Data Visualization – CSE3020
Lab 10 – Integration of R and Tableau

Name: Ayush Sharma
Reg. No: 15BCE1335
Faculty: Dr. Priyadarshini J.

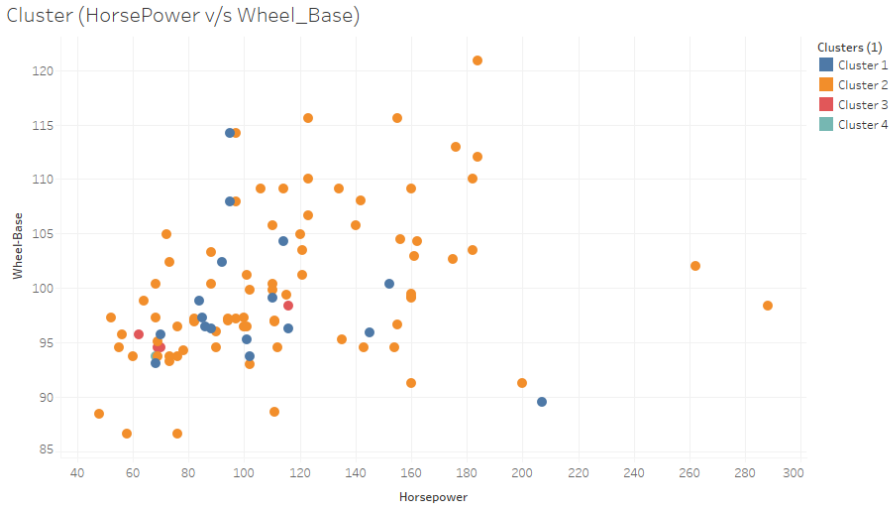
Question: Apply k-means clustering to given URL: https://www.kaggle.com/burgoine/clustering-cars-byspecification/data.
Cluster the cars data into 4 clusters based on:

a. Horsepower and wheel base
b. Horsepower and Length
c. Horsepower and width
d. Horsepower and RPM
e. No of doors and number of cylinders

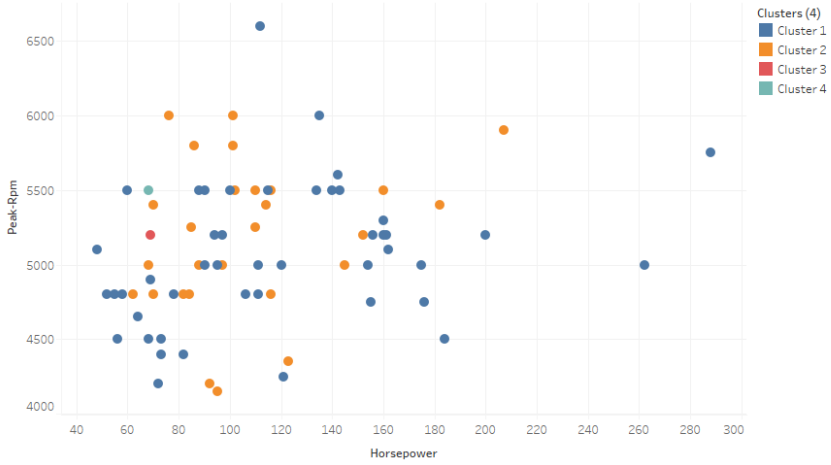
Solution:
Create a calculated field named 'Cluster' in the analysis tab with the following code:

```
SCRIPT_INT('
kmeans(x<-data(.arg1, .arg2, .arg3, .arg4, .arg5, .arg6,.arg7),4)$cluster',
SUM([Horsepower]),SUM([Wheel-Base]),
SUM([Length]),SUM([Width]),SUM([Peak-Rpm]),
COUNT([Num-Of-Cylinders]),COUNT([Num-Of-Doors])
)
```

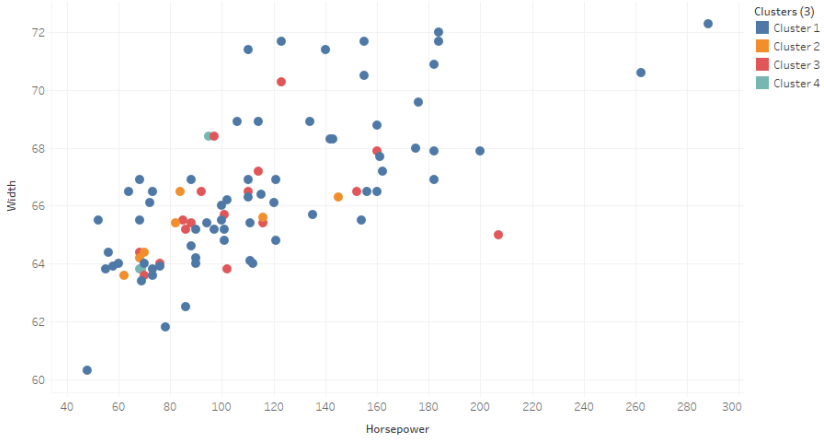
Screenshots:



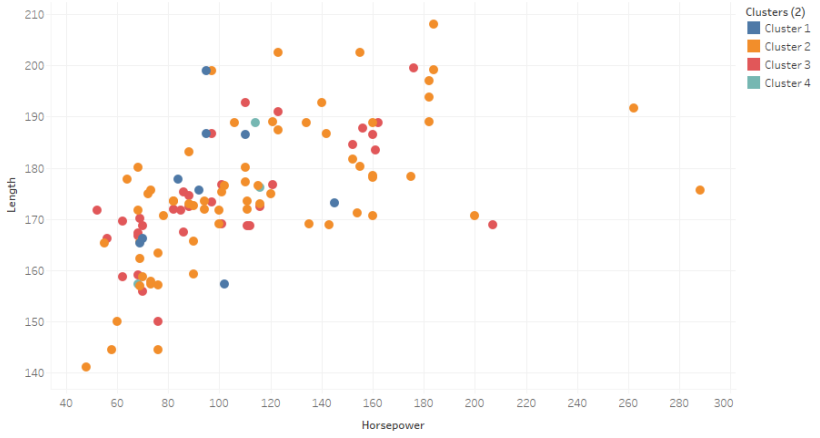
Clusters (HorsePower v/s RPM)



Cluster (HorsePower v/s Width)



Clusters (HorsePower v/s Length)



Clusters (Doors v/s Cylinders)