

## W200 Project 2 Proposal

### Team Members:

- Allie Ayrapetyan
- Shehzad Shahbuddin
- Ahmad Azizi

### Team GitHub Repository:

Project2\_Azizi\_Shahbuddin\_Ayrapetyan

### Primary Dataset:

All time movie box office data (top 1000) that includes movie names, ranks, international gross, domestic gross and worldwide gross.

<https://www.the-numbers.com/box-office-records/worldwide/all-movies/cumulative/all-time>

We plan on merging additional data fields such as movie genre, studio name, movie director and potentially top 3 stars in the movie.

### Variables we plan to explore:

- Worldwide Box Office
- Domestic Box Office
- International Box Office
- Movie Rank
- Movie Release Date
- Movie Genre
- Movie Studio
- Movie Director

Optional:

- Top 3 Stars in the movie

### Research question:

We are curious to see what factors influence a movie's success.

1. How does a genre influence how great a movie does?
2. Is behind the camera talent(director) more impactful than on camera (i.e. having world renown movie stars)?
3. Does movie studio determine movie success?
4. Does year of movie release determine its success?

### Exploratory Analysis:

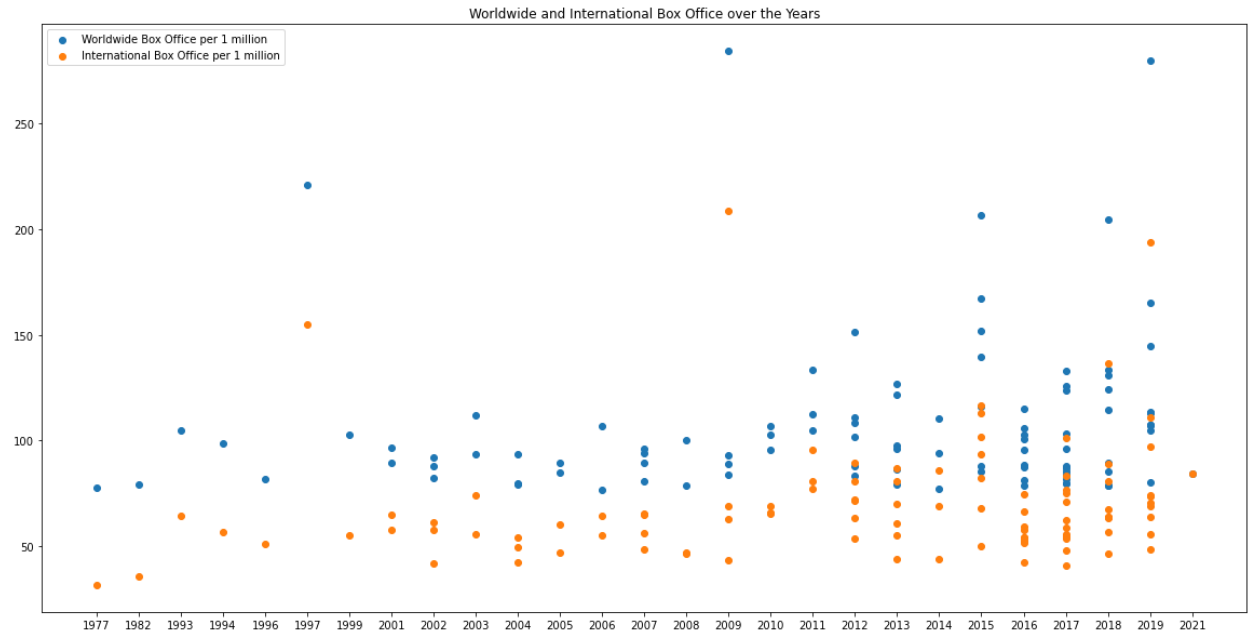
To get a sense of the movies in our data, we looked at the top 20 highest grossing movies of all time and showing them in the below table.

Rank	Rank	Year	Movie	WorldwideBox_Office	DomesticBox_Office	InternationalBox_Office
1	1	2009	Avatar	\$2,845,899,541	\$760,507,625	\$2,085,391,916
2	2	2019	Avengers: Endgame	\$2,797,800,564	\$658,373,000	\$1,939,427,564
3	3	1997	Titanic	\$2,207,986,545	\$659,363,944	\$1,548,622,601
4	4	2015	Star Wars Ep. VII: The Force Awakens	\$2,064,615,817	\$936,662,225	\$1,127,953,592
5	5	2018	Avengers: Infinity War	\$2,044,540,523	\$678,815,482	\$1,365,725,041
6	6	2015	Jurassic World	\$1,669,979,967	\$652,306,625	\$1,017,673,342
7	7	2019	The Lion King	\$1,654,381,934	\$543,638,043	\$1,110,743,891
8	8	2015	Furious 7	\$1,516,881,526	\$353,007,020	\$1,163,874,506
9	9	2012	The Avengers	\$1,515,100,211	\$623,357,910	\$891,742,301
10	10	2019	Frozen II	\$1,446,925,396	\$477,373,578	\$969,551,818
11	11	2015	Avengers: Age of Ultron	\$1,395,316,979	\$459,005,868	\$936,311,111
12	12	2018	Black Panther	\$1,336,494,321	\$700,059,566	\$636,434,755
13	13	2011	Harry Potter and the Deathly Hallows:...	\$1,334,392,544	\$381,193,157	\$953,199,387
14	14	2017	Star Wars Ep. VIII: The Last Jedi	\$1,331,635,141	\$620,181,382	\$711,453,759
15	15	2018	Jurassic World: Fallen Kingdom	\$1,308,334,005	\$417,719,760	\$890,614,245
16	16	2013	Frozen	\$1,286,648,320	\$400,953,009	\$885,695,311
17	17	2017	Beauty and the Beast	\$1,255,080,655	\$504,014,165	\$751,066,490
18	18	2018	Incredibles 2	\$1,242,805,359	\$608,581,744	\$634,223,615
19	19	2017	The Fate of the Furious	\$1,238,703,796	\$225,764,765	\$1,010,939,031
20	20	2013	Iron Man 3	\$1,215,392,272	\$408,992,272	\$806,400,000

20% were debuted in 2018 and 2015,

We also want to get an idea of how year of movie release effects gross so plotting the below scatterplots:

The figure below shows worldwide and international box office over the years.



The histogram below shows the spread of movies released in the different years.

