

The science of movies

by Jamel Dargan

The Project

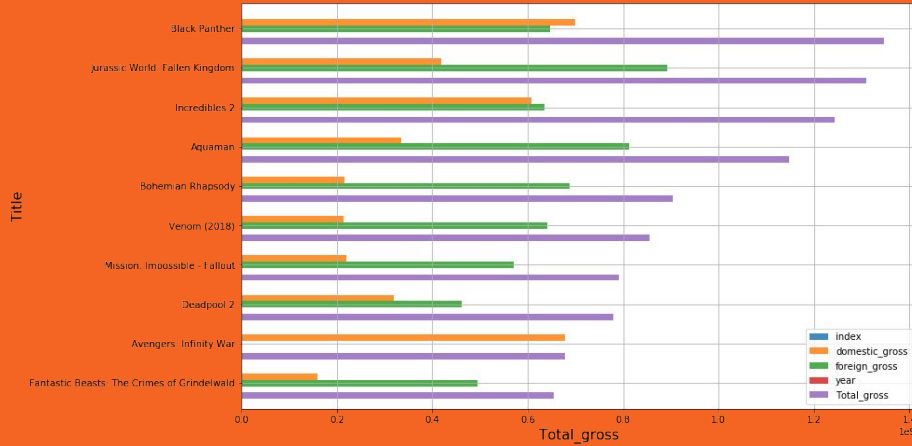
Our objective was to use multiple datasets to explore what type of films are currently doing the best at the box office, to help decide what types of films should be created by a new movie studio.

Sources include:

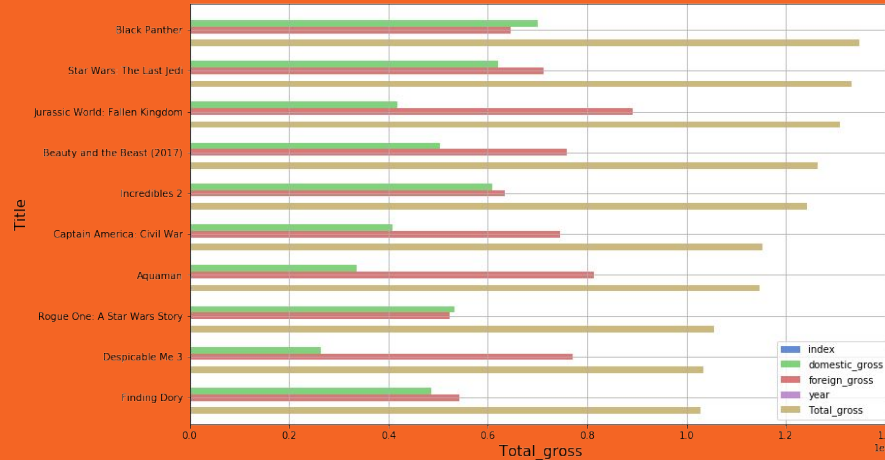
- Box Office Mojo
- IMDB
- Rotten Tomatoes
- TheMovieDB.org



Top 10 Highest Gross Titles, 2018



Top 10 Highest Gross Titles, 2016-18



1. Titles

...Get a sense of what the industry is doing now.

→ Financial Success

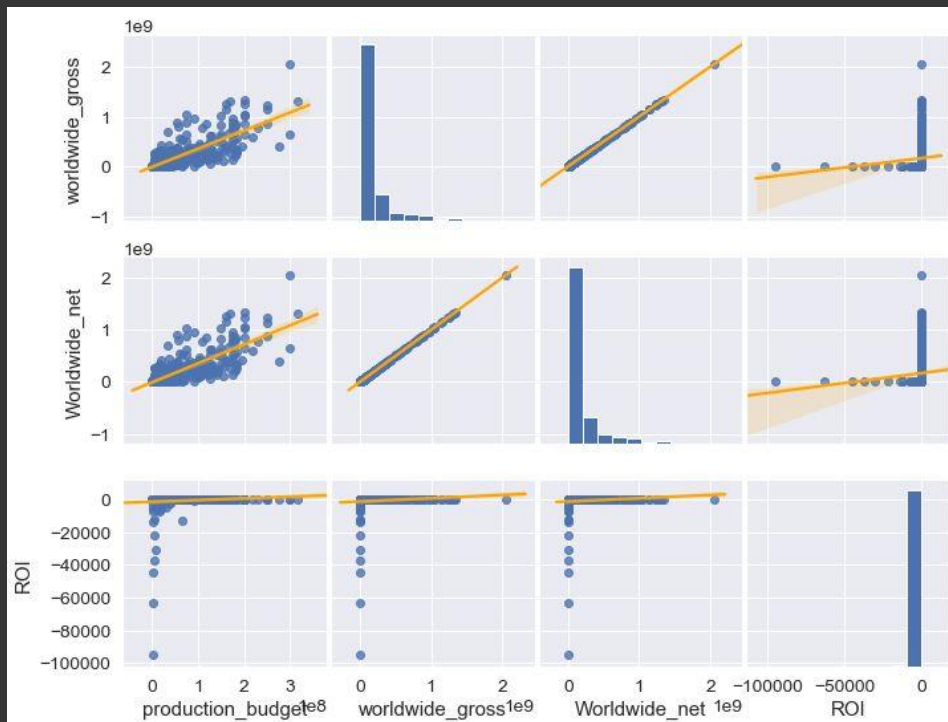
Low-billions and average gross earnings in tens-of-billions

→ Success Metrics

Relationships in the data (budget, gross, net)

→ Potential

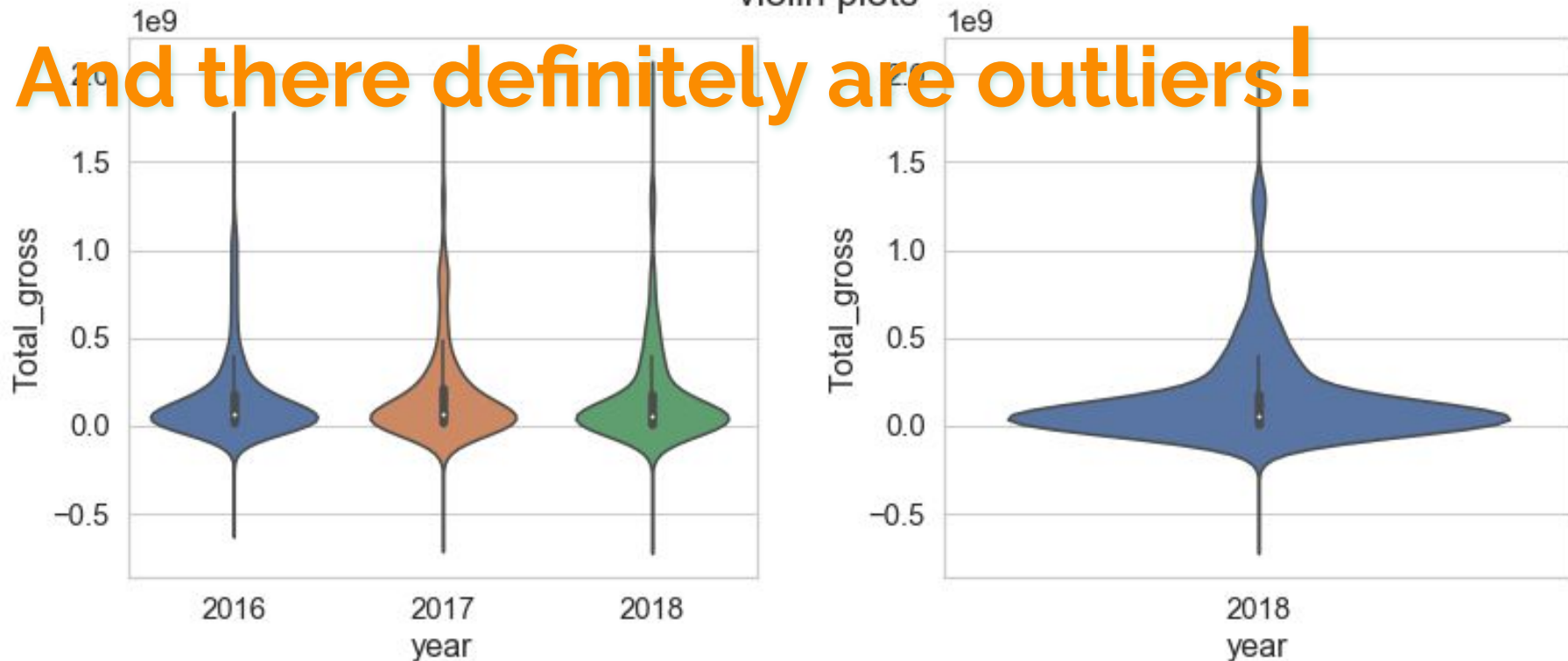
Strong data relationships



Decision

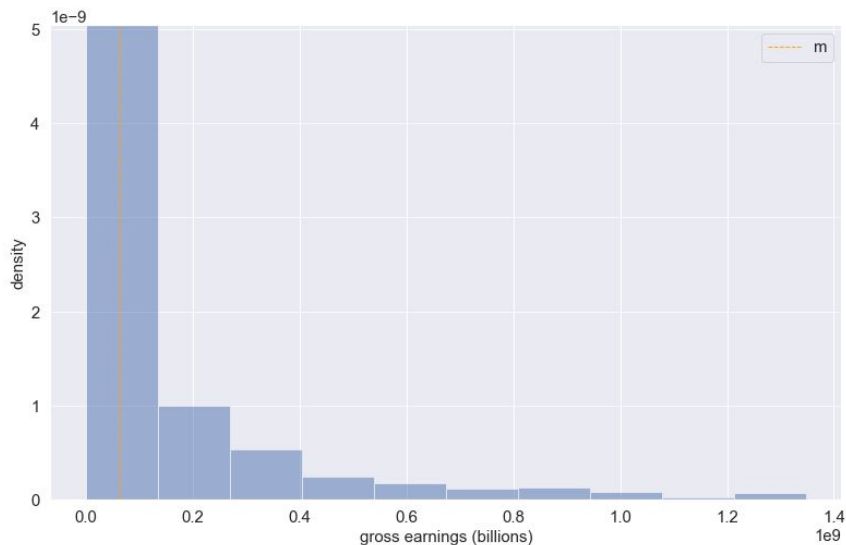
For the most part, as budgets go up, so do earnings. Keeping the outliers (for starters), keeps focus on choices of the most successful.

Gross Earnings Distribution
violin plots



Worldwide Gross

2016-18 Total Gross Earnings (billions),
with median line 'm'



Finding

There is a lot of room in the [middle of the] data.

Median budget films have potential to reap pleasant rewards.

Ooh, Gross!

	index	title	studio	domestic_gross	foreign_gross	year	Total_gross
3075	3080	Black Panther	BV	700100000.0	646900000.0	2018	1.347000e+09
3076	3081	Jurassic World: Fallen Kingdom	Uni.	417700000.0	891800000.0	2018	1.309500e+09
3077	3082	Incredibles 2	BV	608600000.0	634200000.0	2018	1.242800e+09
3078	3083	Aquaman	WB	335100000.0	812700000.0	2018	1.147800e+09
3079	3084	Bohemian Rhapsody	Fox	216400000.0	687200000.0	2018	9.036000e+08
3080	3085	Venom (2018)	Sony	213500000.0	641500000.0	2018	8.550000e+08
3081	3086	Mission: Impossible - Fallout	Par.	220200000.0	571000000.0	2018	7.912000e+08
3082	3087	Deadpool 2	Fox	318500000.0	460500000.0	2018	7.790000e+08
3074	3079	Avengers: Infinity War	BV	678800000.0	1369.5	2018	6.788014e+08
3083	3088	Fantastic Beasts: The Crimes of Grindelwald	WB	159600000.0	494100000.0	2018	6.537000e+08

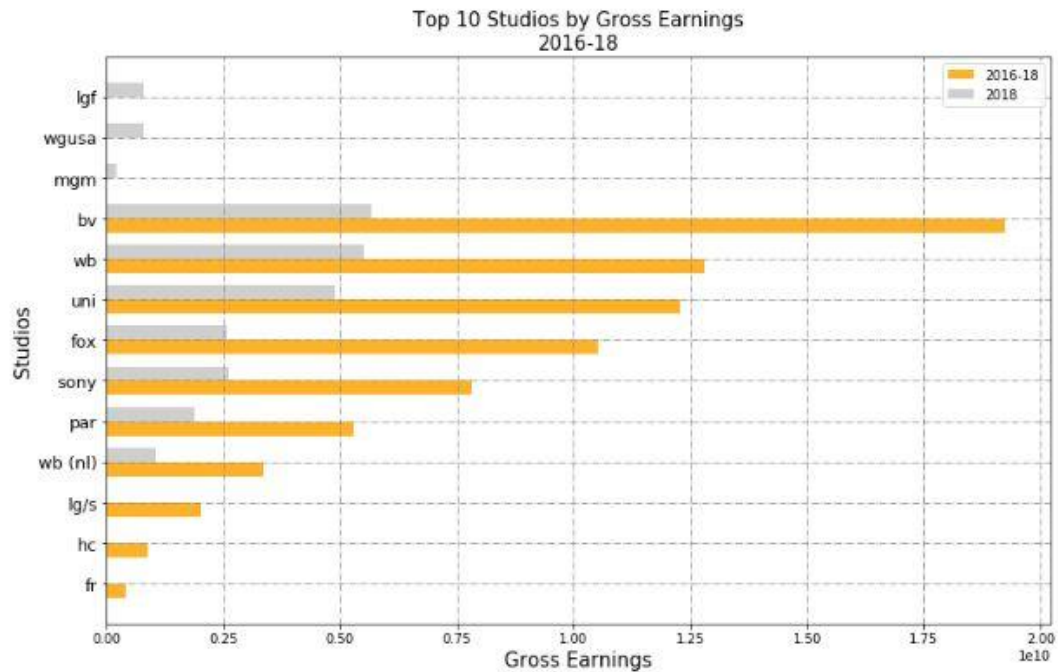


Finding

Even before visual EDA the data begins to tell a story re. action/ adventure, sequels, franchises, remakes, and animation.

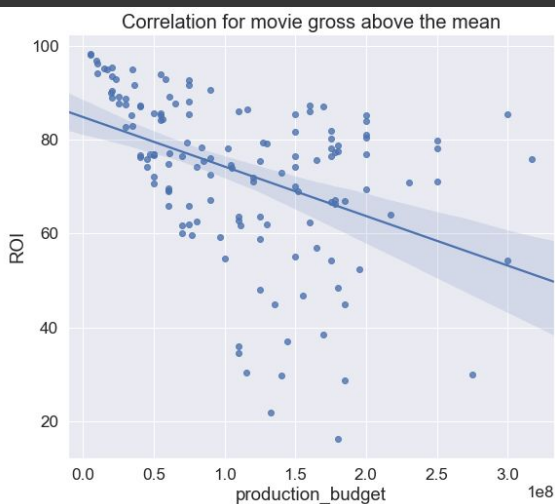
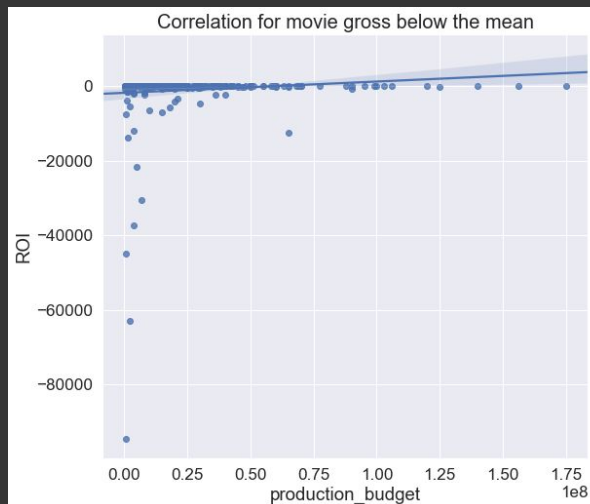
Visual EDA more clearly brings the image into focus.

2. Studios



ROI! Possibly Maybe.

(may be strongly dependent on more factors than we presently have time to explore)



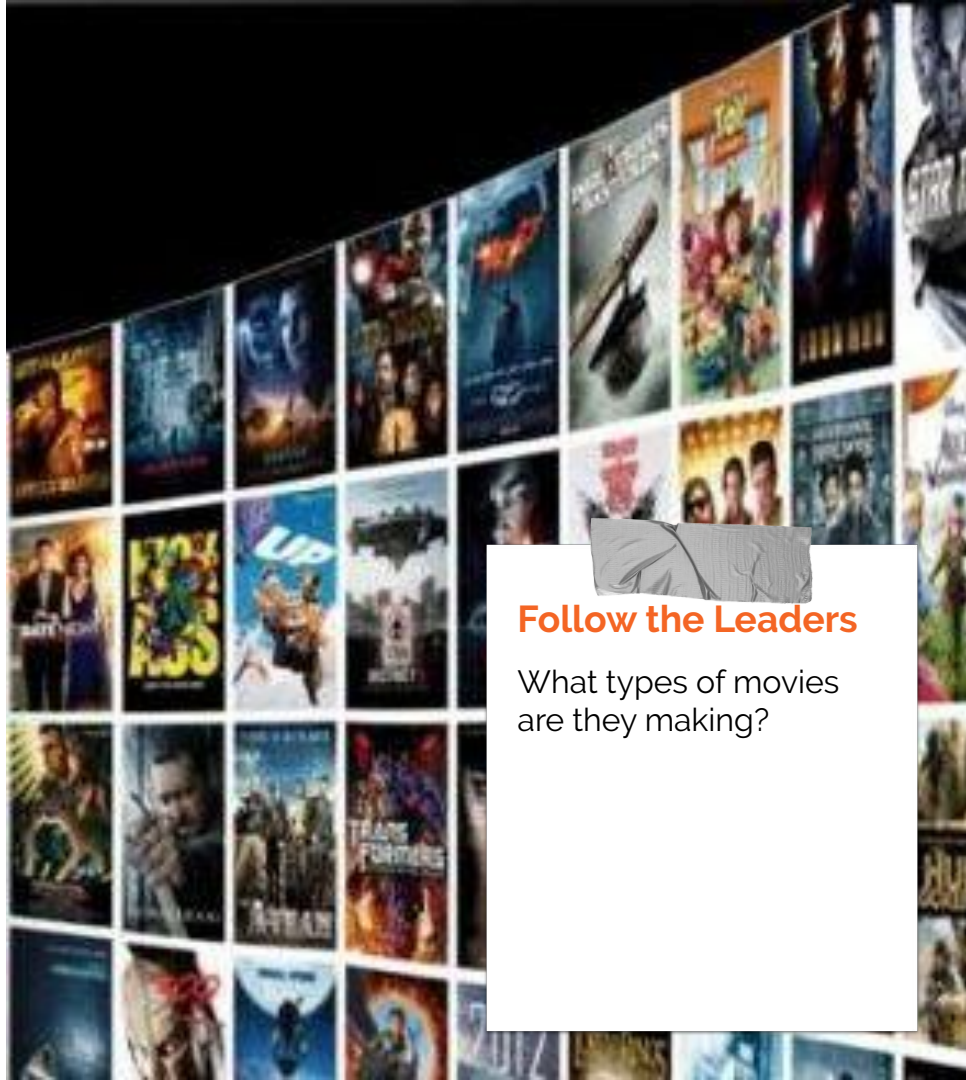
Observe

Big budgets make big dollars, but the percentage return on investment may be relatively small and/ or a long time in the making.

Are there Trends among the Studios?

studio	num_Titles	sum_Gross
BV	10	5.667401e+09
WB	17	5.494300e+09
Uni.	21	4.900200e+09
Sony	15	2.615626e+09
Fox	11	2.583500e+09
Par.	8	1.886200e+09

It happens that the top 6 studios for 2018 are the top 6 studios for 2016-18.



Follow the Leaders

What types of movies
are they making?



3. Genre

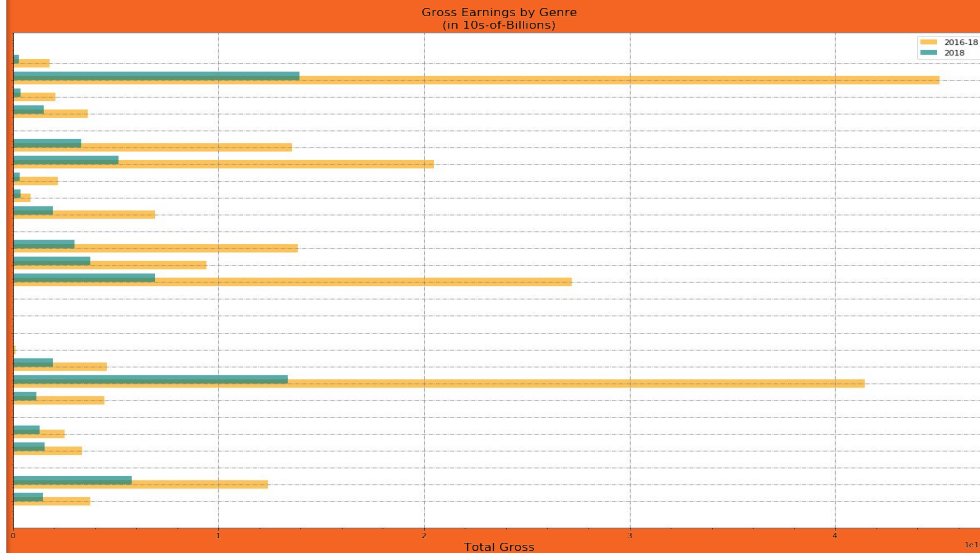
Popularity is not necessarily predictive.:

→ **Most Frequently Used Genres**

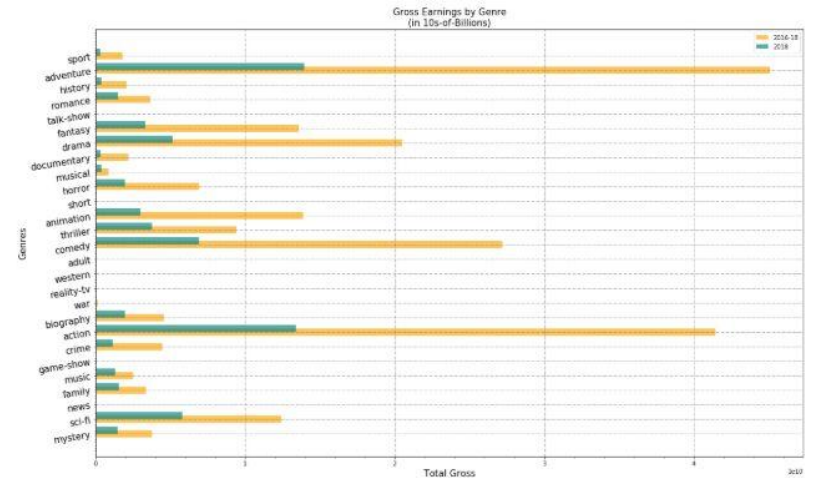
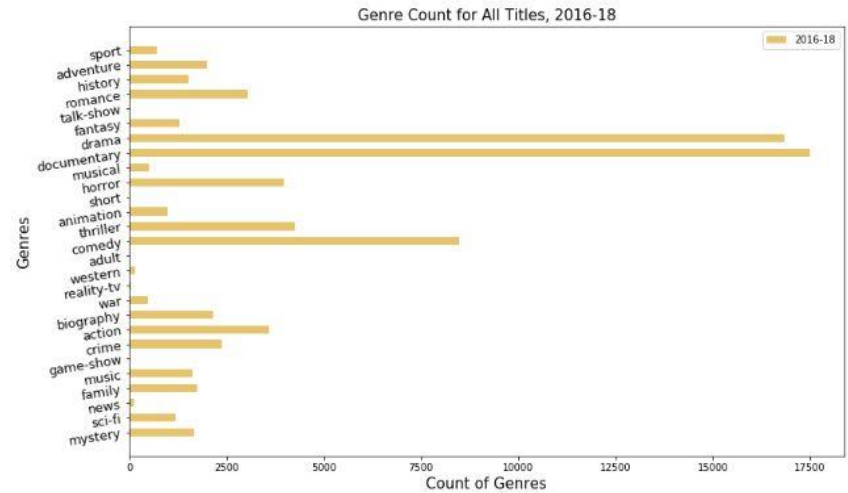
→ **Most Money - making**

→ **Studio Mix**

How do top studios portion the genres they produce?

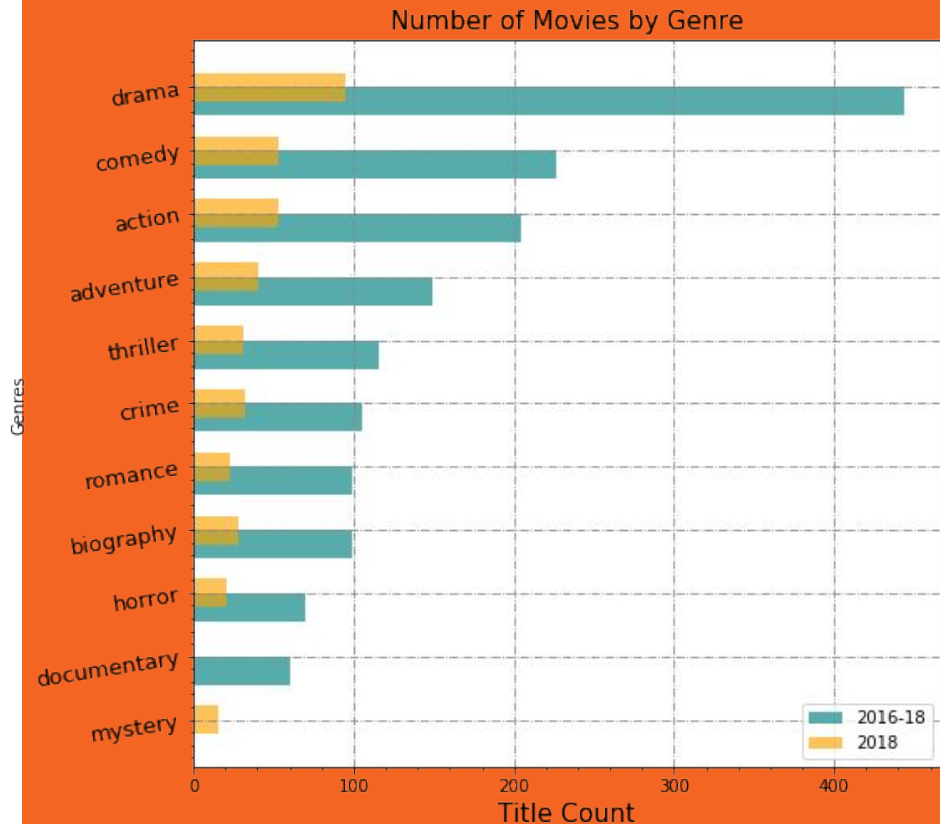


The types of movies
most - often made and
the types of movies
making the most money
are not always the same.



The 6 top studios seem to stick with a formula.

Their top 10 produced genres for **2018** AND **2016-18** are very similar.





4. Closing

→ Recommendation

- ◆ Follow the leaders
- ◆ Their top 10 genres seem like the best place to start.

→ What's next?

- ◆ Look for more predictors
- ◆ Investigate ROI by genre

Thank You.
