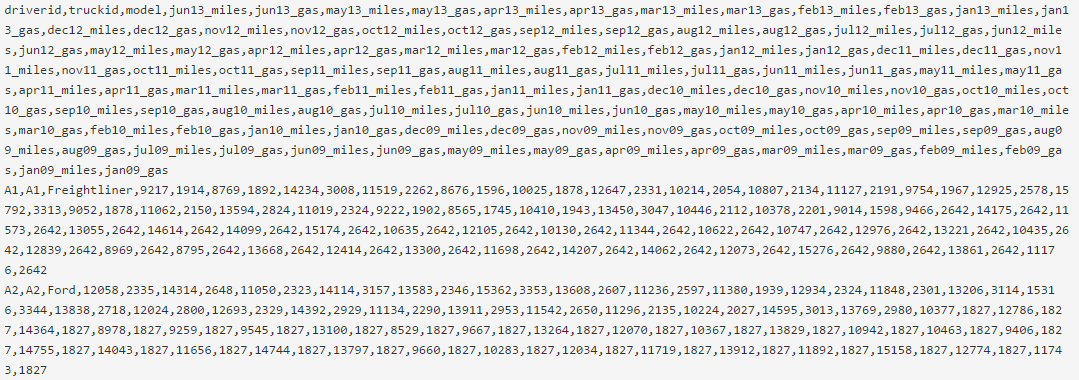
**Geolocation and the wonders it can do!**

Geolocation data is enormous, and that’s the challenge! Hadoop being a cost effective tool to store and process huge amounts of data, benefits us to store and process the geolocation data by providing a business value to us by tracking the assets and predecting the behaviour. This can be applied to enable optimization and hence reducing fuel costs and increasing driver (and asset’s) safety.

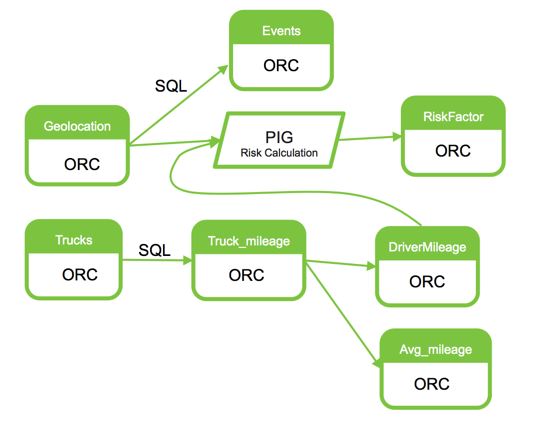
Our sample data looks as following:



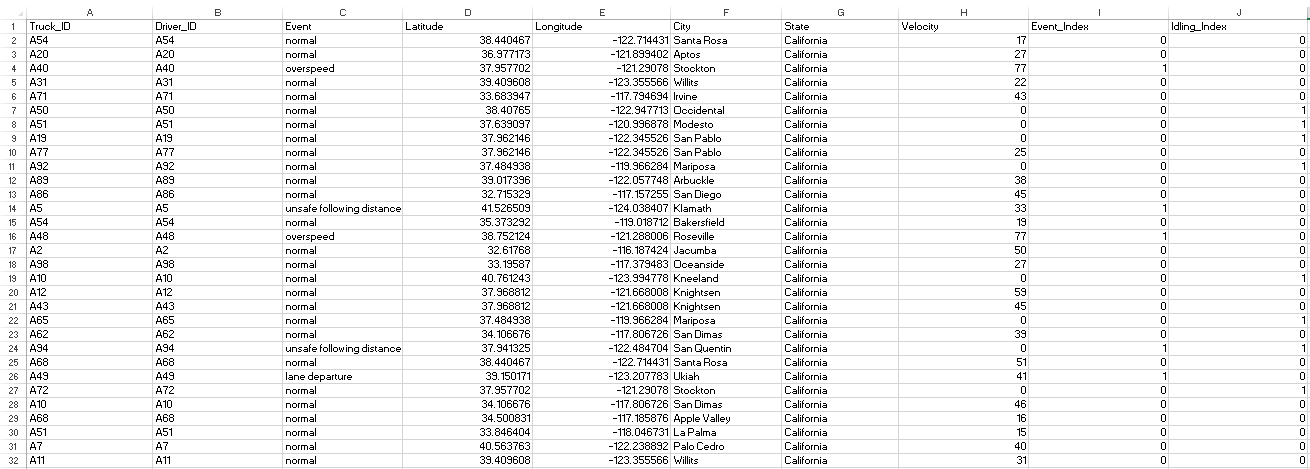
And



We can drill the data using components like Hive, Sqoop, Pig and HCatalog. A simple diagram here shows how the data is being processed:

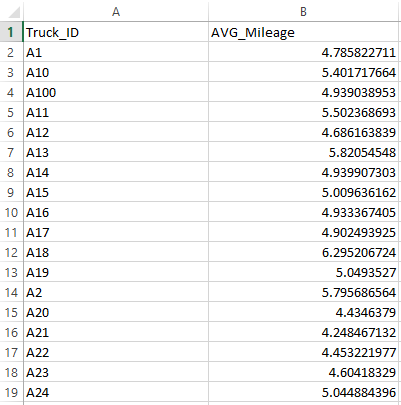


Load the data into excel sheet to visualize:

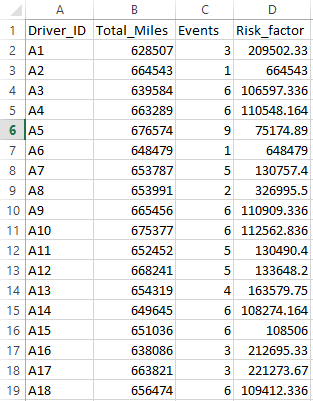


We have categorised the data of 100 different drivers, driving all over California as shown above.

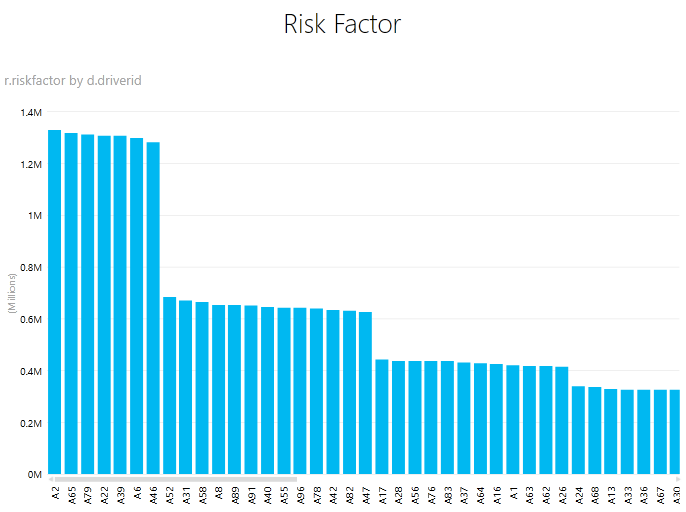
Now, we have a table with average mileage of each driver to predict their individual performances:



We have categorized different events like overspeed, tailing etc. and gave them each an index. The risk factor is calculated for each driver to analyze his performance over few years. We have data ranging from 2005 to 2013.

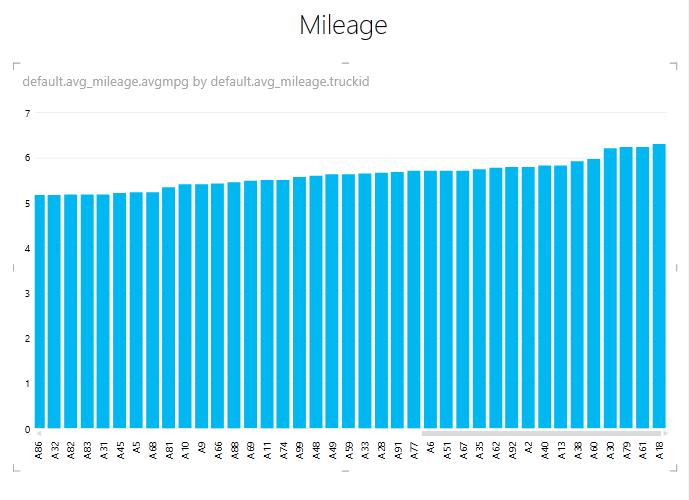


So, now by reflecting the data in Power view, we can arrange the data in order to see that there are 7 drivers with severe risk factor. This indicates that the drivers need special attention. They may be replaced with new and efficient drivers, they may be trained to avoid risks, their work hours could be reduced to increase safe and efficient driving. We want our assets to be safe and hence we would like to reduce risk events.



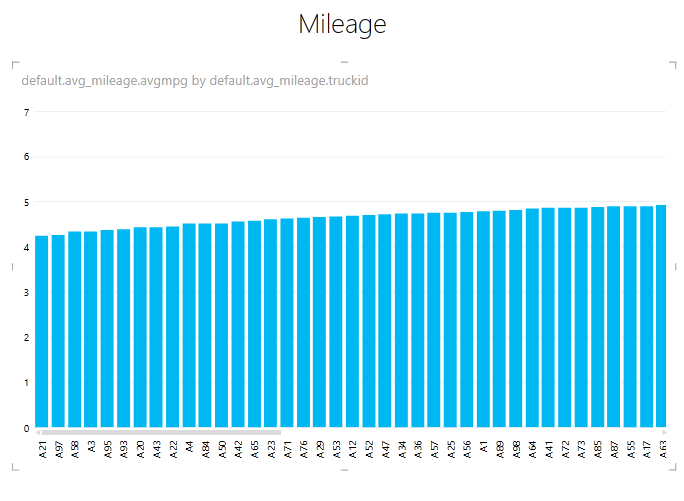
Based on the graph, we can see that drivers A2, A65, A79, A22, A39, A6 and A46 require special attention. This shows that we can narrow our search with precision.

Now, from the calculated mileage table, we develop a graph to view the performance of individual trucks. More mileage indicates better performance. The trucks with good mileage are:

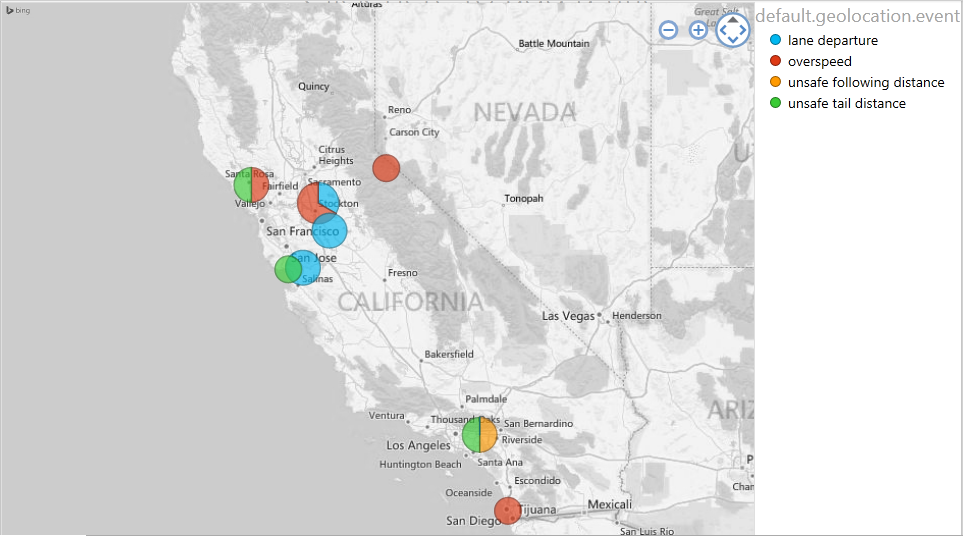


Looks like the trucks A30, A79, A61 and A18 are showing good mileage.

Now on the upside, we can see which trucks are showing bad mileage so that we can get them fixed.

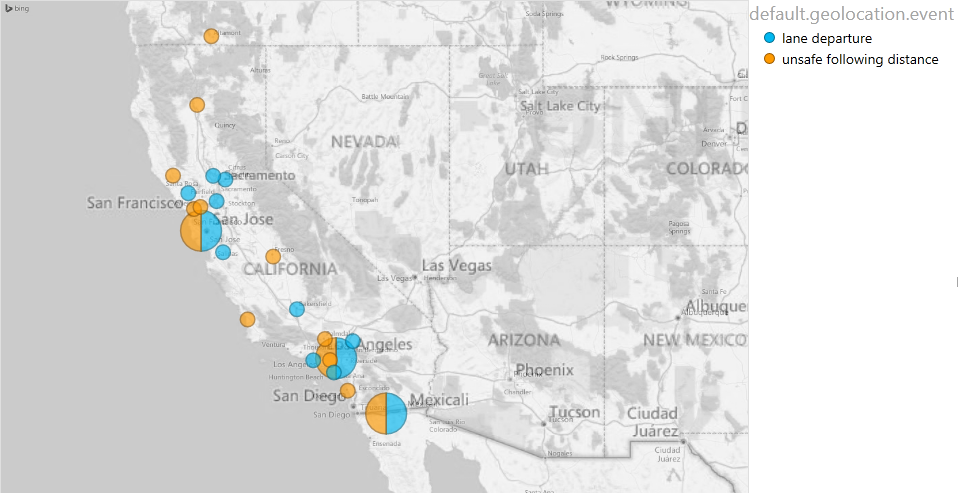


Based on the geolocation coordinates, we can view the performance of the drivers across California as shown:

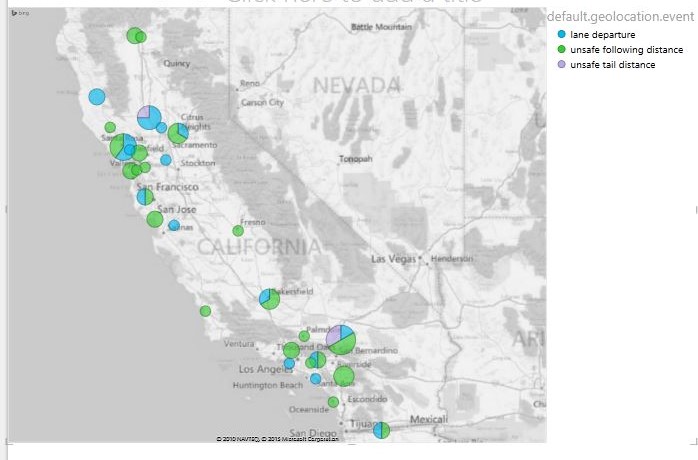


We can see the unsafe events happening across the state. We can observe that our drivers are over speeding in certain locations. We care about our drivers and more importantly our assets. Safety of our customer’s assets is of a higher priority.

Next, we can narrow down our search to a single driver and view his overall performance over time across the state. By reducing the work hours we can ensure that our driver has good sleep and time to relax before he picks up work the next day. This way, we can improve individual performance of our drivers. Good sleep leads to good and relaxed driving and thus ensures safe driving.



We can also find which trucks are idling and wasting the fuel:



**Further:** We can deploy real-time location service in hospitals. This will enable medical staff to locate medical equipment within the hospital in real-time and save time trying to track equipment down.