

SI649 Fall 2016 – Lab 1-answer-sheet for group project

Name: Akio Kakishima

Unique Name: akiokaki

Lab Partner's Name:

- Please answering the following questions with **screenshots** and descriptions (Yes, you don't need to repeat what we have done in the class here, just answer questions included here, e.g. 3.3, 5.2)
- When taking screenshots, please include the ENTIRE tableau window.
- Bonus tasks are optional and give you extra credit.
- When answering question 10, 11 and 12, please document steps that you used to generate the visualization. Also describe your solutions in words.

You will upload the following items to the canvas:

- ☐ the pdf version of the answer sheet
- ☐ the pdf version of the individual project

-----Group Task-----

Question 3: Which is the most profitable movie?

- ☐ 3.3 Which has the largest difference between worldwide gross and U.S. gross?

Movie name: Avatar at \$2.5 billion

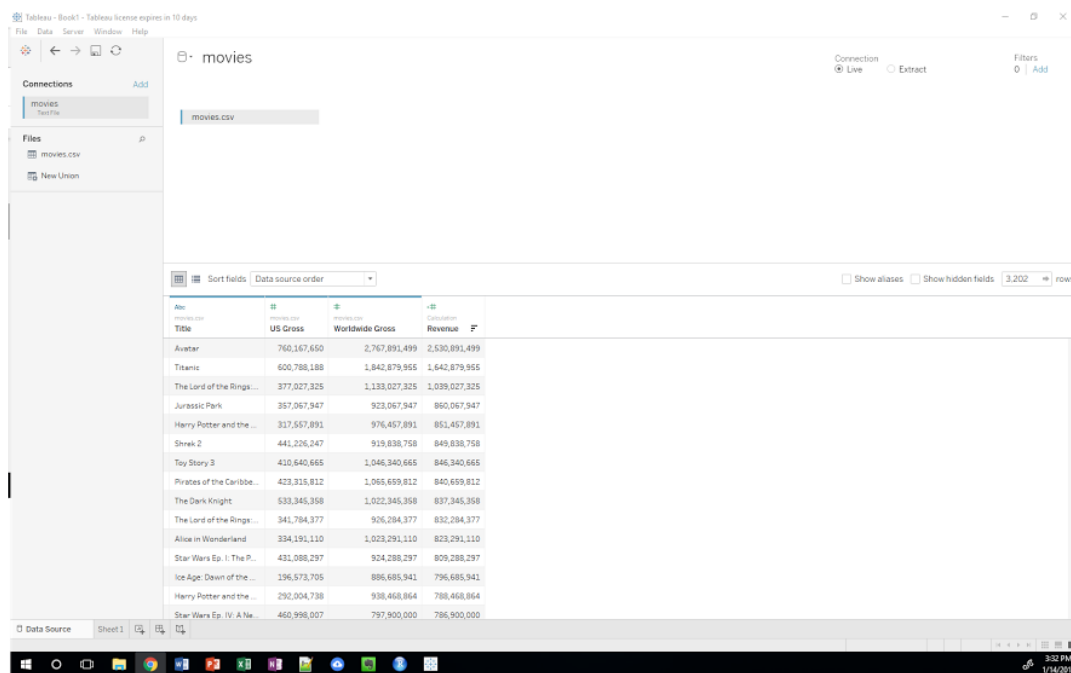


Tableau - Book1 - Tableau license expires in 10 days

Connections: movies (Text File)

Files: movies.csv, New Union

Sort fields: Data source order

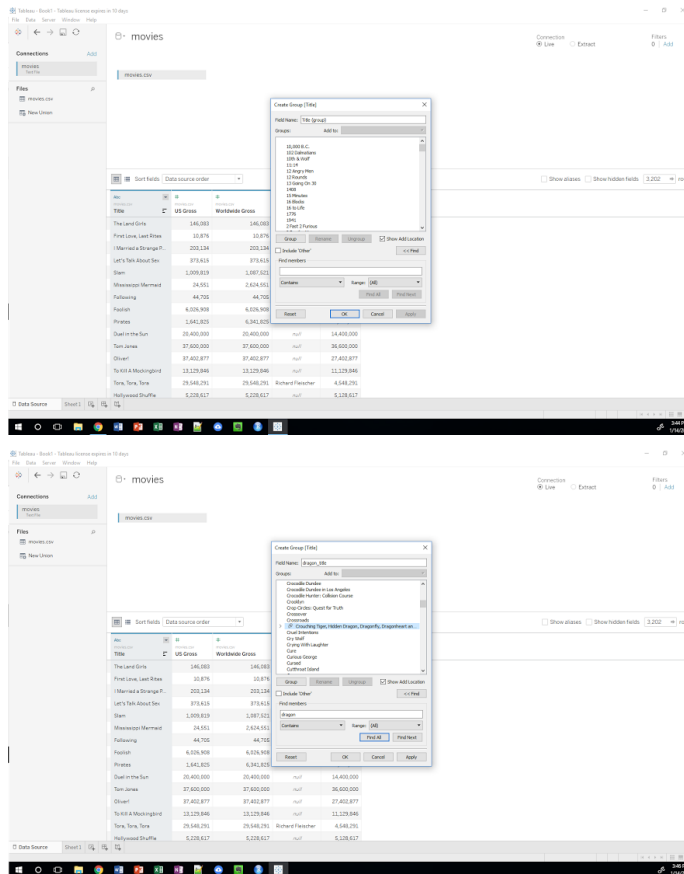
3,202 rows

Row	Title	US Gross	Worldwide Gross	Revenue
1	Avatar	760,167,650	2,767,891,499	2,530,891,499
2	Titanic	600,788,188	1,842,879,955	1,642,879,955
3	The Lord of the Rings: The Two Towers	377,027,325	1,133,027,325	1,039,027,325
4	Jurassic Park	357,067,947	923,067,947	860,067,947
5	Harry Potter and the Chamber of Secrets	317,557,891	976,457,891	851,457,891
6	Shrek 2	441,226,247	919,838,758	849,838,758
7	Toy Story 3	410,640,665	1,046,340,665	846,340,665
8	Pirates of the Caribbean: The Curse of the Black Pearl	423,315,812	1,065,659,812	840,659,812
9	The Dark Knight	533,345,358	1,022,345,358	837,345,358
10	The Lord of the Rings: The Fellowship Ring	341,784,377	906,284,377	832,284,377
11	Alice in Wonderland	334,191,110	1,023,291,110	823,291,110
12	Star Wars Ep. I: The Phantom Menace	431,088,297	904,288,297	809,288,297
13	Ice Age: Dawn of the Dinosaurs	196,573,705	886,685,941	796,685,941
14	Harry Potter and the Prisoner of Azkaban	292,004,738	938,468,864	788,468,864
15	Star Wars Ep. IV: A New Hope	460,998,007	797,900,000	786,900,000

Question 5: How do I find all movies that have contain specific string in their names?

☐ 5.2 Create a group with all titles that contains the string “dragon”

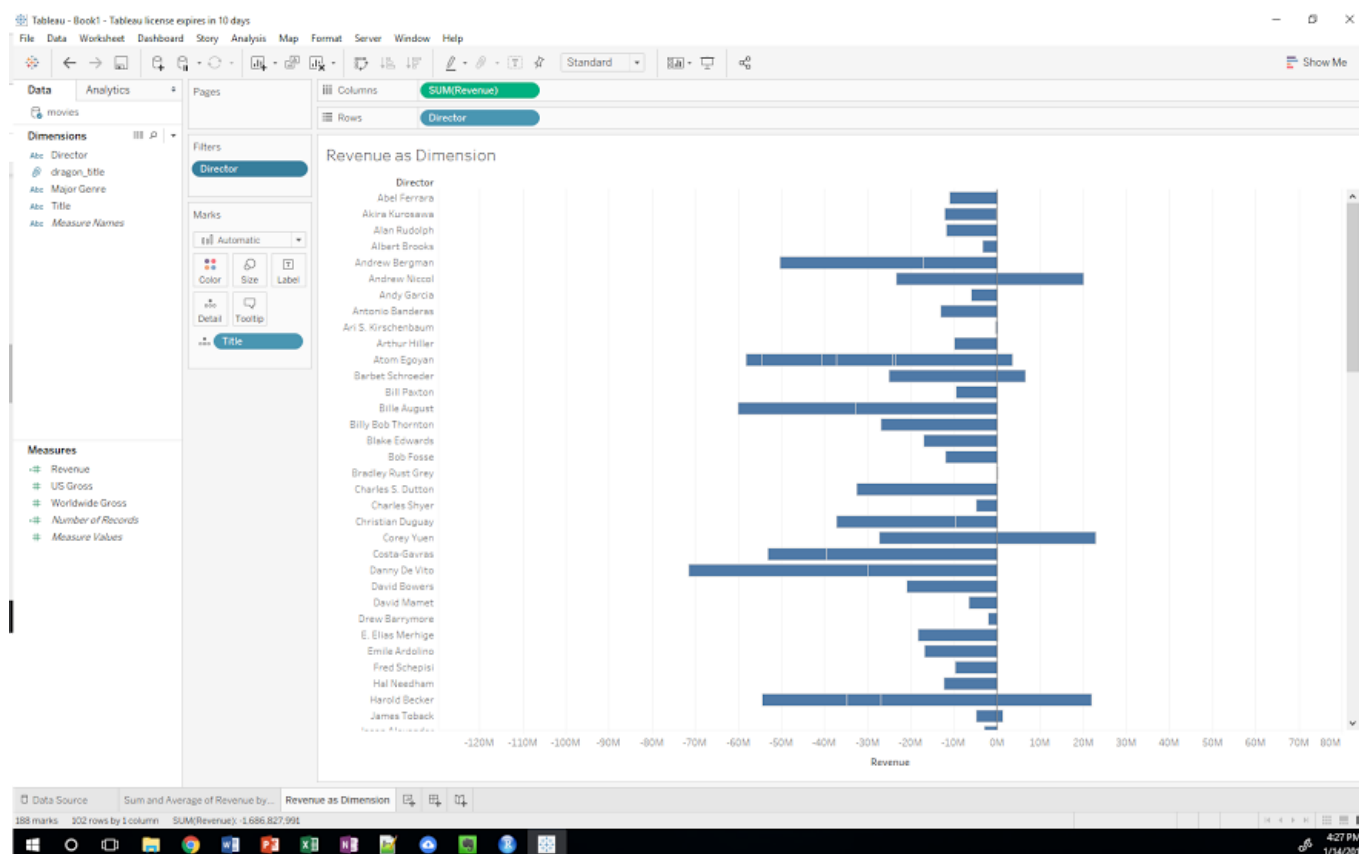
Paste screenshot here of you creating the group (we don't need the full list):



Question 7: Is there any director who has never directed any profitable movie?

☐ 7.6 Are there directors who have never directed profitable movies? Who are they? (Hint1 : you can use filters to complete this task. Hint2: never directed profitable movie =all movies directed have revenues less than zero) (You don't need to write down the whole list, just show us how to find them.)

Screenshot:



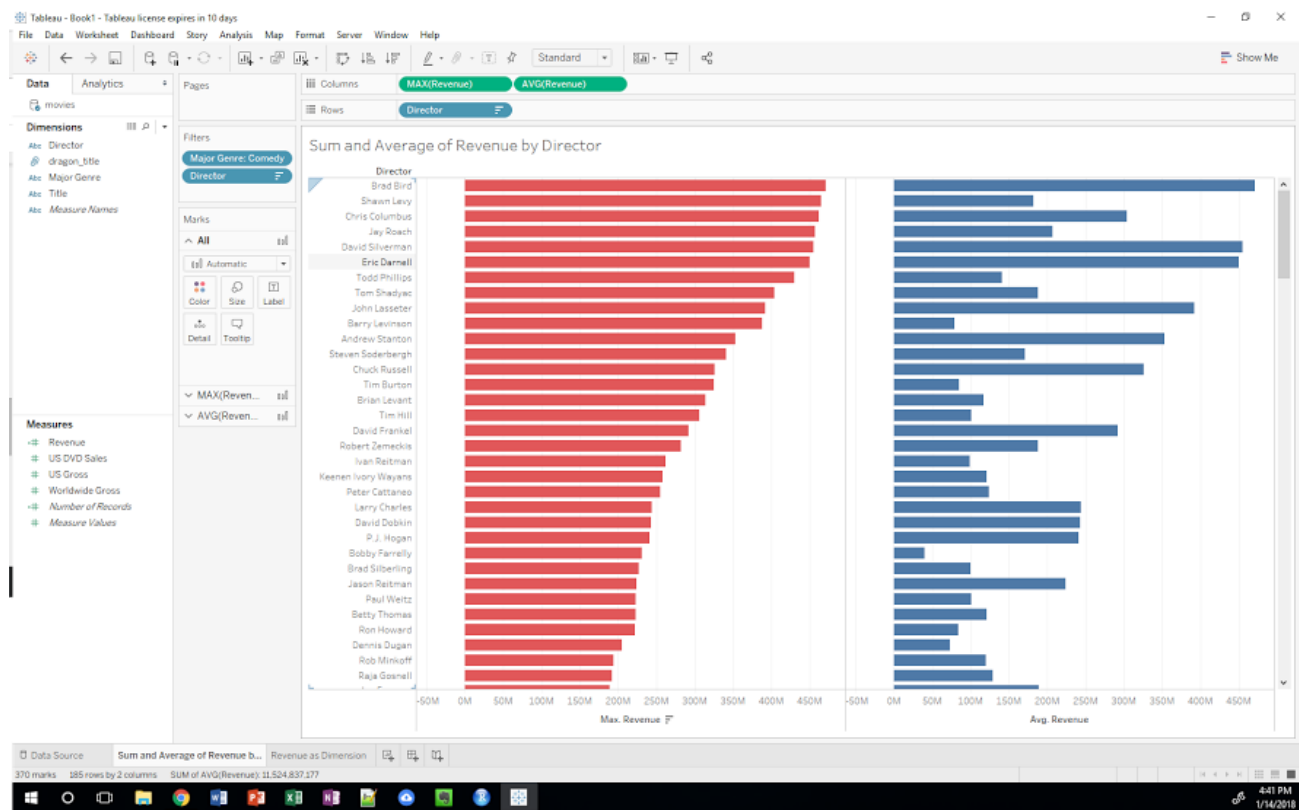
Answer: Put Sum of Revenue as column and director as row, and filter by director using a conditional where Sum(Revenue) is ≤ 0 . Find the directors where the bars are only available to the left of the 0 line.

Question 8: Can I get more information about a specific genre?

☐ 8.3 Which director has directed the most profitable Comedy movie? (Hint: use

Columns: MAX(Revenue) and filters/pages shelf)

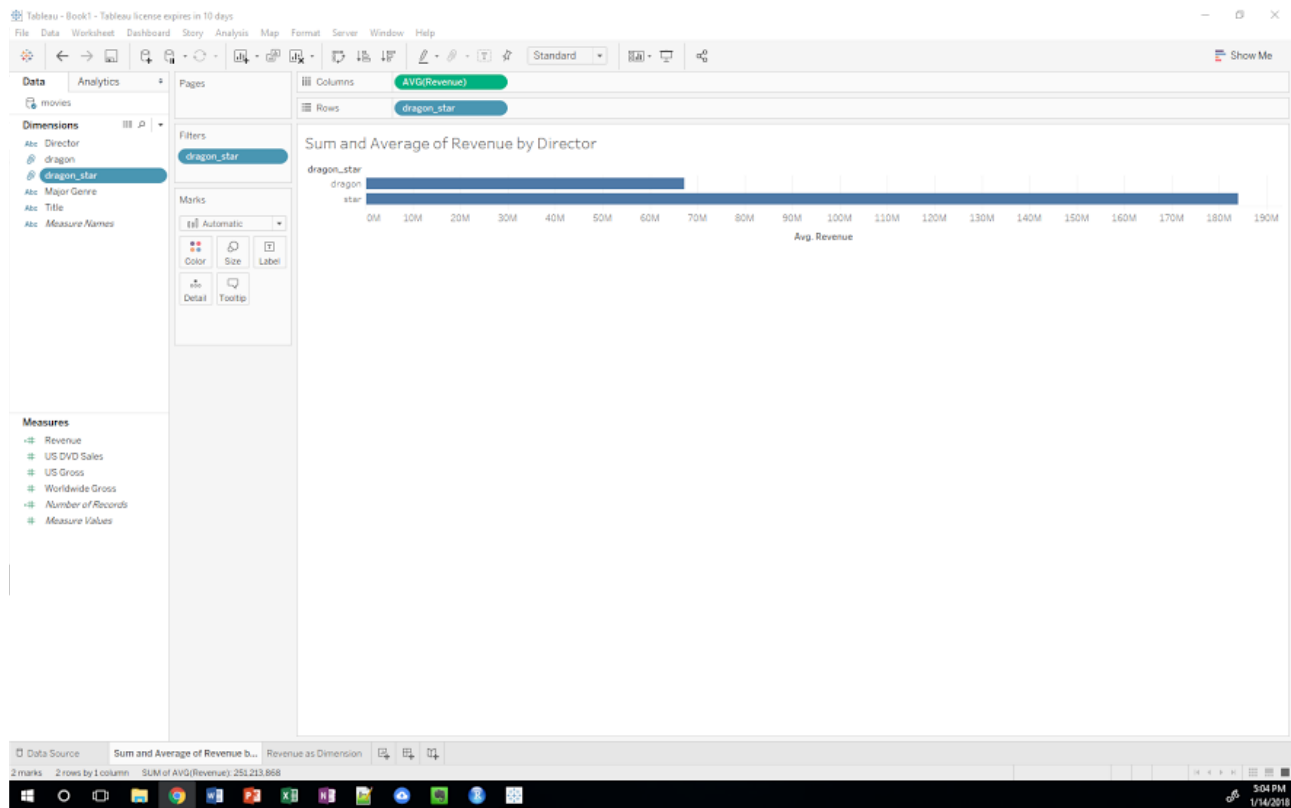
Paste your screenshot:



Director name: Brad Bird

☐ 8.4 Bonus Question: do movies that contain the name “star” have higher average revenue than movies contain the name “dragon”?

Screenshot of visualization:



Answer: “Star” seems to earn more on average than “dragon”

Question 11: Is it better to invest in action movies or in comedy movies? Which director would you invest in?

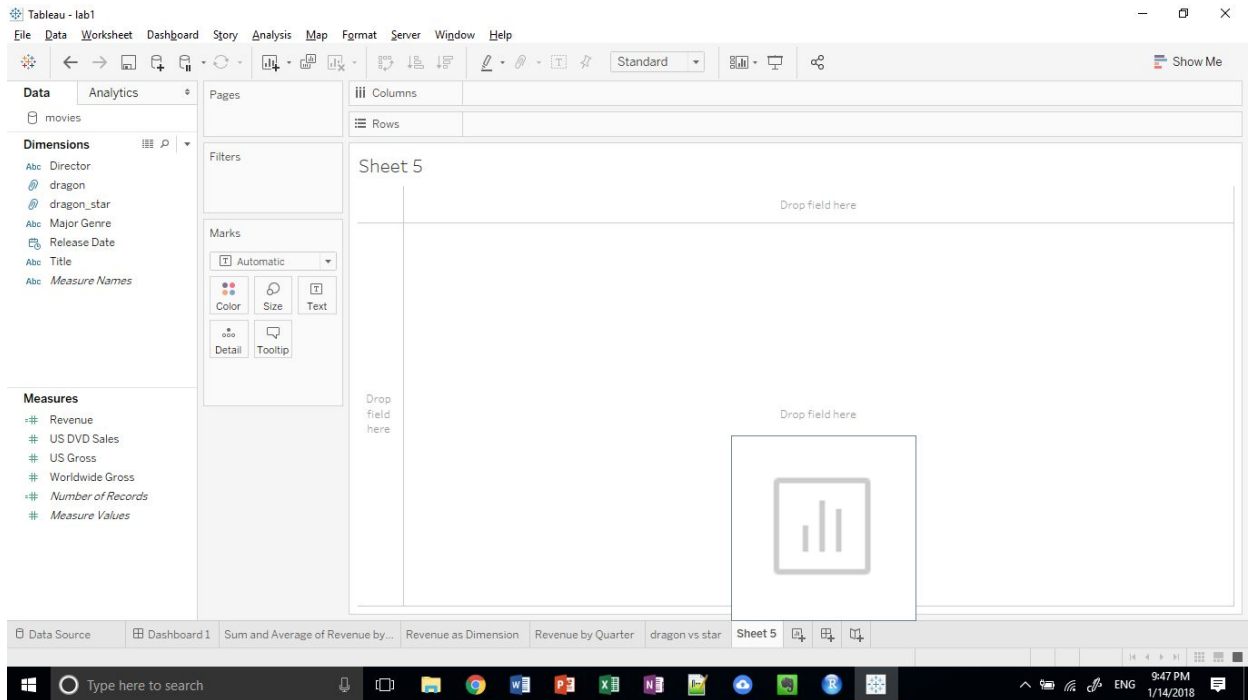
(This is an open-ended question. Please write down the steps that you used to create the visualizations, and describe how these visualizations help you to make your decision)

This is a difficult question because of the way the data is broken down in our dataset. For reasons unknown, the genre for comedy is split into three categories, while action is its own standalone category. What’s even more counterintuitive is that the given genre of “Comedy” is smaller than the combination of Black/Romantic Comedy.

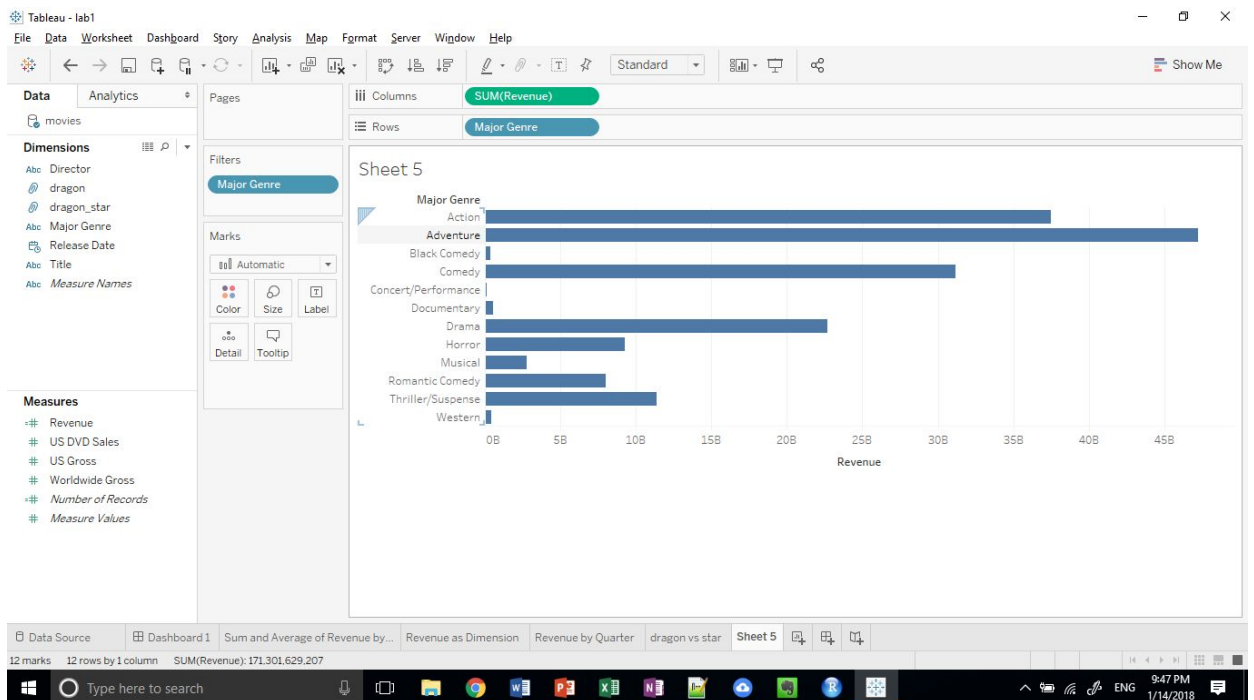
But if I had to pick only one specific genre, I would pick action, because it has the highest average over comedy.

I obtained this solution by completing the following steps:

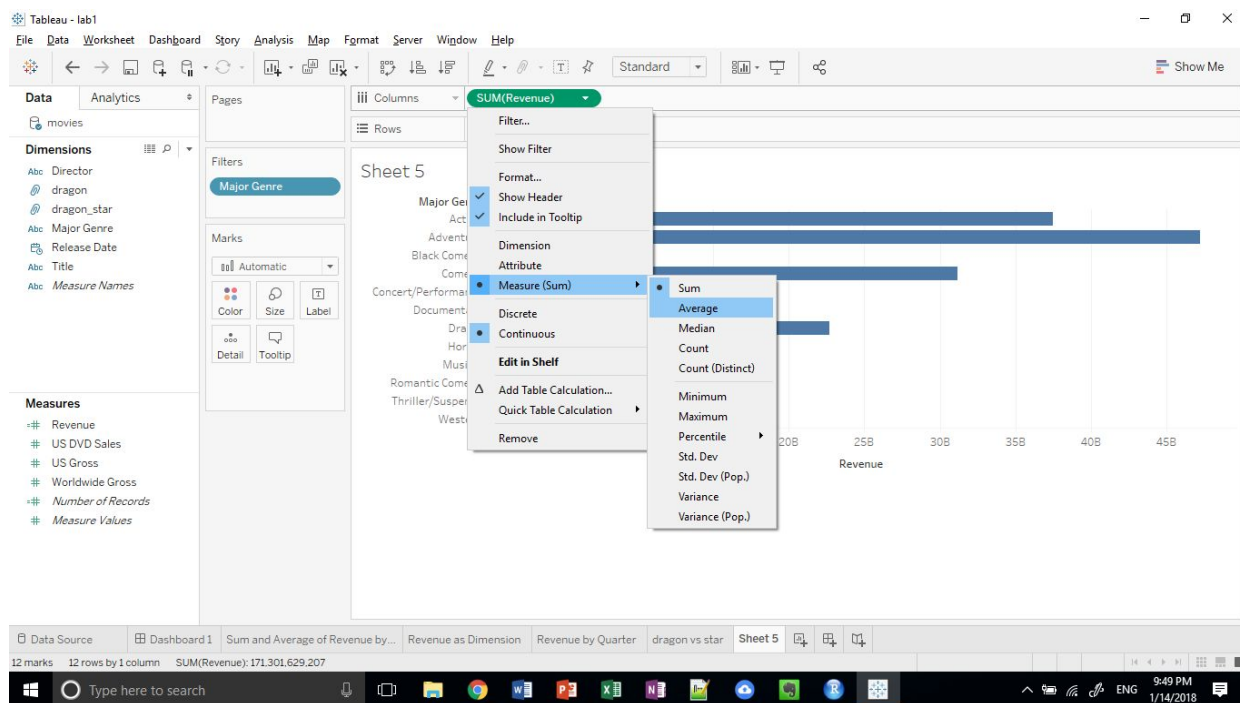
- ☐ Step1: Opened up a new worksheet



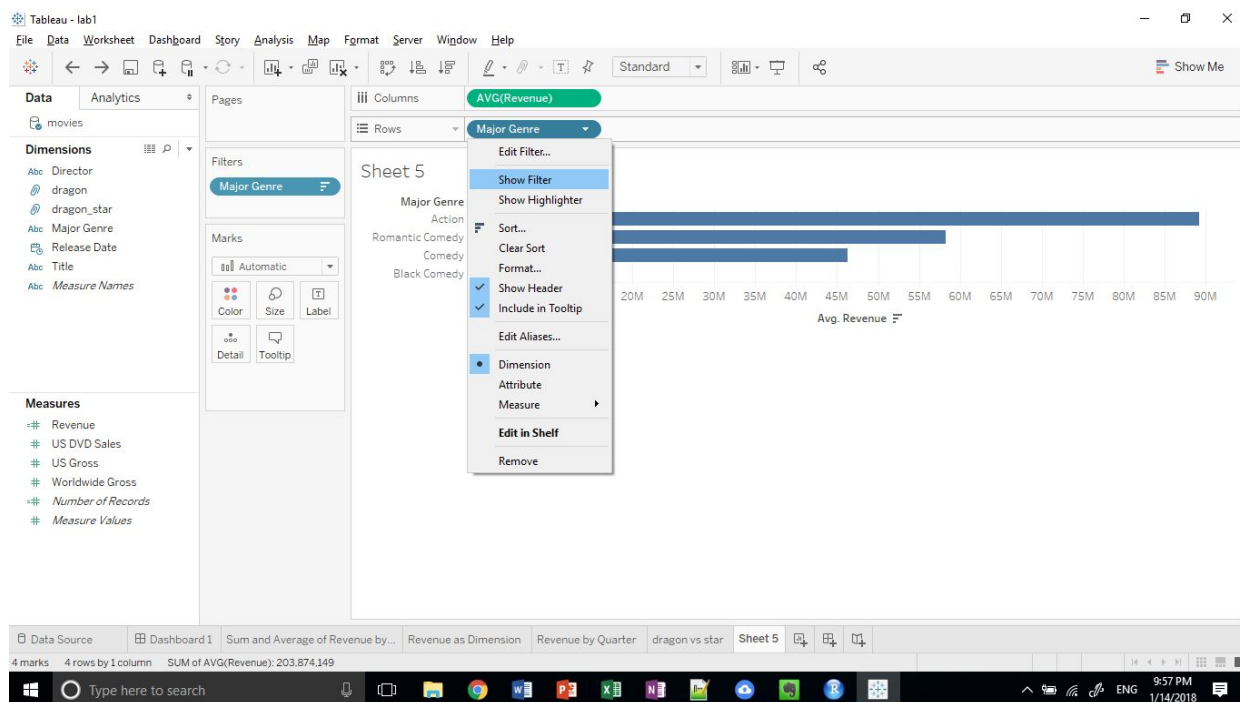
☐ Step2: Added “Major Genre” to Rows and Revenue to Columns. Got rid of null.



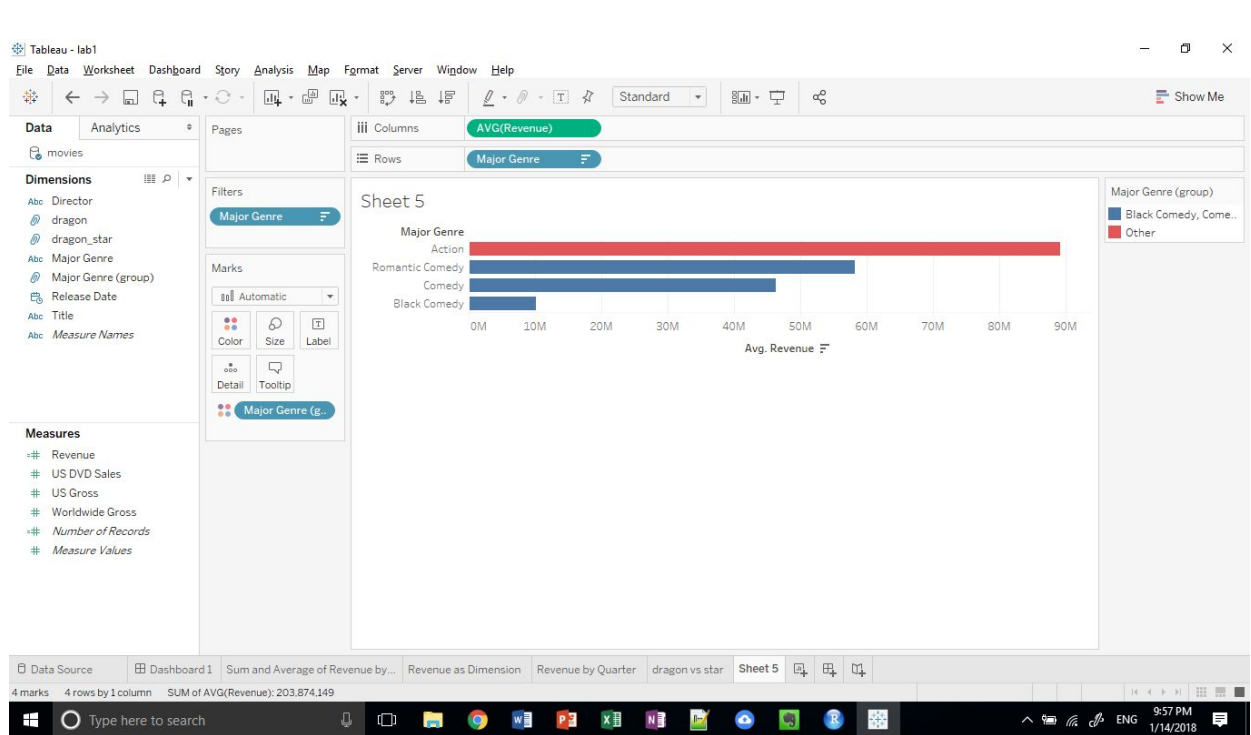
☐ Step 3: Changed Revenue to AVG



□ Step 4: Show filter to filter down to Comedy or Action



□ Step 5: Group all comedy and color code

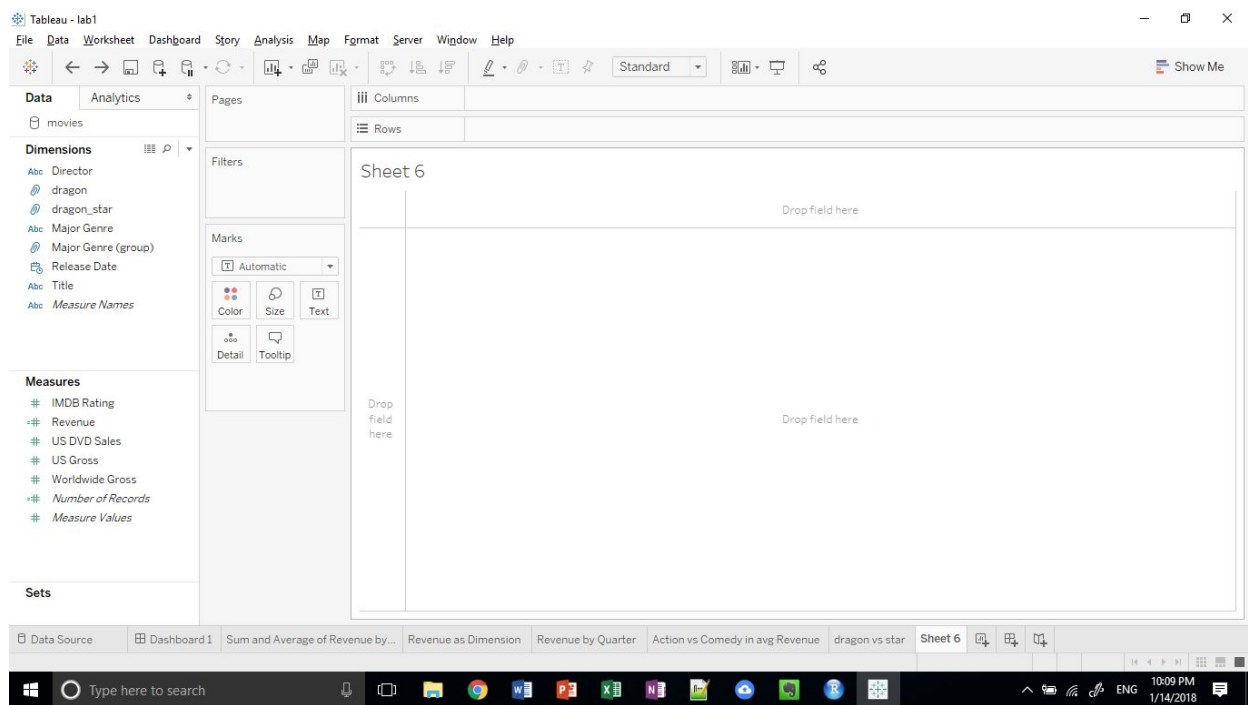


Question 12: What's the relationship between IMDB ratings and the revenue? Are highly-rated movies more profitable? (Please write down the steps that you used to create the visualizations, and describe how these visualizations help you to make your decision)

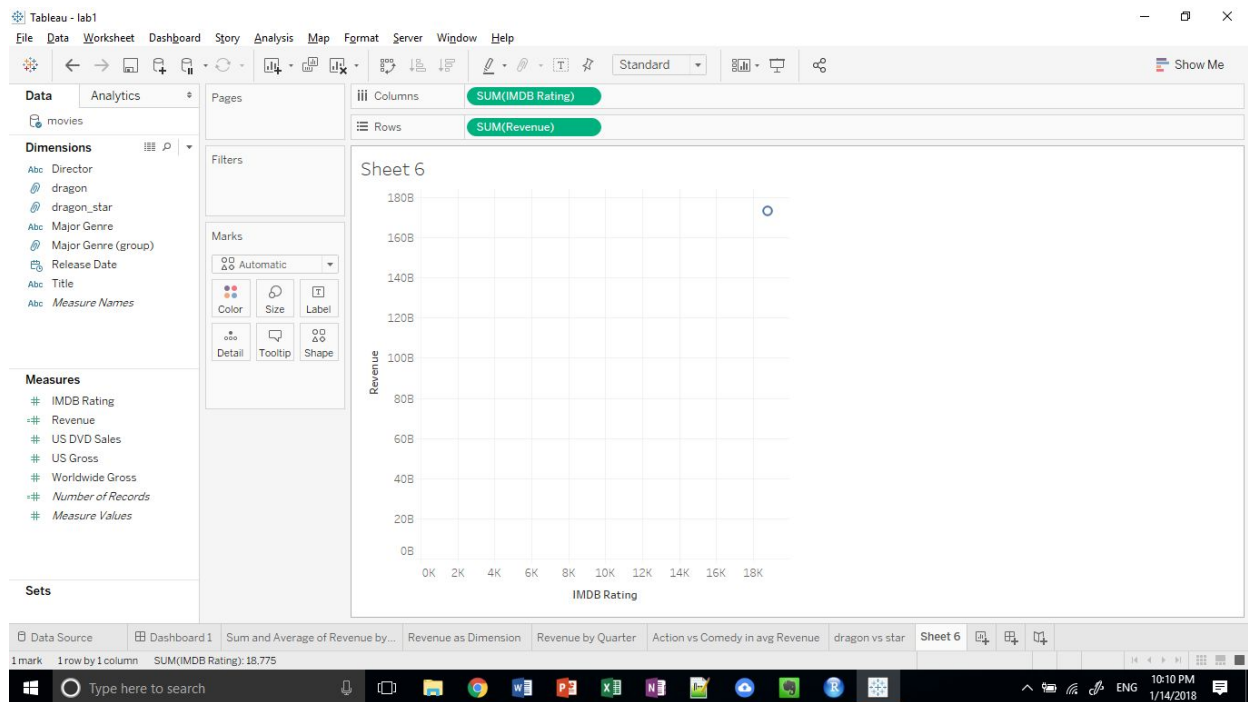
The relationship between IMDB rating and revenue seem to be positively correlated. Higher the rating on IMDB, the more money it makes. However, statistical analysis should be done to determine how strong this upward slope is.

I obtained this solution by completing the following steps:

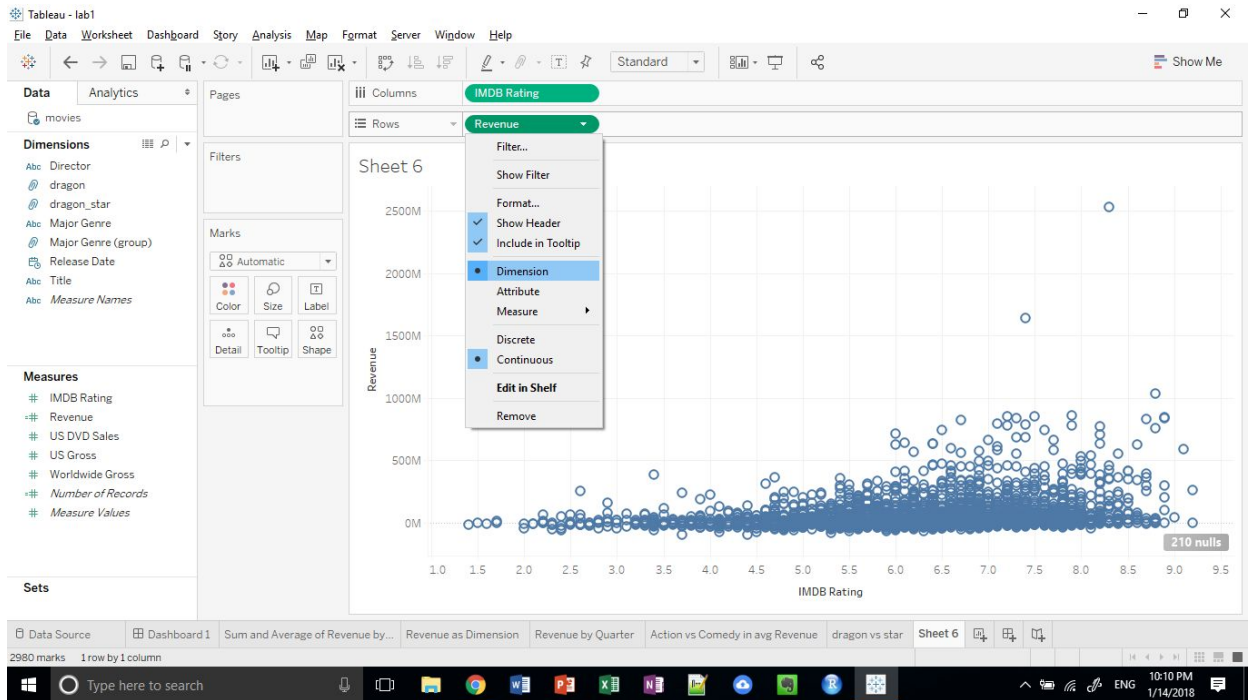
☐ Step1: Create new new worksheet



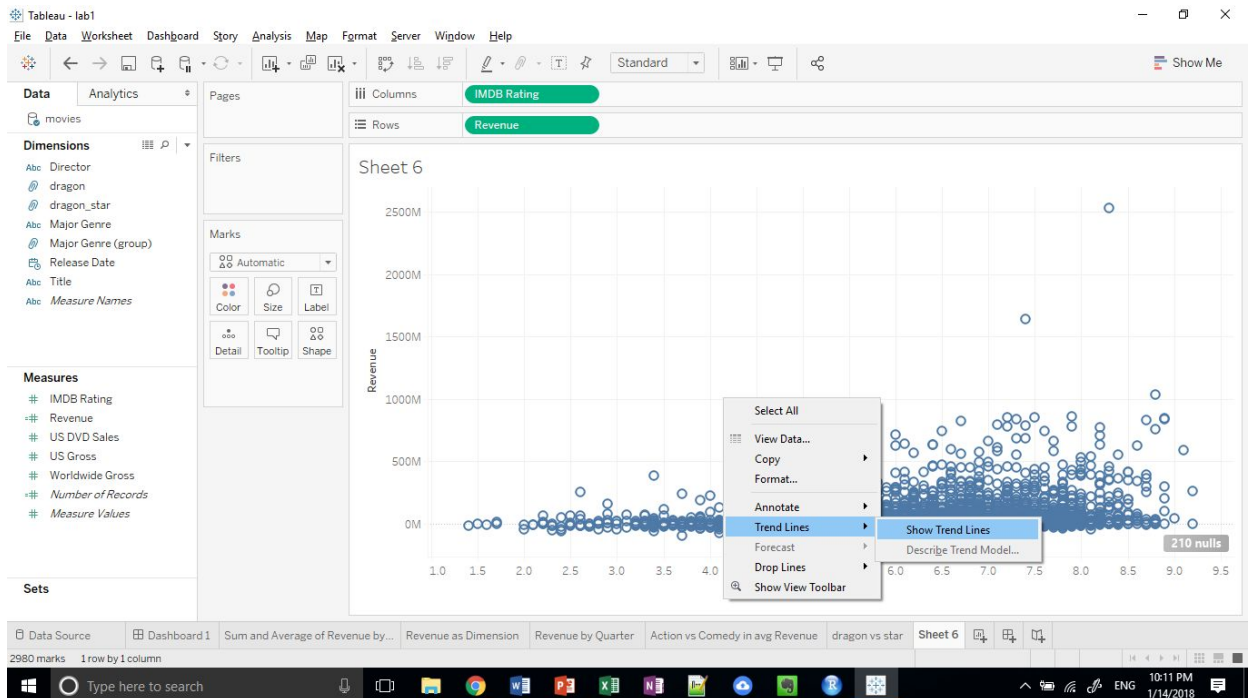
☐ Step2: Add IMDB and revenue into rows and columns, respectively



☐ Step3: Change each variable in to dimension



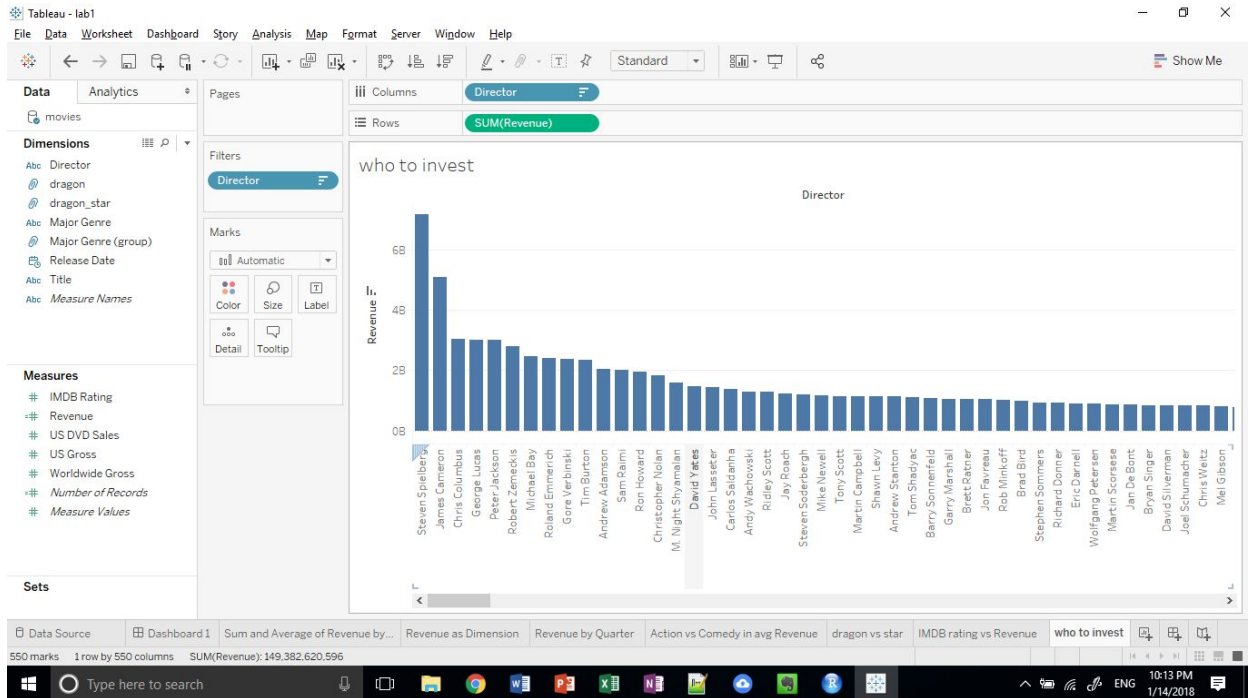
□ Step4: Show trendlines



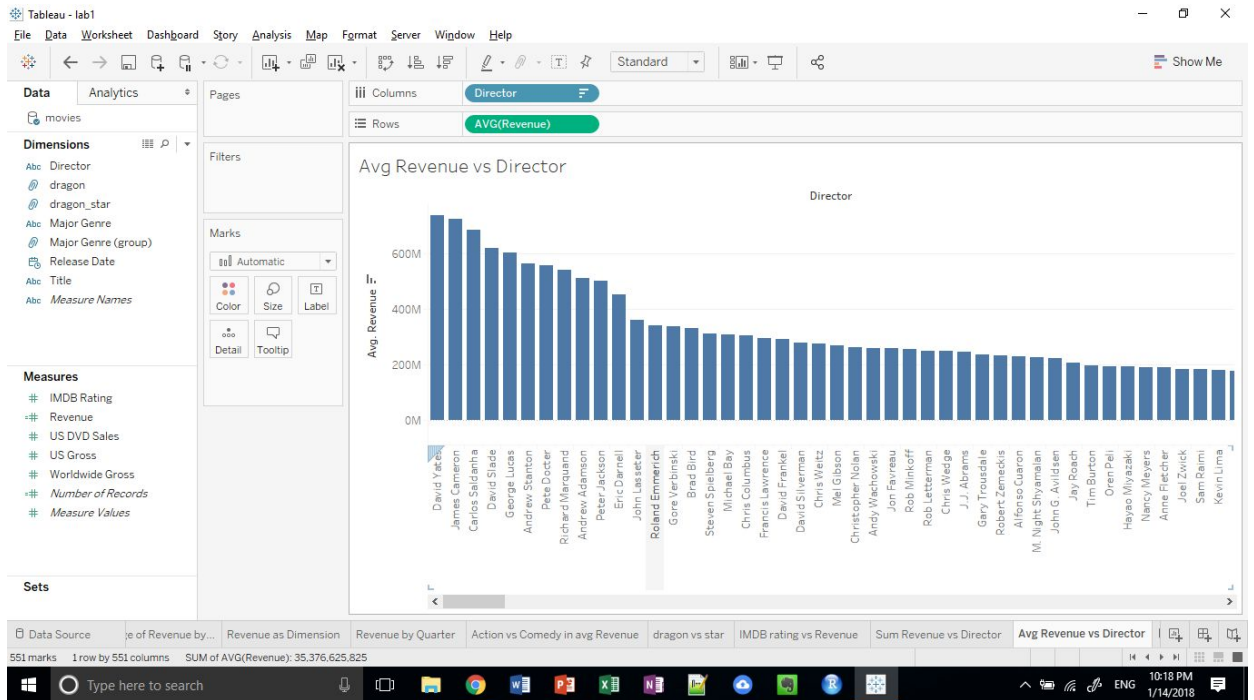
Question 13.1: Who would you invest? Create some visualizations to justify your decision.

Solution: I decide to invest in Spielberg because he earns the most amount of money. When looking at the sum revenue, Spielberg has the top, but when looking at the average revenue, Yates is has the top. Taking into account success rate defined as the ratio of the number of movies with a positive revenue over the number of movies with a negative revenue, Spielberg and Yates both are at 1. In thinking about maximizing my investment, putting money into Spielberg will earn me the most money.

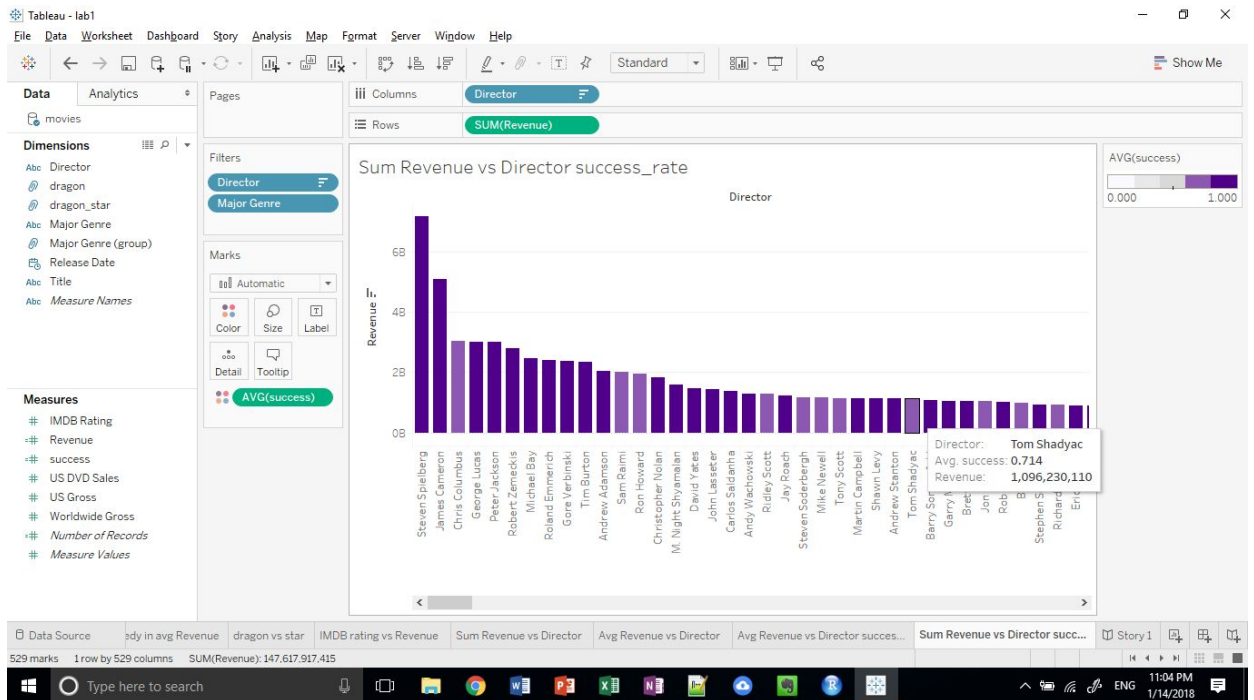
☐ visualization 1 tells me that Steven Spielberg rakes in the highest sum revenue.



☐ visualization 2 David Yates raked in the most revenue on average.



visualization 3 Highest sum of revenue by success rate.



visualization 4 Highest avg of revenue by success rate.

