

AI academy



Final capstone
project

03/03/2023.



Integrants



Miguel Cruz Le Hardy



Gerardo Yañez



**Araceli Paola Malagón
Sánchez**



**Maximiliano García
Puerto**



**Jackson, Daniel
Chimdiebube**

1.Introduction

- Overview
- Business Understanding

2.Analysis

- Data Understanding
- Data Analysis
- Statistical Inference

3.Summary

- Recommendations
- Next steps
- Thank You

Overview

In this project we were exploring data sets using data analysis, statistical methods and generating insights to solve the needs of Computing Vision.

Business Understanding

Our team is currently tasked with generating valuable insights as computer vision ventures into the film industry:

- The movies who have the higher earnings and its categories
- The best categories to develop a movie in
- The most likely successful directors within these categories to make a successful movie
- Find out if the duration of a movie has an influence in the earnings and engagement

Analysis

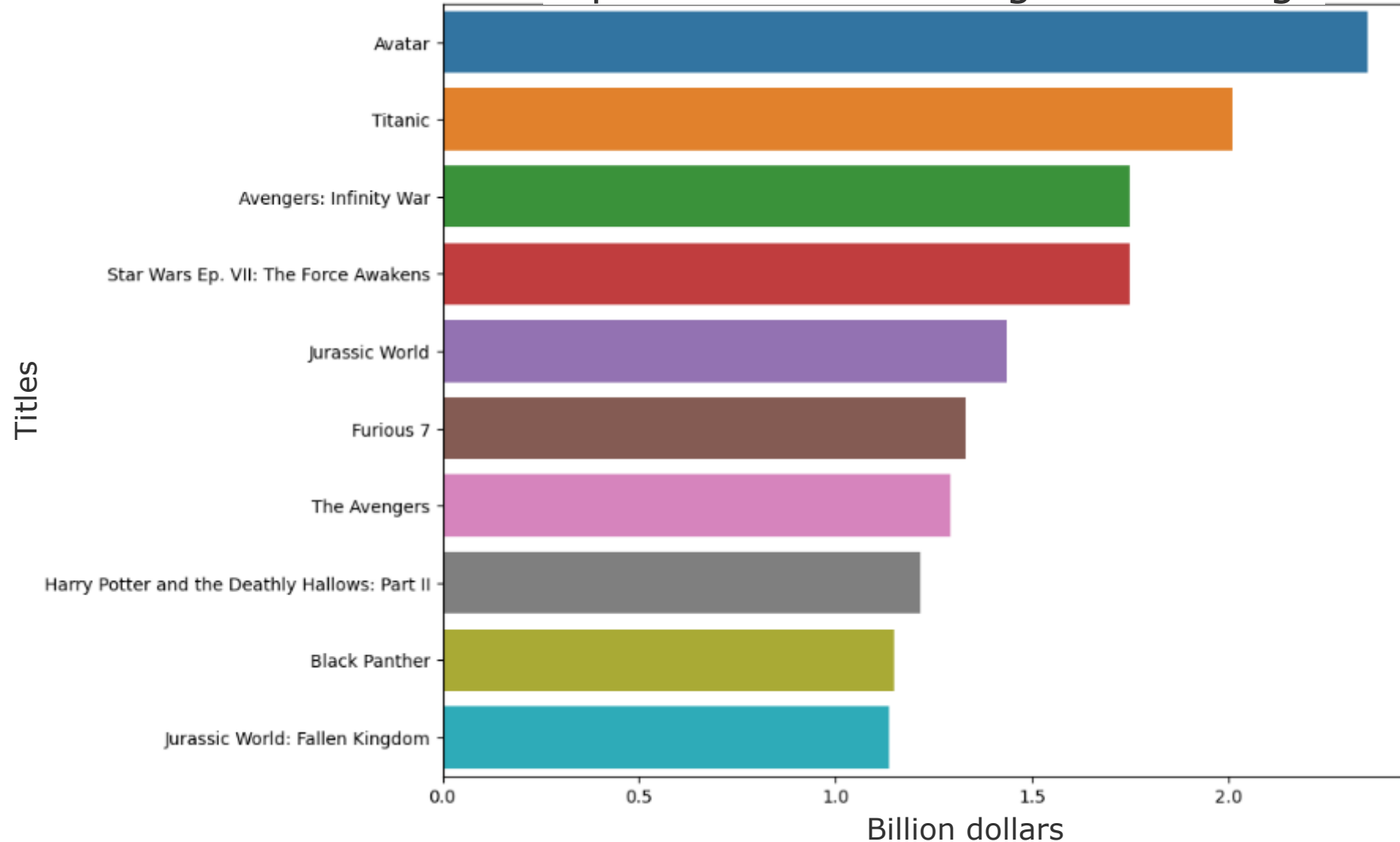


Data Understanding

- The data that we used is from pages about movies such as the moviedb.org and rotten tomatoes. We used them due to their records where we got info such as movie ratings, genres, directors, etc
- We used data sets with more than 140,000 entries and using statistical measures such as mean or standard deviation to understand the behavior of data
- The conclusions and graphs we're getting from the information of the datasets, such as the graph of the most popular genres in movies using columns from im.db dataset
- We had to adjust our analysis due to limitations with some datasets that weren't up to date and other ones had missing data

Data Analysis

Top 10 movies with highest earnings

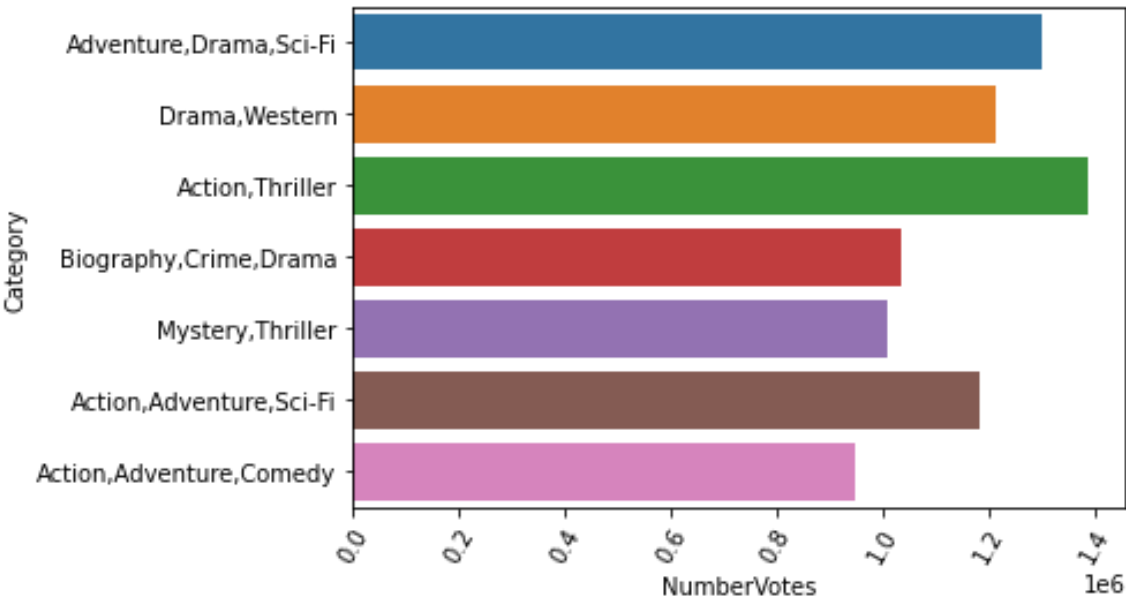


Data Analysis - Earnings

Top 10 movies with highest earnings and it's categories

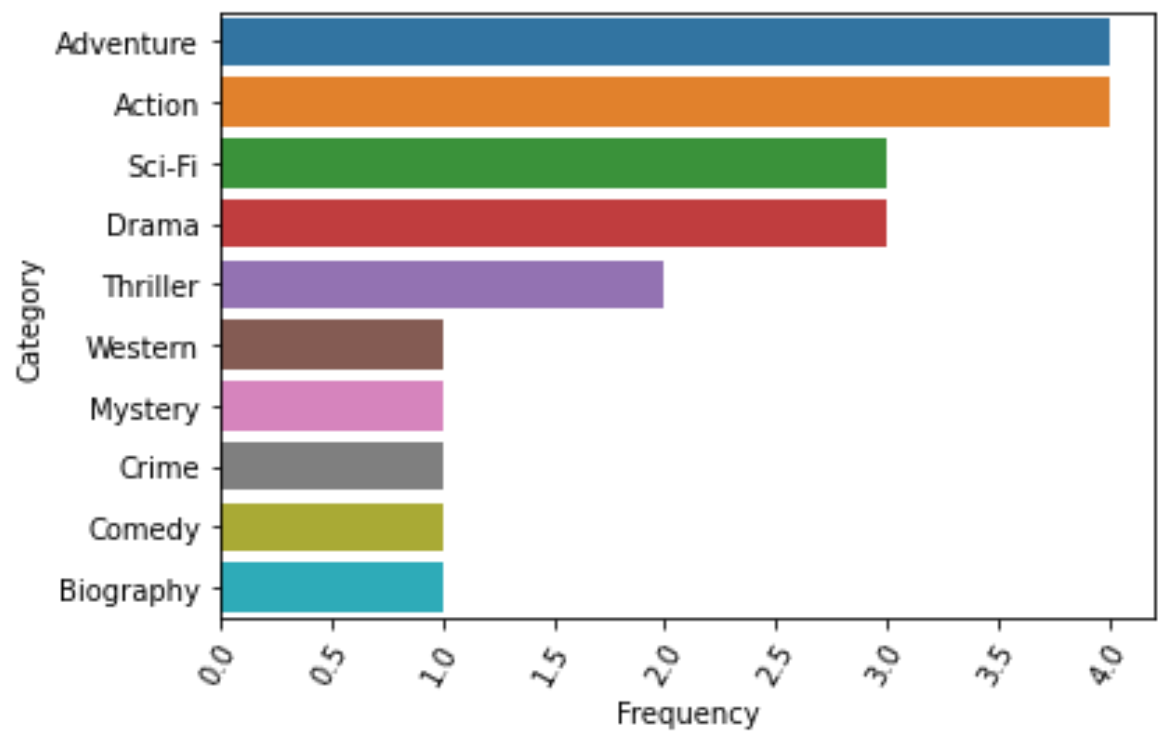
Movie	Genre
Avatar	Action, Adventure, Sci-fi, Fantasy
Titanic	Drama
Avengers: Infinity War	Action, Sci-fi
Star Wars Ep. VII: The Force Awakens	Action, Sci-fi-, Adventure, Fantasy
Jurassic World	Action, Sci-fi-, Thriller, Adventure
Furious 7	Action, Adventure
The Avengers	Action, Fantasy, Adventure, Sci-fi
Harry Potter and the Deathly Hallows: Part II	Adventure, Fantasy, Mystery, Drama
Black Panther	Action, Fantasy, Adventure, Sci-fi
Jurassic World: Fallen Kingdom	Action, Adventure, Sci-Fi

Best categories with the most votes



	Category	Rating	NumberVotes
0	Adventure,Drama,Sci-Fi	8.6	1299334
1	Drama,Western	8.4	1211405
2	Action,Thriller	8.4	1387769
3	Biography,Crime,Drama	8.2	1035358
4	Mystery,Thriller	8.1	1005960
5	Action,Adventure,Sci-Fi	8.1	1183655
6	Action,Adventure,Comedy	8.1	948394

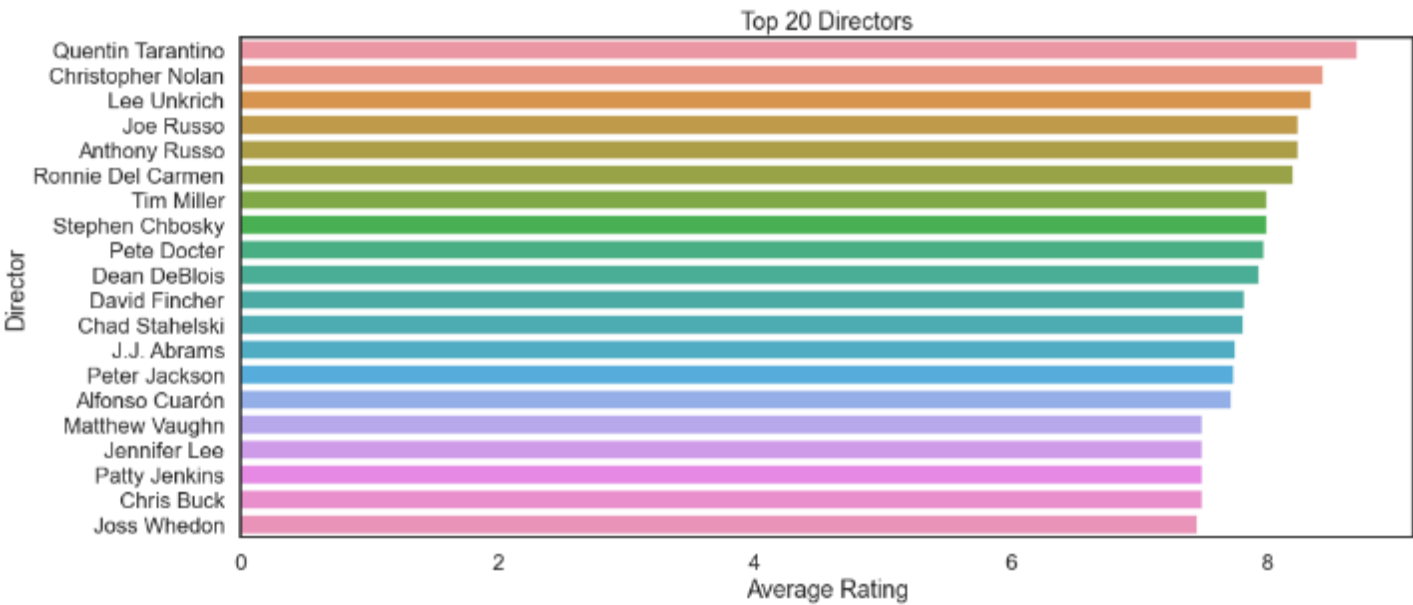
Best categories to make a movie



The top 5 categories with the most votes match with the categories of table top 10 higher earnings

Data Analysis-Directors

Top 20 Directors with their average ratings

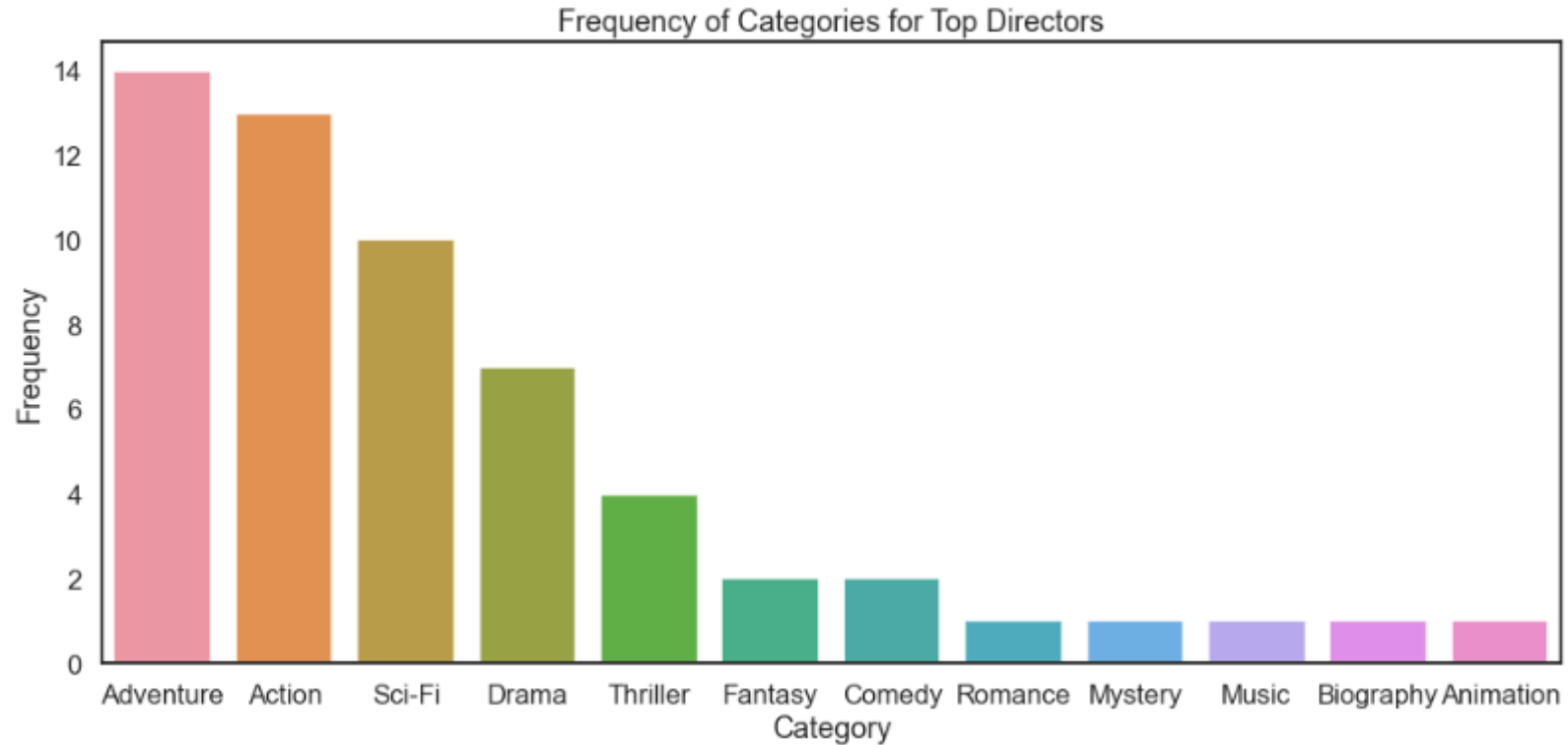


Directors organized by the movies they made and categories they worked

	Category	Rating	movie_id	Directors	primary_title
0	Action,Adventure,Biography	8.0	tt1663202	Alejandro G. Iñárritu	The Revenant
1	Drama,Sci-Fi,Thriller	7.7	tt1454468	Alfonso Cuarón	Gravity
2	Action,Adventure,Sci-Fi	7.8	tt1843866	Anthony Russo	Captain America: The Winter Soldier
3	Action,Adventure,Sci-Fi	8.5	tt4154756	Anthony Russo	Avengers: Infinity War
4	Action,Adventure,Sci-Fi	8.0	tt1877832	Bryan Singer	X-Men: Days of Future Past
5	Adventure,Drama,Sci-Fi	8.6	tt0816692	Christopher Nolan	Interstellar
6	Action,Thriller	8.4	tt1345836	Christopher Nolan	The Dark Knight Rises
7	Action,Adventure,Sci-Fi	8.8	tt1375666	Christopher Nolan	Inception
8	Drama,Music	8.5	tt2582802	Damien Chazelle	Whiplash
9	Drama,Thriller	8.0	tt0947798	Darren Aronofsky	Black Swan
10	Drama,Mystery,Thriller	8.1	tt2267998	David Fincher	Gone Girl
11	Comedy,Drama,Romance	7.7	tt1045658	David O. Russell	Silver Linings Playbook
12	Adventure,Drama,Fantasy	8.1	tt1201607	David Yates	Harry Potter and the Deathly Hallows: Part 2
13	Action,Adventure,Animation	8.1	tt0892769	Dean DeBlois	How to Train Your Dragon
14	Action,Adventure,Sci-Fi	7.2	tt1392170	Gary Ross	The Hunger Games
15	Action,Adventure,Sci-Fi	8.1	tt1392190	George Miller	Mad Max: Fury Road
16	Action,Adventure,Fantasy	8.0	tt2488496	J.J. Abrams	Star Wars: Episode VII - The Force Awakens
17	Action,Adventure,Comedy	8.1	tt2015381	James Gunn	Guardians of the Galaxy
18	Action,Adventure,Sci-Fi	6.9	tt0458339	Joe Johnston	Captain America: The First Avenger
19	Action,Adventure,Sci-Fi	7.0	tt1228705	Jon Favreau	Iron Man 2

Data Analysis - Directors

Frequency of categories individually for these top directors



We found that **adventure**, **action** and **sci-fi** were the three most highly rated categories for **directors**

Statistical Inference

The relevance of runtime

Z-test

H_a: The mean of the number of votes when a movie has a runtime between 90 and 130 is significantly bigger than the mean of all the movies in our dataset

H₀: There is no significant difference between the sample mean and the population mean

P-value = 0.0096 **α** = 0.05

Movies with runtimes between 90 and 130 minutes have a high number of votes

Summary



Recommendations

- To make a movie with high earnings **the category that could have most impact on people will be Action** because it has the most number of votes
- **Select a director due to their history of highly rated movies with a high number of votes** in the category of either **Adventure, Action, or Sci-Fi**
- We found that movies **with runtimes between 90 and 130 minutes have a high number of votes**

Next Steps and Limitations

- Some datasets aren't updated and we couldn't match information amongst them. So since we don't have a recent collection of industry data we would work to update datasets
- We link datasets but sometimes we had to deal with missing data. We avoided working with columns with 30% or more missing data
- We can see the tendency of the genres you choose through time and which of them have the best relation to number of votes
- In further iterations we would have liked to compare profits with top directors with updated data

Thank You

Thank you for your attention, for any questions please contact us:

- mcruzlehardy@deloitte.com
- mgarciapuerto@deloitte.com
- amalagonsanchez@deloitte.com
- gyanezarvizu@deloitte.com
- daniejackson@deloitte.com