

MIS 6346.502
BIG DATA PROJECT

TRUCKS RISK FACTOR ANALYSIS

Sushma Reddy Vutla

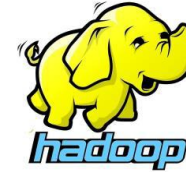


PROBLEM STATEMENT

- To utilize the provided HDP tools to process truck fleet data, including geographic information, vehicle metrics, and event logs.
- Analyze the dataset to identify high-risk factors such as driver behavior, geographic locations, and vehicle conditions.
- Develop strategies to mitigate identified risks and improve overall trucking safety to address the objective of reducing accidents caused by large trucks.
- Generate analytics reports and visual presentations from the analyzed data to provide informed decisions with actionable insights to enhance trucking safety and identify dangerous commercial truck drivers nationwide.



DATA PROCESS FLOW



- Transfer *geolocation.csv* into Cloudera VM.
- Load the imported file into Impala table and integrate HDFS instance.
- Load the tables into Tableau via ETL transactions.



WORKFLOW SCHEMA



geolocation (default.geolocation)+ (default)

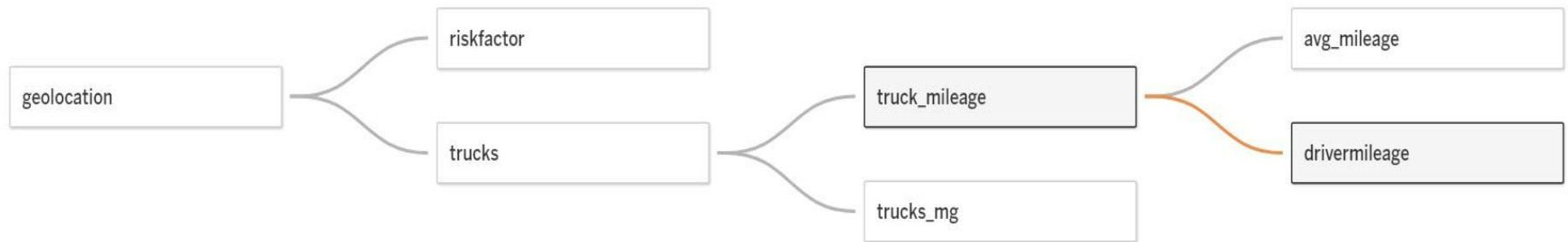
Connection

☒ Live

☐ Extract

Filters

0 | [Add](#)

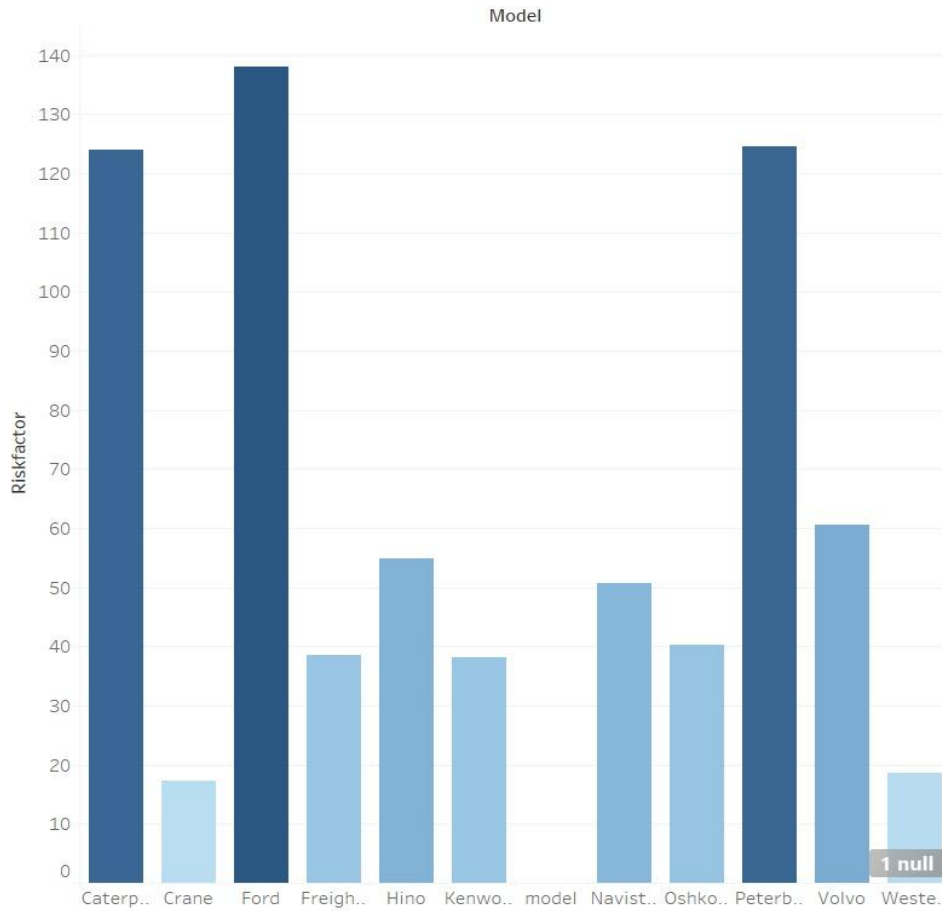


Risk Visualization – 1

Sum of risk by Model



Risk Factor by Model



SUM(Riskfactor)

17.3 138.0

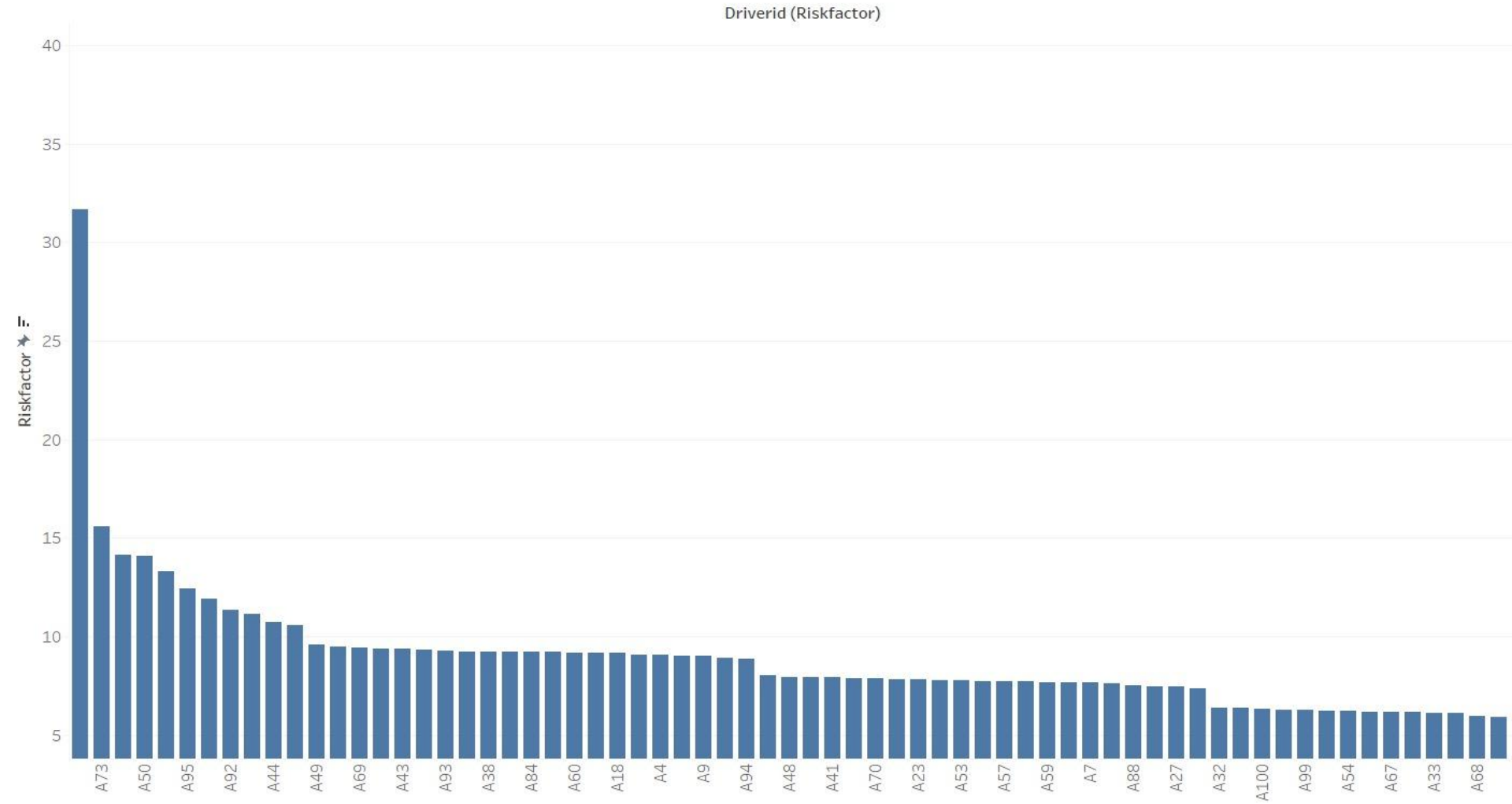


Risk Visualization-2

Drivers and Risk factor



Risky Drivers

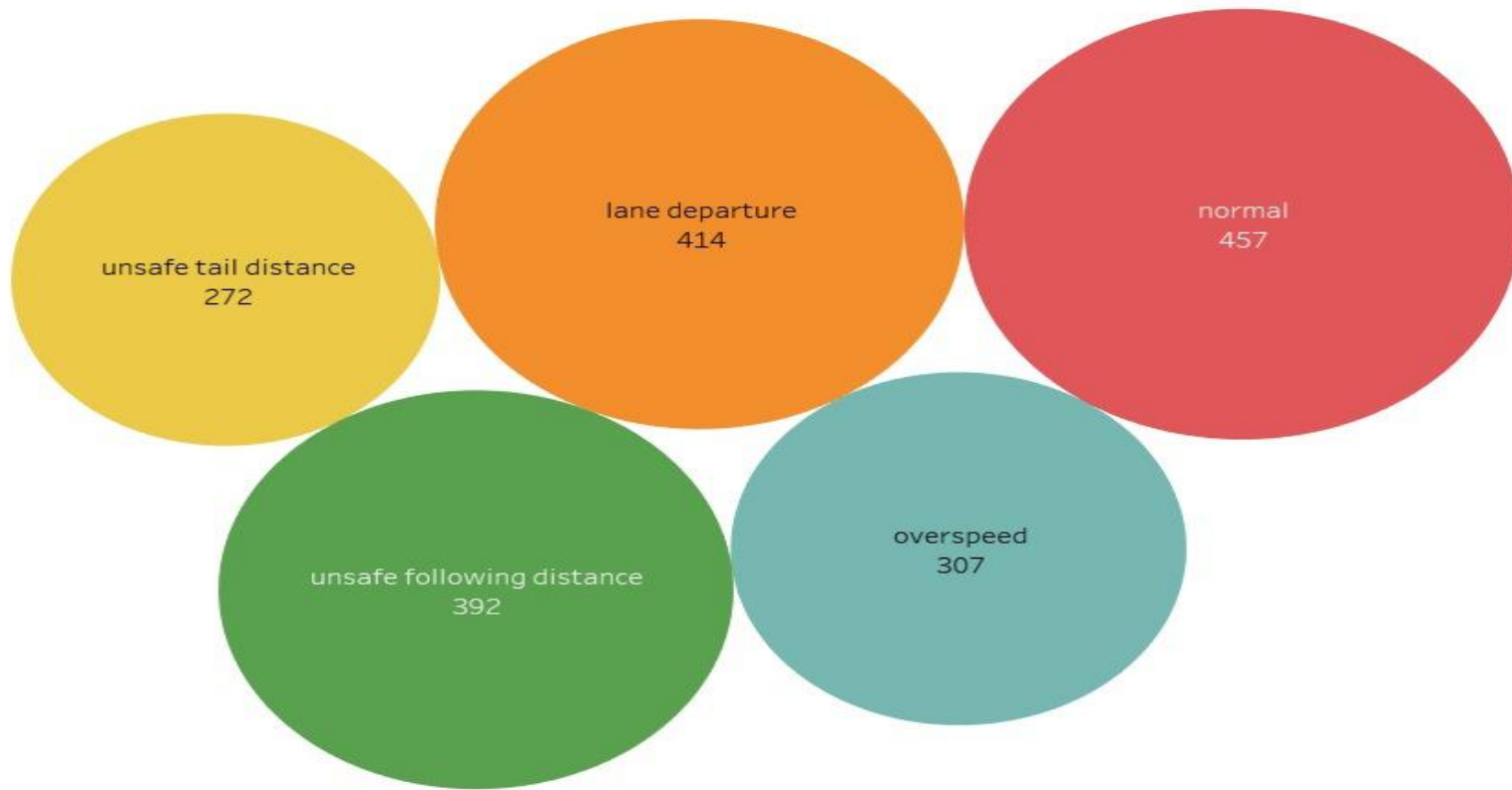


Risk Visualization – 3

Risk due to events



Count of Events

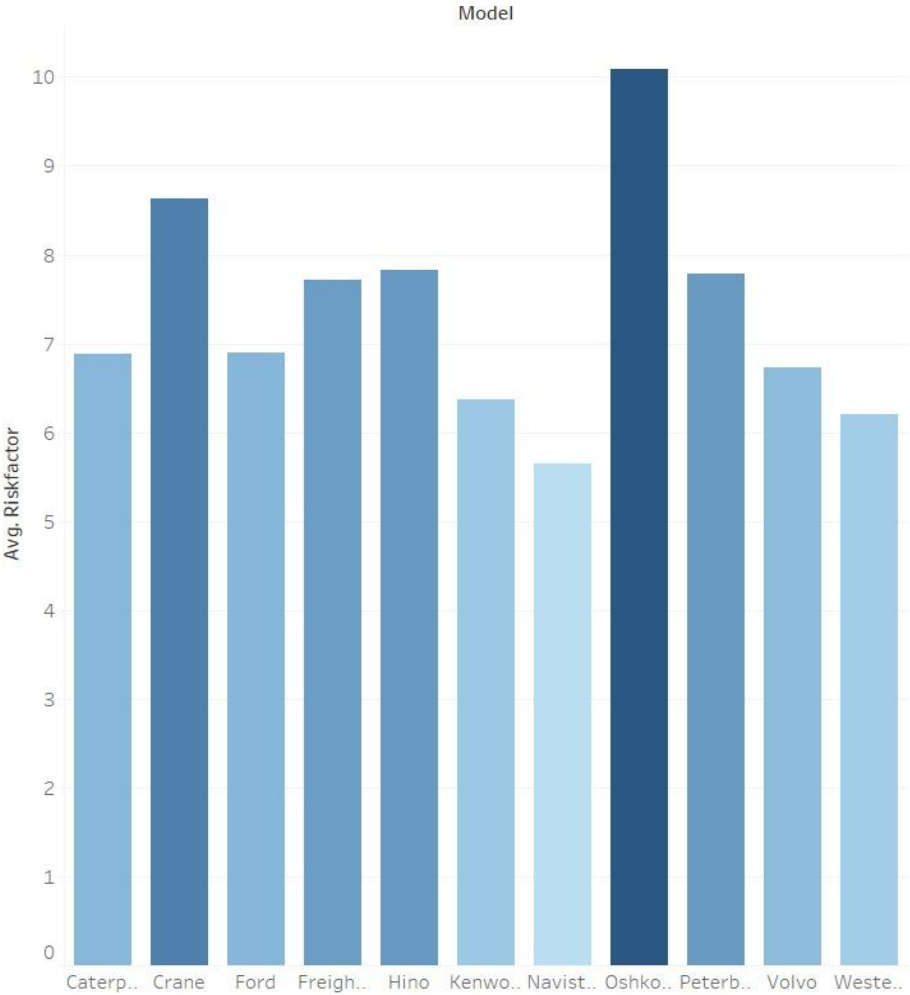


Risk Visualization – 4

Average Risk by each Model



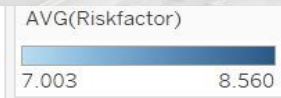
Average Risk Factor vs Model



Risk Visualization – 5 Cities with Risk Factor > 7



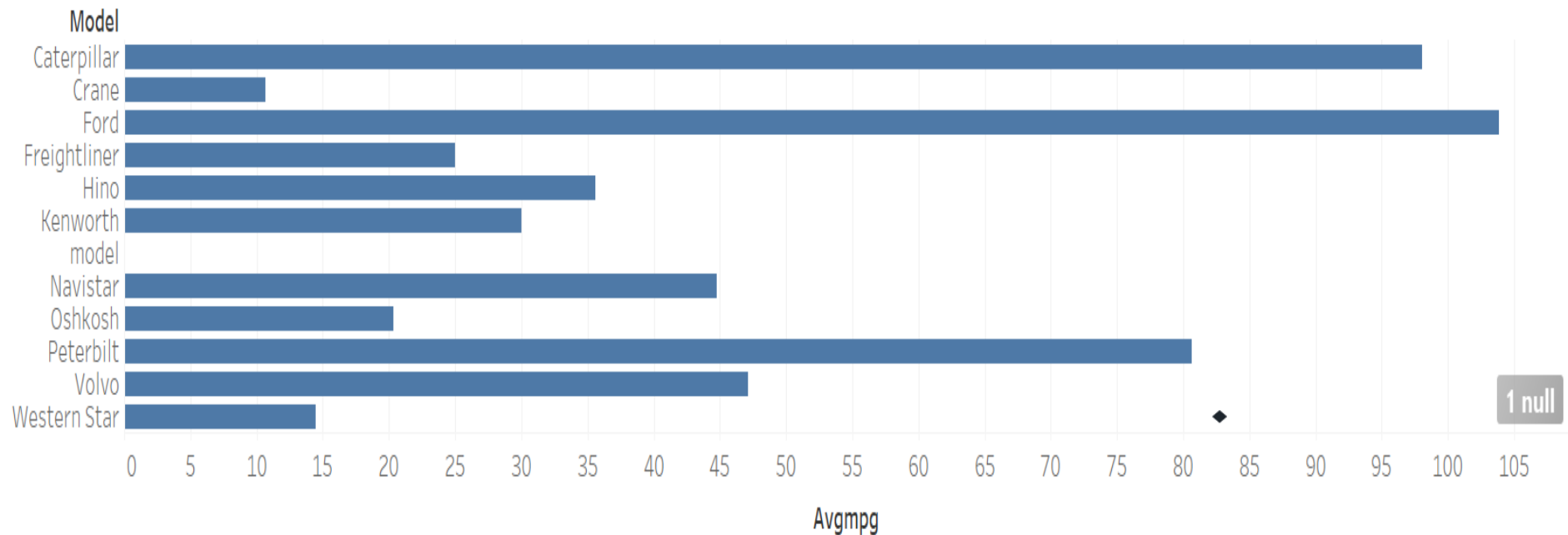
Cities in California with Avg Risk factor > 7



TRUCK MODEL vs AVERAGE MILEAGE



Sheet 1

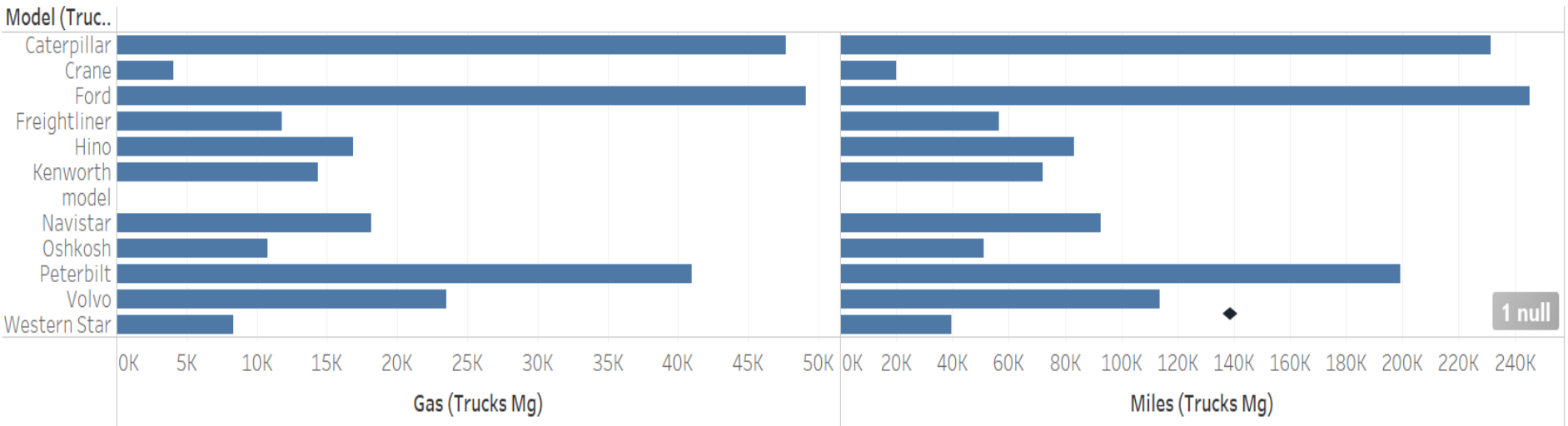


Fuel Efficiency – Mileage covered by truck model per amount of gas loaded



Columns	SUM(Gas (Trucks Mg..	SUM(Miles (Trucks M..
Rows	Model (Trucks Mg)	

Sheet 2

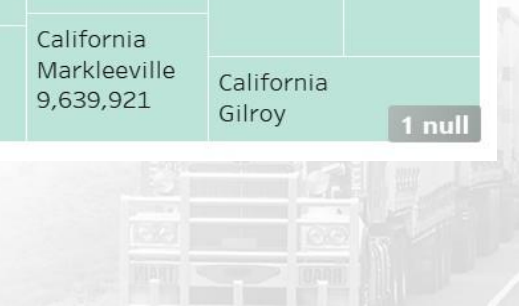


Total miles covered by trucks in each city of California



California Santa Rosa 60,920,756	California Antelope 29,140,838	California La Palma 19,414,451	California San Pablo 16,206,803	California Lodi 16,136,672	California Modesto 16,094,902					
		California San Diego 19,321,328								
	California Arbuckle 23,309,738	California San Quentin 16,315,476								
		California Jacumba 16,284,868								
California Willits 52,597,816	California Aptos 23,243,267	California Irvine 16,266,776	California La Puente	California Santa Maria	California Klamath 9,703,893	California Rough and Ready	California Knights Landing	California Martinez	California Redding	
	California San Fernando 19,563,562		California Kneeland							
	California Stockton 29,267,260	California San Francisco 19,482,477	California Isleton 16,236,401	California Glendora	California Occidental 9,700,586	California Klamath 9,703,893	California Rough and Ready	California Knights Landing	California Martinez	California Redding
			California Bakersfield 16,230,108	California La Quinta	California Mariposa 9,700,273					
California Apple Valley 29,148,107	California San Dimas 19,431,150	California Marysville	California Ojai 9,698,797	California Markleeville 9,639,921	California Napa 9,658,568	California Oceano 9,636,491	California Hollister 9,613,256	1 null		

1 null





THANK
YOU!!

