

Sami Eljabali
Stat133-Hw1
s133bq
2/5/2008

0) I used read.csv because its default delimiter is the comma, and it also reads it into the table format and creates a dataframe from it.

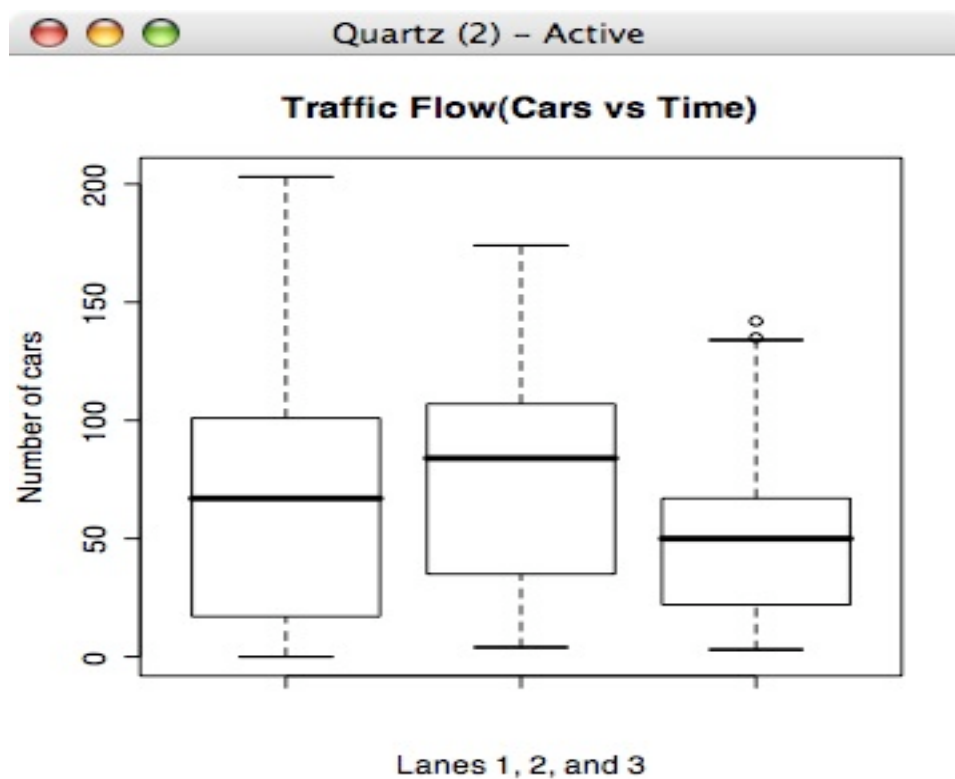
```
>file = file("http://www.stat.berkeley.edu/users/nolan/stat133/data/flow-occ-table.txt")  
>data = read.csv(file)  
>typeof(data)  
→ "list"
```

1) Is the flow in lane 2 is typically about 50% higher than in lane 3?

```
>f1 = data$Flow1  
>f2 = data$Flow2  
>f3 = data$Flow3  
>o1 = data$Occ1  
>o2 = data$Occ2  
>o3 = data$Occ3
```

Here I decided to have each column assigned to a variable since I am going to be using these vectors often.

➤ boxplot(f1, f2, f3, main = "Traffic Flow", xlab = "Lane 1, 2, and 3")



```

> summary(f2)
Min. 1st Qu. Median Mean 3rd Qu. Max.
 4.0  35.0  84.0  74.8 107.0 174.0
> summary(f3)
Min. 1st Qu. Median Mean 3rd Qu. Max.
 3.0  22.0  50.0  48.1  67.0 142.0

```

Median for Lane2 is 84 and 48 for Lane3, so the statement is correct

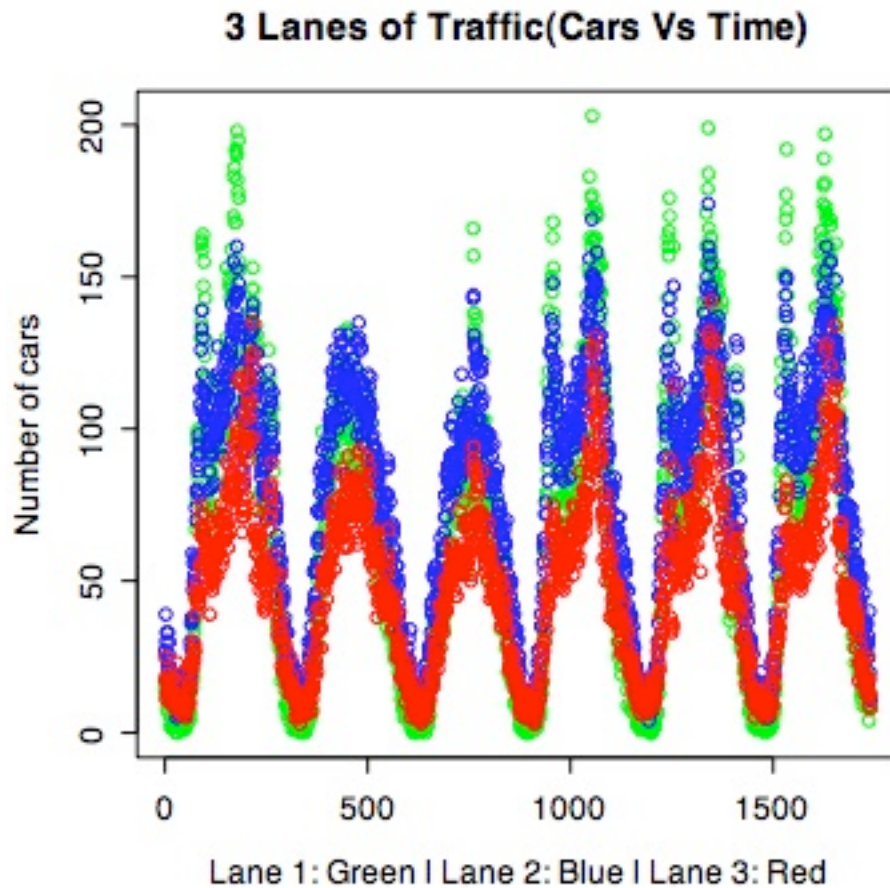
2) Consider the flows in the three lanes using time series plots. Which lane typically serves the most traffic? Is the statement, “When one lane is congested, the others are too,” an accurate description of the relationship you found?

To tackle this problem I thought it would be best if I were to have all lanes and their data to be plotted on the same graph to clearly see their differences.

```

>plot(f1, col="green", xlab = "Lane 1: Green | Lane 2: Blue | Lane 3: Red", ylab = "Number of cars",
main = "3 Lanes of Traffic(Cars Vs Time)")
>points(f2, col="blue")
>points(f3, col="red")

```

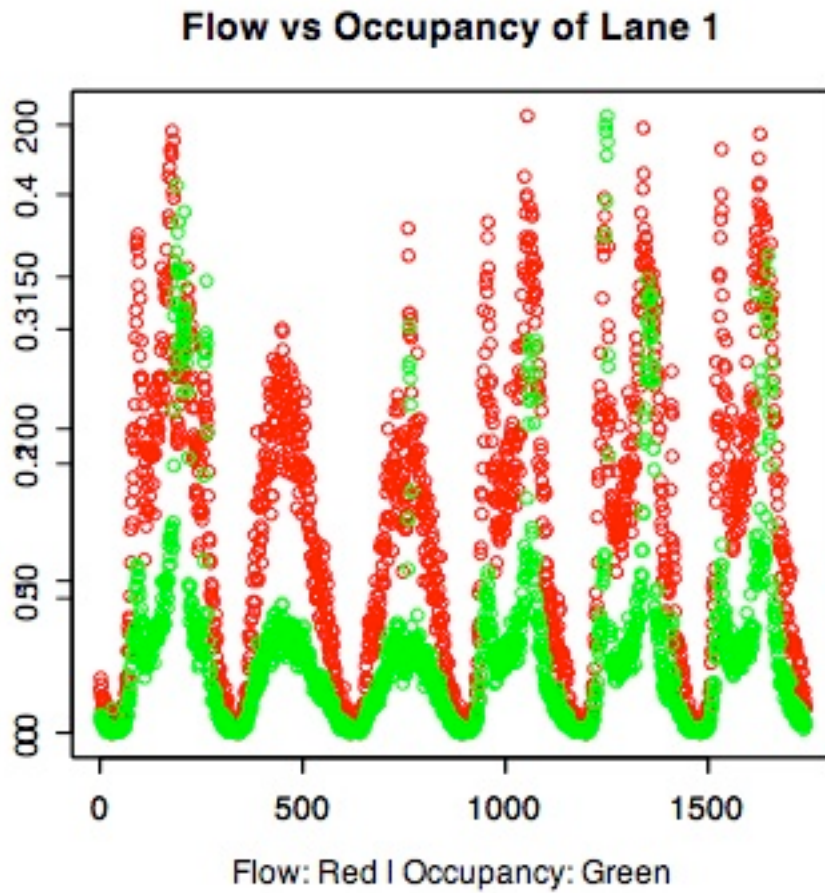


According to the graph Lane 1, which is colored in green, is the most traffic congested, and yes the other lanes were also congested when lane 1 was congested.

3) Consider the relationship between flow and occupancy. Plot flow against occupancy for lane 1. In one sentence, describe the shape of the points. Relate the shape you see to a property of traffic. Is this relationship born out in the other two lanes? Explain any similarities and differences that you find.

```
> plot(f1,col="red", main = "Flow vs Occupancy of Lane 1", xlab="Flow: Red | Occupancy: Green",ylab="")
> par(new=T)
> plot(o1,col="green", xlab="",ylab="")
```

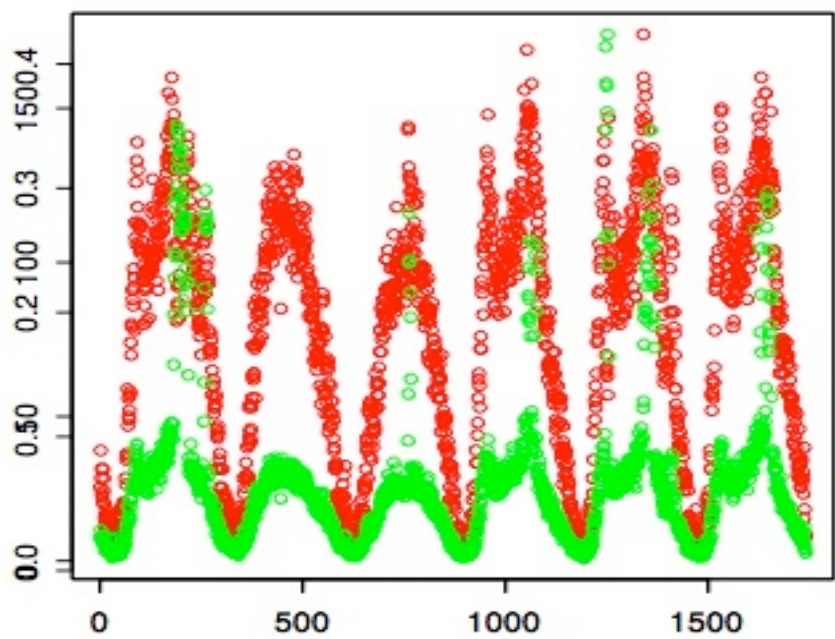
The same was applied with Lane 2 and 3, but with changing their respective variables and labels.



When the cars are going slow there are more cars and more occupancy, which can be less seen in the 2nd and 3rd hump. Traffic jams cars, can be seen in the 1st, 4th, 5th, and 6th hump.

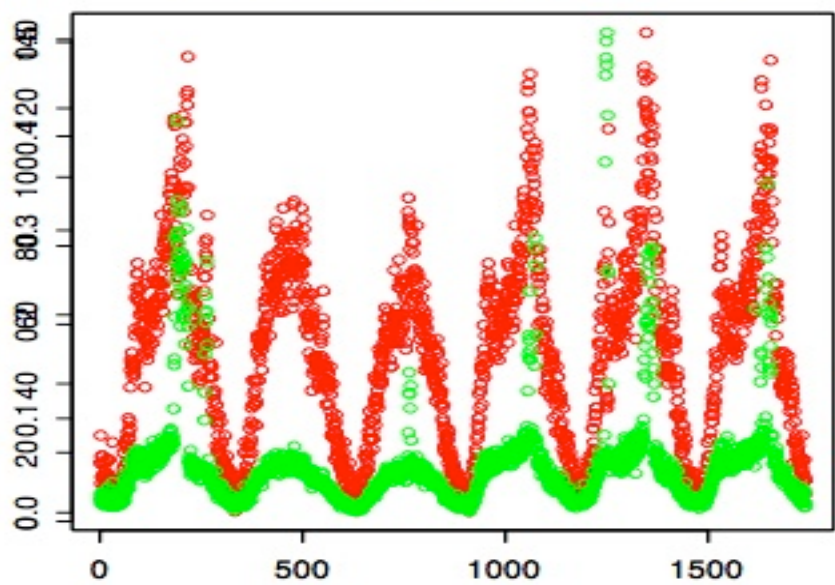
The patterns on the other lanes are the same in the other lanes, which eliminates the question whether that highway has carpool or not, for it looks like there isn't.

Flow vs Occupancy of Lane 2



Flow: Red | Occupancy: Green

Flow vs Occupancy of Lane 3



Flow: Red | Occupancy: Green