

Tableau Handout Solutions

Table of Contents

Tableau Basics	1
Tableau Setup & Resources	2
Tableau Terms	2
The Movie Dataset	3
PART 1: Data set-up and manipulation	3
PART 2: Creating Simple Visualizations with Tableau	7
PART 3: Putting it all together	14

Tableau Basics

Tableau is used to generate interactive data visualizations and makes very common visualizations very easy. Our focus today is on creating simple visualizations using Tableau.

In this tutorial you will learn:

Data	Know how to import data
	View and describe data within Tableau
	Sort Data
	Select data types (Date, Number, String)
	Make Calculated Fields
	Change the # of rows displayed
	Create Groups
Sheet	Differentiate measure and dimensions
	Create simple visualization and change chart type
	Adjust encodings
	Adjust color
	Adjust aggregation methods (e.g. sum, average, count)
	Sort axis
Organize Sheets	Build dashboard
	Build stories
Action	Create annotations on the chart
	Create data filters
	Link sheets together with actions.
	Add analytics enhancements (e.g., constant line, regression)
	Edit legend and filter by legend.

Tableau Setup & Resources

Please make sure you have Tableau installed and running. You can get a copy at:

<https://www.tableau.com/academic/students>

Make sure you can run Tableau. If you need more information on Tableau take a look at:

- Videos:
 - <http://www.tableau.com/learn/training>
- Books:
 - <https://link.springer.com/book/10.1007/978-1-4842-1934-8>
- Other Recommended Resources:
 - Youtube!

Tableau Terms

Sheet: A sheet is a singular chart/map in Tableau.

Dashboard: A dashboard is a canvas for displaying multiple sheets at a time and allowing them to interact with each other.

Workbook: A workbook is the entire Tableau file containing your sheets and dashboards.

Measure: A variable from the dataset that is meant to be aggregated. (This means it should be a number that it makes sense to do math with: sum, average, and so on.) Measures are often continuous data. Examples include GPA, sales, quantity, quota, height, and salary.

Dimension: A categorical variable from the dataset that is used to slice and dice the data into different categories. Dimensions are often discrete data. Examples include country, gender, student ID, and name.

Rows and Columns Shelves: The Rows shelf and the Columns shelf is where you determine which variables will go on what axis. Put data you want displayed along the X-axis on the Columns shelf and data you want displayed on the Y-axis on the Rows shelf.

Filter: A filter is used to limit what data is being displayed on the sheet. Visible controls for a filter on a sheet or dashboard are called Quick Filters.

Tooltip: Tooltips are text boxes that appear when hovering over a mark on a sheet in order to give more information. The text and text formatting in them are easily edited through the Marks card.

Marks card: The Marks card is the tool used to create a sheet that controls most of the visual elements in a sheet. Using the Marks card, you can switch between different chart types (bar, line, symbol, filled map, and so), change colors and sizes, add labels, change the level of detail, and edit the tool tips.

The Movie Dataset

Scenario: you want to invest in the movie industry. However, you are not sure which director, which movie, or what movie genre that you want to invest your money in. In order to make a wise investment, you obtained some data about movies. You want to create visualizations to help you to make decisions.

PART 1: Data set-up and manipulation

Download the movies dataset [here](#) with your Stanford credential, open the Tableau app, and load the data.

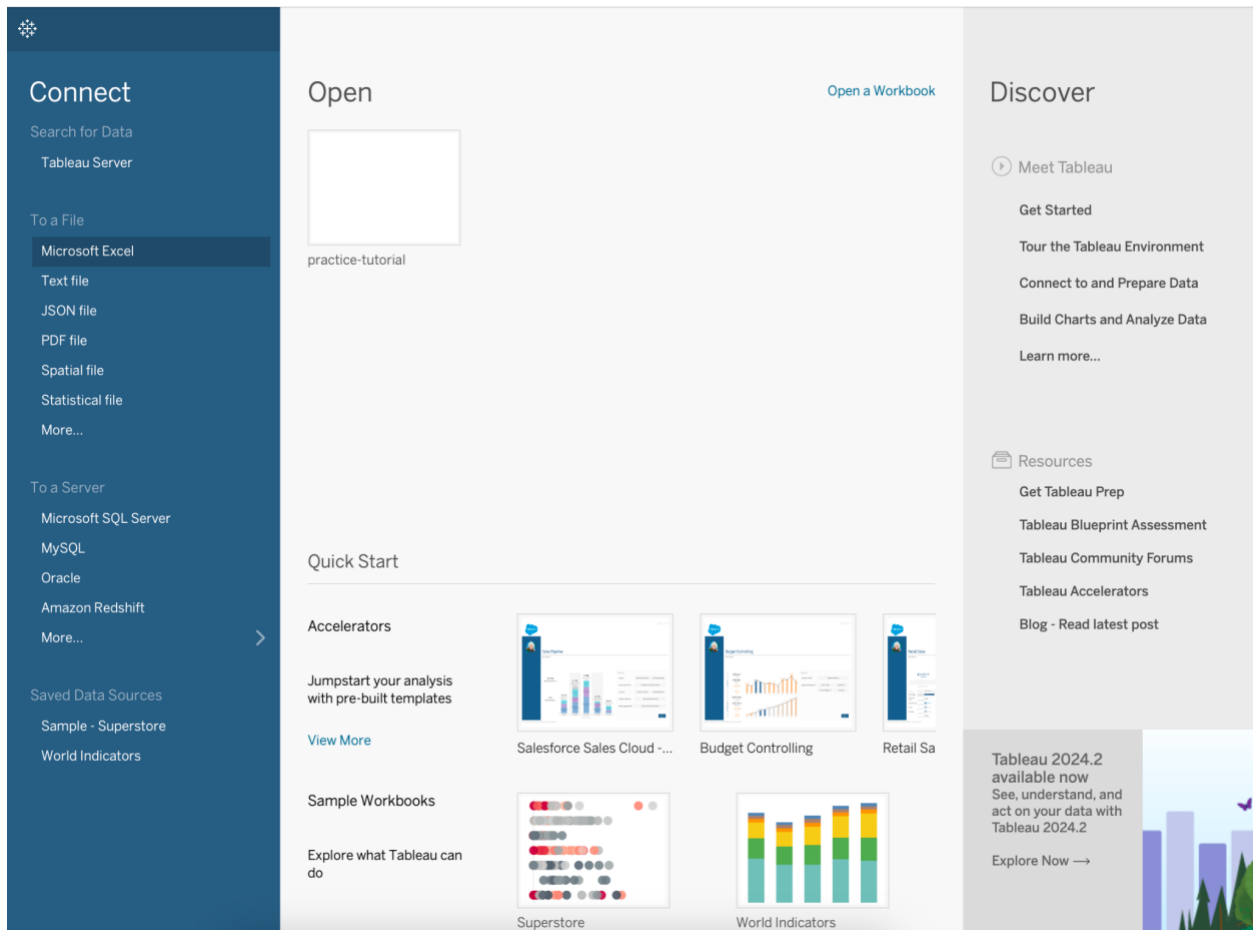


Figure 1

You can load in the data by either clicking to “Open a Workbook” or clicking the “Microsoft Excel option.” Once you load it, you will see something like this:

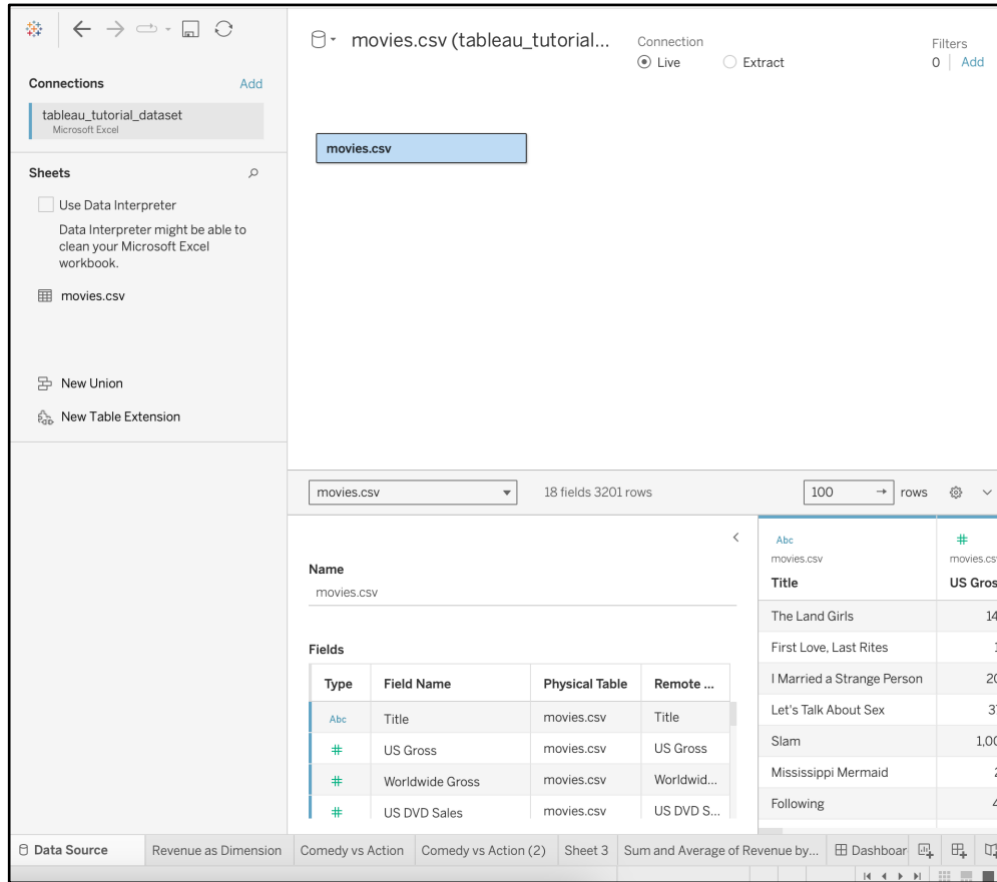


Figure 2

You can view and edit the data under the “Data Source” Tab. Make sure that you are displaying all rows in the data, or your results might look a little different from ours. (Figure 2.1)

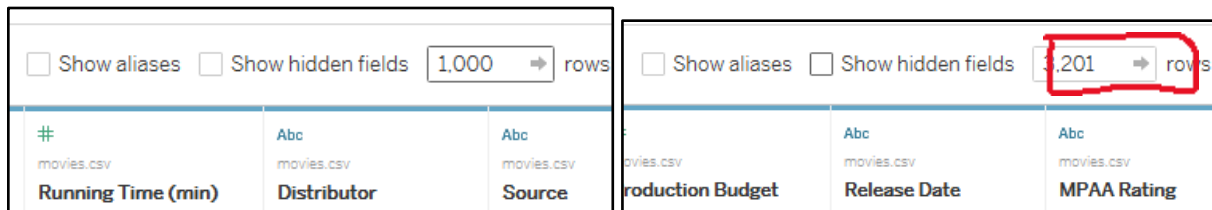
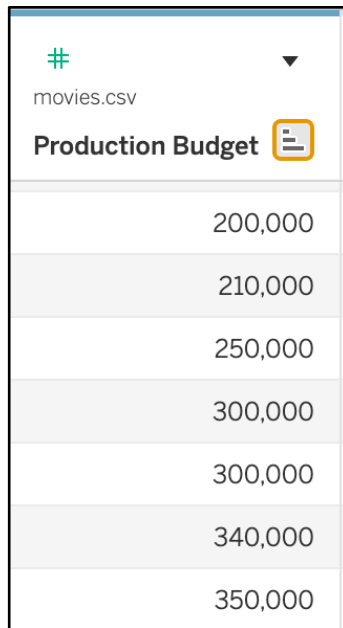


Figure 2.1

Question 1: Which movie has the highest production budget? Which movie has the lowest production budget?

1.1 Sort the dataset by production budget (Figure 3)



The screenshot shows a data table with a header row containing a green hash symbol (#) and a dropdown arrow. Below the header, the text "movies.csv" is visible. The main header of the table is "Production Budget" with a yellow icon of a list. The table contains eight rows of numerical values representing production budgets.

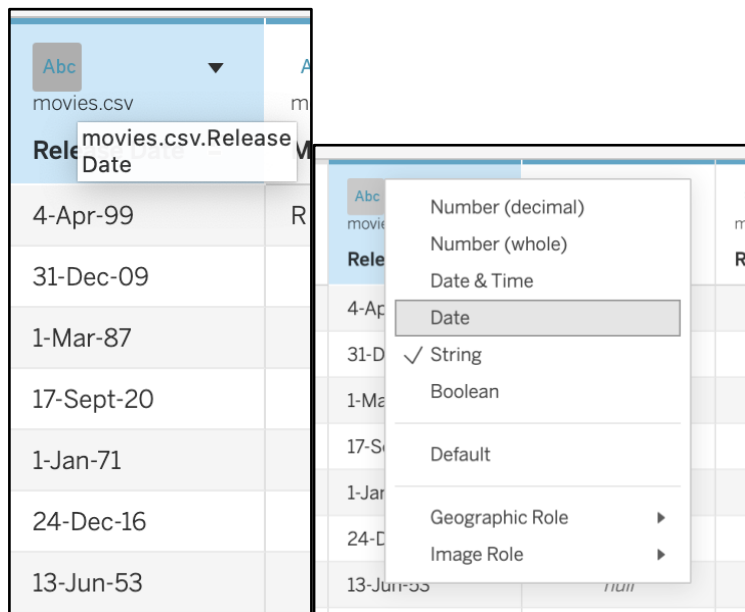
Production Budget
200,000
210,000
250,000
300,000
300,000
340,000
350,000

Figure 3

Question 2: How do I sort the dataset by releasing date?

2.1 Convert the "Release Date" from a String to a Date. Click the upper left hand corner to access the list of data types. (Figure 4.1, and 4.2)

2.2 Sort the table by "Release Date".



The screenshot shows a data table with a header row containing "Abc" and a dropdown arrow. Below the header, the text "movies.csv" is visible. The main header of the table is "Release Date" with a yellow icon of a list. The table contains eight rows of date values. A dropdown menu is open over the "Release Date" header, showing a list of data types: Number (decimal), Number (whole), Date & Time, Date (highlighted), String (checked), Boolean, Default, Geographic Role, and Image Role.

Release Date
4-Apr-99
31-Dec-09
1-Mar-87
17-Sept-20
1-Jan-71
24-Dec-16
13-Jun-53

Figure 4.1 and 4.2

Question 3: Which is the most profitable movie?

3.1 Sort the dataset by the “worldwide gross”

3.2 Create a calculated field called “Revenue” that shows the actual profit that these movies make (hint: Revenue= Worldwide Gross-production budget, if we only care about movies that target the international market) (Figure 5)

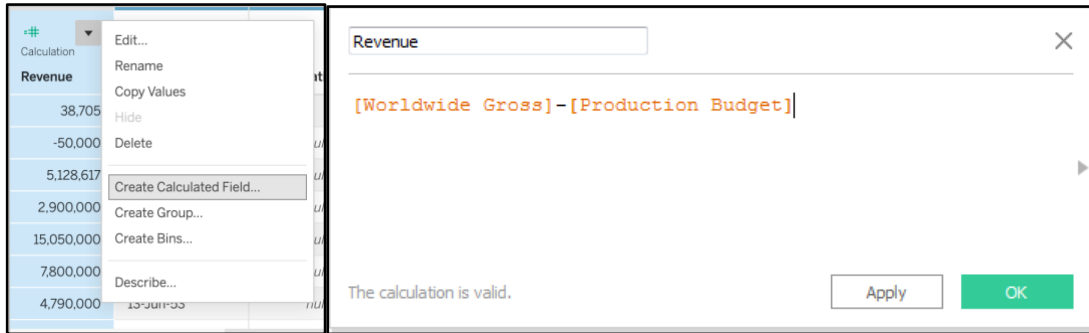


Figure 5

Question 4: How many directors are included in the dataset?

4.1 Click on the triangle next to “Director”-> select “Describe”. The “Describe Field” window shows the answer. (Figure 6)

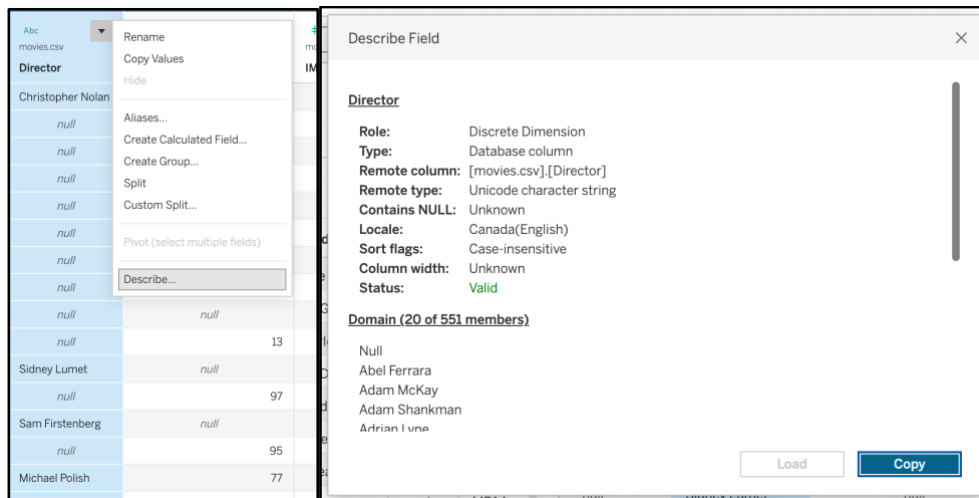


Figure 6

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

5.1 Create a group of all movies that contain “star” in their names. (Figure 7)

Hint: Triangle Click-> create group-> enter search criteria-> “Find ALL”-> Group ->Change group name (optional) -> Include Others(optional) -> Change Field name(optional)

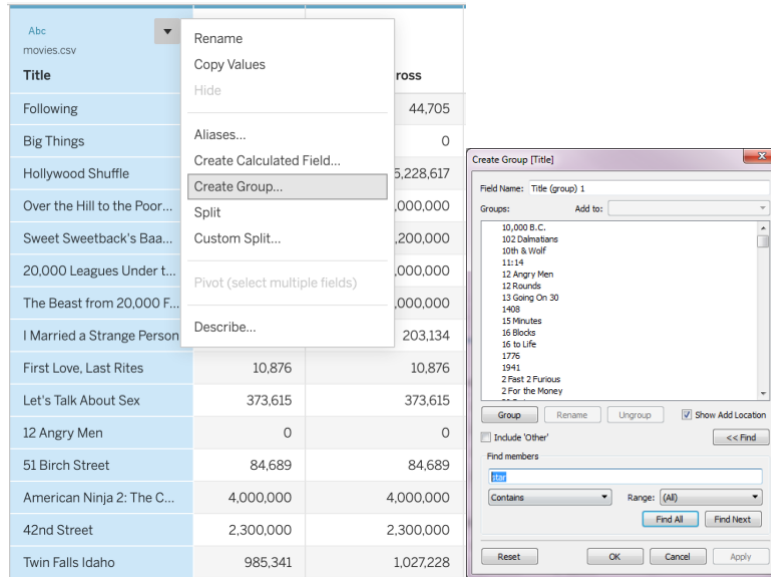


Figure 7

PART 2: Creating Simple Visualizations with Tableau

After doing basic data manipulations, we are going to create some simple visualizations. Click on the “Sheet1” tab and you will see a layout similar to Figure 8.

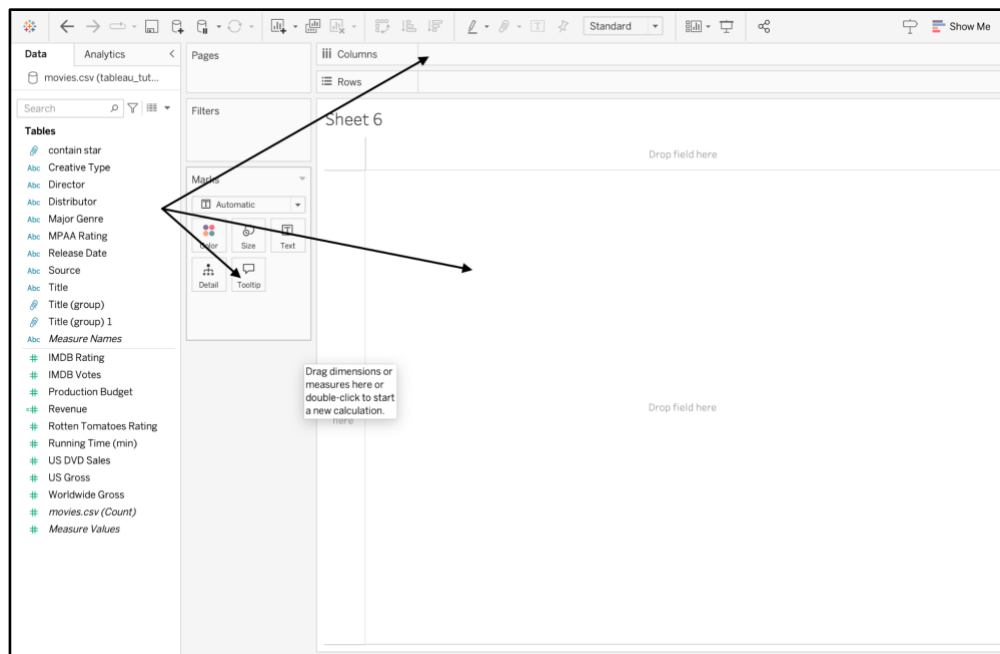


Figure 8: the arrows are indicating that you can drag and drop any of the measures or dimensions into any of the 3 areas shown.

The side bar has two important blocks, **measures** and **dimensions**.

Dimensions are values that determine the **level of detail at which measures are aggregated**. Basically, you will have a “mark” for each dimension (e.g., states in the US). You can think of them as slicing the

measures or creating groups into which the measures fit. The combination of dimensions used in the view defines the view's basic level of detail.

Measures are **values that are aggregated**. They can still be summed, averaged, counted, or they can have a minimum or maximum.

Question 6: Who is the most successful director?

6.1 Drag “Director” to the row, drag “Revenue” to the column. The “Show Me” menu let you create different types of charts (Figure 9).

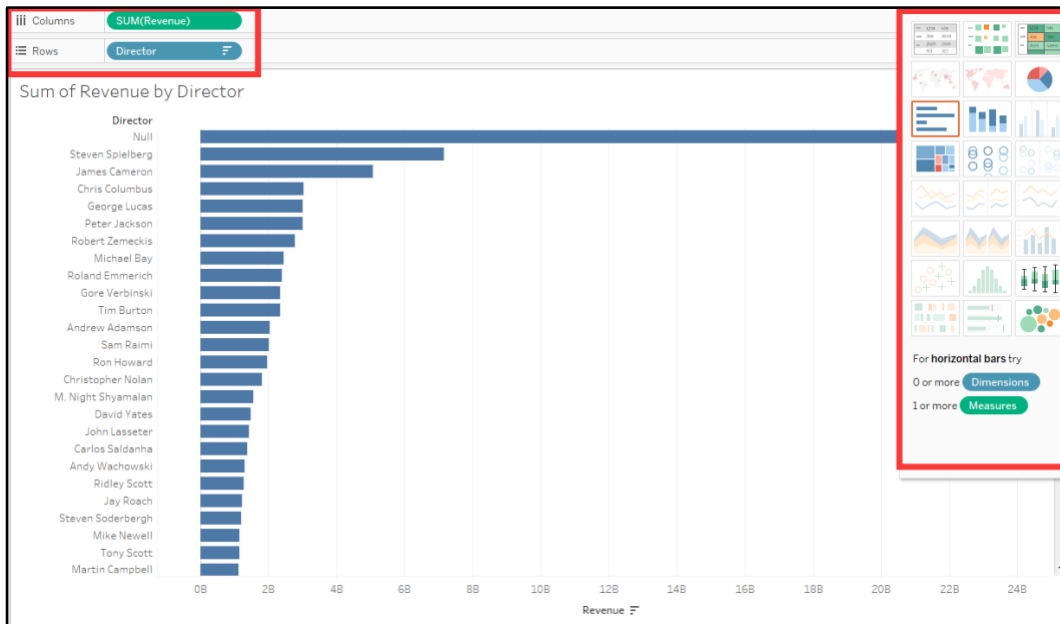


Figure 9

6.2 Which director has the highest sum of revenue? (Try to create a different visualization and sort the revenue) (Figure 10)

6.3 Filter Null Values in the “Director” field by right-clicking on the “Null” value and excluding the values. (notice that a new filter is created)

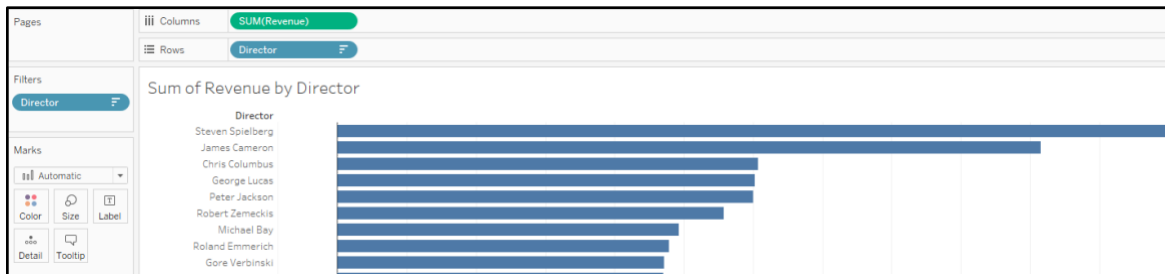


Figure 10

6.4 Drag “Revenue” to column and convert it to the measure of average (i.e.) (Figure 11).

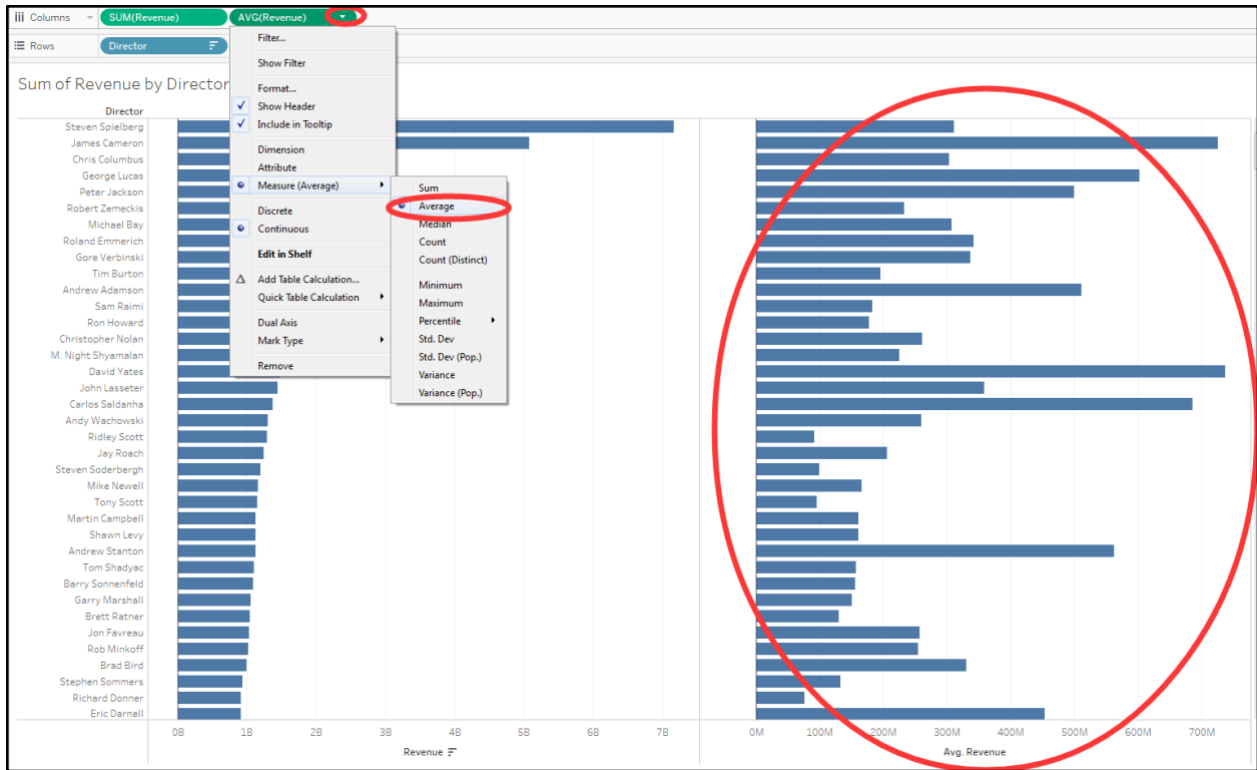


Figure 11

6.5 Add annotations (Figure 11.1)

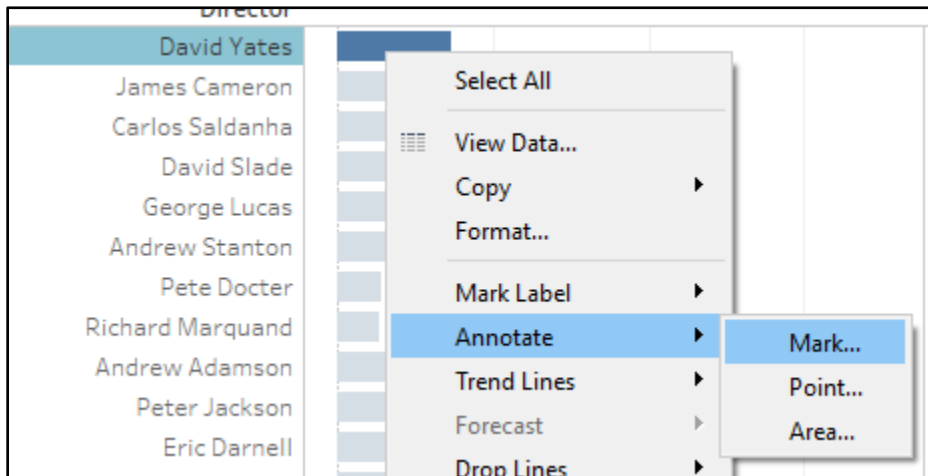


Figure 11. 1: Right click the field/object that you want to annotate

6.6 Edit the color of the bar by clicking on the “Color” option in the “Mark” shelf.

6.7 Edit the Sheet1 title to be “Sum and Average of Revenue by Director”.

Question 7: Are there any directors who have never directed a profitable movie?

7.1 Create a new worksheet and name it as “Revenue as Dimension”.

7.2 Drag “Director” to the Rows. Drag “Revenue” to Columns, and convert it to Dimension (Figure 12).

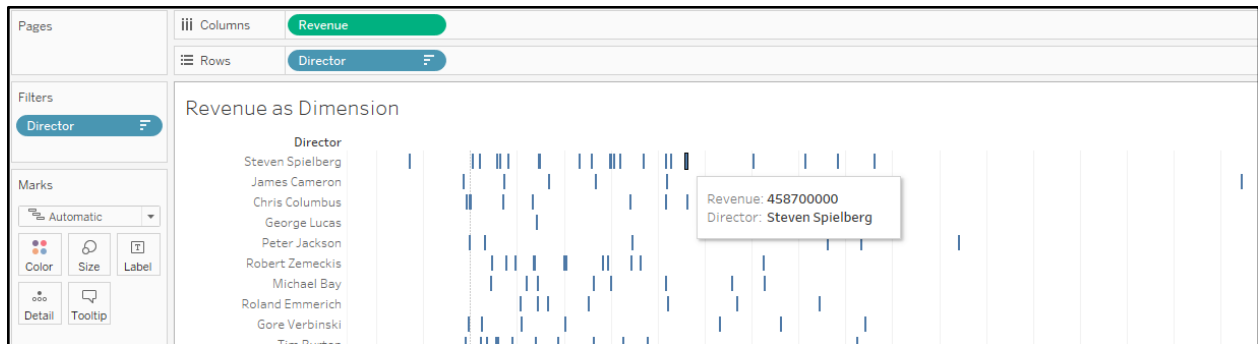


Figure 12

7.3 Add “Title” to the “Mark” shelf as tooltip so that each revenue mark is associated with the movie title (Figure 13)

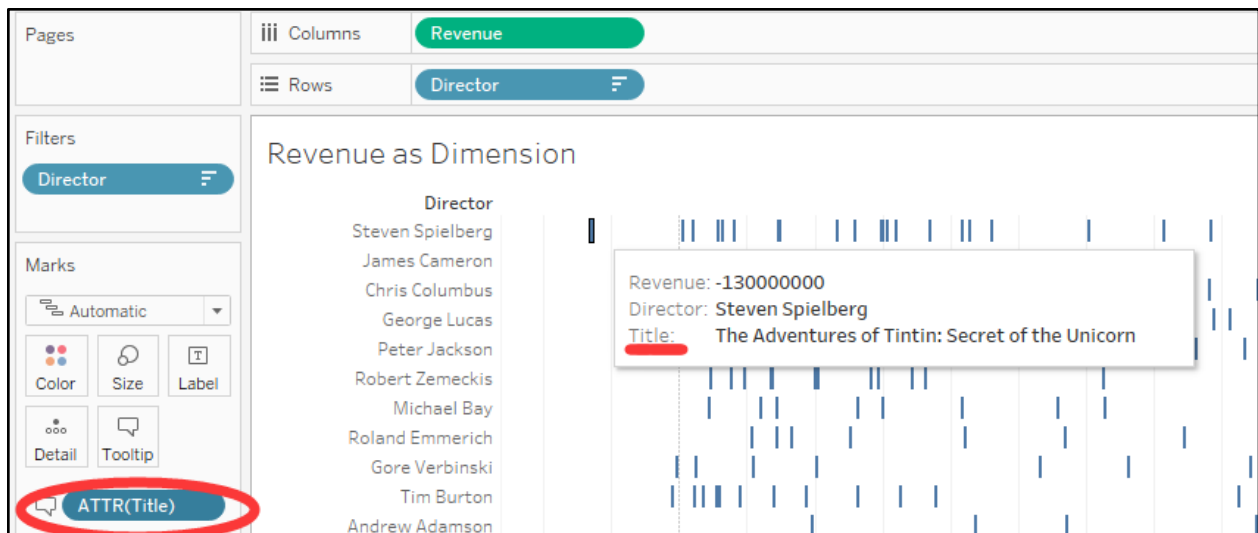


Figure 13

7.4 Add a constant line to indicate movies that have negative revenue (Hint: Analytics-> Summarize->Constant Line-> Value=0) (Figure 14)



Figure 14

7.5 Drag “Major Genre” to “Marks” shelf and encode it using color. (Figure 15)

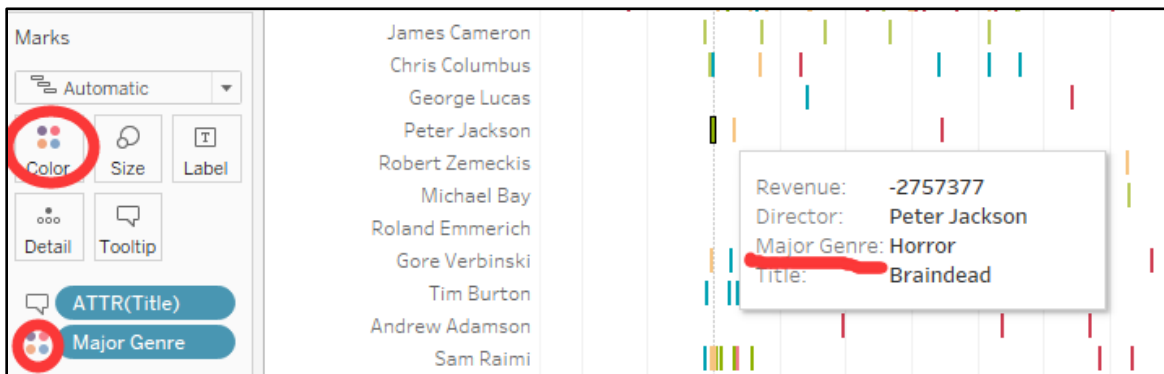


Figure 15

7.6 Are there directors who have never directed profitable movies? Who are they?

Hints:

1. You can use filters to complete this task.
2. Never directed profitable movie = all movies directed have revenues less than zero

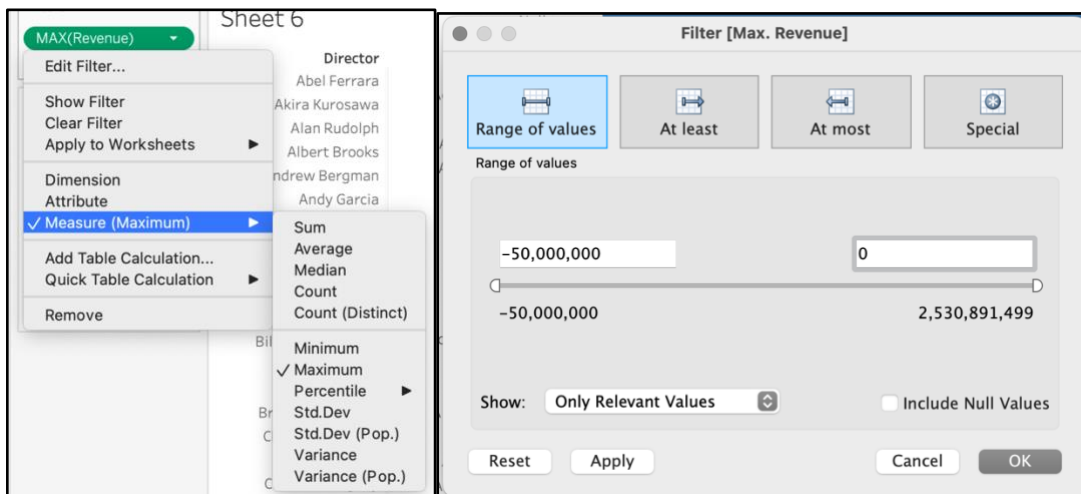


Figure 16

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

8.1 Go back to the “Sum and Avg of Revenue” worksheet. Drag “Major Genre” to the “Filter” shelf and display only movies from the genre “Action”. (Figure 16)

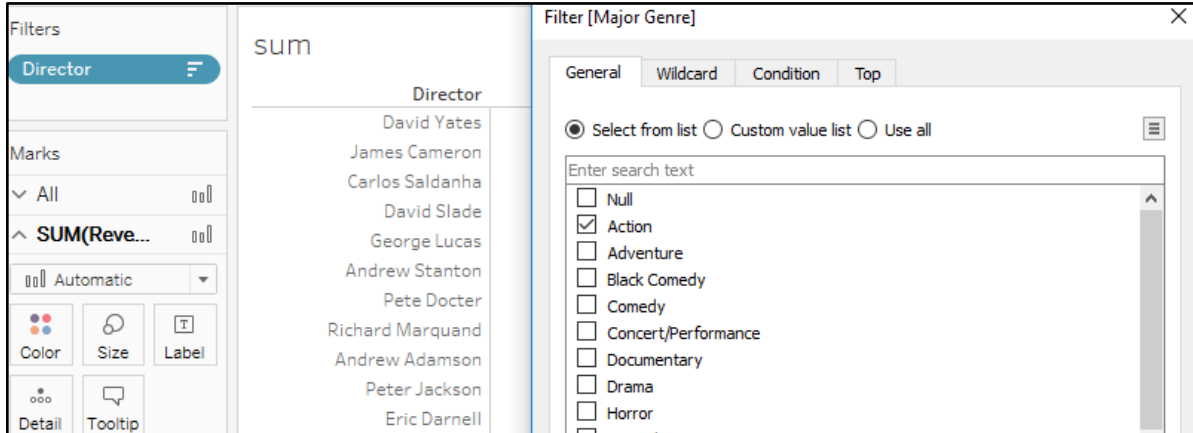


Figure 17

8.2 Find Directors who have the top 10 DVD sales by adjusting the director filter (Figure 17).

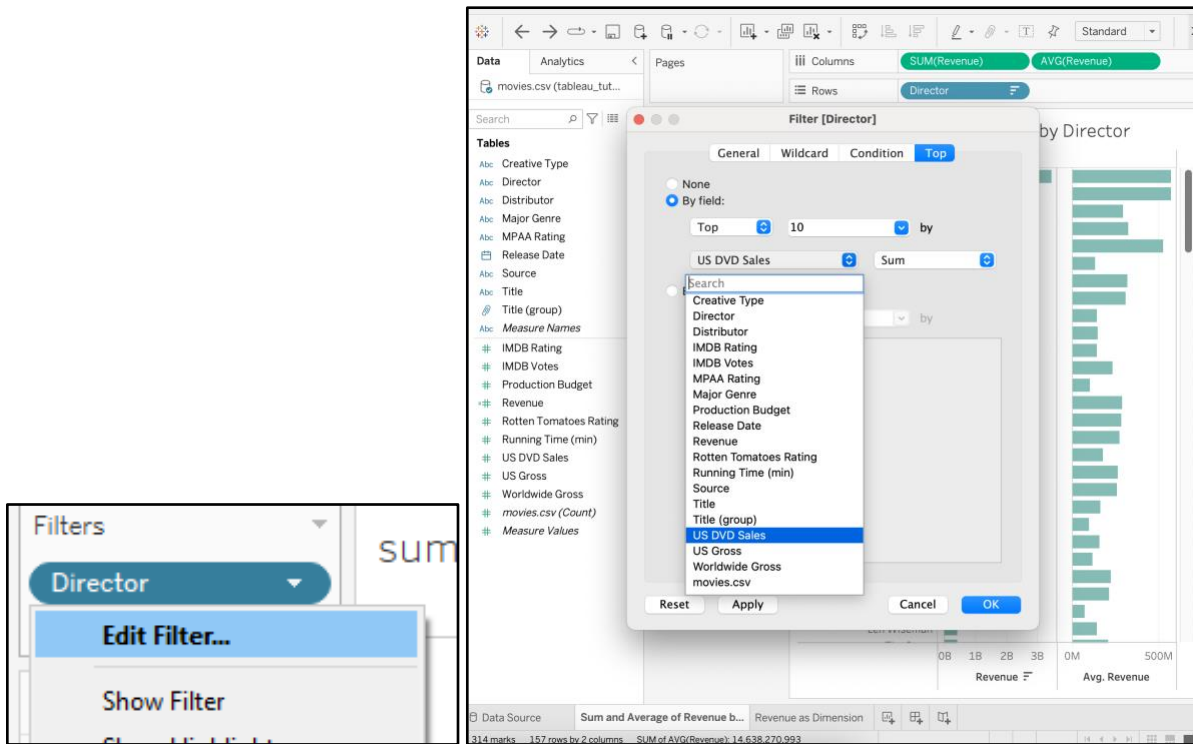



Figure 18

8.3 Which director has directed the most profitable Comedy movie?

(Hint: use  and filters/pages shelf)

Question 9: When is the best time to release a movie?

Create a visualization of revenue by quarters.

Hint: Click on the plus sign next to the YEAR (Figure 18)

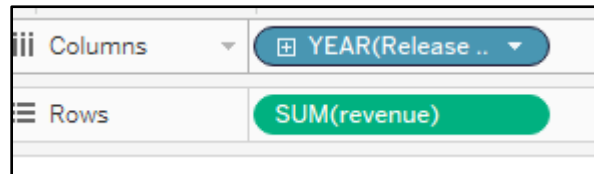


Figure 19

PART 3: Putting it all together

Question 10: How do I compare two visualizations in the same view?

10.1 Create a dashboard to host both worksheets that we have created (e.g. “Sum and Avg of Rev” and “Rev as Dimension”) (Figure 19).

10.2 Select “Action” from the legend so that only action movies are displayed in the (Revenue as Dimension visualization) (Figure 19).

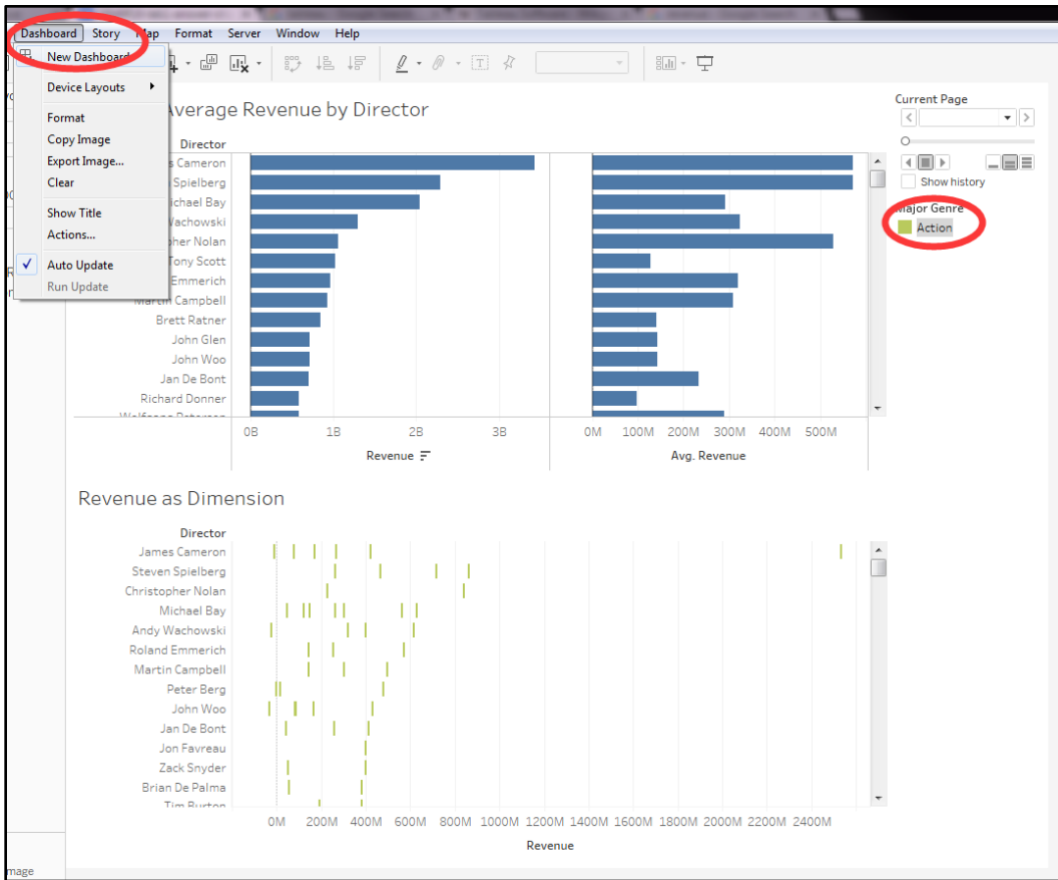


Figure 19

10.3 Add a highlight action: “Dashboard”-> “Action”-> “Add filter”. In the pop-up window, add a “Hover” action for the field of “Director”. (Figure 20).

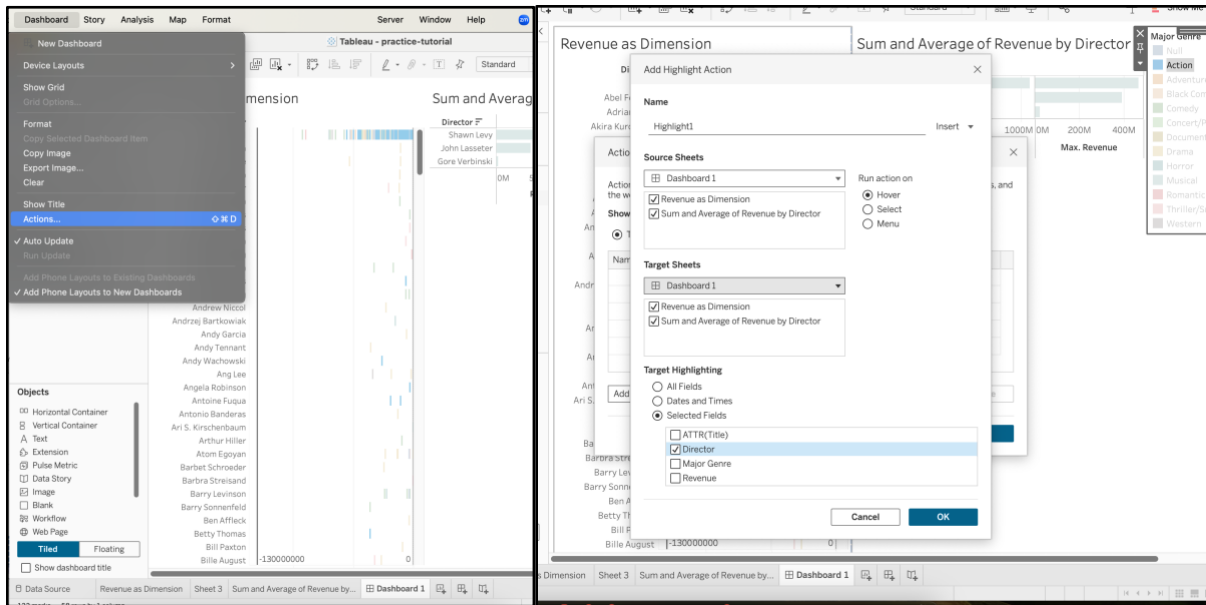
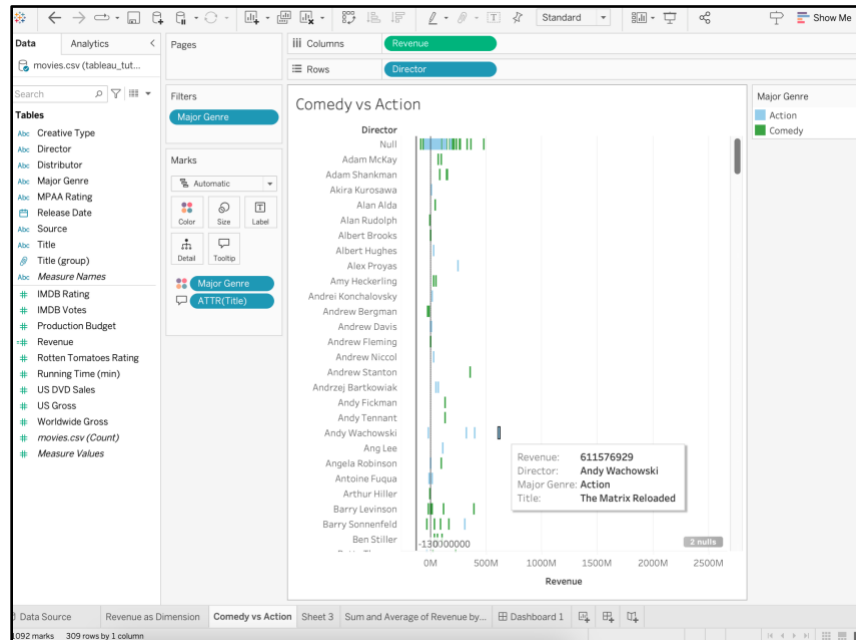
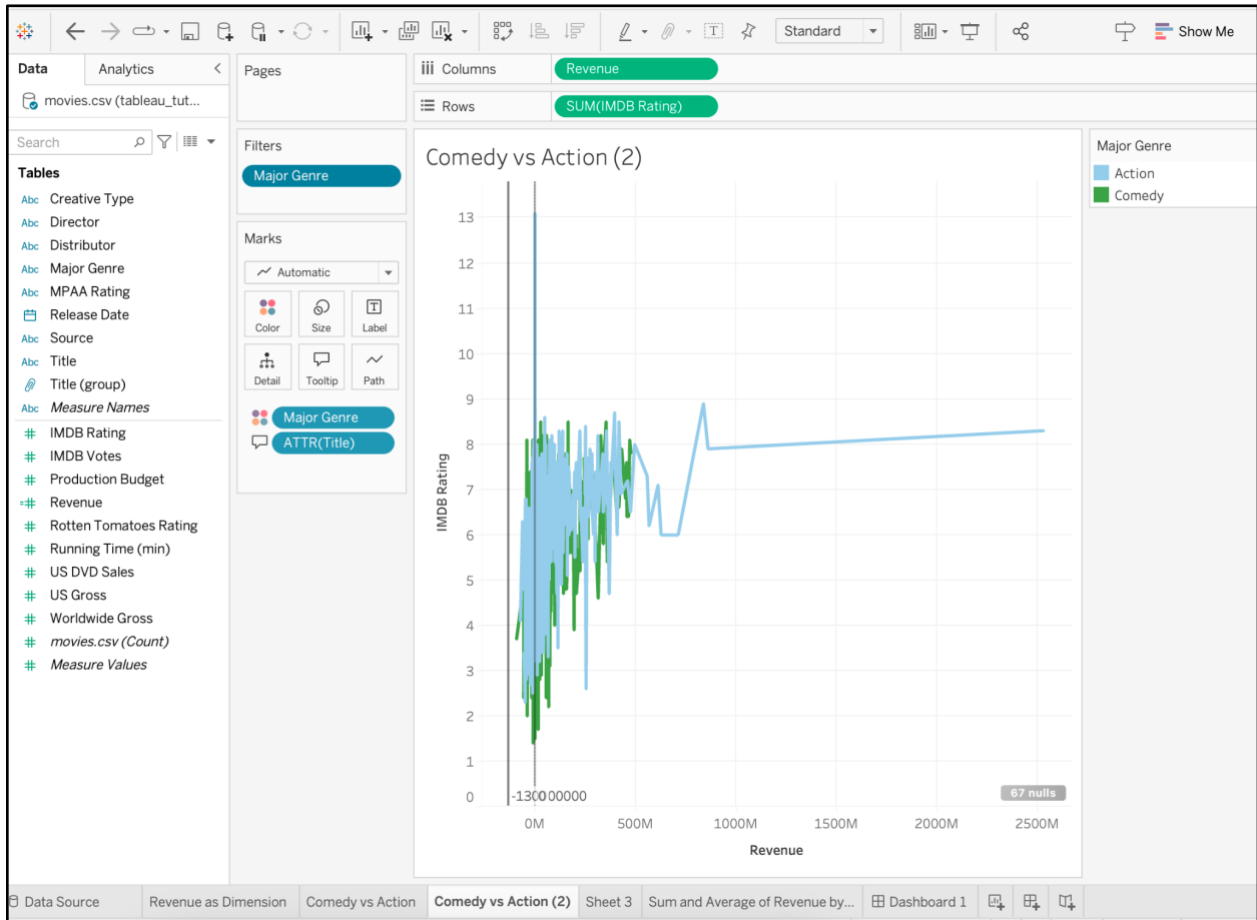


Figure 20


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

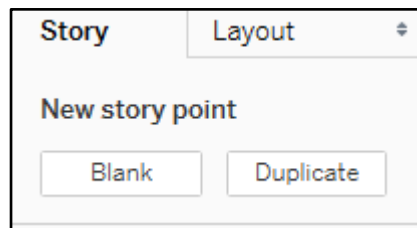


Question 12: What's the relationship between IMDB ratings and the revenue? Are highly-rated movies more profitable?



Question 13: I am going to invest in

13.1 Make a story by clicking on the "add story" button  (figure 21)



13.2 Annotate the story and add new story points. (figure 22)

13.3 Who would you invest? Create some visualizations to justify your decision.

PART 4: Exporting your visualizations

Interactive Tableau workbooks can only be uploaded to Tableau online, Tableau Server and Tableau servers. However, you can embed these visualizations to your own website: see the instruction here: https://onlinehelp.tableau.com/current/api/js_api/en-us/JavaScriptAPI/js_api.htm

You can export to image and pdf also.

For image export, you can only export worksheets: (figure 23)

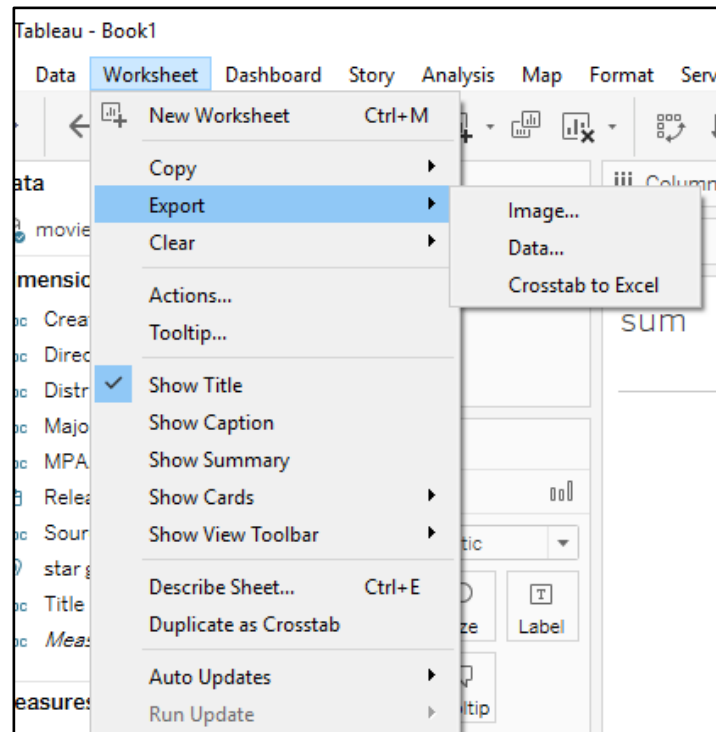


Figure 23: export by clicking worksheet-> export->image

For pdf export (which you can edit on editors such as illustrator), go to file-> print to pdf (figure 24)

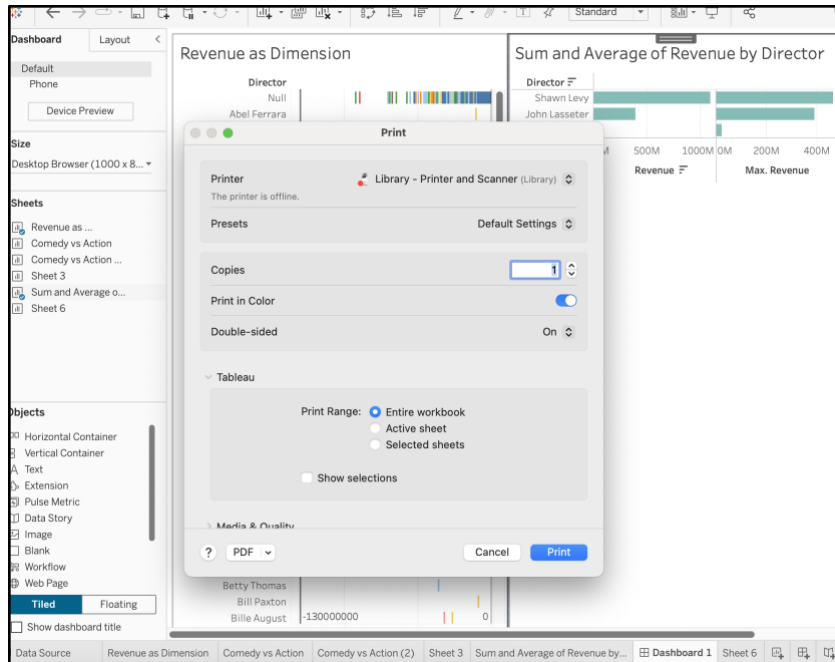


Figure 24

Note: Make sure that you check out the tableau public gallery <https://public.tableau.com/en-us/s/gallery>