

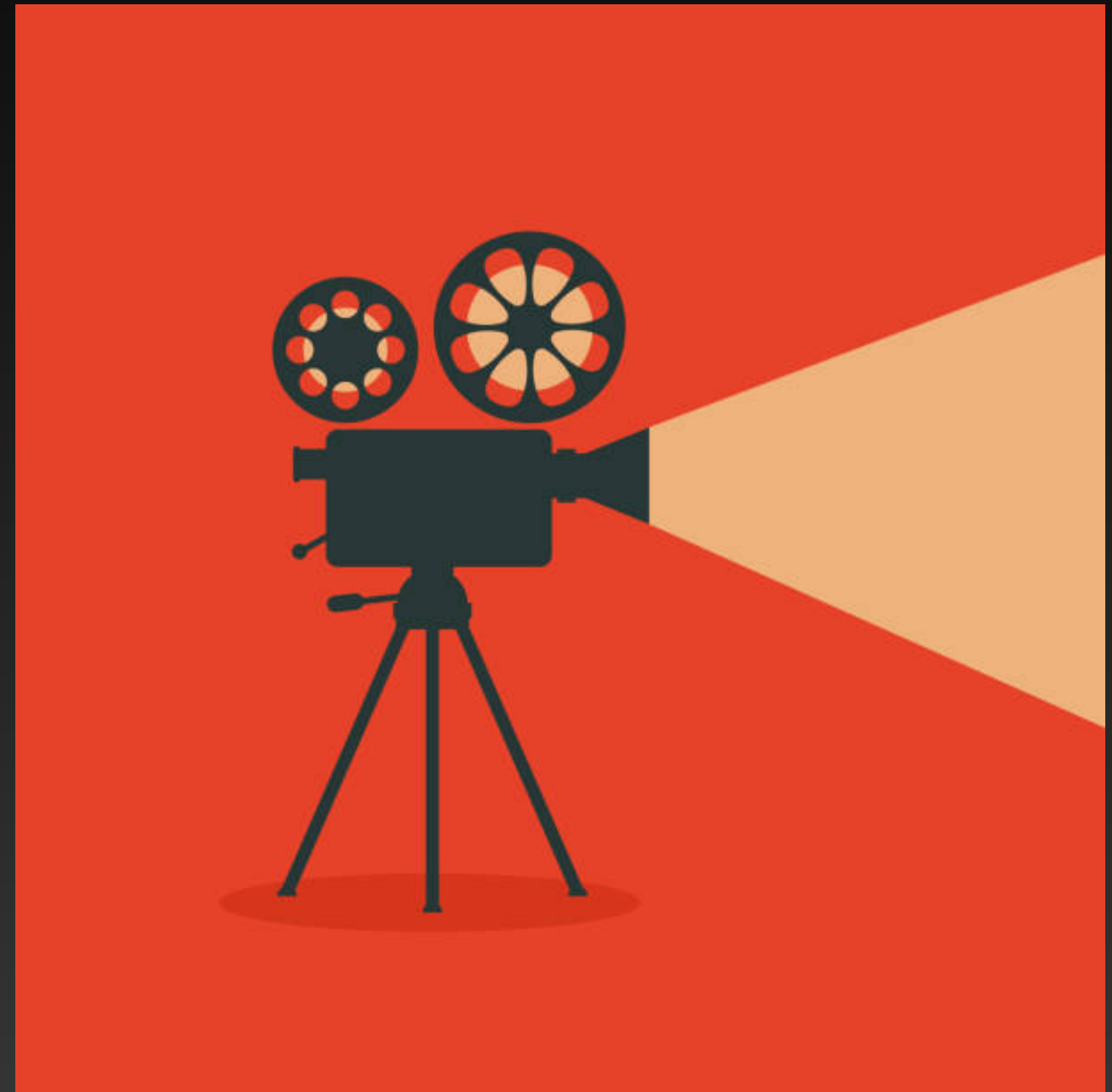
Movie Industry

Flatiron School

Kamile Yagci 07/01/2021

Content

- Overview
- Business Problem
- Data
- Method
- Analysis and Result
 - Profit Definition
 - Preliminary Filters
 - Business Questions 1, 2, 3, 4
- Conclusion
- Future Work
- Extra Slides



Overview

- This project is the analysis of the Movie Industry.
- The results of this study will advise Microsoft on making business decisions in movie making.

Project Details:

- <https://github.com/kamileyagci/dsc-phase-1-project>
- Blog:

Business Problem

In order to determine the most profitable approach in movie making, these four business questions are studied:

1. Is there a correlation between the production budget and the profit?
2. Is there a correlation between the ratings and the profit?
3. Which directors, writers, actors and actresses make the most profit?
4. Which genres make the most profit?

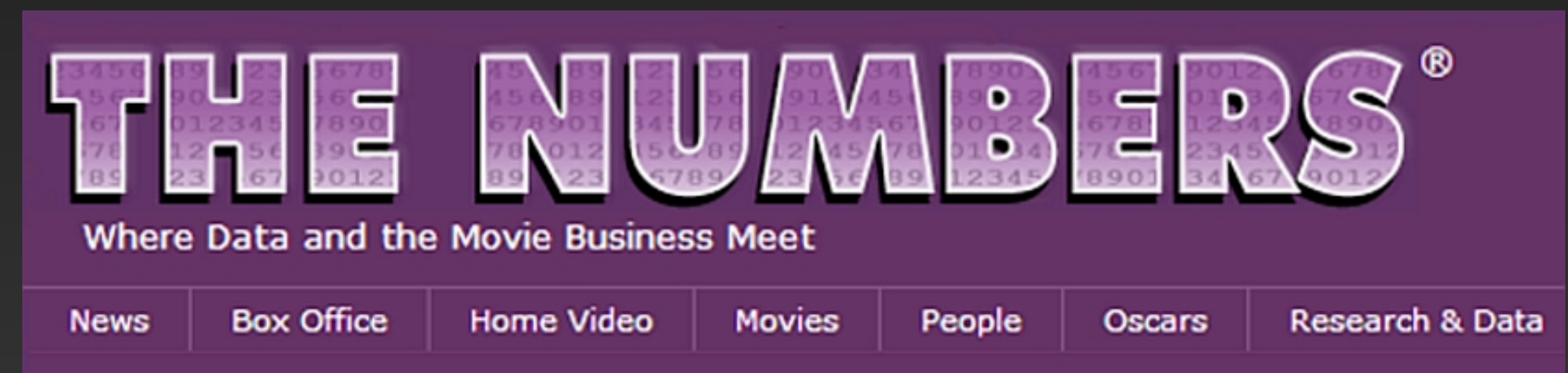
Data

Data Sources:

- IMDB
- The Numbers

Data files:

- tn.movie_budgets.csv.gz
- imdb.title.basics.csv.gz
- imdb.title.ratings.csv.gz
- imdb.title.principals.csv.gz
- imdb.name.basics.csv.gz



Profit Definition

Two types of profit is defined for this study:

- Profit:

$$\text{profit_gross} = \text{worldwide_gross} - \text{production_budget}$$

- Profit Rate:

$$\text{profit_rate} = \text{worldwide_gross} / \text{production_budget}$$

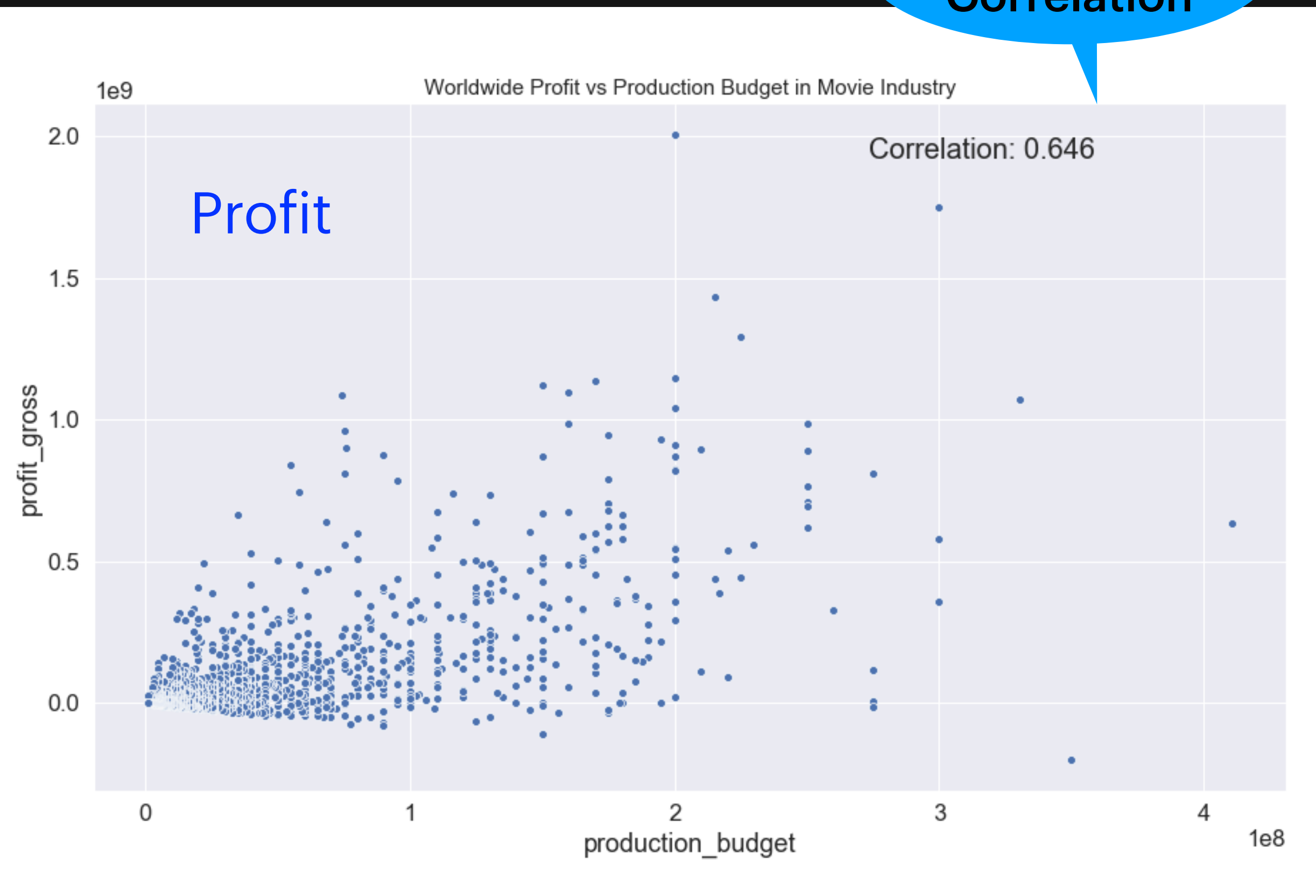
Preliminary Filters on Data

- Remove old movies with release year before 1990
- Remove outlier movies with
 - very low production budget ($< \$10,000,000$) and
 - very high profit rate (> 30).

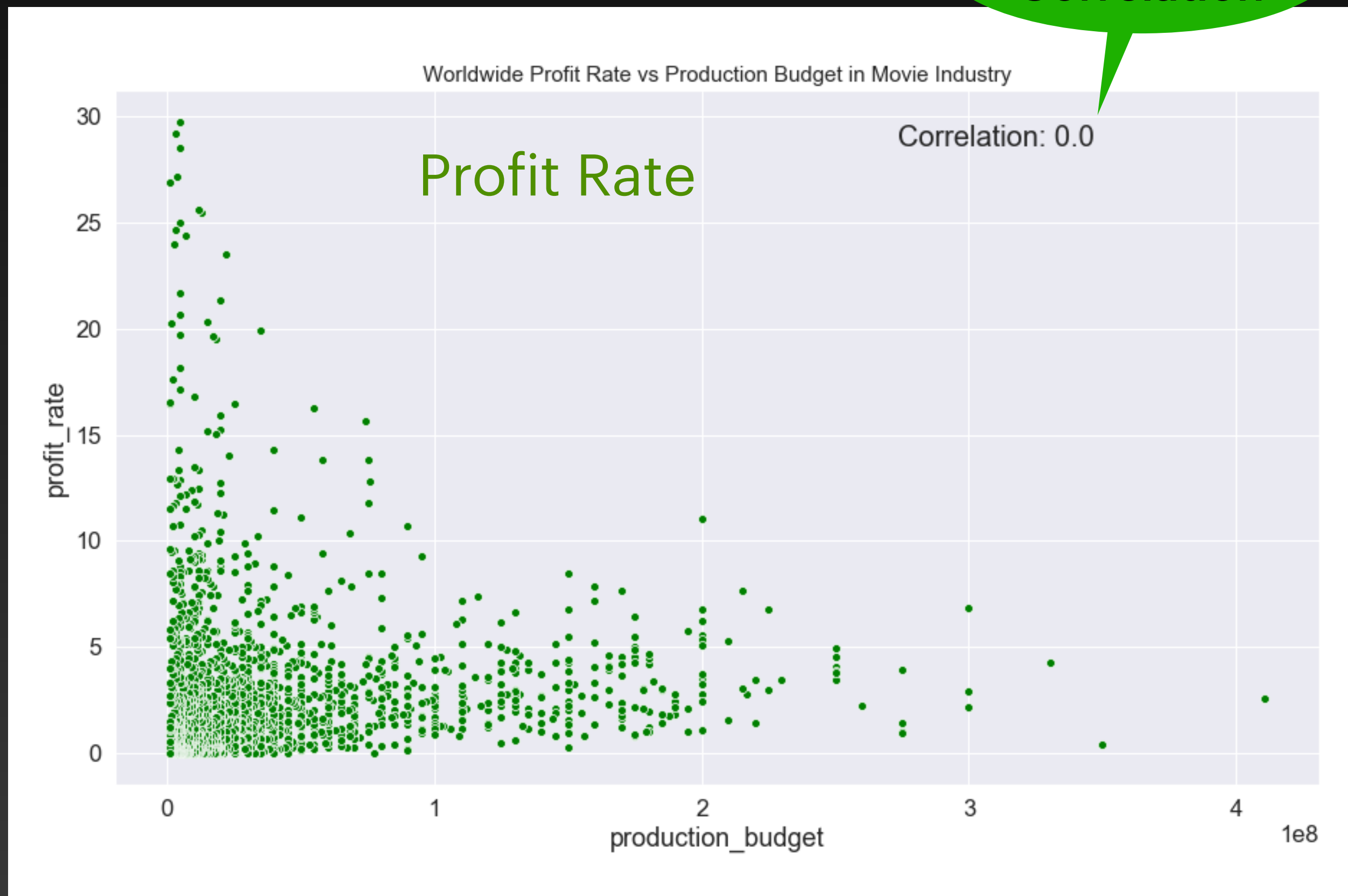
Business Question 1

Is there a correlation between the production budget and the profit?

High
Correlation



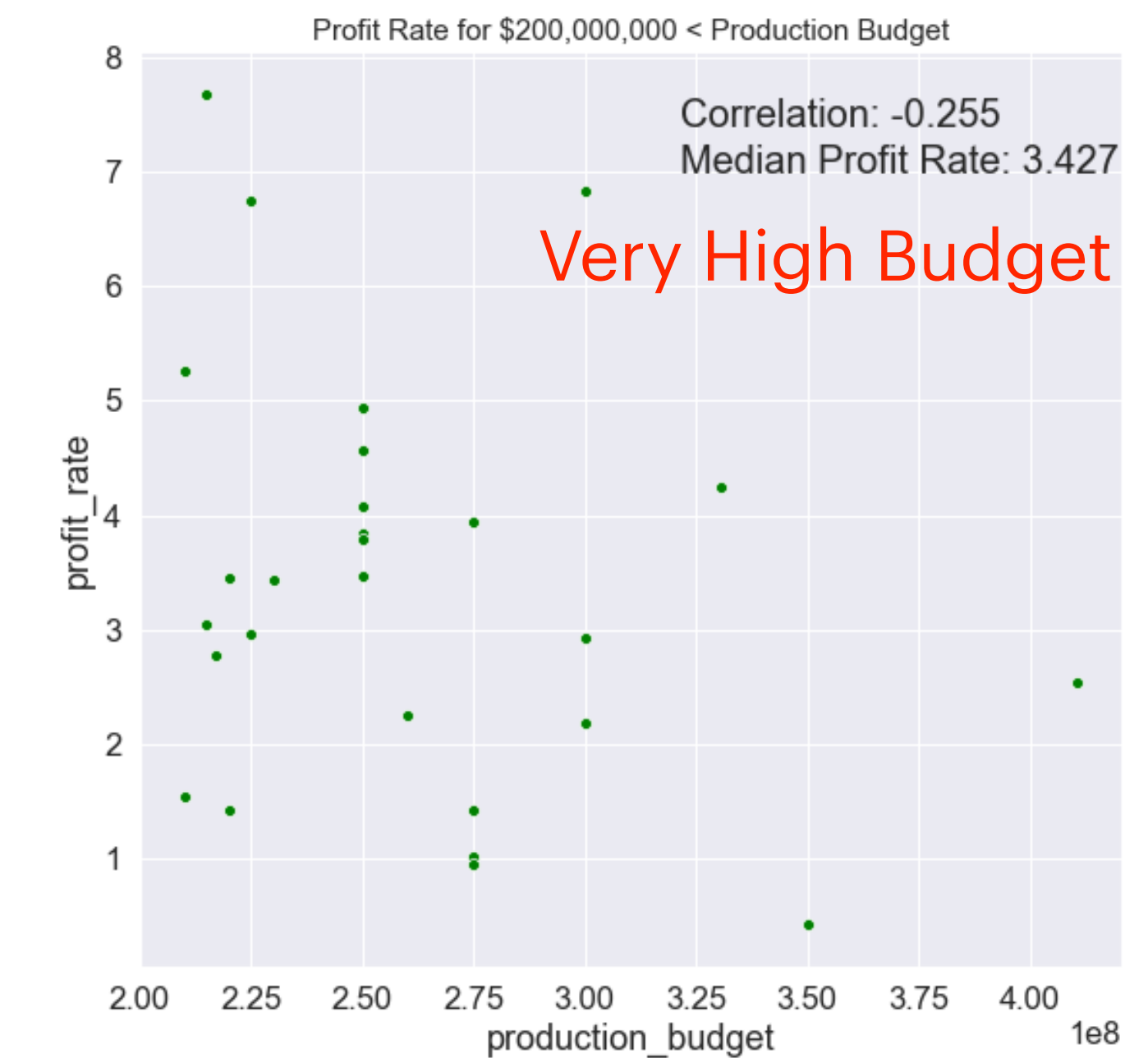
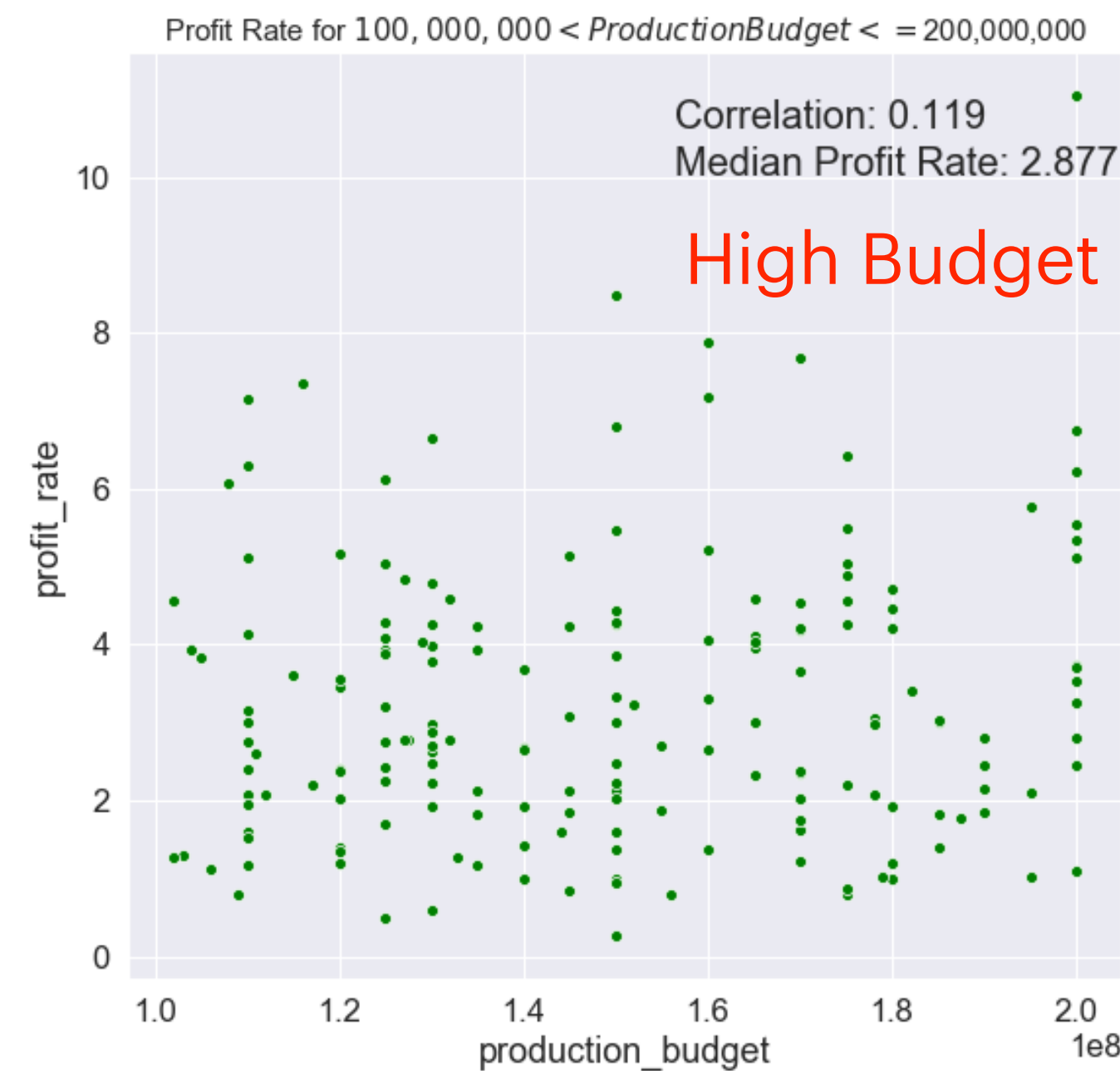
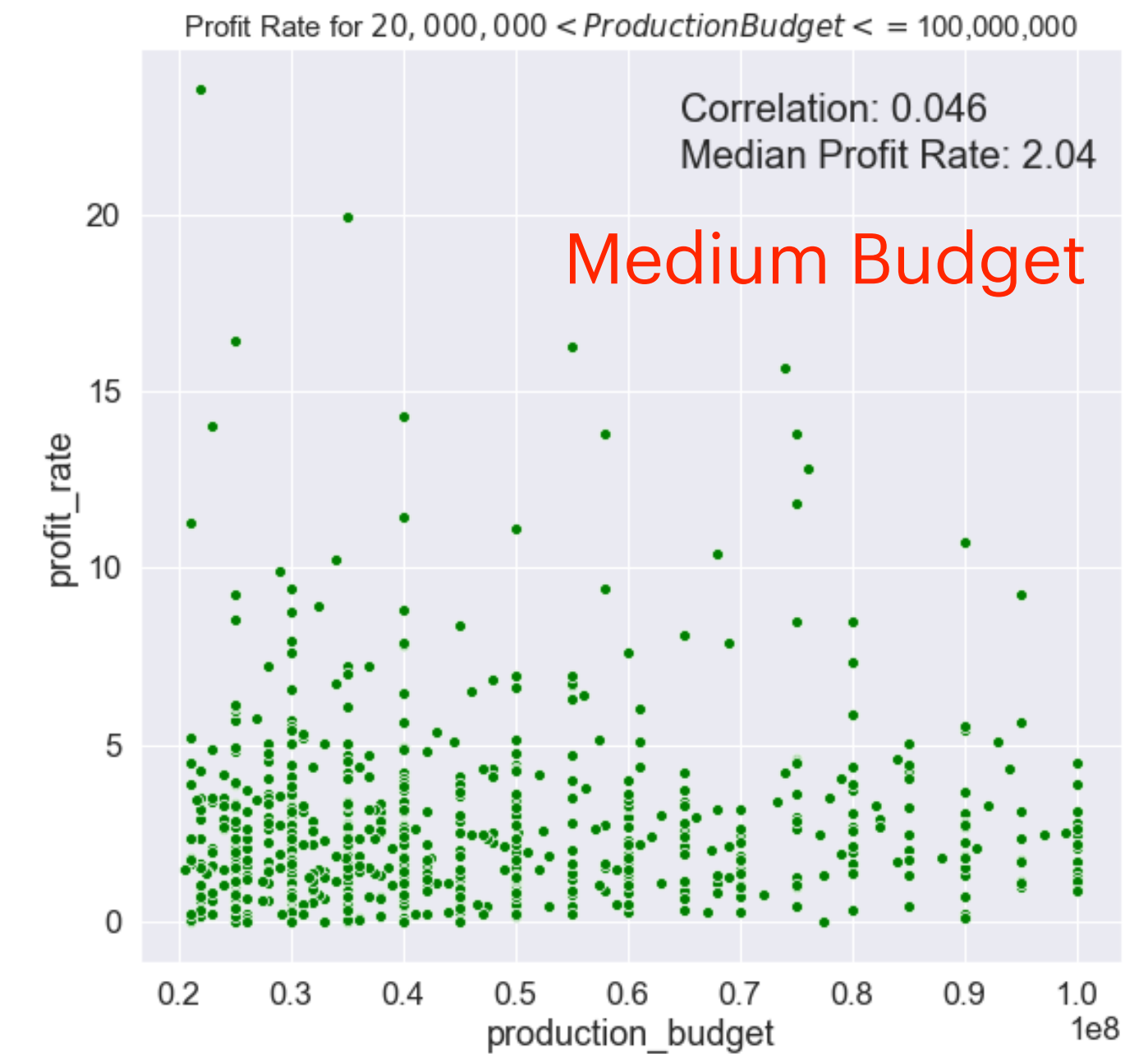
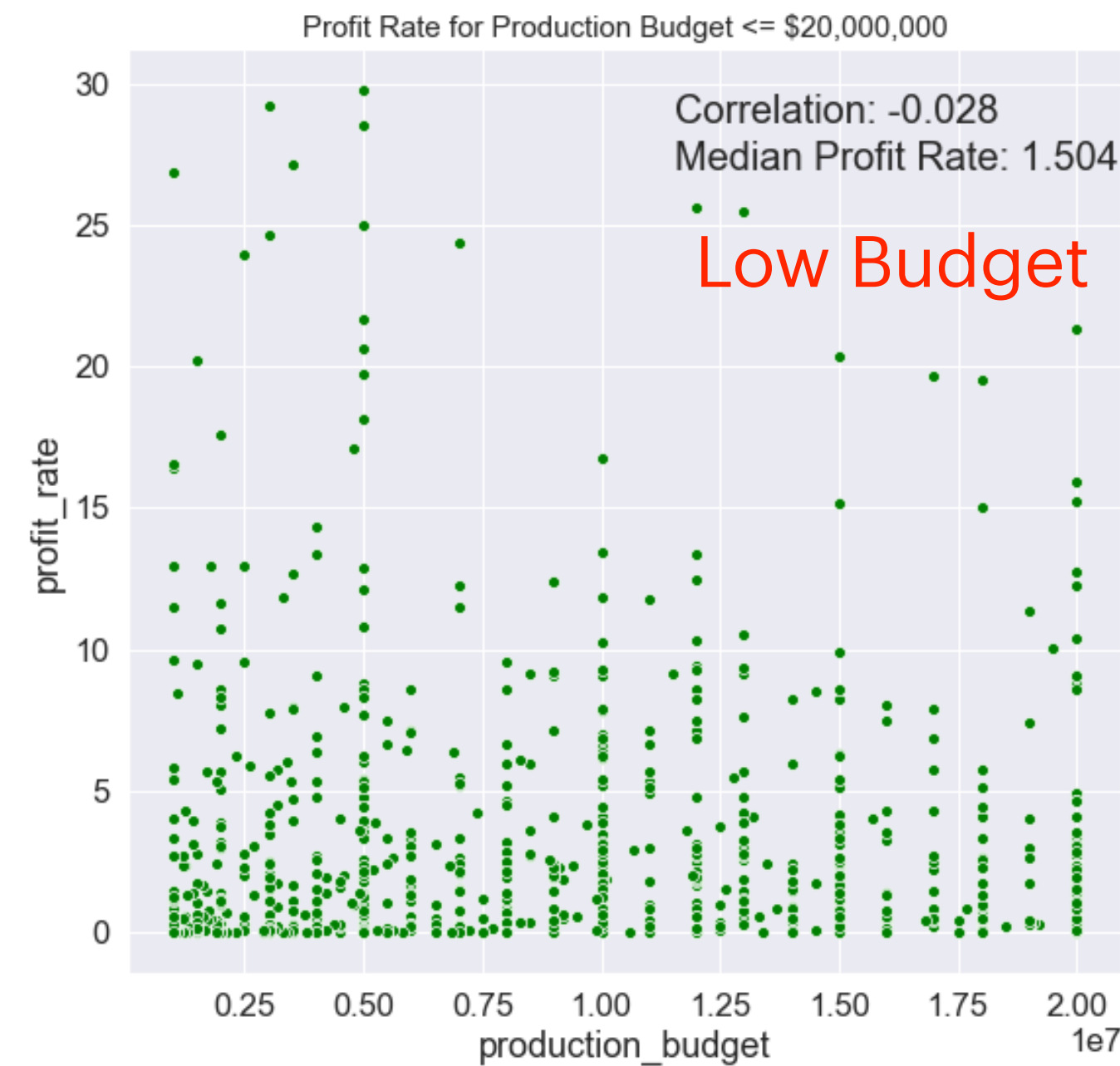
No
Correlation



Business Question 1 (cont'd)

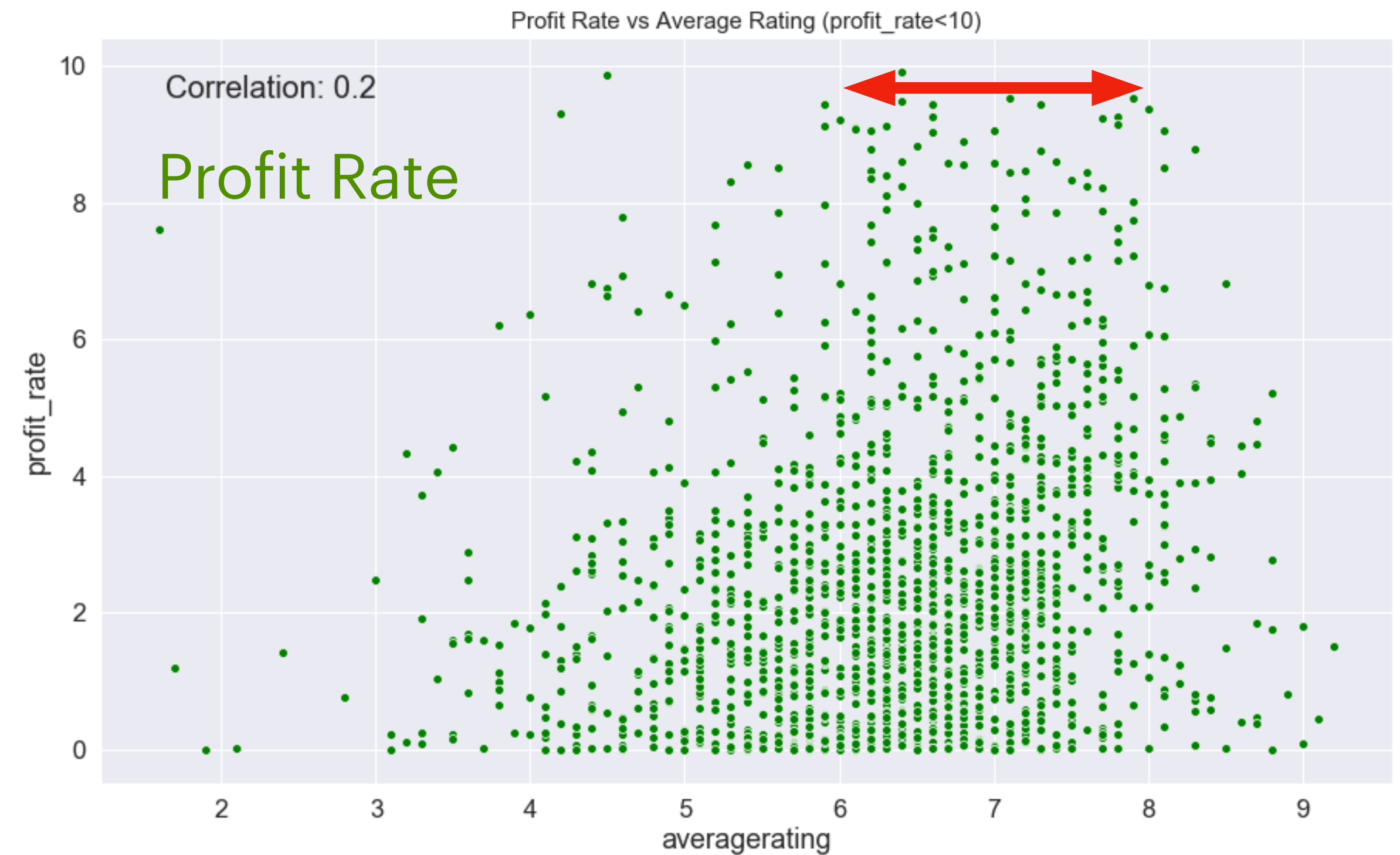
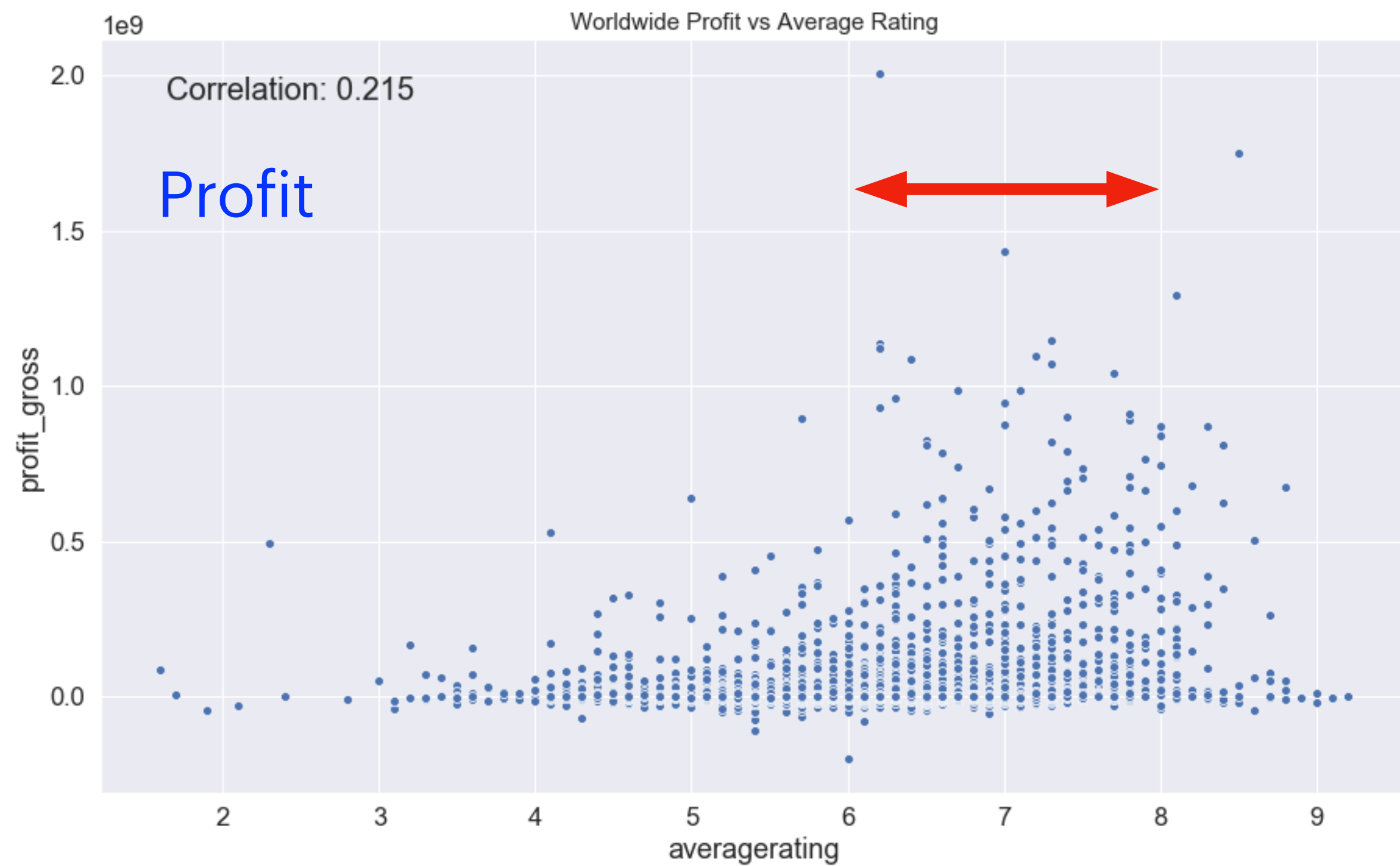
Investment Recommendations:

- Low budget: $\text{production_budget} \leq \$20,000,000$:
 - Not recommended.
- Medium budget: $\$20,000,000 < \text{production_budget} \leq \$100,000,000$:
 - Acceptable.
- High budget: $\$100,000,000 < \text{production_budget} \leq \$200,000,000$:
 - Highly recommended
- Very high budget: $\text{production_budget} > \$200,000,000$:
 - Risky



Business Question 2

Is there a correlation between the ratings and the profit?



- Good correlation up to rating 8
- Recommendation: Aim for making movies with average rating between 6 and 8

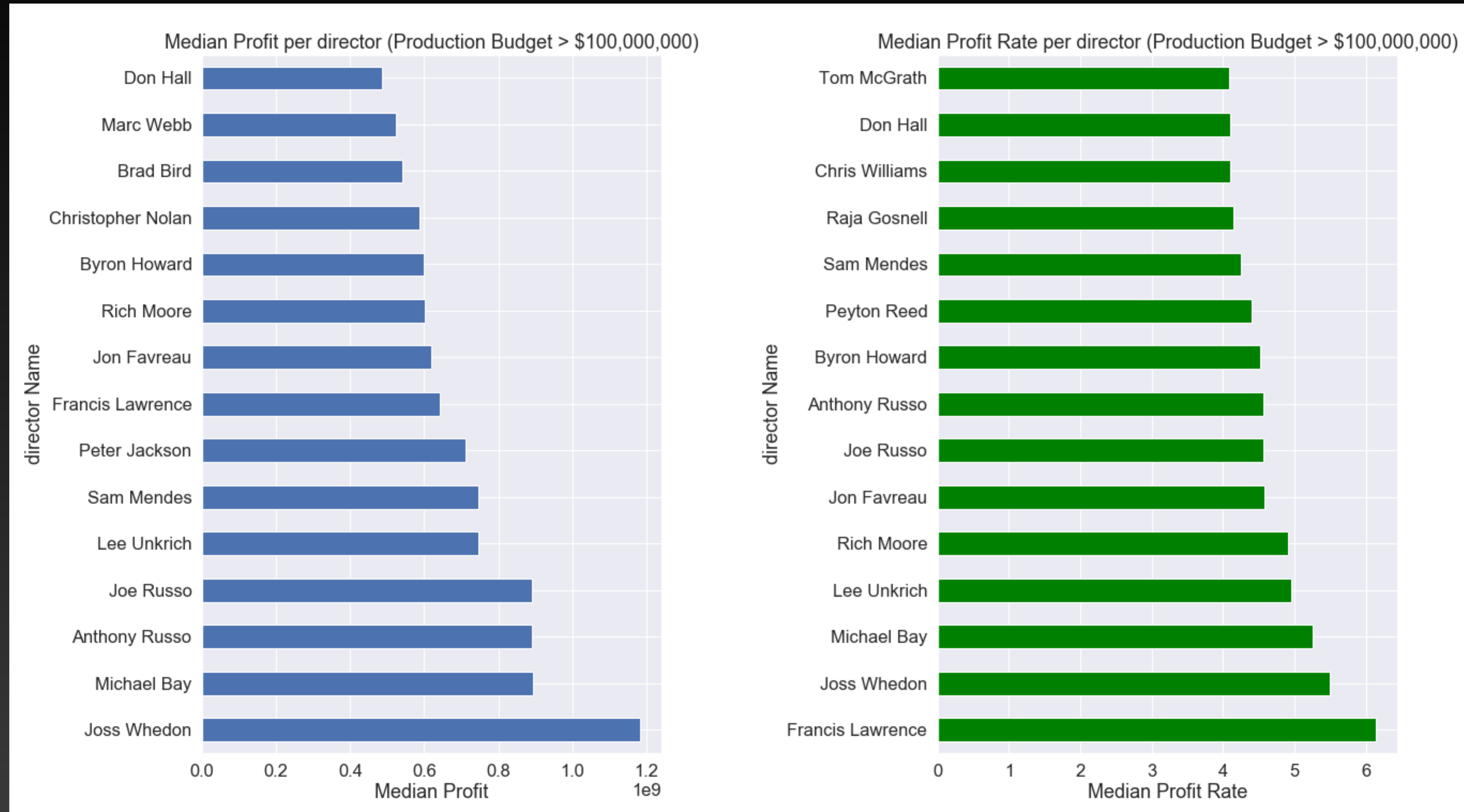
Business Question 3

Which directors, writers, actors and actresses make the most profit?

- Business Questions 3 and 4, are analyzed in two separate budget ranges:
 - High budget: $\text{production_budget} > \$100,000,000$
 - Medium budget: $\$20,000,000 < \text{production_budget} \leq \$100,000,000$
- The name of the deceased are removed from the data
- The median profit and median profit rate are calculated for each name

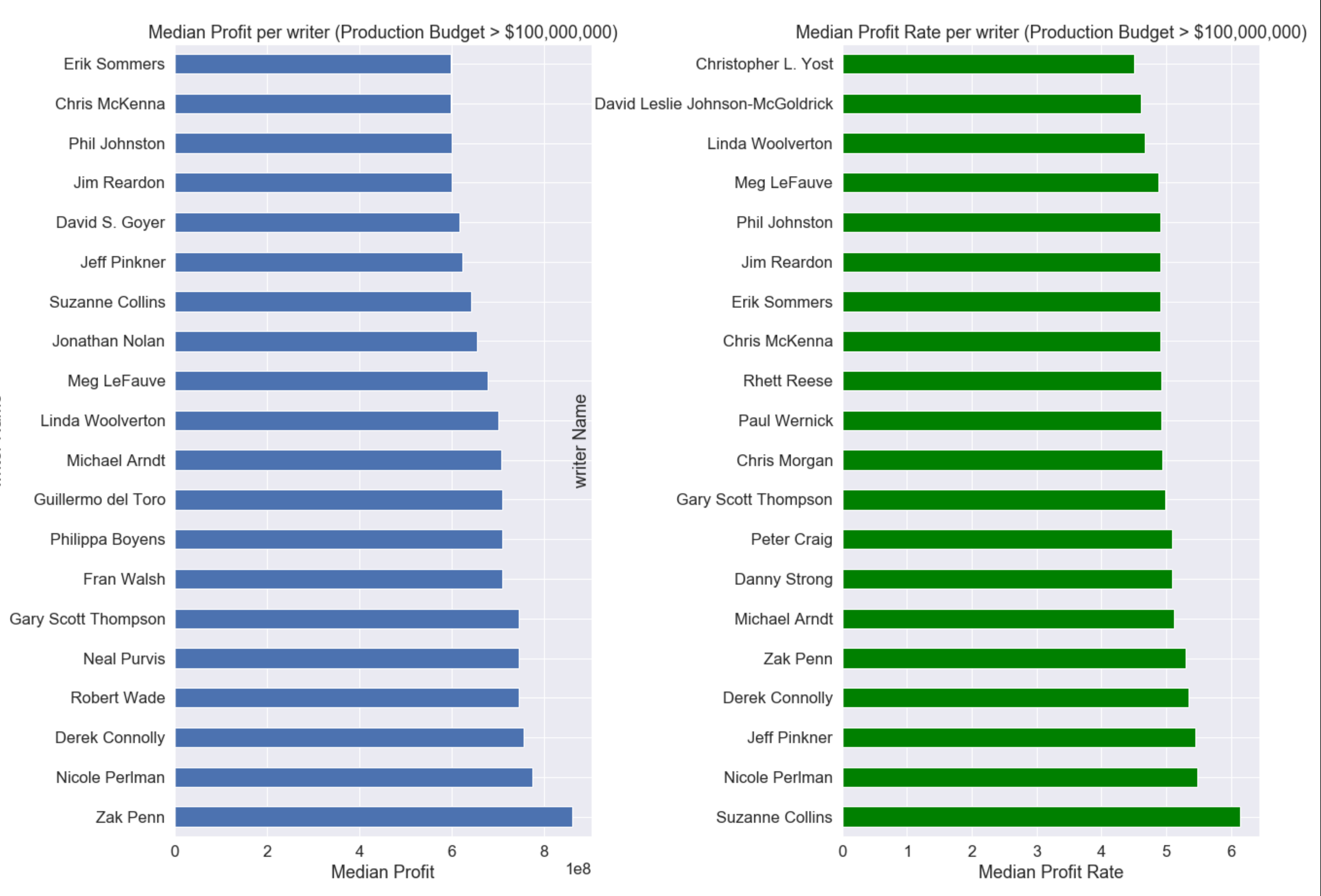
Business Question 3: Most Profitable Directors at High Budget

- 'Lee Unkrich',
- 'Michael Bay',
- 'Jon Favreau',
- 'Rich Moore',
- 'Anthony Russo',
- 'Sam Mendes',
- 'Joe Russo',
- 'Byron Howard',
- 'Joss Whedon',
- 'Francis Lawrence',
- 'Don Hall'



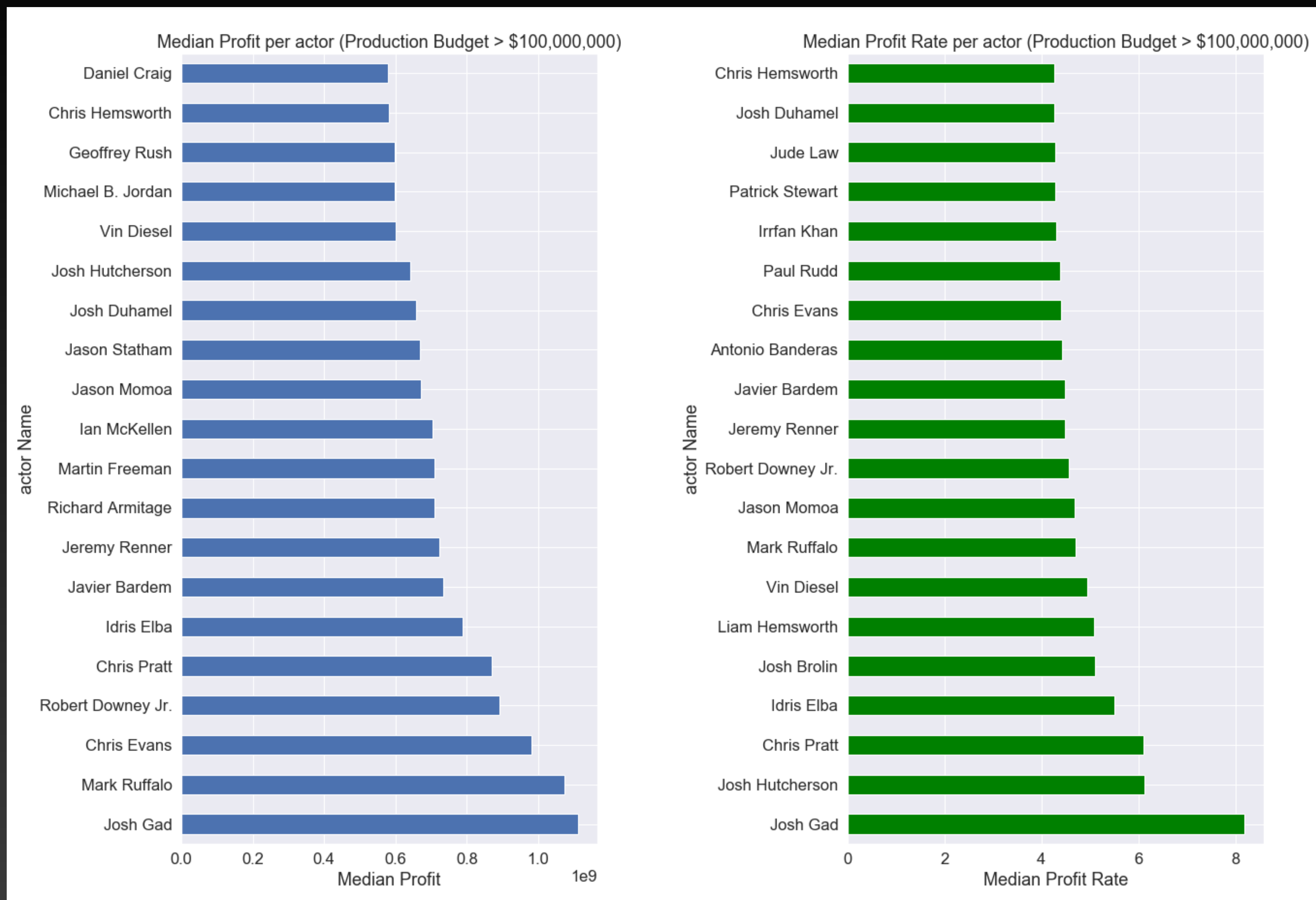
Business Question 3: Most Profitable **Writers** at High Budget

- 'Linda Woolverton',
- 'Phil Johnston',
- 'Gary Scott Thompson',
- 'Meg LeFauve',
- 'Jim Reardon',
- 'Nicole Perlman',
- 'Jeff Pinkner',
- 'Erik Sommers',
- 'Chris McKenna',
- 'Derek Connolly',
- 'Zak Penn',
- 'Michael Arndt'



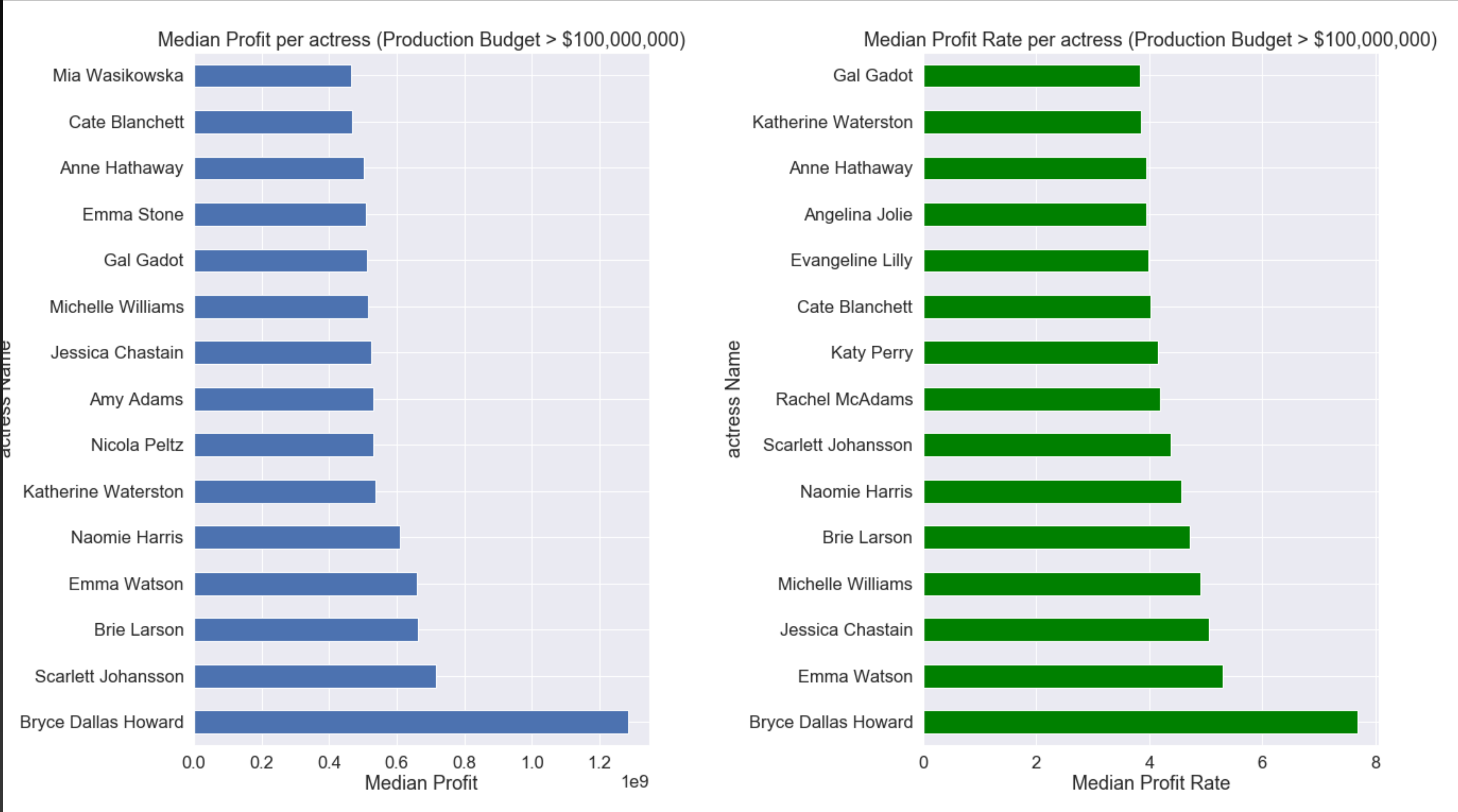
Business Question 3: Most Profitable **Actors** at High Budget

- 'Josh Duhamel',
- 'Vin Diesel',
- 'Mark Ruffalo',
- 'Chris Hemsworth',
- 'Josh Gad',
- 'Robert Downey Jr.',
- 'Javier Bardem',
- 'Chris Evans',
- 'Idris Elba',
- 'Josh Hutcherson',
- 'Chris Pratt',
- 'Jason Momoa',
- 'Jeremy Renner'



Business Question 3: Most Profitable **Actresses** at High Budget

- 'Katherine Waterston',
- 'Emma Watson',
- 'Gal Gadot',
- 'Scarlett Johansson',
- 'Anne Hathaway',
- 'Jessica Chastain',
- 'Naomie Harris',
- 'Bryce Dallas Howard',
- 'Michelle Williams',
- 'Cate Blanchett',
- 'Brie Larson'

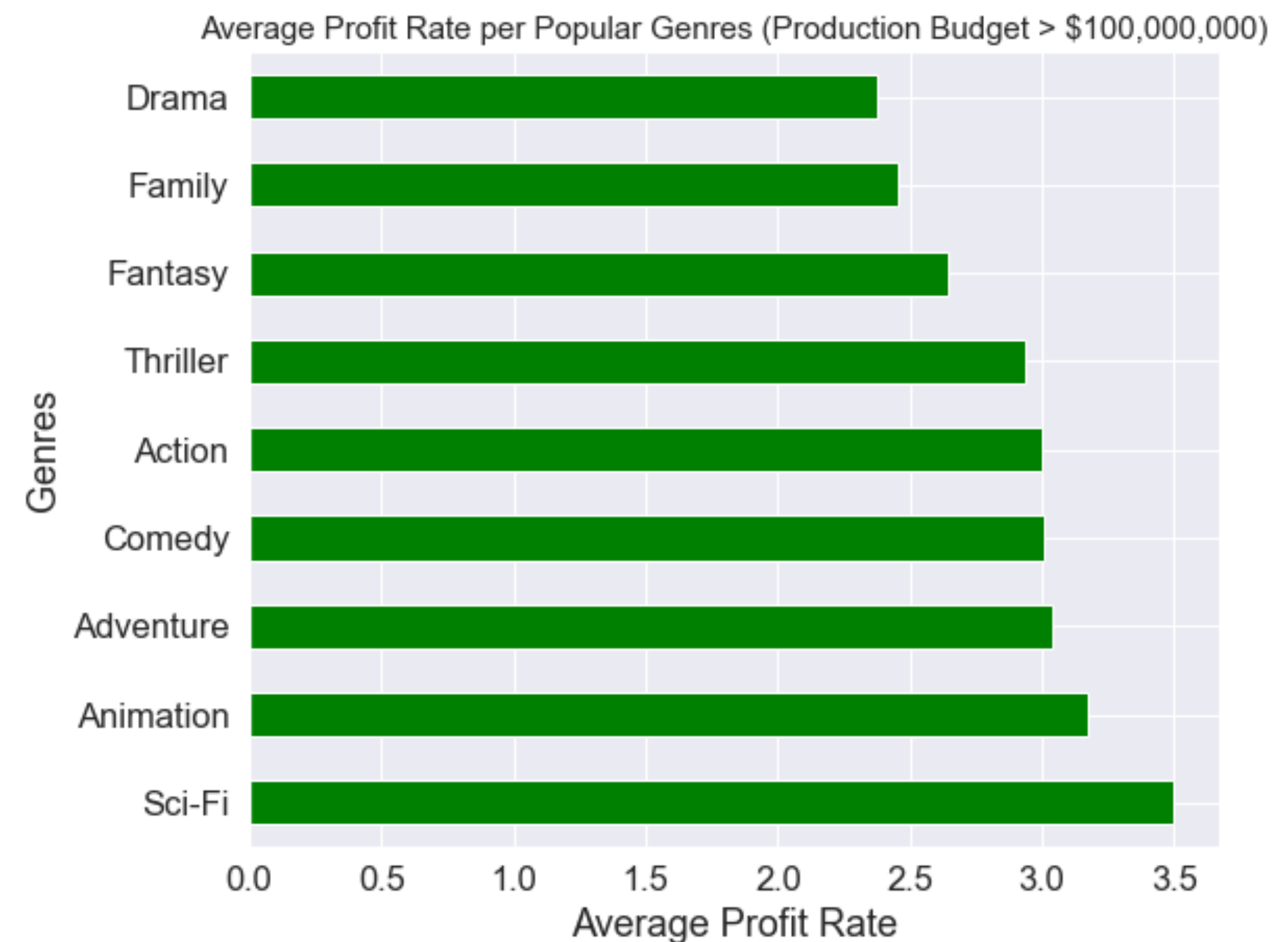
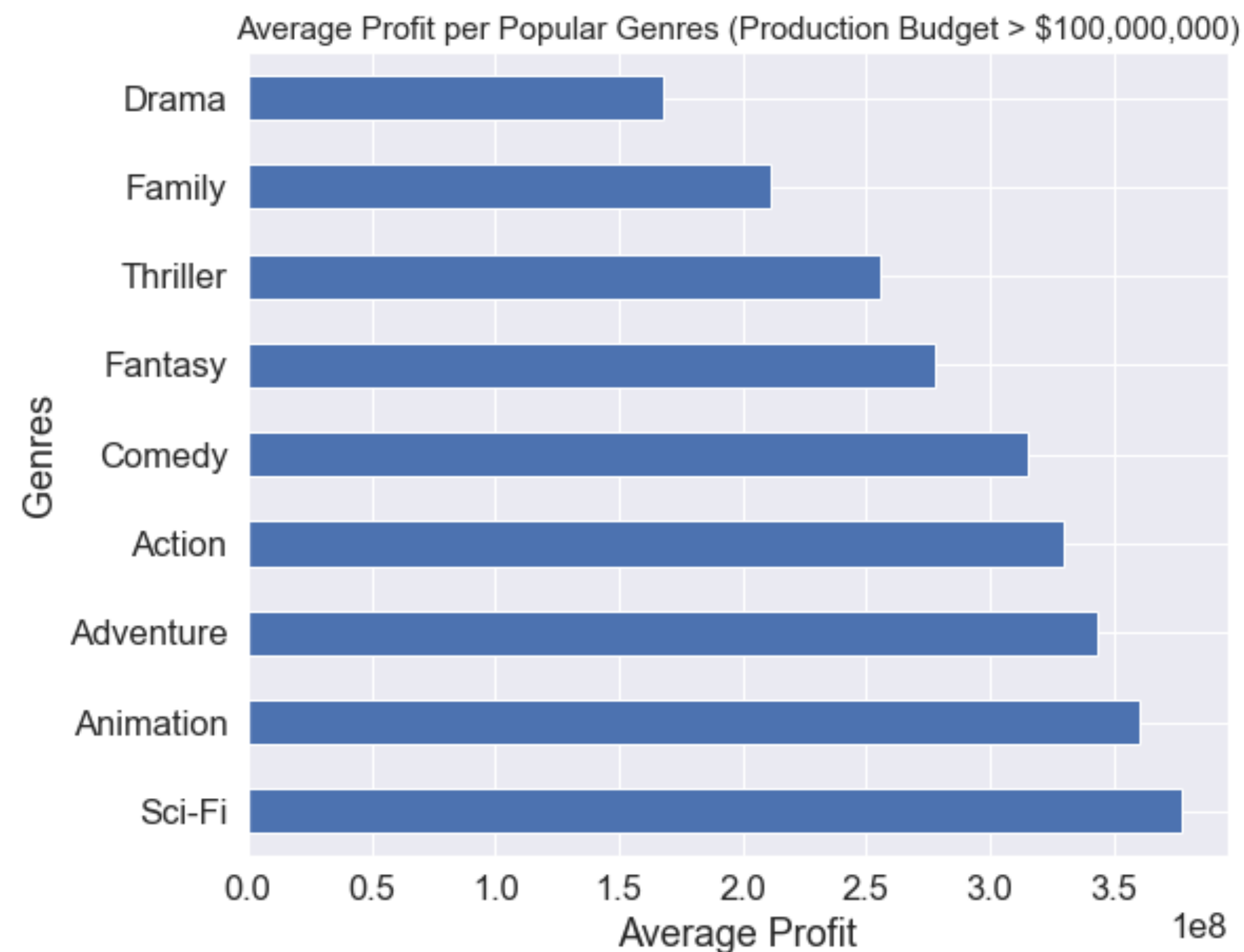


Business Question 4

Which genres make the most profit?

Most profitable genres for High Budget:

- Sci-Fi, Animation, Adventure, Comedy, Action



Business Question 3 & 4 at Medium Budget

- Profitable directors, writers, actors and actresses at medium budget:
 - Directors: 'Chris Renaud', 'Christopher Miller', 'Genndy Tartakovsky', 'Kyle Balda', 'Pierre Coffin', 'James Foley', 'Steve Martino', 'Jake Kasdan', 'Phil Lord', 'Paul Feig', 'Olivier Megaton'
 - Writers: 'Ken Daurio', 'Stephenie Meyer', 'Kevin Hageman', 'Dan Hageman', 'Erik Sommers', 'Cinco Paul', 'Wellesley Wild', 'Anthony McCarten', 'Brian Lynch', 'E.L. James', 'Robert Mark Kamen', 'Craig Mazin', 'Alec Sulkin', 'Billy Ray', 'Chris McKenna', 'Niall Leonard'
 - Actors: 'Eric Johnson', 'Jamie Dornan', 'Michael Keaton', 'Taylor Lautner', 'Eric Stonestreet', 'Aziz Ansari', 'Neil Patrick Harris', 'Bradley Cooper', 'Thomas Brodie-Sangster', 'Dylan O'Brien', 'Seth MacFarlane', 'Andy Samberg', 'Benjamin Bratt', 'Marwan Kenzari', 'Ed Skrein', 'Denis Leary'
 - Actresses: 'Ali Larter', 'Hailee Steinfeld', 'Jennifer Lawrence', 'Dakota Johnson', 'Eloise Mumford', 'Sandra Bullock', 'Jennifer Ehle', 'Kristen Wiig', 'Miranda Cosgrove', 'Famke Janssen', 'Selena Gomez'
- Profitable genres at medium budget:
 - Animation, Music, Family, Comedy, Adventure

Conclusion

Business recommendations in movie making for the most profit:

- Invest on movies with high production budget between \$100,000,000 and \$200,000,000.
- The movies with medium production budget, \$20,000,000 and \$100,000,000, are acceptable for investment; but not as profitable as the high budget movies.
- Aim for making movies with average rating between 6 and 8.
- Hire the listed directors, writers, actors and actresses for the movies in the recommended budget range. See Results for Question 3.
- At high budget, invest on Sci-Fi, Animation, Adventure, Action, Comedy genres
- At medium budget, invest on Animation, Music, Family, Comedy, Adventure genres

Future Work

- Repeat the analysis in more budget ranges: \$10,000,000 - \$50,000,000 - \$100,000 - \$200,000,00 - up.
- Study the most profitable directors, actors, actresses and actors for each genre separately.
- This study used the data of the movies after 1990. Repeat the study for the movies after 2000, even though it will cause the reduction of the entries and increase the standard deviation in the results.

Questions?

Extra Slides

Method

- Explore Data
- Determine the Business Questions
- Define Profit variables
- Merge, Clean and Save Data
- Analyze Business Question 1
- Analyze Business Question 2
- Analyze Business Question 3
- Analyze Business Question 4

Merge, Clean and Save Data

- Merge Data:
 - The data from 5 files are merged into one Data Frame
- Clean Data:
 - Remove the unnecessary columns
 - Remove rows with \$0 worldwide_gross
 - Change the currency columns to float
 - Separate genres
 - Slice the release year from the release date
- Save Data:
 - Merged and cleaned data Frame is saved into a new CSV file. This is used for analysis.
 - This step is done after profit variables added to the data Frame