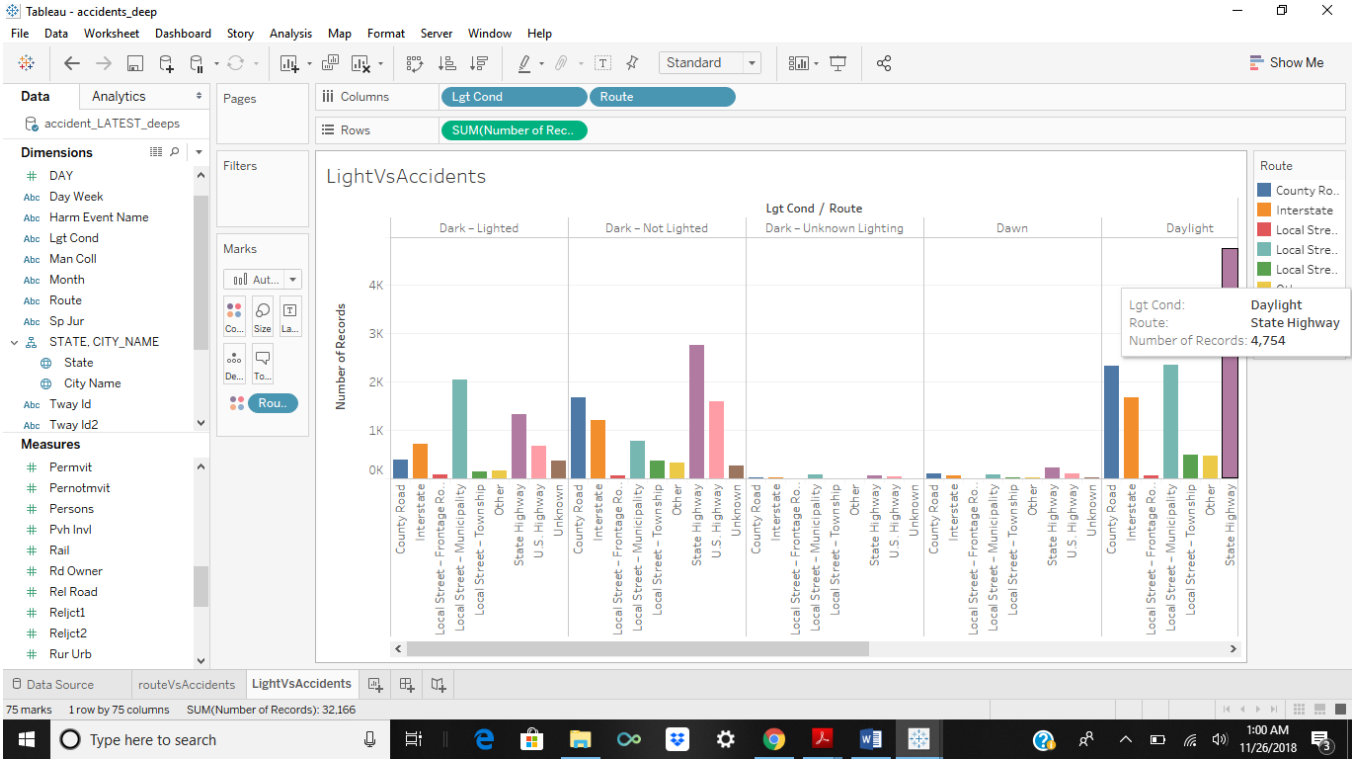
Data Visualization with Tableau

Meetu
#801035153

Task 1: Data is taken from Kaggle - https://www.kaggle.com/nhtsa/2015-traffic-fatalities

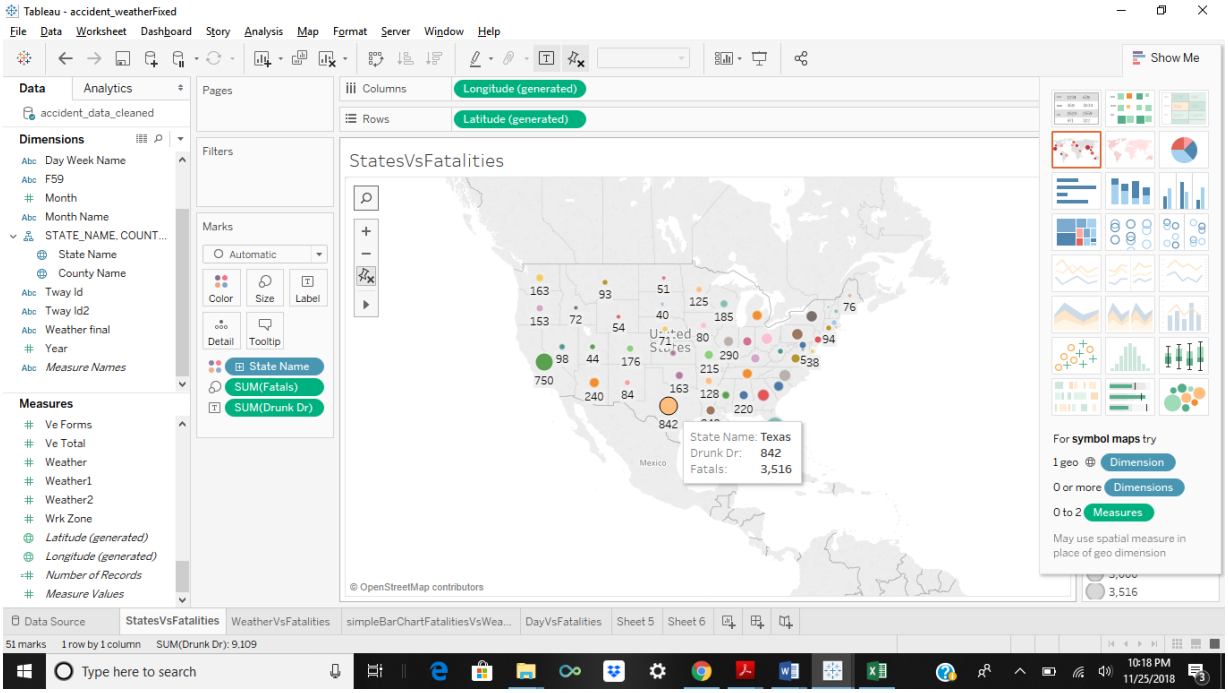
We will look for the answers to below questions via visualizations in Tableau.

Q1: How light condition, route and number of accidents are related?

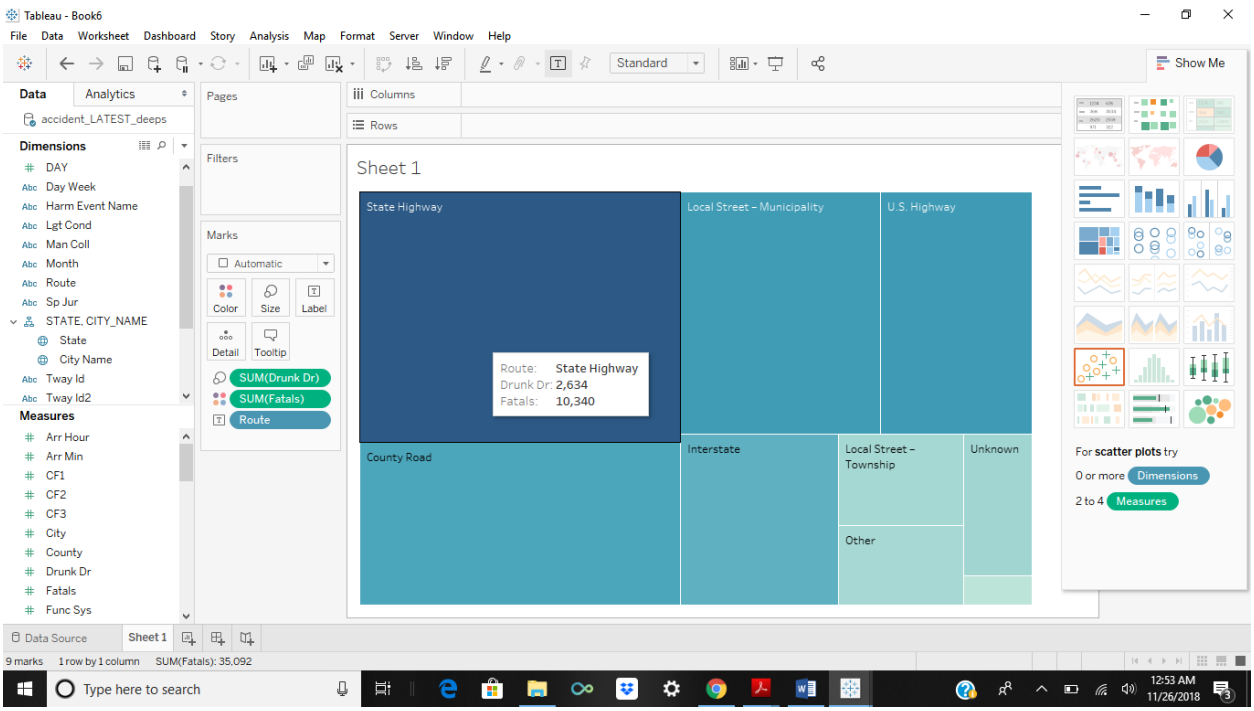


The above graph shows number of accidents as per the light amount and the route. Surprisingly, number of accidents are more during the daylight hours. This could be due to negligent driving involving use of mobile phones, tiredness, not taking enough break time etc.

Q2: What areas are more persistent to traffic accidents?

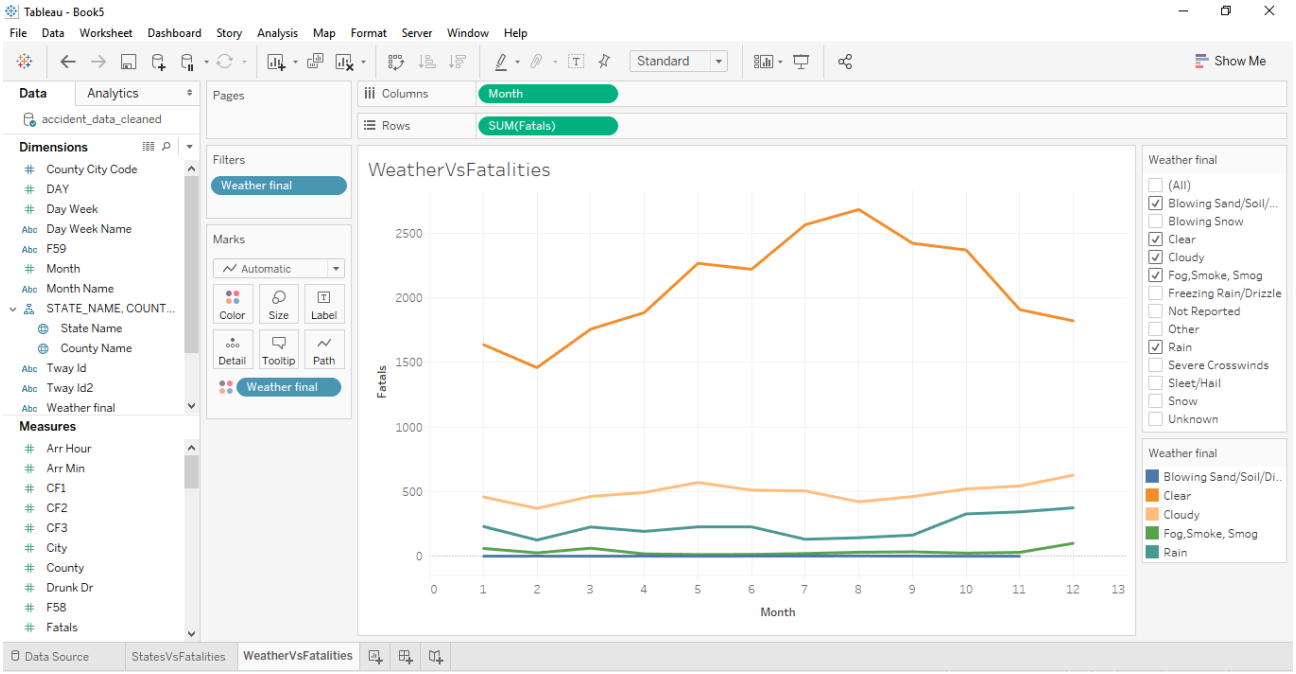


From the map above, we observe that the area Texas has maximum fatalities (3516) and Drunk Driving cases of 842. Different colors are used to show the states, Size of circles represents the number of fatalities and we have used text to display the number of drunk driving cases.

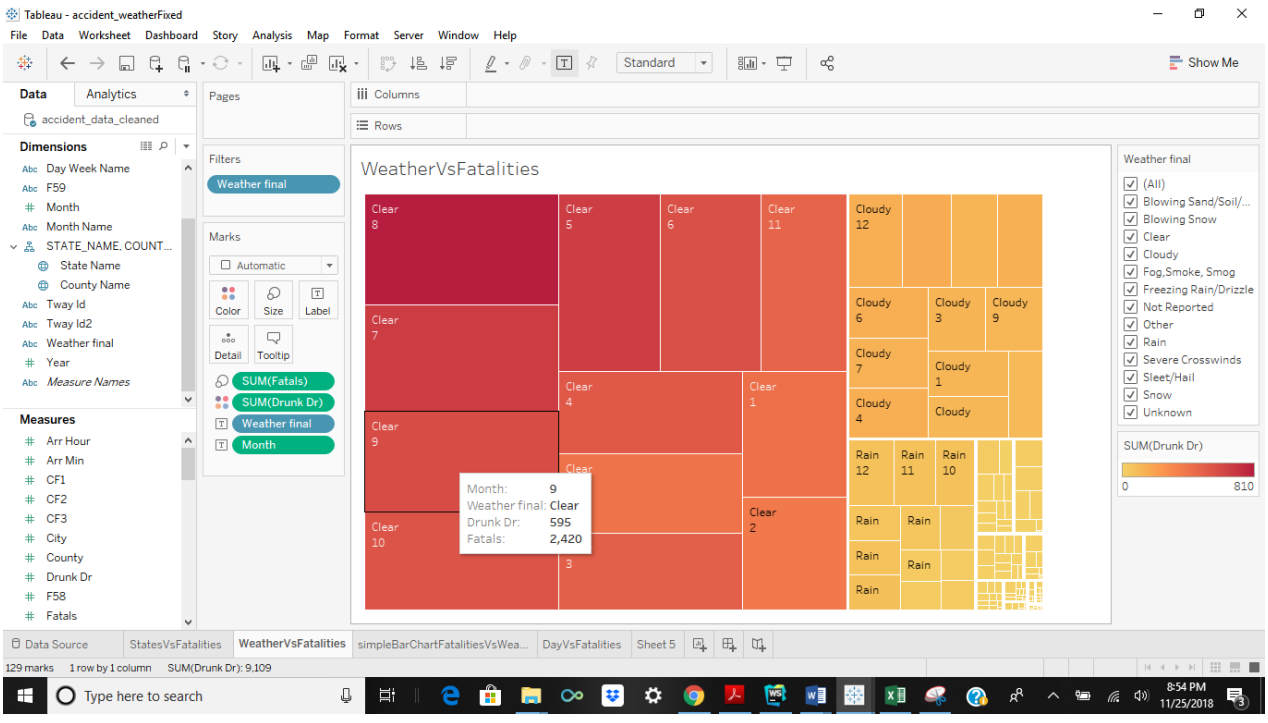


Here in the tree map we observe that on the state highways number of accidents are more as compared to any other kind of route.

Q3: Identify if there is trend between weather and the accidents.

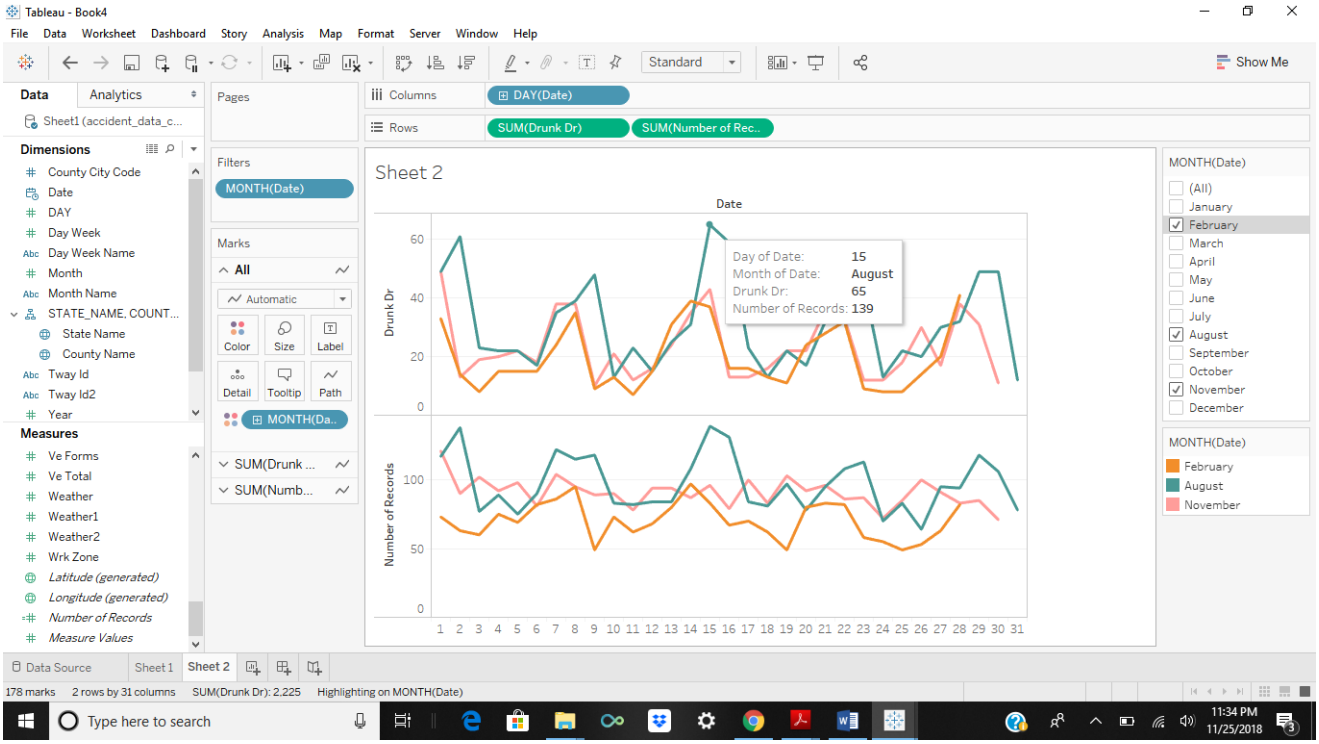


We can use line chart also. We used Color and filter to show various weather conditions. Another dimension is Month. So, the above graph shows line chart for the fatalities occurring in different months, under different weather condition. Surprisingly, most of the fatalities due to accident occur in clear weather condition.



We can use Tree map also. Dark green color shows maximum fatalities and lighter color shows lesser fatalities. The label is for the month (represented in numbers) and the size of the rectangular box depicts the fatalities. From the above tree map, we clearly see that there are more drunk driving and fatalities in August under clear weather.

Q4: Which day of the month most crashes occur?



The line chart above shows that the maximum accident cases occur on 15th August – total number of accidents are 139, out of which Drunk driving cases are 65.We can also drill down to hours or can go up to year or quarters level information.

Task 4: We had thought of creating a parallel plot, which we could not generate straight forward from its "Show me" display panel. In Show me display panel various kinds of line charts, pie charts, Maps, Scatter plots and many more visuals are shown but not the drag and drop feature for parallel plot.

Task 5: We have used accidents.csv dataset. Once the data is loaded easily to the tableau, it helps to create easy visualizations with interesting findings or the underlying trends in dataset. With Tableau I was able to create visuals with simple drag and drop of columns from dataset and using filters, color schemes, text labels etc. This helped me to gain better understanding of the data and figure out interesting insights. For example, in our dataset we observed that the maximum number of accidents occur in clear weather instead of rainy or snowy weather. We also noticed that surprisingly most accidents occur in daylight and that too not due to drunken driving.

Task 6:

- Tableau can connect with any kind of data file such as txt, csv, xls etc. It can also integrate with multiple datafiles using joins as we do in Sql. Tableau can be easily integrated with other software such as R using Rserve library from R and with other data sources such as Hadoop using collection of APIs in WDC – Web Data Connector.
- Once the data is loaded, then it's just a matter of drag and drop of columns to see the relationship between various columns of dataset. Different visuals are supported in Tableau in its "show me" window panel according to the type of variable. Using those visuals, one can create an interactive dashboard and can add storytelling feature too. This helps business users to understand the data quickly and make decisions effectively.
- With Tableau server, there is no version control over the published reports. Once the report is published on the server and overwritten, there is no way to recover the old version of it.