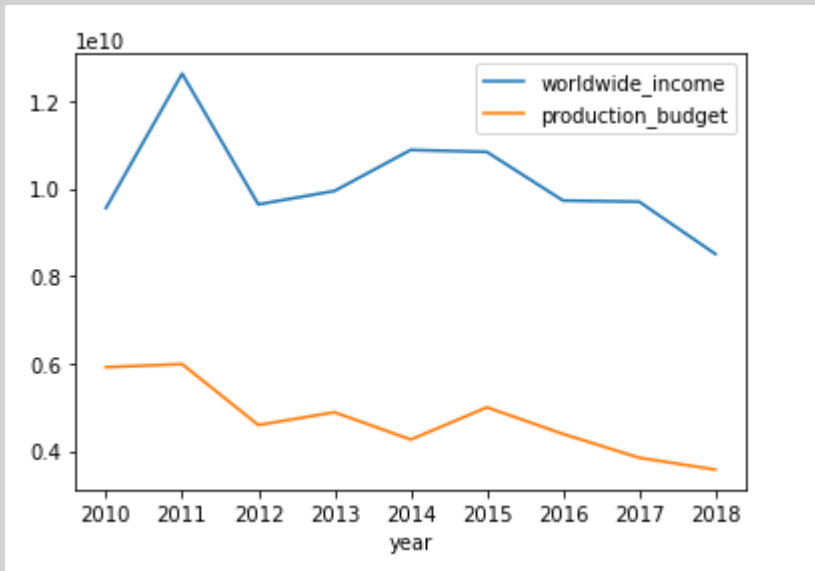


Movie Analysis

Giancarlos Morales

Problem Statement

Microsoft is looking to get into the movie business and needs some recommendations to get started.



Data Sources

The following sources were used

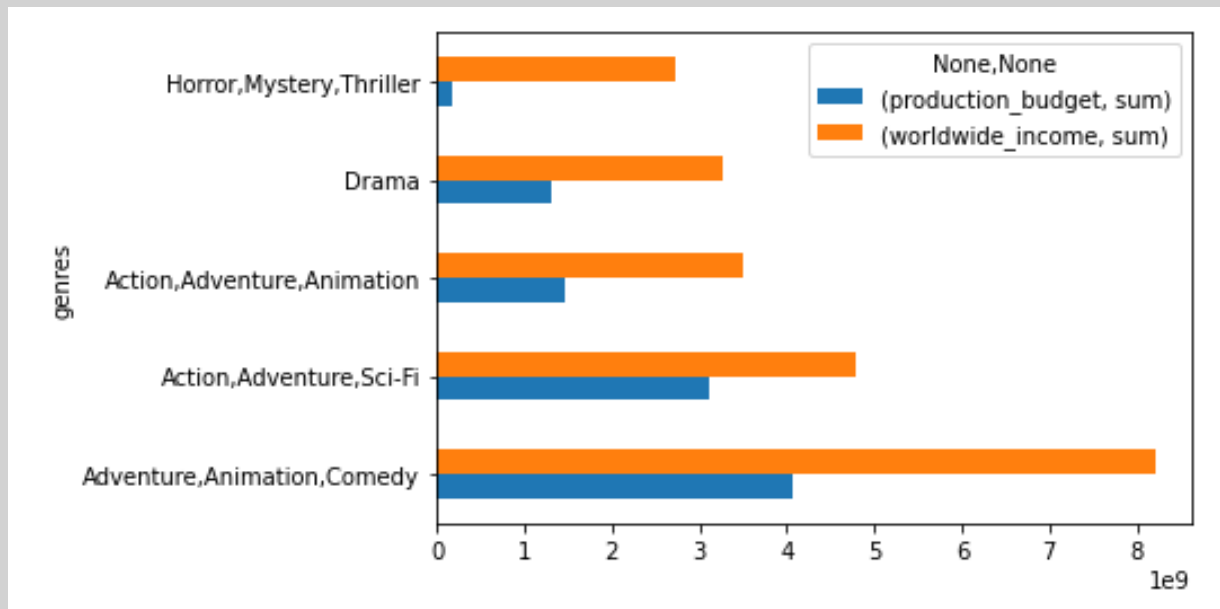
- The Numbers - budget by movie title
- IMBD – genre, writer, and director
- Box Office Mojo – studio

Methodology

- Used the following methods
 - Import Data– Glob
 - Data Cleaning – Pandas and Numpy
 - Data normalization - scipy
 - Visualizations – Mathplotlib and Seaborn

What is the Most Profitable Genre

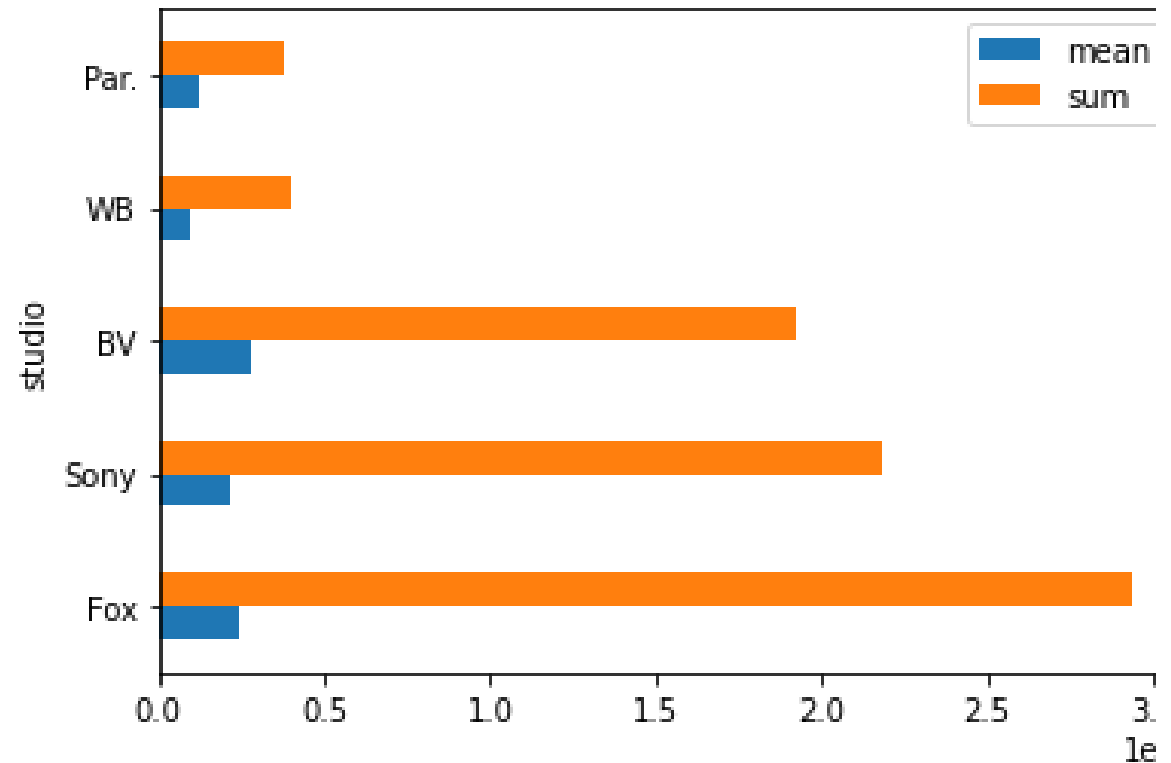
- Highest income movie genre is Adventure/Animation/Comedy
- Looks like production budget and worldwide income are closely related



Which Studios make the most profitable adventure/animation/comedy movies?

Fox and Sony would both be good candidates.

studio	mean	count	sum
Fox	2.443953e+08	12	2932743236
Sony	2.181866e+08	10	2181866221
BV	2.748213e+08	7	1923748814
WB	9.869397e+07	4	394775864
Par.	1.254218e+08	3	376265432



Which Directors Make
the highest ROI for the
best performing genre

directors	mean	count	sum
nm0757858	3.438663e+08	3	1.031599e+09
nm0331532	3.895658e+08	2	7.791317e+08
0261,nm0615780,nm2320658	5.193839e+08	1	5.193839e+08

Recommendations

- 1.) Adventure/Animation/Comedy is the best genre by income generated
- 2.) Fox is the studio which has the most financial success creating this genre.
- 3.) Director id nm0757858 has created the highest number of movies and generated most income.

Next Steps

Look at relationship between vote count / popularity and worldwide income.

Predict the total income per vote.

	production_budget	domestic_gross_x	worldwide_gross	runtime_minutes	popularity	vote_average	vote_count	year	worldwide_income	production_budget_Z	worldwide_income_Z	worldwide_gross_Z	gross_return
production_budget	1.0	0.6	0.8	0.2	0.4	-0.0	0.5	-0.0	0.5	0.6	0.2	0.4	-0.2
domestic_gross_x	0.6	1.0	0.9	0.2	0.4	0.1	0.6	-0.0	0.9	0.2	0.3	0.5	0.1
worldwide_gross	0.8	0.9	1.0	0.2	0.5	0.1	0.7	0.0	1.0	0.4	0.5	0.6	0.0
runtime_minutes	0.2	0.2	0.2	1.0	0.3	0.3	0.3	0.0	0.2	0.1	0.1	0.1	-0.1
popularity	0.4	0.4	0.5	0.3	1.0	0.3	0.6	0.2	0.4	0.2	0.2	0.3	-0.1
vote_average	-0.0	0.1	0.1	0.3	0.3	1.0	0.4	0.0	0.2	0.1	0.1	0.1	-0.1
vote_count	0.5	0.6	0.7	0.3	0.6	0.4	1.0	-0.0	0.6	0.2	0.3	0.4	-0.0
year	-0.0	-0.0	0.0	0.0	0.2	0.0	-0.0	1.0	0.0	-0.0	0.0	0.0	0.0
worldwide_income	0.5	0.9	1.0	0.2	0.4	0.2	0.6	0.0	1.0	0.3	0.5	0.6	0.1
production_budget_Z	0.6	0.2	0.4	0.1	0.2	0.1	0.2	-0.0	0.3	1.0	0.3	0.5	0.0
worldwide_income_Z	0.2	0.3	0.5	0.1	0.2	0.1	0.3	0.0	0.5	0.3	1.0	0.9	-0.0
worldwide_gross_Z	0.4	0.5	0.6	0.1	0.3	0.1	0.4	0.0	0.6	0.5	0.9	1.0	-0.0
gross_return	-0.2	0.1	0.0	-0.1	-0.1	-0.1	-0.0	0.0	0.1	0.0	-0.0	-0.0	1.0



Thanks!