Name: Husain Rangwala

Date: 27th February 2021

Title: Module 6 Project

Executive Summary Report 6

Key findings

Part I

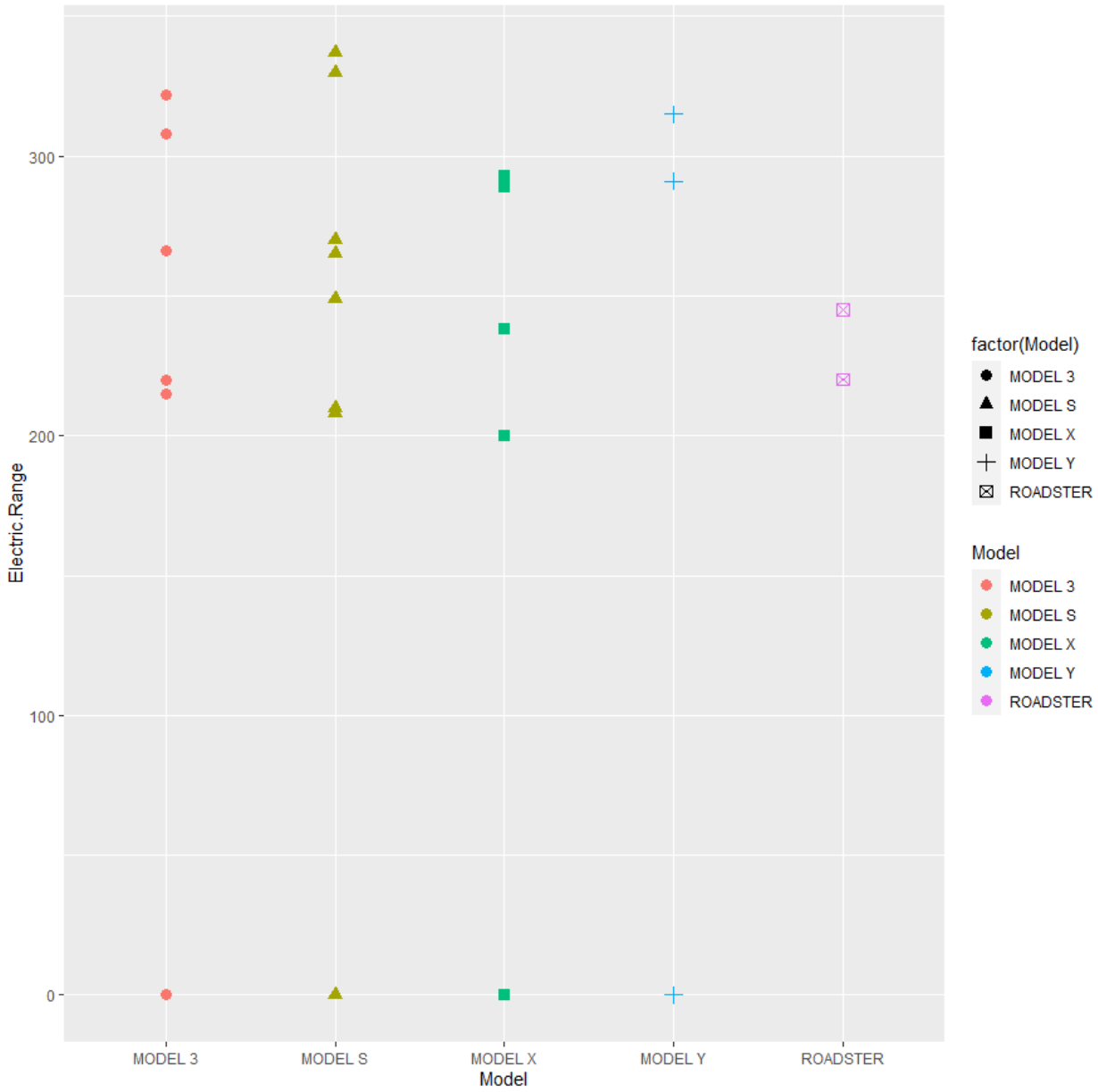
1. Summarize the data in a table.

The dataset I have chosen for this report is “Electrical Vehicle Population Data”. The data set contain a list of all the electrical vehicle that is currently being used in the united states of America. The Dataset Year Range is from 1997 to 2021. The data set shows a variety of make and their various model released over the time.

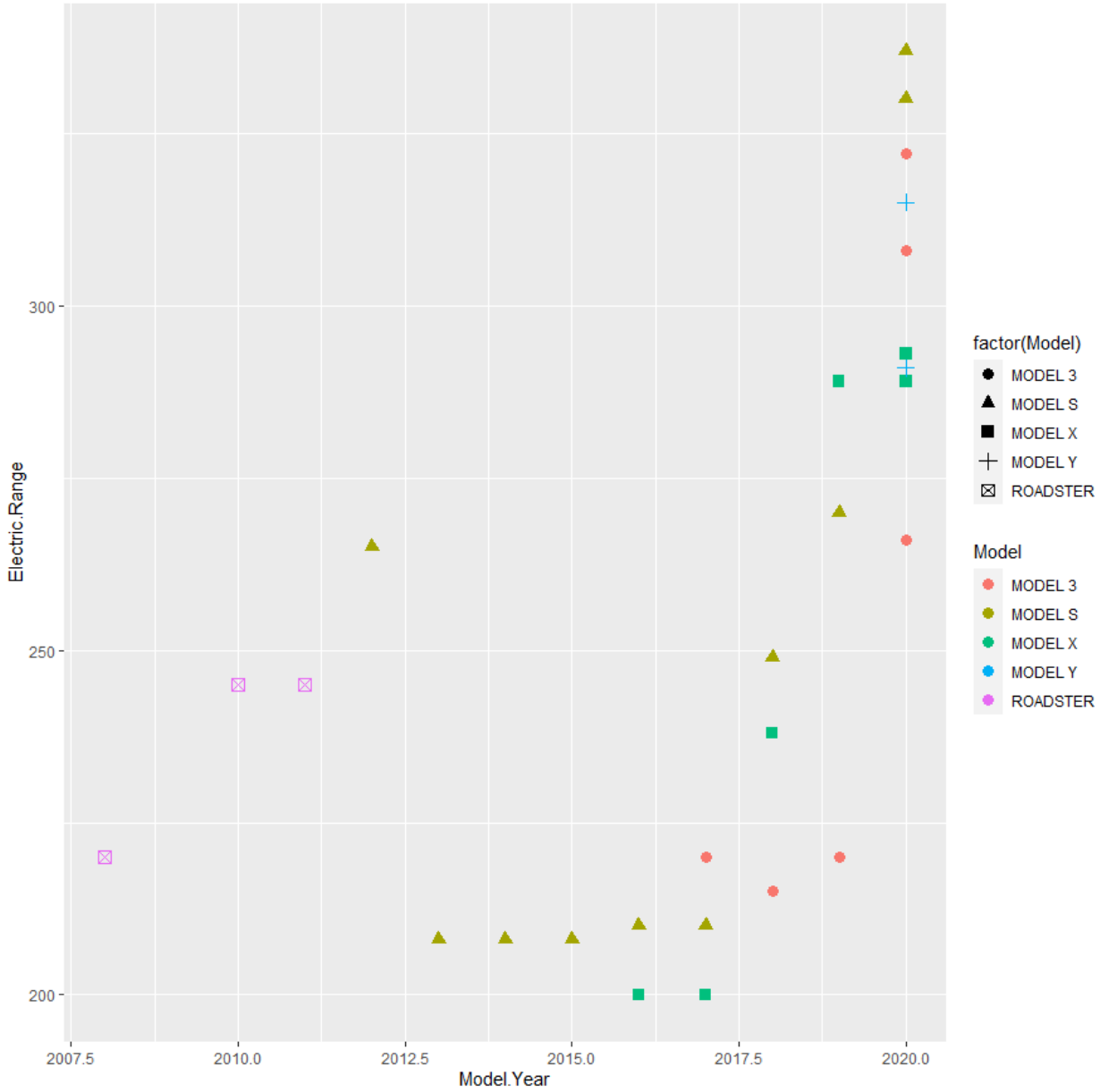
Model Year	Make	Model	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID
2018	TESLA	MODEL 3	215	36000	48	3.06E+08
2020	TESLA	MODEL 3	322	0	37	1.29E+08
2016	MERCEDES-BENZ	B250E	87	41450	36	4.77E+08
2016	FORD	C-MAX ENERGI	19	31770	25	3.21E+08
2018	MITSUBISHI	OUTLANDER	22	34595	24	1.15E+08
2017	BMW	I3	97	42400	47	1.91E+08
2020	HONDA	CLARITY	47	0	6	1.1E+08
2020	KIA	NIRO	239	0	41	1.39E+08
2020	FORD	FUSION	26	0	32	1.29E+08
2018	BMW	I3S	114	44450	48	3.5E+08
2020	NISSAN	LEAF	215	0	44	1.26E+08
2020	KIA	NIRO	239	0	34	1.03E+08
2020	CHEVROLET	BOLT EV	259	0	33	1.26E+08
2018	KIA	NIRO	26	27900	36	2.69E+08
2020	NISSAN	LEAF	149	0	12	1.38E+08
2018	TESLA	MODEL 3	215	36000	4	1968858
2016	KIA	SOUL	93	31950	1	2.71E+08
2020	TESLA	MODEL 3	266	0	25	1.11E+08
2019	TESLA	MODEL 3	220	35000	5	2280099

2. Graphs that help visualize the data. These can be bar charts, histograms, pie charts, etc. Be sure the chosen graph best represents the information you want to highlight.

Below graph highlights the various tesla makes and the electrical range they provide for each model. The graph represents each of the model (Model 3, Model S, Model X, Model Y, Roadster) and their maximum distance per one electrical charge.



The below graph is a comparison of Model year to the distance per one electrical charger for all Tesla models. Tesla roadster in the year 2008 could travel 220 kms in one electrical charge and after 2008, Tesla upgrade the charge circle 230-240 volts per one charge. Similarly, the model S initial model could only travel 208 km in one electrical charge but with substantial improvement over the year, model S could now travel to more than 310 kms. Model X with the lowest charge cycle of 200 kms was the least before improvement, but with enhancement model X can travel to close 300 kms. Likewise, to model X, Model 3 also made upgrades to increase their charge cycle to 305 kms. Model Y has kms per one charge cycle in 290-305 Kms.



3. Explain the story the data is telling you.

o What business question do your descriptive analyses answer? Provide a brief discussion of the findings.

1. Which electrical vehicle provides the greater electrical range among the makers?
2. Which Legislative District and State has the most electronic vehicle registered?
3. Which vehicle has the lowest and highest MSRP?

Below snippet presented the business question answer.
Tesla make show the greatest electrical range among the rest of the competitors who have Clean Alternative Fuel Vehicle Eligible enabled.

The legislative district with the maximum electrical vehicle belongs to 48 and the state is Washington.

The lowest MSRP belong to the Mitsubishi I-Miev with $22995 and the maximum MSRP belong to Porsche 918 Spyder with $845000

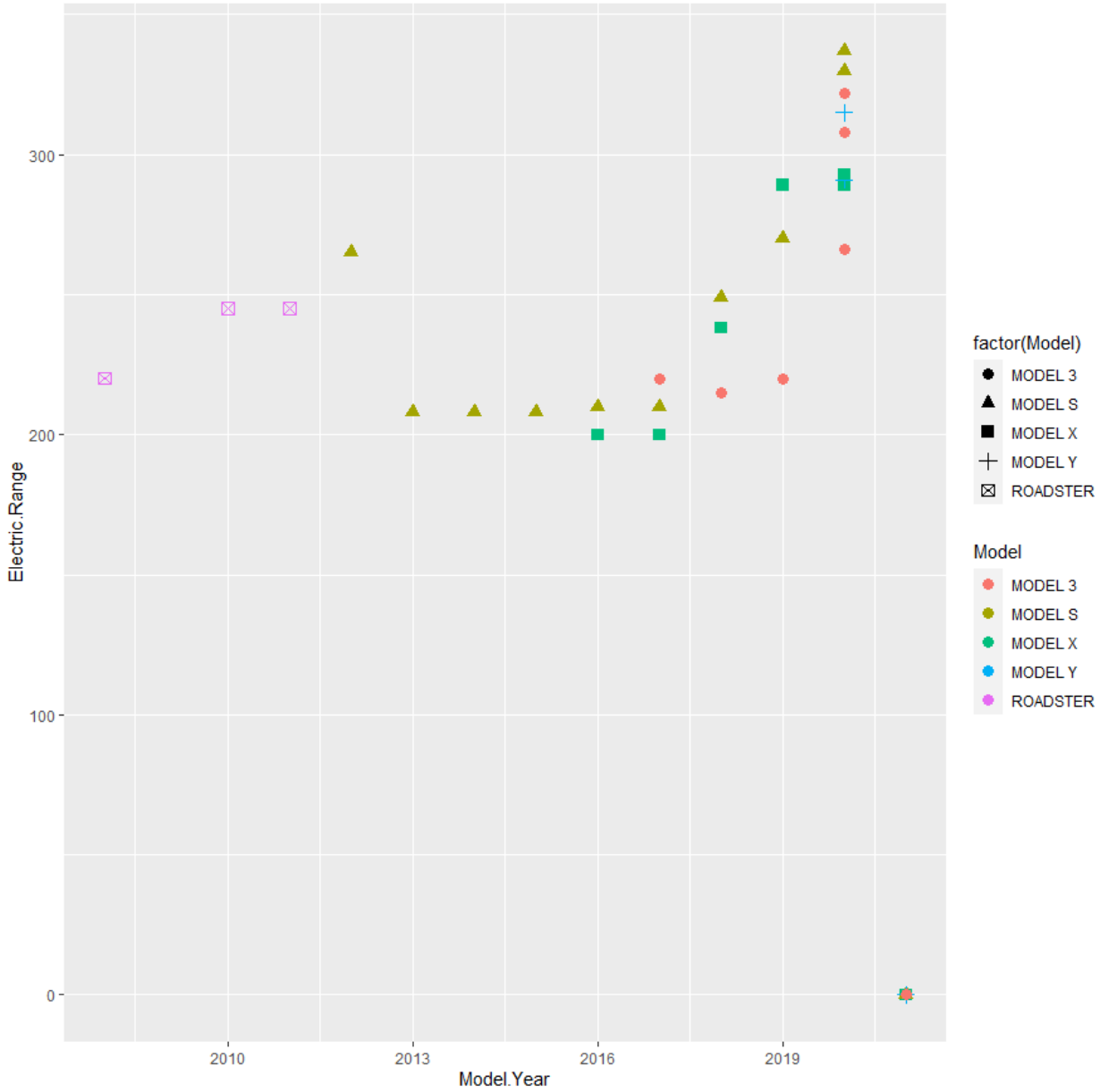
o If there are any unusual values, discuss them. If data values are “out of range,” clean the data as needed. Delete the out-of-range values and run the analysis again.

While observing the Tesla dataset, there were some out of range and some missing data that were creating the graph to go askew and not provide proper output.

o If you remove out of range values for any of the variables, present both the analysis with the out-of-range values and the analysis without the out-of-range value(s).

One of us out-of-range values was the data that was collected in the year 2021, as dataset is updated on regular bases the data could not be collected and update. For this reason, the data from 2021 was remove.

Before removal of the outliers, the data presented as below. All the models were delivered to their respected owner, but the data for those models could not collected and updated.



o Identify additional questions that the data is leading you to ask. What new attributes are needed to answer those questions?

Business Question:

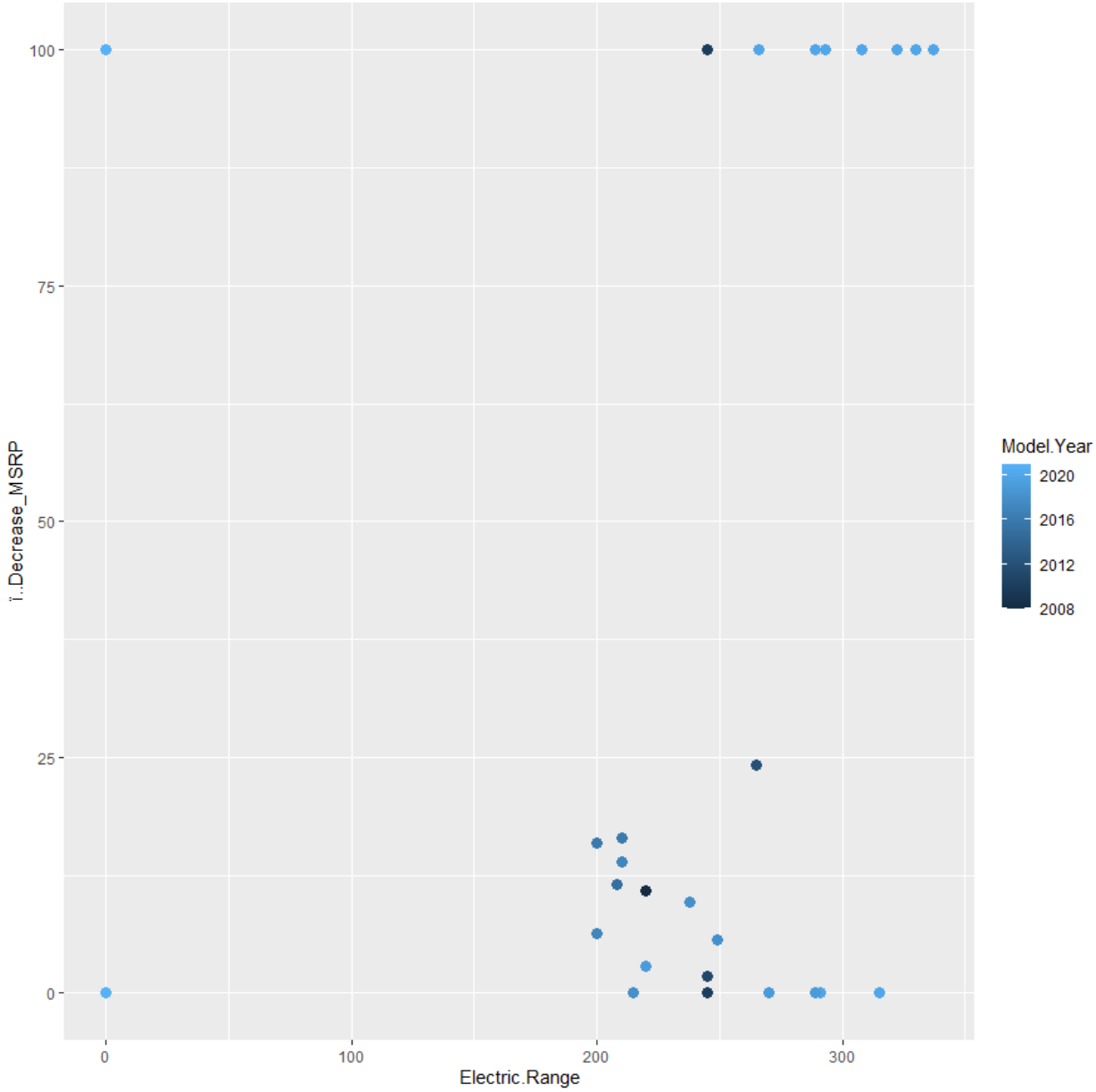
1. What would the resell value for the car?
2. Are Battery Electric Vehicle (BEV) better than Plug-in Hybrid Electric Vehicle (PHEV)?

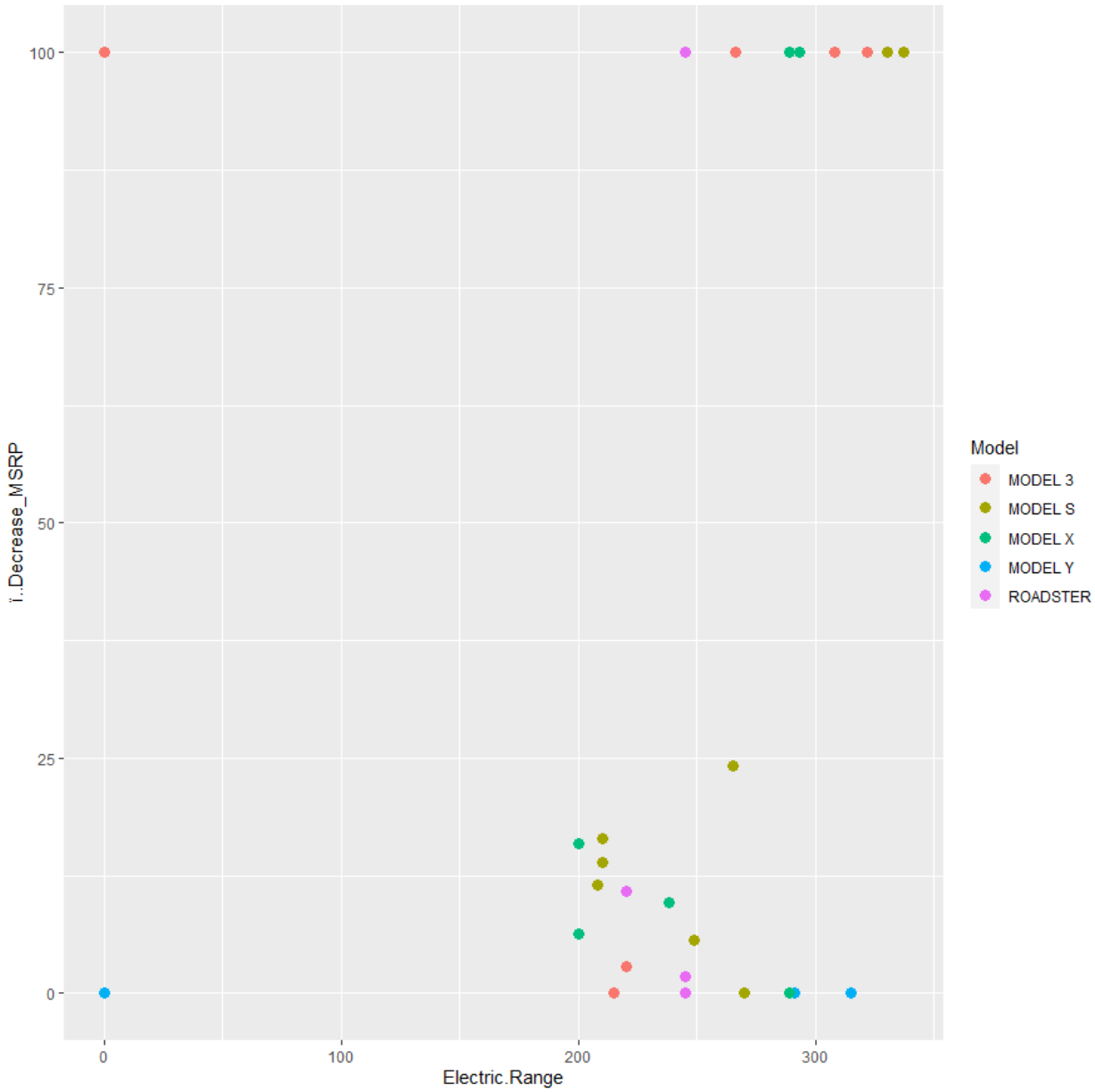
Part II

Create new attributes based on the data and the questions you identified in Part 1.

From the above question, a new attribute can be created that could calculate the percentage decrease in MSRP or the resell values of an old Tesla model.

Below is the graph that represents the same.





The graph is between electric range versus the resell value and the Year being the deciding factor. The second graph represent the range versus the resell value as per the model.
Using these two graphs, some of the car has 100 % decreased chance of reselling while some of them have some resell value in the market depending on the model and the year of the purchase.

Part III

Now that you have worked with the data, what is the data saying to you?

After understanding and working with the data, I could conclude that some of the hybrid vehicle are costlier and pure electric vehicles. Secondly, we can use the DOL Vehicle ID and ZIP Code to locate the vehicle location.

What have you learned about the attributes?

From working on the data, I could collect the resell, location, legislative district, and state from which the car owner belongs to for each vehicle.

What are some follow-up questions you would like to have answered?

One of the follow up question, I observed was the resell value of Tesla vehicle and how the year, model and range used can determine the resell value of the vehicle.

Identify 3-5 observations or follow-up questions that you have.

Using the data, could we track the location of the car?

Difference between using Battery Electric vehicle and Plug-in Hybrid Vehicle in longer timespan?

Can the legislative district and state be responsible for the vehicle Market price rise?

Bibliography

Below link were used to help solving any blocker in R scripting.

1. https://www.geeksforgeeks.org/adding-elements-in-a-vector-in-r-programming-append-method/
2. https://discuss.analyticsvidhya.com/t/how-to-remove-value-from-a-vector-in-r/2975
3. https://www.dummies.com/programming/r/how-to-work-with-variable-names-in-r/
4. https://www.youtube.com/watch?v=rgAvJmvfA2c
5. R in Action: Data Analysis and Graphics with R By Robert Kabacoff
6. https://www.statmethods.net/advgraphs/parameters.html

Appendix

```
#-----------------------------------#
#      Plotting Basics: Rangwala    #
#-----------------------------------#

#. Print your name at the top of the script and load these libraries: FSA, FSAdata, magrittr,
#dplyr, tidyr plyr and tidyverse#
setwd("~/College_work/extras")

#install.packages(c("FSA","FSAdata","magrittr","dplyr","tidyr","plyr","tidyverse"))
#help(package="dplyr")
#help(package="magrittr")
#help(package="tidyr")
#help(package="plyr")
#help(package="tidyverse")



library(FSA)
library(dplyr)
library(magrittr)
library(tidyr)
library(plyr)
library(tidyverse)
library(ggplot2)

df=read.table("Electric_car.csv", header = TRUE, sep = ",")
headtail(df)

df1=subset(df, Make=="TESLA",
```

```
        select=c(Model.Year, Model, Electric.Range, Base.MSRP, Legislative.District ,DOL.Vehicle.ID))
summary(df1)
headtail(df1,n=10)
write.csv(headtail(df1,n=10), file="output.csv")
#plot(df1)



#plot(df1$Model, df1$Electric.Range)
ggplot(df1, aes(x=Model,y=Electric.Range, color=Model))+geom_point(stat =
"identity",size=3,aes(shape=factor(Model)))



df1 %>%
 filter(Model.Year<=2020)%>%
 ggplot(aes(x=Model.Year,y=Electric.Range, color=Model))+
 geom_point(stat = "identity",size=3,aes(shape=factor(Model)))



ggplot(df1, aes(x=Model.Year,y=Electric.Range, color=Model))+
 geom_point(stat = "identity",size=3,aes(shape=factor(Model)))

Decreased_MSRP = read.csv("New table.csv", header = TRUE)
is.data.frame(Decreased_MSRP)
df2 = cbind(df1,Decreased_MSRP)
summary(df2)

ggplot(df2, aes(x=ï..Decrease_MSRP, color=Model))+
 geom_point(aes(y=Model.Year),size=3,color="green")
```

```
ggplot(df2, aes(x=Electric.Range, y=ï..Decrease_MSRP, color=Model))+
  geom_point(stat = "identity",size=3)
```