HW1: Stakeholders and Observation

Kuhu Gupta

1 Domain

Data visualization is viewed by many disciplines as a modern equivalent of visual communication. It involves the creation and study of the visual representation of data. To communicate information clearly and efficiently, data visualization uses statistical graphics, plots, information graphics and other tools. This is something that I am passionate about.

2 Technological Product

Tableau is a powerful business intelligence and data visualization tool that has a very intuitive user interface. You don't need any coding knowledge to work with Tableau. It is very useful in drilling-down data, creating insightful reports and garner actionable business insights. Therefore, I would like to focus on one this tool for this assignment.

3 Stakeholders

- **Students:** Students are potential stakeholders of Tableau as it helps them in project reports, exploring data for machine learning. In the classroom and in the lab, Tableau enriches the learning experience with analytics anyone can use and understand. Tableau is free for active students.
- **Corporate Sector/Companies:** Several companies use Tableau as a business intelligence (BI) software that helps people see and understand their data.And now, everybody from the local business owner to big corporations to your doctor want to analyze all this data. They want to get their hands on it and make sense of it and try to predict consumer behavior.
- **Developers:** The developers of the Tableau software are one of the important stakeholders. The developers are concerned with the quality of what they produce, in this case Tableau software as a whole. Since professionally, quality is a matter of concern, and because they can reduce stress in their future lives if the interface doesn't have any bugs and is easy and straightforward to use.
- **Competitors:** There are lots of other attractive Tableau alternatives that make good use of business data assets through cloud-based and self-service business intelligence, like, QlikSense, Microsoft Power BI, Sisense BI,etc. Apart from this there are open-source libraries used for data visualization like D3.js and P5.js used in the data journalism and academics for data visualization.

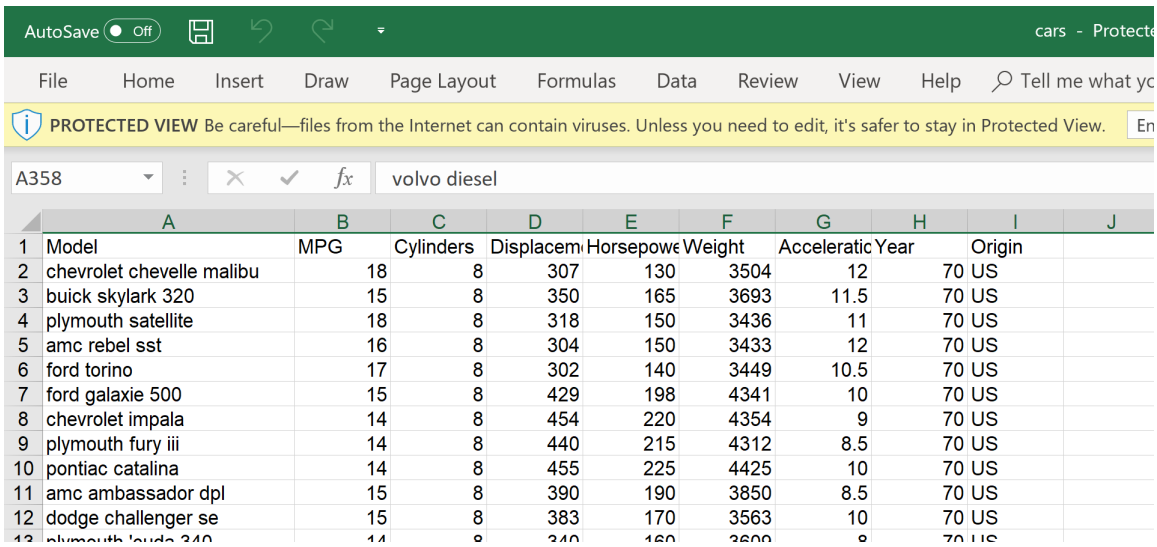
4 Observations, Observational Setting and Interactions

To record my observations, I chose to do it with a student who is a beginner in using Tableau. She is a student at Georgia Tech pursuing Biomedical engineering and has taken two business courses this semester where she has to make business reports and dashboards. Therefore, I thought it would be a good idea to observe someone from whom the tool is relatively new. Moreover, see the usability of the tool from the perspective of a person who is using it for the first time.

I asked the user to analyze conventional "cars" data set using Tableau. Here is the snapshot of the data-

I thought the best way to observe a user using a visual analytics software or any other software would be to watch the user perform specific tasks and record the time they take to finish them. In visualization, there are tasks like

- Retrieve Value
- Filter
- Sort



	A	B	C	D	E	F	G	H	I	J
1	Model	MPG	Cylinders	Displacem	Horsepowe	Weight	Acceleratic	Year	Origin	
2	chevrolet chevelle malibu	18	8	307	130	3504	12	70	US	
3	buick skylark 320	15	8	350	165	3693	11.5	70	US	
4	plymouth satellite	18	8	318	150	3436	11	70	US	
5	amc rebel sst	16	8	304	150	3433	12	70	US	
6	ford torino	17	8	302	140	3449	10.5	70	US	
7	ford galaxie 500	15	8	429	198	4341	10	70	US	
8	chevrolet impala	14	8	454	220	4354	9	70	US	
9	plymouth fury iii	14	8	440	215	4312	8.5	70	US	
10	pontiac catalina	14	8	455	225	4425	10	70	US	
11	amc ambassador dpl	15	8	390	190	3850	8.5	70	US	
12	dodge challenger se	15	8	383	170	3563	10	70	US	

Figure 1: A snapshot of "Cars" data set

- Compute Derived Value
- Find Extremum
- Sort
- Determine Range
- ... Many More

Thus, I have created certain tasks for 2 sessions of observations on similar lines as mentioned above-

Session 1:

- Name the car with the highest value of weight. (Retrieve Value)
- Show me the cars having the origin as the USA. (Filter)
- Sort the cars based on cylinders. (Sort)
- Show me a chart avg value of mpg and color them based on origin. (Compute derived Value)

Session 2:

- Name the car with the highest horsepower. (Retrieve Value)
- Show me the cars with cylinders greater than 4. (Filter)
- Sort the cars based on MPG. (Sort)
- Show me the chart with the sum of acceleration. (Compute derived value)

I followed a procedural approach to observe the user.

Step 1: For the initial 20 seconds, explore and try to see all the features of the tool.

Step 2: Upload the data file.

Step 3 : Do the above specified tasks of Session 1. (I was observing as well as recording the time taken to do the tasks.)

Step 4: Give 10 minutes break.

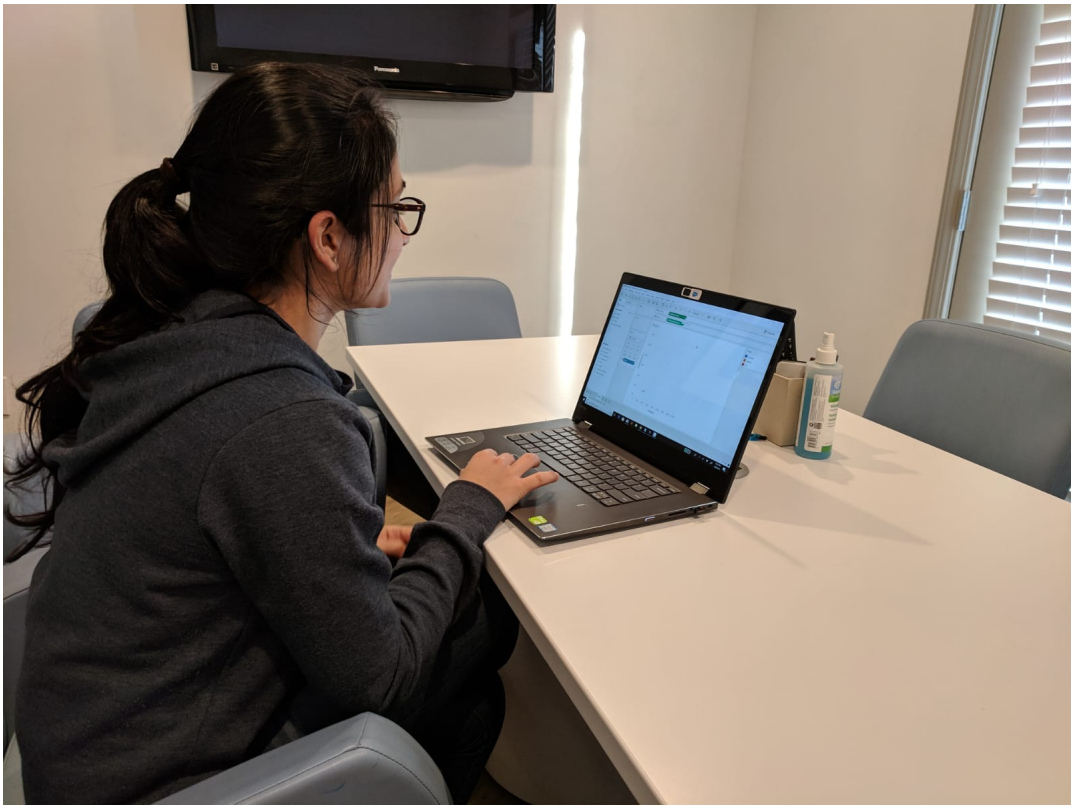


Figure 2: The user working in a conference room

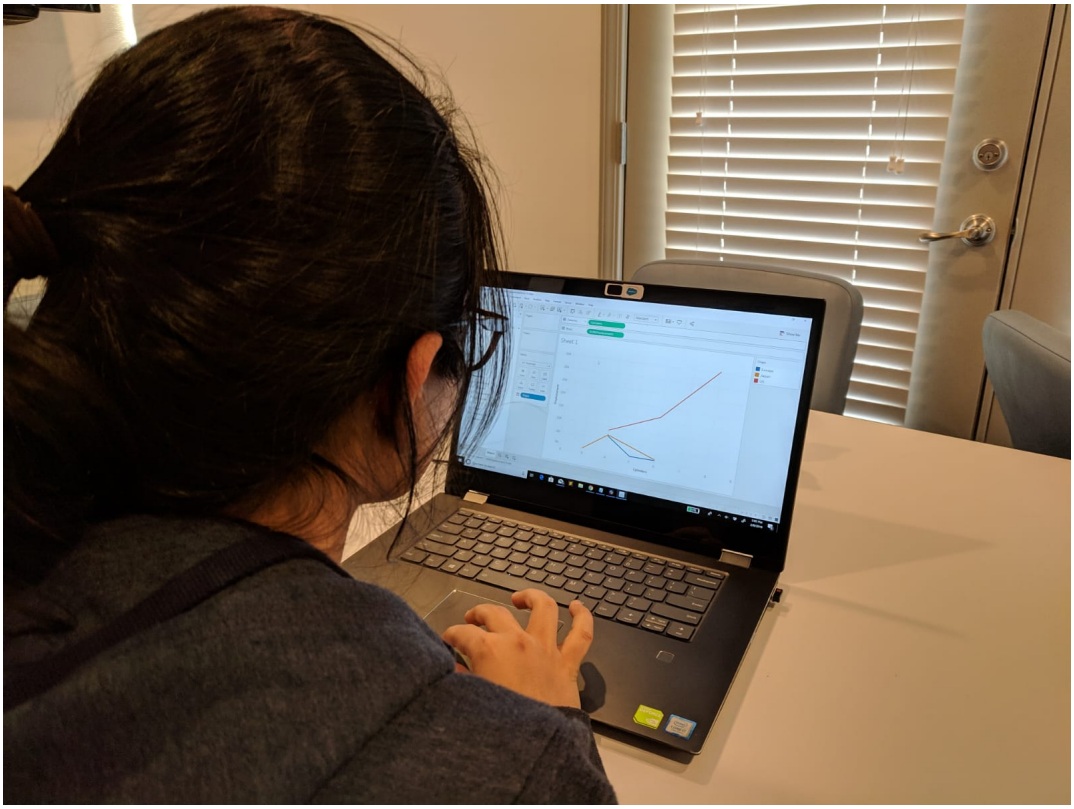


Figure 3: User working with Tableau

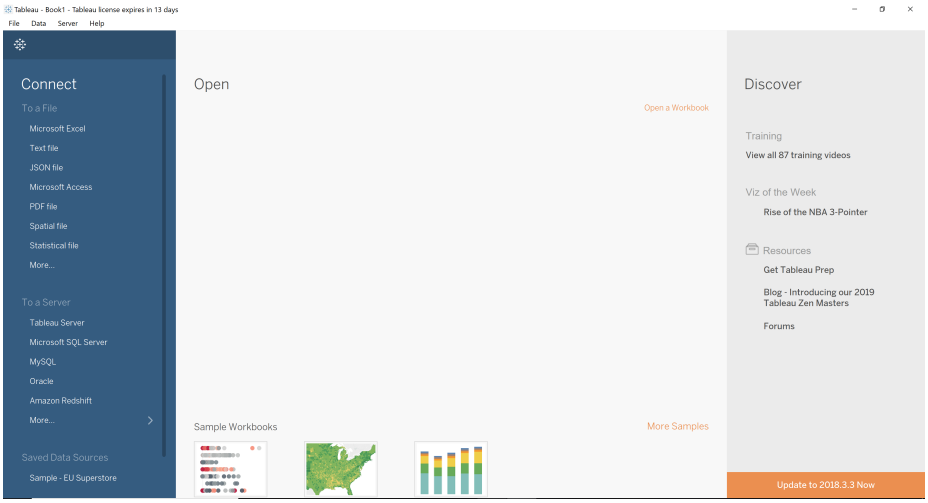


Figure 4: Upload Data File page

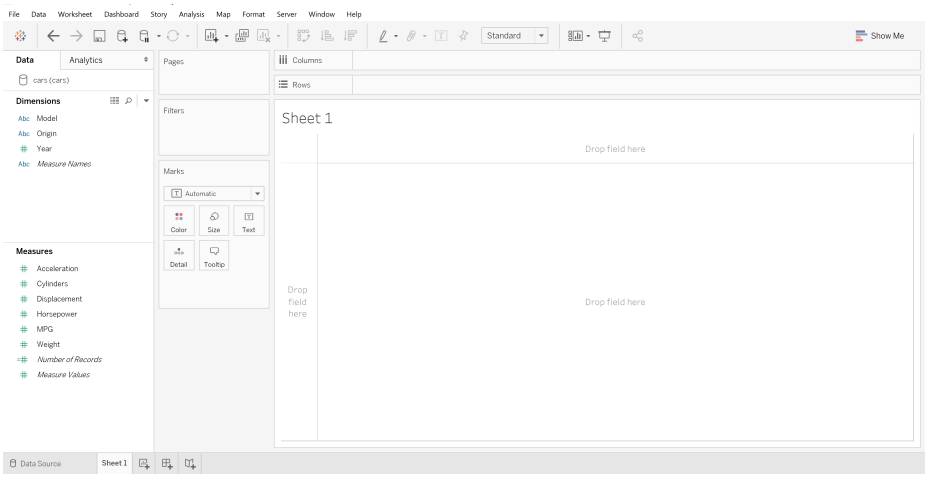


Figure 5: Dashboard Workbook

Step 5: Do the above specified task of Session 2. (I was observing as well as recording the time taken to do the tasks.)

On running session 1 and session 2, it could provide me with some insights while the user was performing tasks. On uploading the data file, the user struggled for a bit and try to upload data by using the conventional file menu from the top rather than going for options given on the side as seen in Figure 4.

Secondly, the transition to workbook after uploading data was slow as the user seemed to be a bit lost about how to get to the worksheet after uploading the data file. On reaching the dashboard I asked her to tell me about her views and what she understands from several buttons. According to her, "dimensions" were independent variables, "measures" were dependent variables, "color" allows to show different categories, "size" to change the width of bars (she was making a bar chart) and "show me" to give different representations.

While doing the tasks of session 1, firstly tried using options on clicking the variables rather than dragging them. For the Task 1, the user gave a wrong answer at first because of the default operation of sum on weights. A lit later she realized that it was the wrong answer. The user keeps scrolling up and down to find the highest value and mostly preferred a horizontal bar representations. I observed she didn't try to use any other chart except for bar charts. For Task 3 of sorting, she was switching between horizontal and vertical representations continuously to give the answer. While doing Task 4, she seemed to struggle a lot and could not complete it. She was looking for options by right clicking on the variable but could not find anything by which she could color the bars based on country. After a ten minute break, I started session 2. In terms of time, it was much better as in Session 1 for Task 1, the user took 49.64 seconds and for task 2 took 59.32 seconds. However in session 2, for task 1 user took 44.75 seconds and 2.45 seconds.

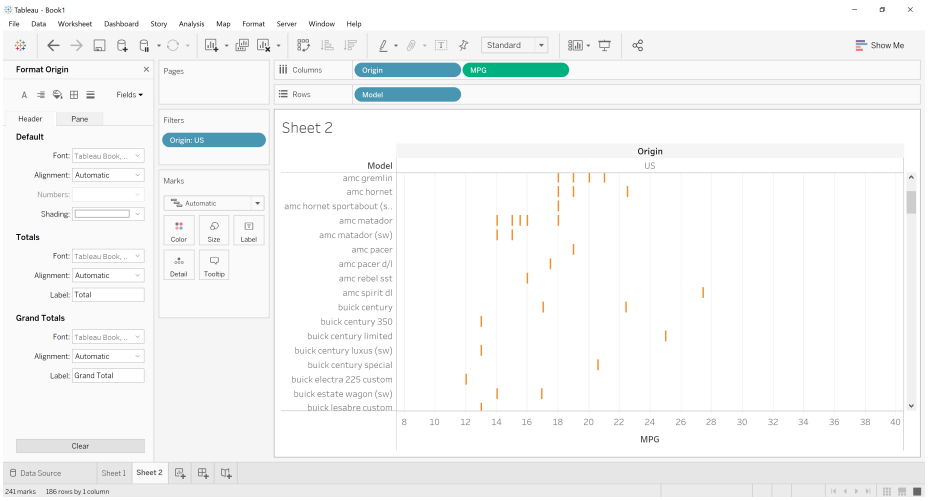


Figure 6: Snapshot of a chart made by the user

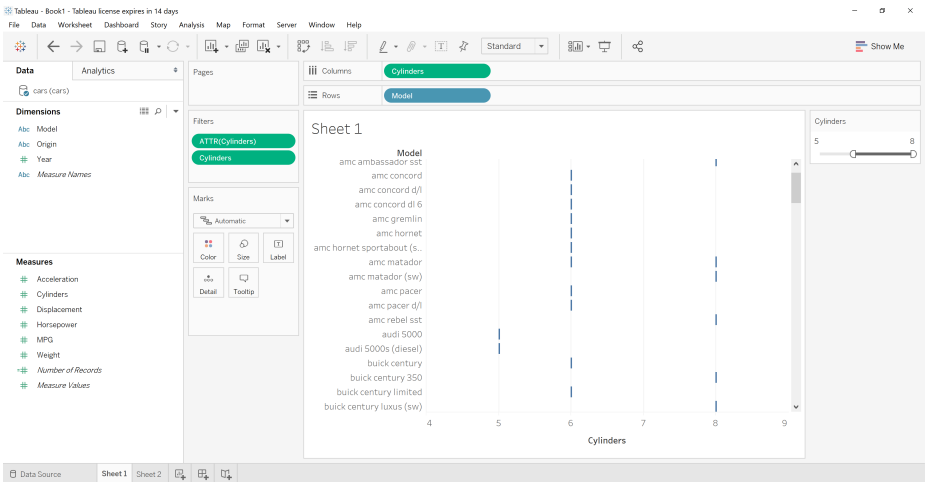


Figure 7: Dashboard Workbook

5 Design Implications

On observing Session 1 and Session 2, I realized on an abstract level that it is very important to teach uses of different types of visualization as the user was stuck with Bar charts and didn't try to use other options. The meaning of "size" and " color" could be different in terms of visualization. The user could not make that size could be a visual encoding for the data but took it as an aesthetic addition. Secondly, while uploading the data file, the user has follows conventional style of uploading file by going to the "File" Option and clearly misses out the different data source options given at the side. I feel it could have been more clear if the options would have been placed in such a way that it more visible. "Drag and Drop" is not something that comes initially in the mind of the users.