

Presented by :  
Yash Kumar Jain  
Thi Truc Linh Le  
Qinghua Ye

# Late Delivery Analysis in the US

Big Data Infrastructure and Cloud Computing  
Professor : Patrick BURY  
Paris 7 July 2022



# Hypothesis

In America, Late deliveries are high due to traffic and bad road conditions.

# Dataset

## **1. DataCo SMART SUPPLY CHAIN FOR BIG DATA ANALYSIS**

A DataSet of Supply Chains used by the company DataCo Global was used for the analysis. Dataset of Supply Chain , which allows the use of Machine Learning Algorithms and R Software.

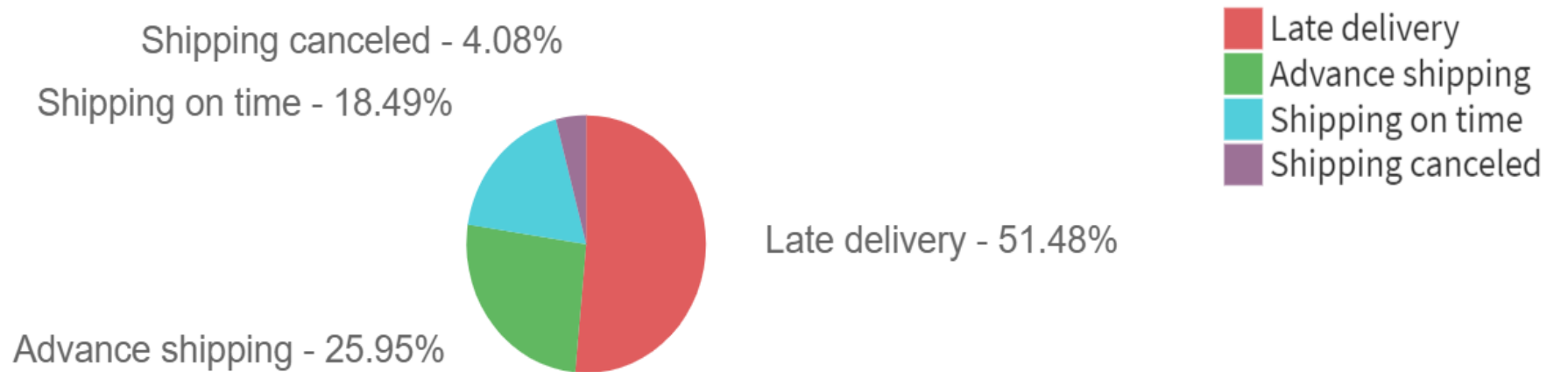
## **2. Road Surface Condition**

A Dataset of some index about the car and road surface, traffic and driving style

## **3. Traffic Condition**

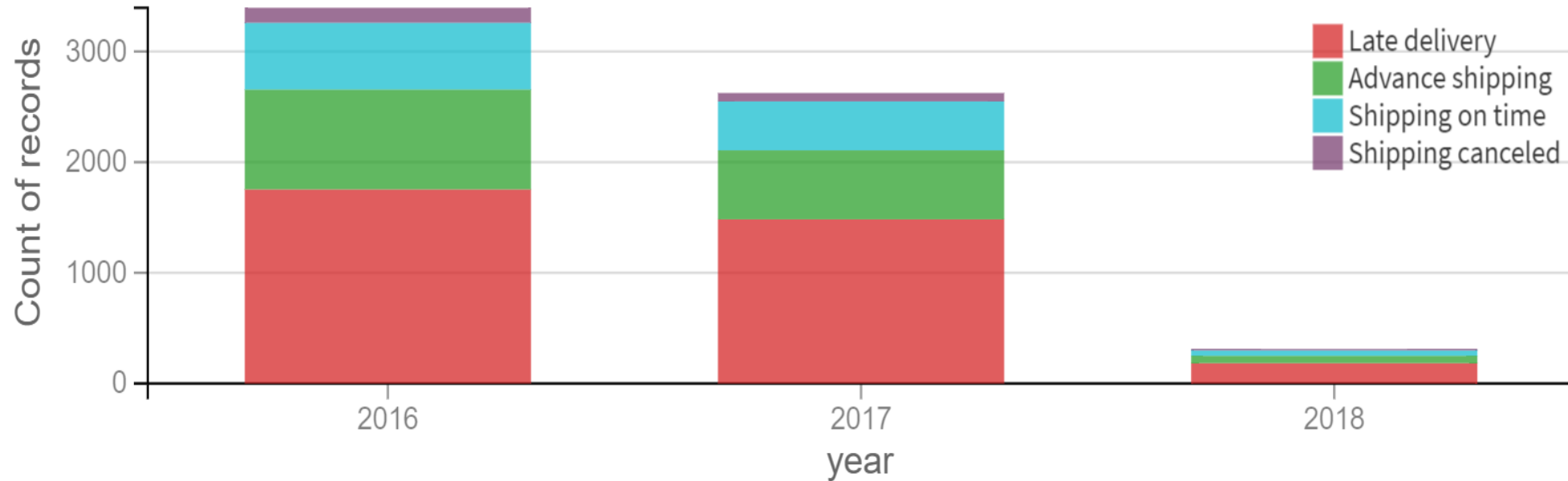
**Tool:** Dataiku, excel.

## Count by Delivery Status



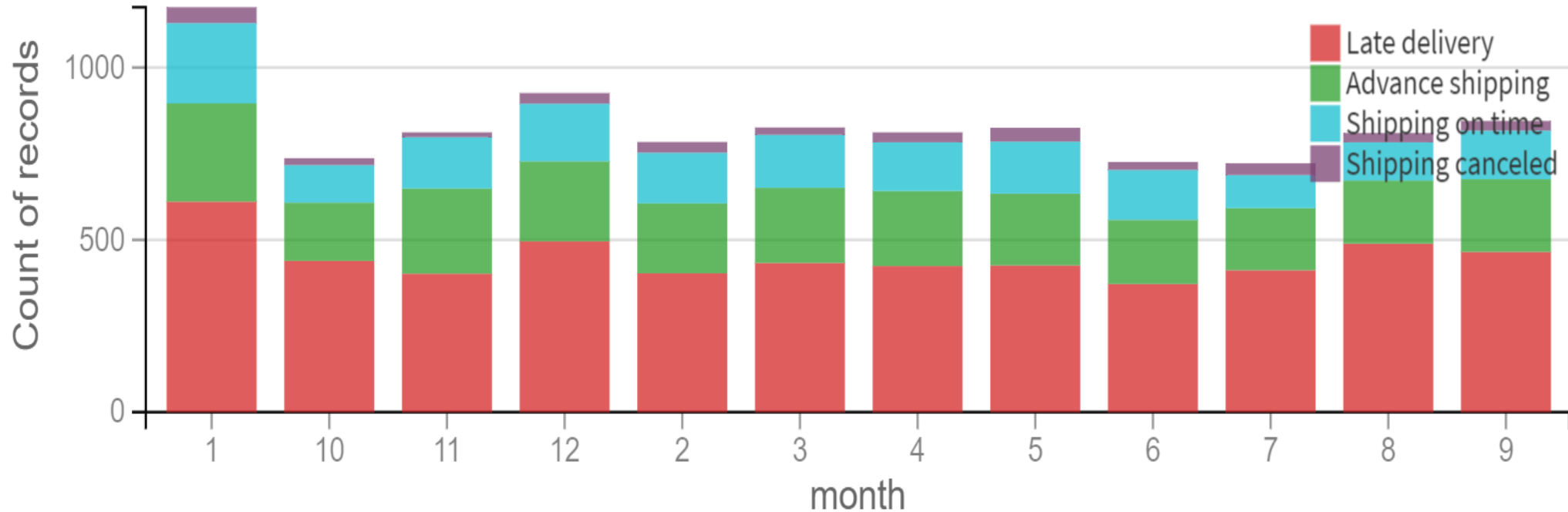
This pie chart shows that there is a problem with the delivery in USA. The percentage of late delivery is very large. It accounts for 51.48%.

Count by year and Delivery Status



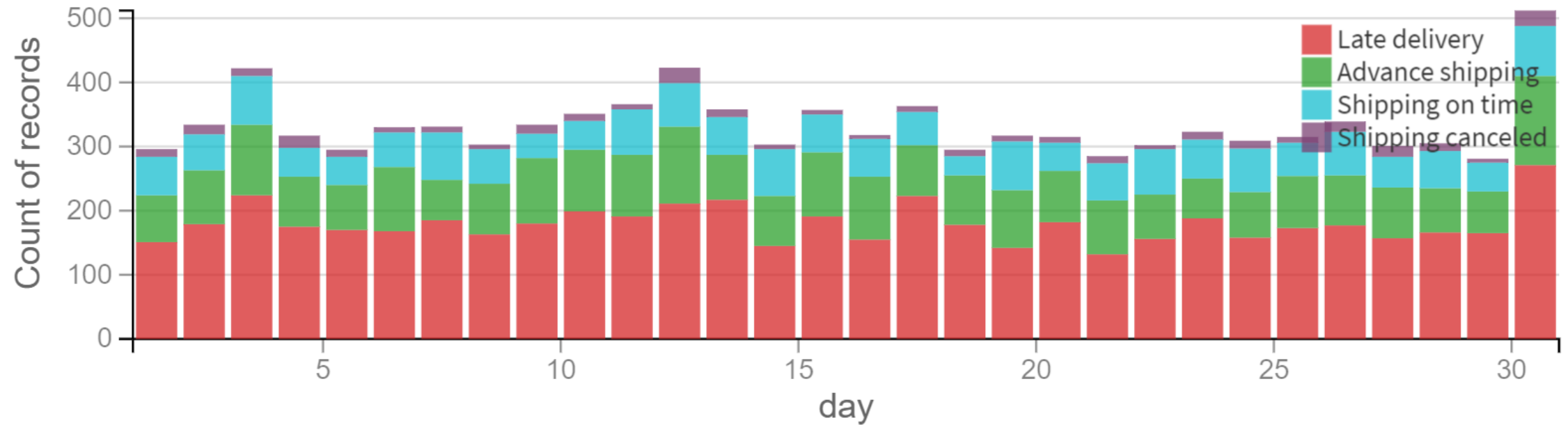
In 3 years 2016, 2017 and 2018 have a very high late delivery rate.

Count by month and Delivery Status



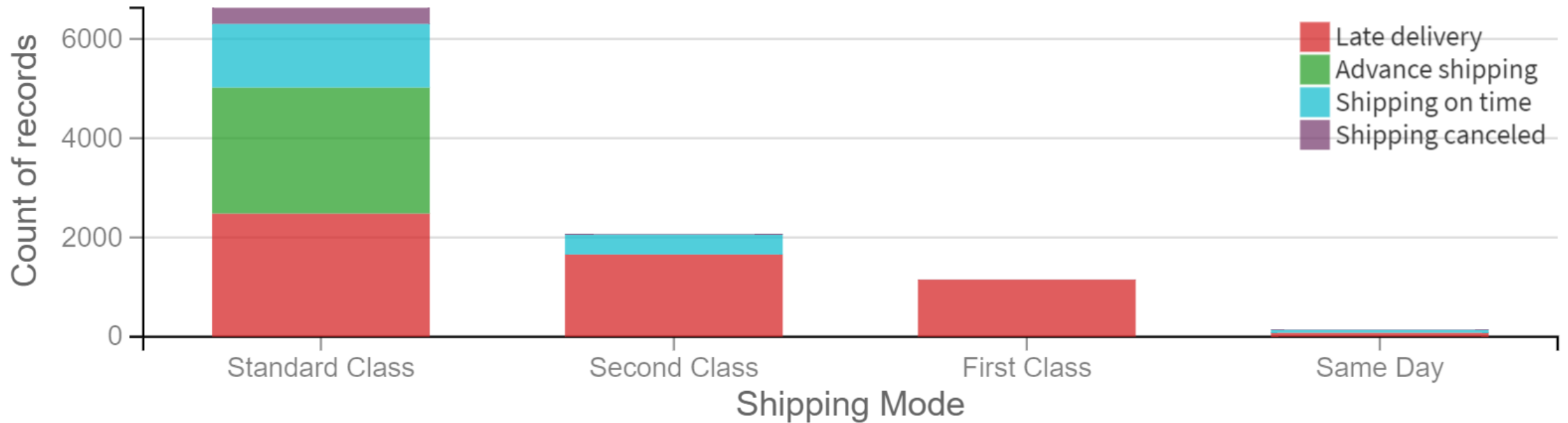
All months have a high late delivery rate. January, December, August, and September these months have the highest late rate.

# Count by day and Delivery Status



We can see that Late delivery rates appear every day of the month.

Count by Shipping Mode and Delivery Status



According to the Graph, We can see that the rate of late delivery is almost 100% if the customer chooses the first class and 80% in the second class. Same day is the best option to avoid the risk of late delivery.



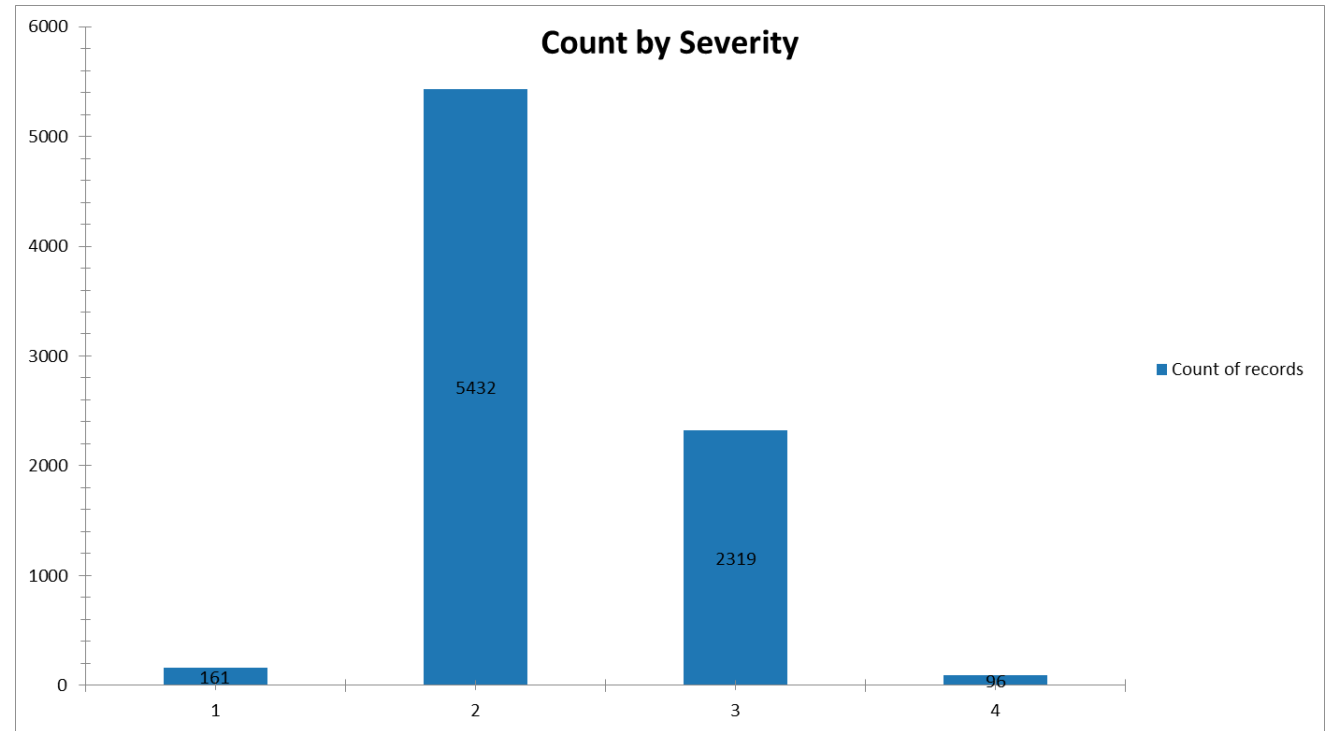


# Traffic and Road Conditions

Let's Assume there might be high traffic and road condition was not good. We took Traffic control dataset and road condition dataset .

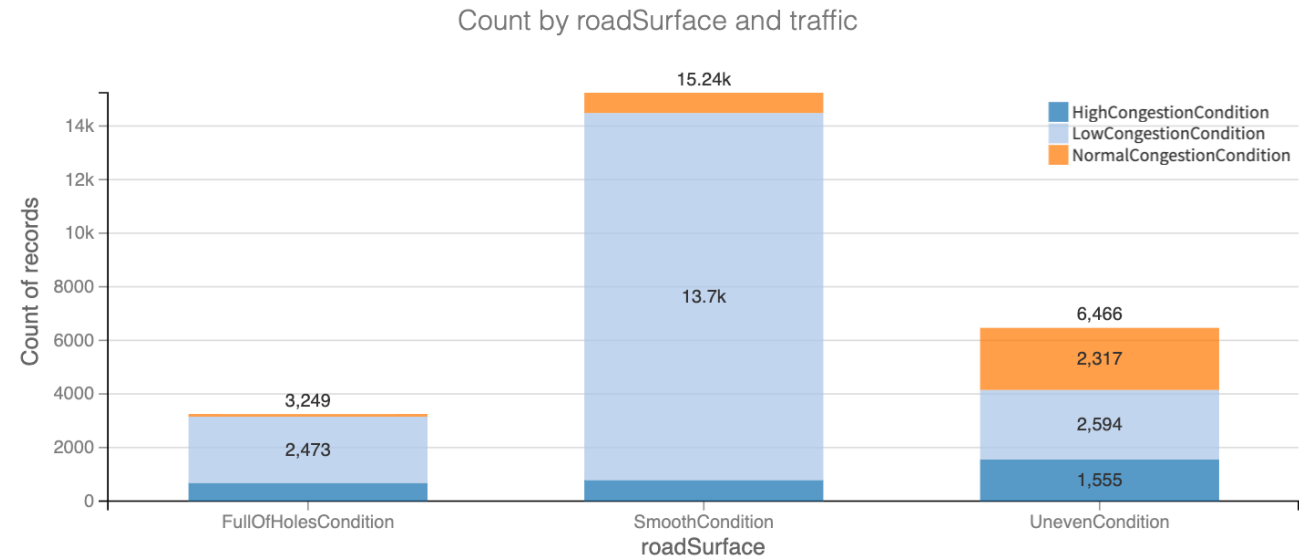
# Traffic Status

- The overall traffic is bit high.



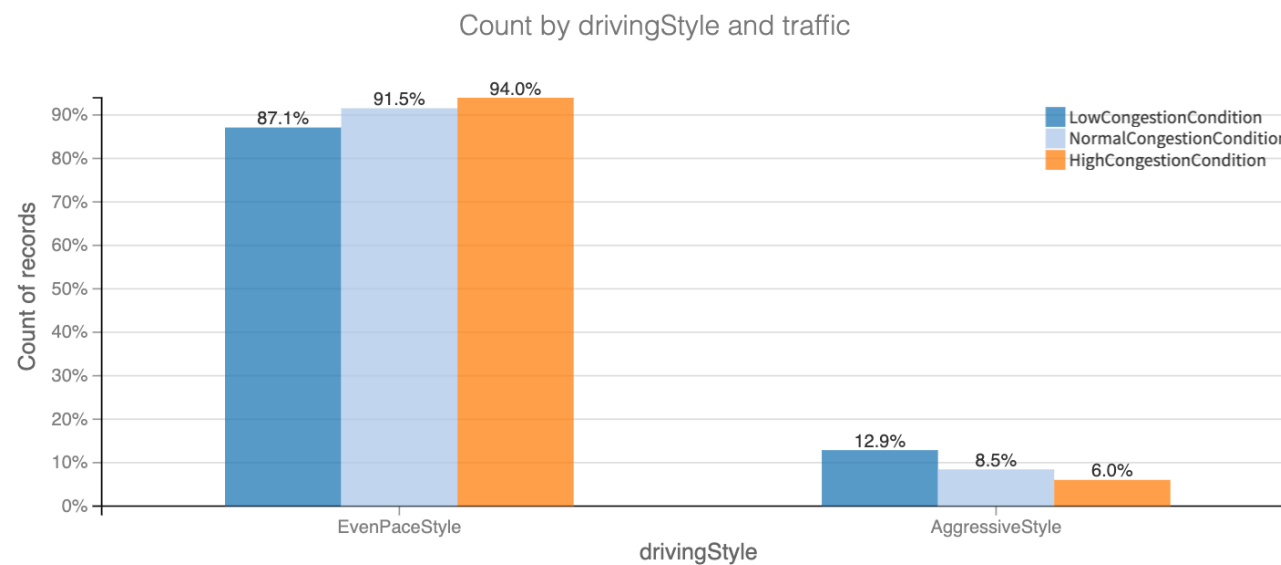
# Traffic condition Over Road Surface

- According to the graph, we can see that the traffic condition is much more better which is in low congestion when the road surface is smooth. That means road surface have influence on traffic condition.



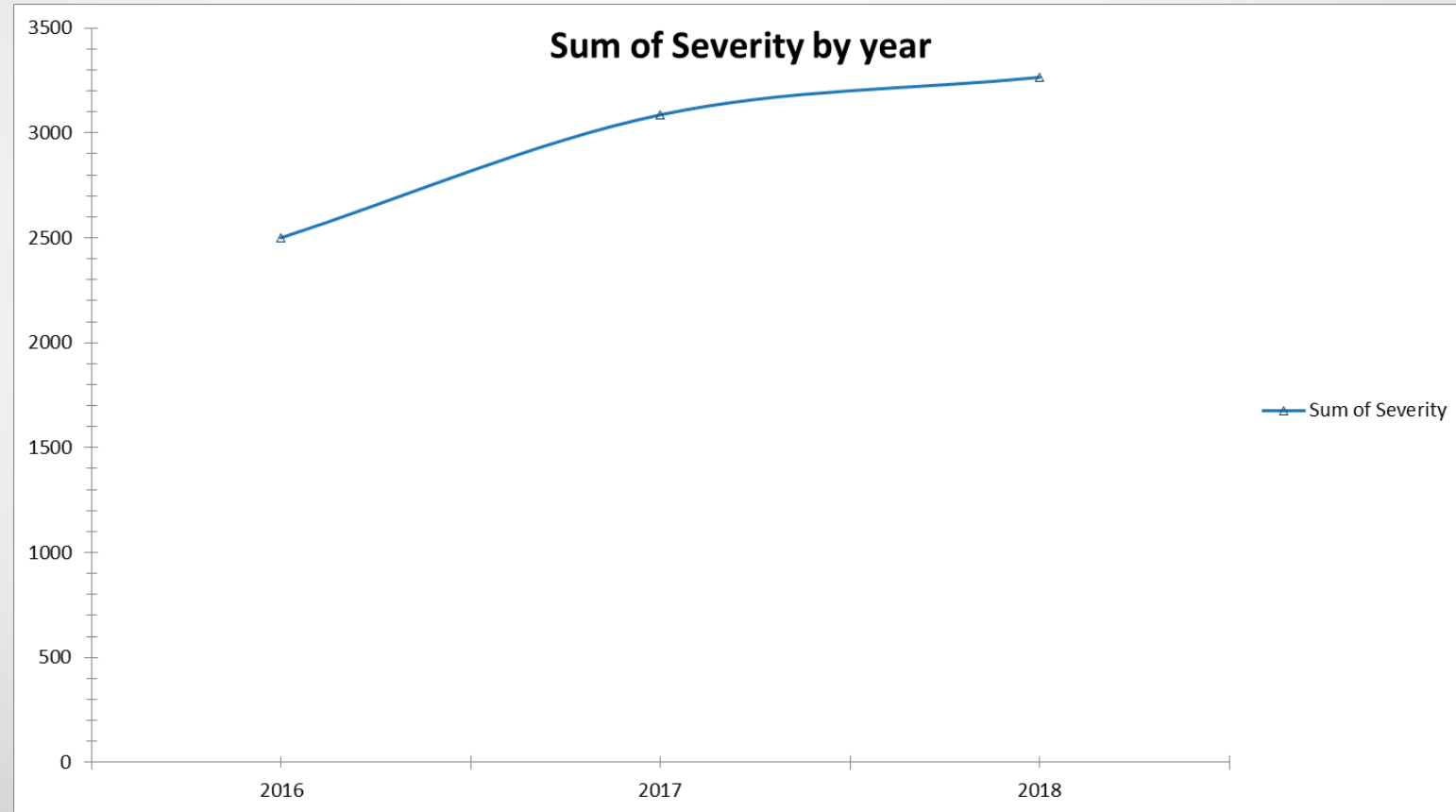
# Traffic Condition Over Driving Style

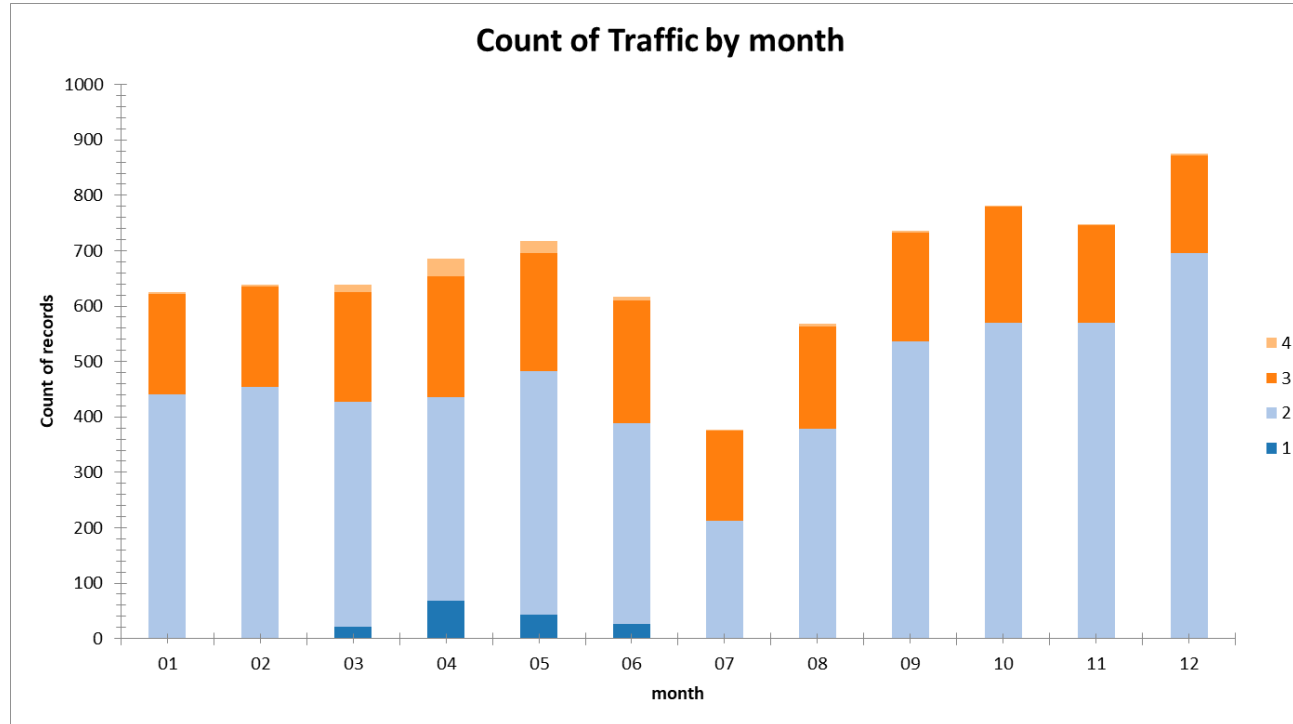
- From the graph, we can see that each traffic condition has nearly same percentages among the two kinds of driving styles. Hence, driving style has almost no influence on traffic condition.



## Traffic condition over years

The highest traffic was found between 2017 and 2018.



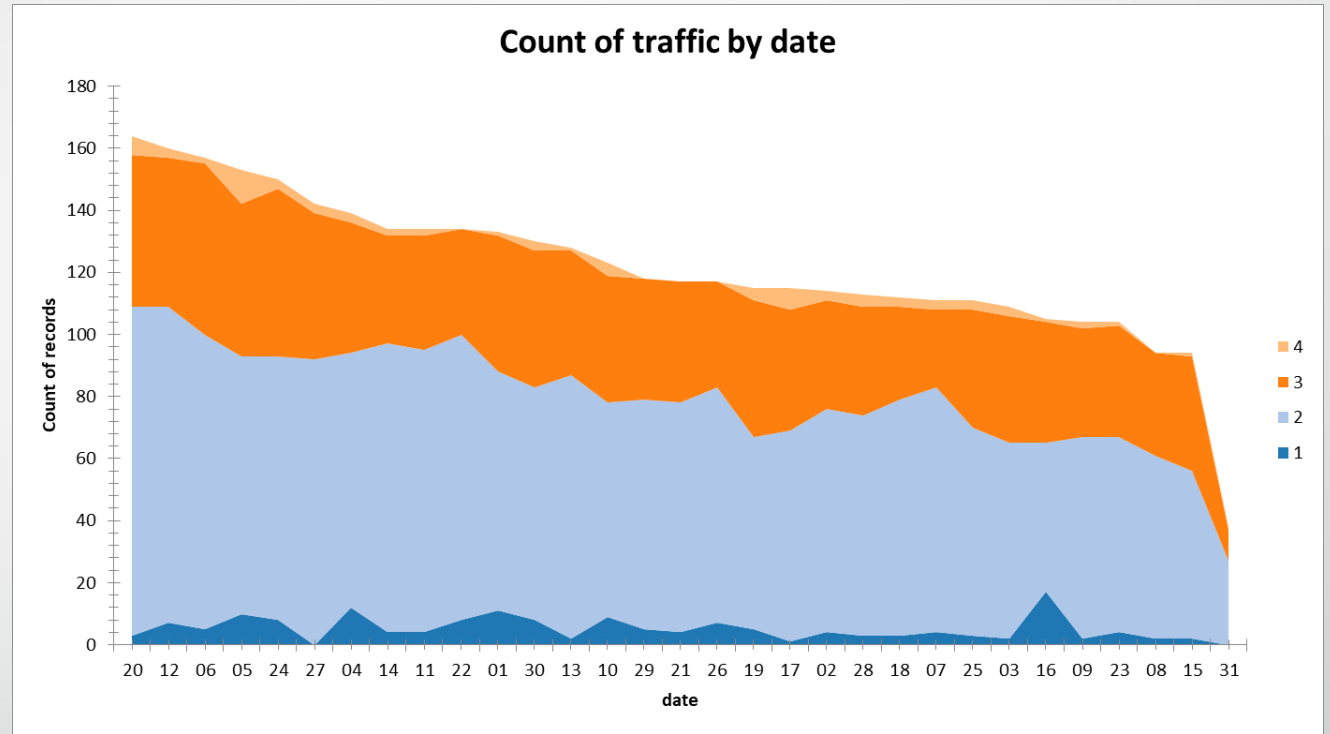


# Traffic Condition by month

March, April, May, June and December has highest traffic

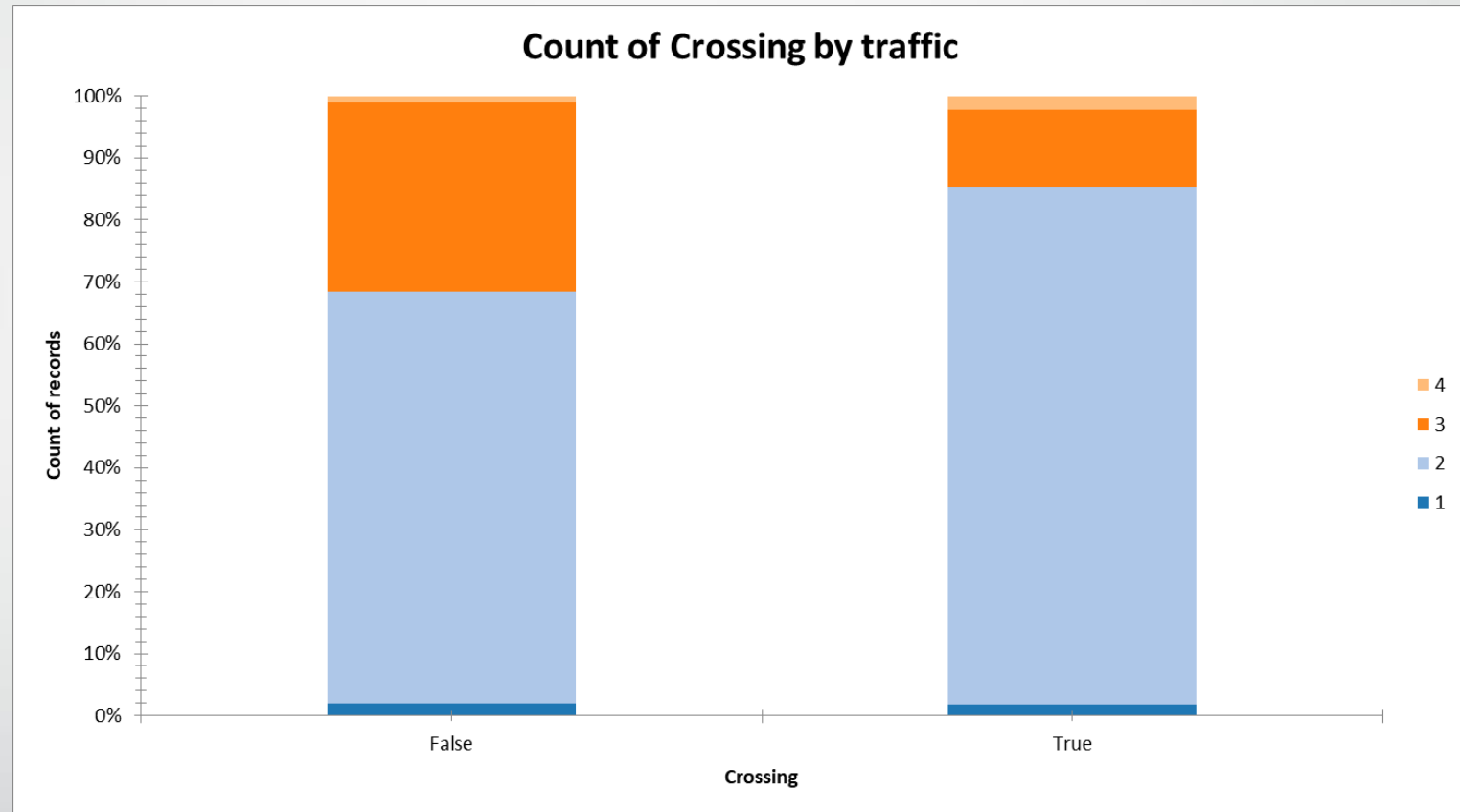
## Traffic condition By date

20, 12, 06, 05 and 24 were the busiest day in the month of march, april, may, june and december



Traffic condition when  
there is Crossing

Traffic is highest and intense when  
there is no crossing.





# Conclusion

- On the days when there were Late deliveries and shipping cancellation, there as an increase in traffic which resulted in a slowing down of the delivery.