile&utm_medium=ios_app)

