

Final Kaggle Project

Renzo Tanaka-Wong

2/28/2021

Download Useful Packages

```
library(readr)
library(corrplot)

## corrplot 0.84 loaded

library(ggplot2)
library(car)

## Loading required package: carData

library(leaps)
library(GGally)

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg      ggplot2

library(caret)

## Loading required package: lattice
```

Download data

```
cars.train <- read_csv("carsTrain.csv")

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   Manufacturer = col_character(),
##   Model = col_character(),
##   Type = col_character(),
##   AirBags = col_character(),
##   DriveTrain = col_character(),
##   Cylinders = col_character(),
##   Man.trans.avail = col_character(),
##   Origin = col_character(),
##   Make = col_character()
## )

## See spec(...) for full column specifications.
```

```
cars.test <- read_csv("carsTestNoY.csv")

## Parsed with column specification:
## cols(
##   .default = col_double(),
##   Manufacturer = col_character(),
##   Model = col_character(),
##   Type = col_character(),
##   AirBags = col_character(),
##   DriveTrain = col_character(),
##   Cylinders = col_character(),
##   Man.trans.avail = col_character(),
##   Origin = col_character(),
##   Make = col_character()
## )
## See spec(...) for full column specifications.

dim(cars.train)

## [1] 1500 25

dim(cars.test)

## [1] 500 24
```

Exploratory Data Analysis

Categorical Variables

First let's see which of our predictors are numerical and which are categorical.

```
head(cars.train)

## # A tibble: 6 x 25
##   Ob Manufacturer Model Type MPG.highway AirBags DriveTrain Cylinders
##   <dbl> <chr>      <chr> <chr>    <dbl> <chr>    <chr>    <chr>
## 1     1 Volkswagen Corr~ Spor~    25 None    Front     6
## 2     2 Buick      Rivi~ Mids~    27 Driver~ Front     6
## 3     3 Infiniti  Q45   Mids~    22 Driver~ Rear      8
## 4     4 Mazda     626   Comp~    34 Driver~ Front     4
## 5     5 Chevrolet Corv~ Spor~    25 Driver~ Rear     8
## 6     6 Lincoln   Cont~ Mids~    26 Driver~ Front     6
## # ... with 17 more variables: EngineSize <dbl>, Horsepower <dbl>, RPM <dbl>,
## #   Rev.per.mile <dbl>, Man.trans.avail <chr>, Fuel.tank.capacity <dbl>,
## #   Passengers <dbl>, Length <dbl>, Wheelbase <dbl>, Width <dbl>,
## #   Turn.circle <dbl>, Rear.seat.room <dbl>, Luggage.room <dbl>, Weight <dbl>,
## #   Origin <chr>, Make <chr>, PriceNew <dbl>

# Number of numeric predictors
sum(unlist(lapply(cars.train, is.numeric)))

## [1] 16

unlist(lapply(cars.train, is.numeric))

##           Ob           Manufacturer           Model           Type
##           TRUE             FALSE             FALSE             FALSE
##   MPG.highway       AirBags       DriveTrain       Cylinders
```

```
##           TRUE           FALSE           FALSE           FALSE
##      EngineSize      Horsepower           RPM      Rev.per.mile
##           TRUE           TRUE           TRUE           TRUE
##      Man.trans.avail Fuel.tank.capacity      Passengers      Length
##           FALSE           TRUE           TRUE           TRUE
##           Wheelbase      Width      Turn.circle      Rear.seat.room
##           TRUE           TRUE           TRUE           TRUE
##      Luggage.room      Weight      Origin      Make
##           TRUE           TRUE           FALSE           FALSE
##      PriceNew
##           TRUE
```

```
# Number of categorical predictors
sum(unlist(lapply(cars.train, is.character)))
```

```
## [1] 9
```

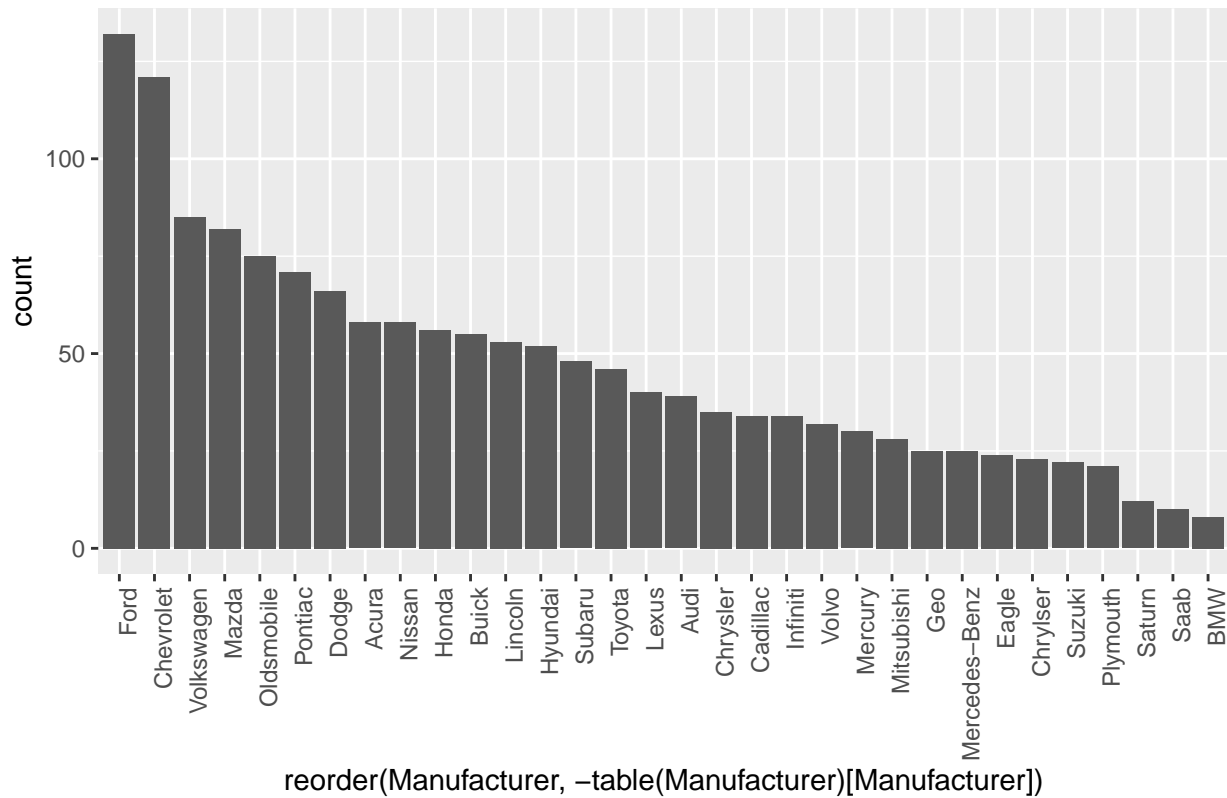
```
unlist(lapply(cars.train, is.character))
```

```
##           Ob      Manufacturer      Model      Type
##           FALSE      TRUE      TRUE      TRUE
##      MPG.highway      AirBags      DriveTrain      Cylinders
##           FALSE      TRUE      TRUE      TRUE
##      EngineSize      Horsepower      RPM      Rev.per.mile
##           FALSE      FALSE      FALSE      FALSE
##      Man.trans.avail Fuel.tank.capacity      Passengers      Length
##           TRUE      FALSE      FALSE      FALSE
##           Wheelbase      Width      Turn.circle      Rear.seat.room
##           FALSE      FALSE      FALSE      FALSE
##      Luggage.room      Weight      Origin      Make
##           FALSE      FALSE      TRUE      TRUE
##      PriceNew
##           FALSE
```

Let's start by graphing the categorical variables and explore their relationships with our response variable PriceNew.

```
# Count by manufacturers
ggplot(data=cars.train, aes(x=reorder(Manufacturer, -table(Manufacturer)[Manufacturer]))) + geom_bar()
```

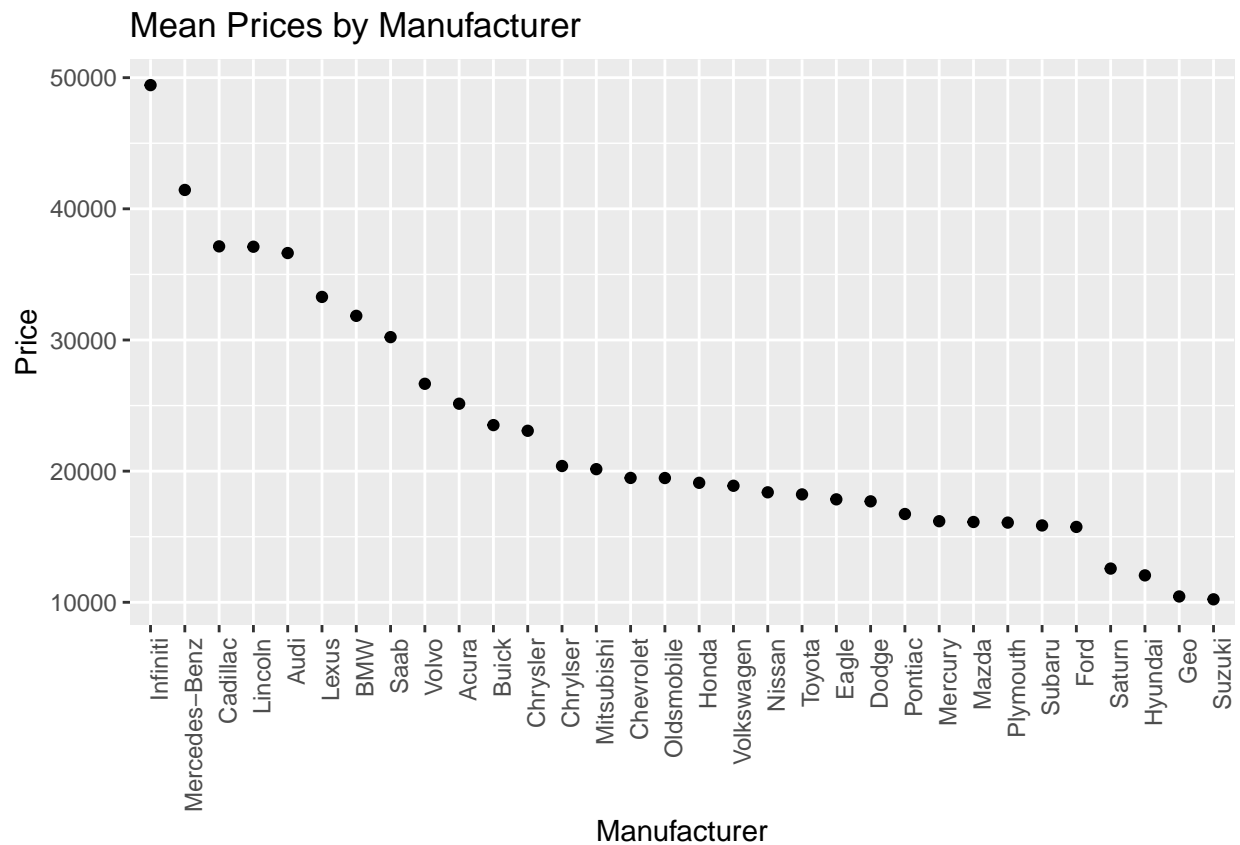
Barplot of Manufacturers



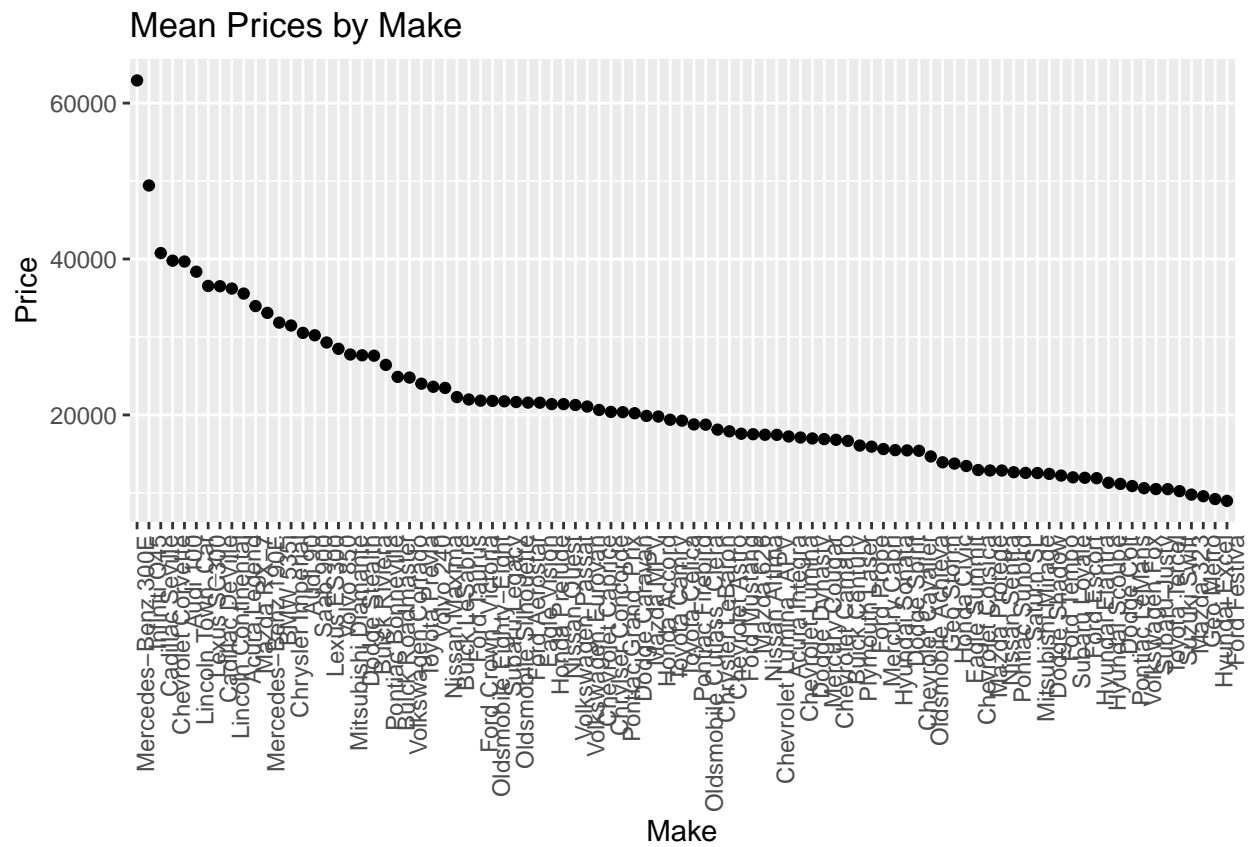
Plot of mean prices by manufacturer

```
mean_prices_manufacturer <- tapply(cars.train$PriceNew, cars.train$Manufacturer, mean)
```

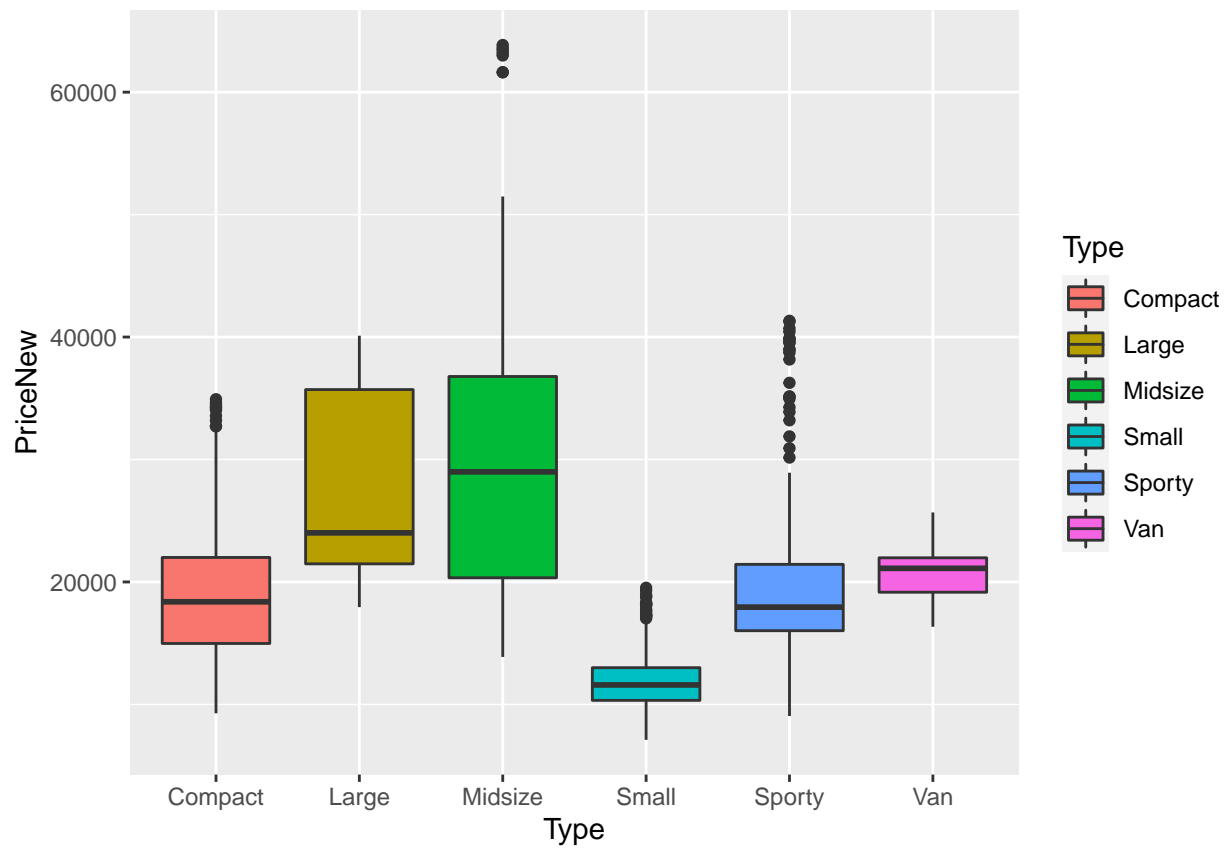
```
qplot(reorder(names(mean_prices_manufacturer), -mean_prices_manufacturer), mean_prices_manufacturer) +
```



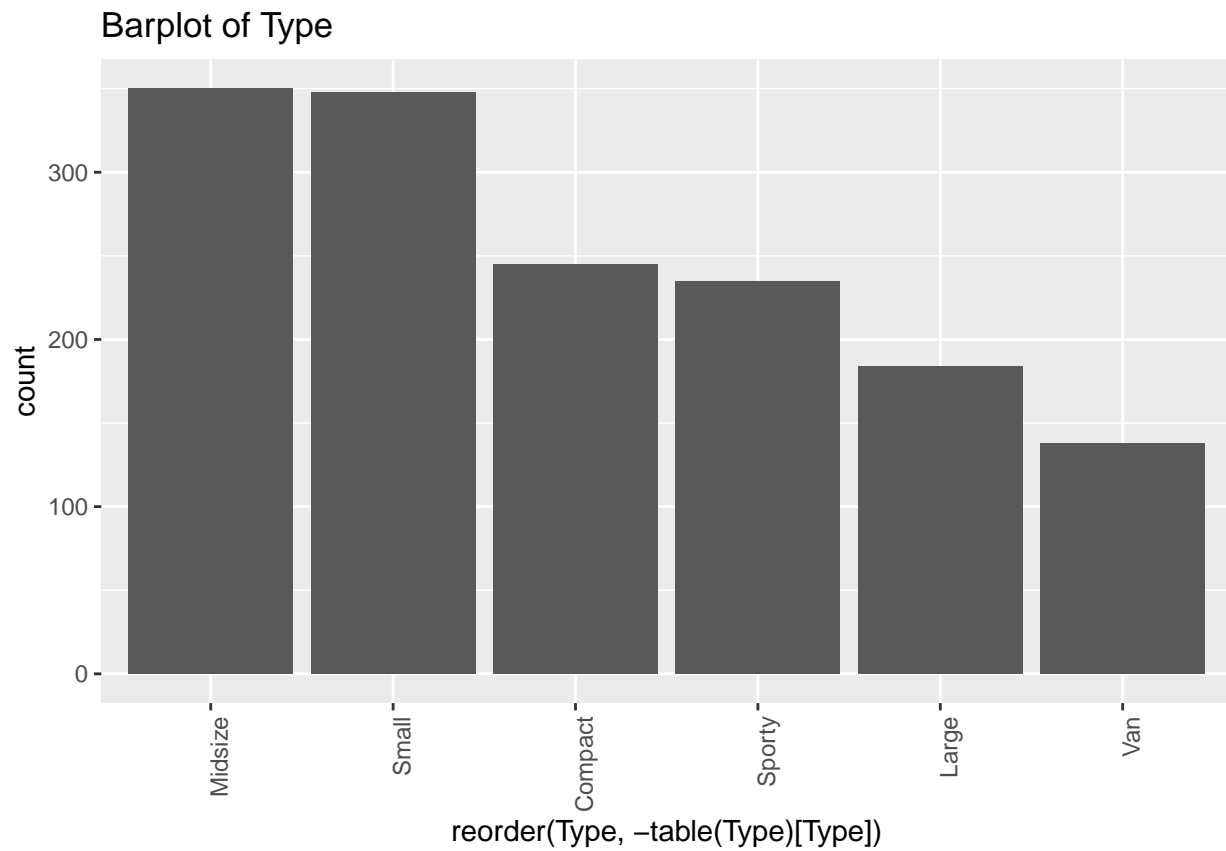
```
# Mean prices by model
mean_prices_model <- tapply(cars.train$PriceNew, cars.train$Model, mean)
qplot(reorder(names(mean_prices_model), -mean_prices_model), mean_prices_model) + theme(axis.text.x = e
```

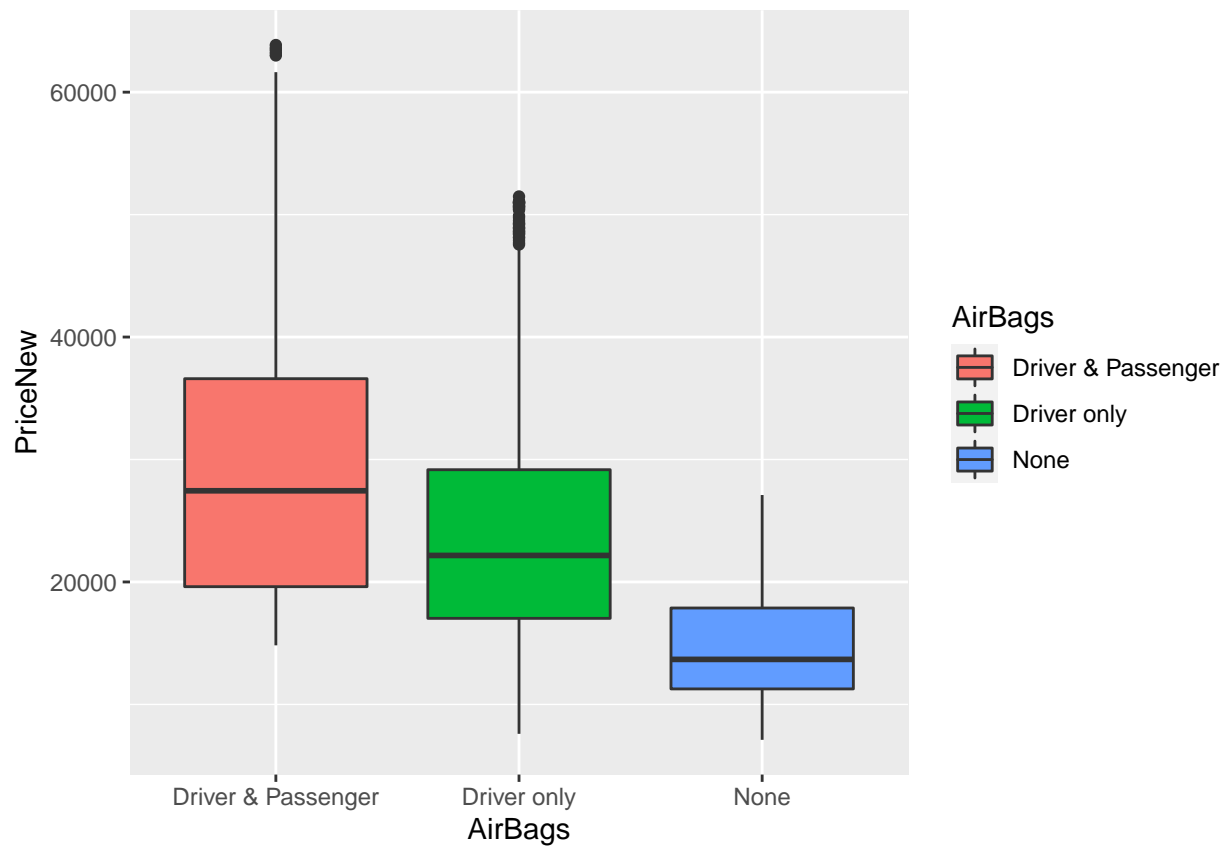
```
# Type
ggplot(cars.train, aes(x=Type, y=PriceNew)) +geom_boxplot(aes(fill= Type,shape = Type))
```



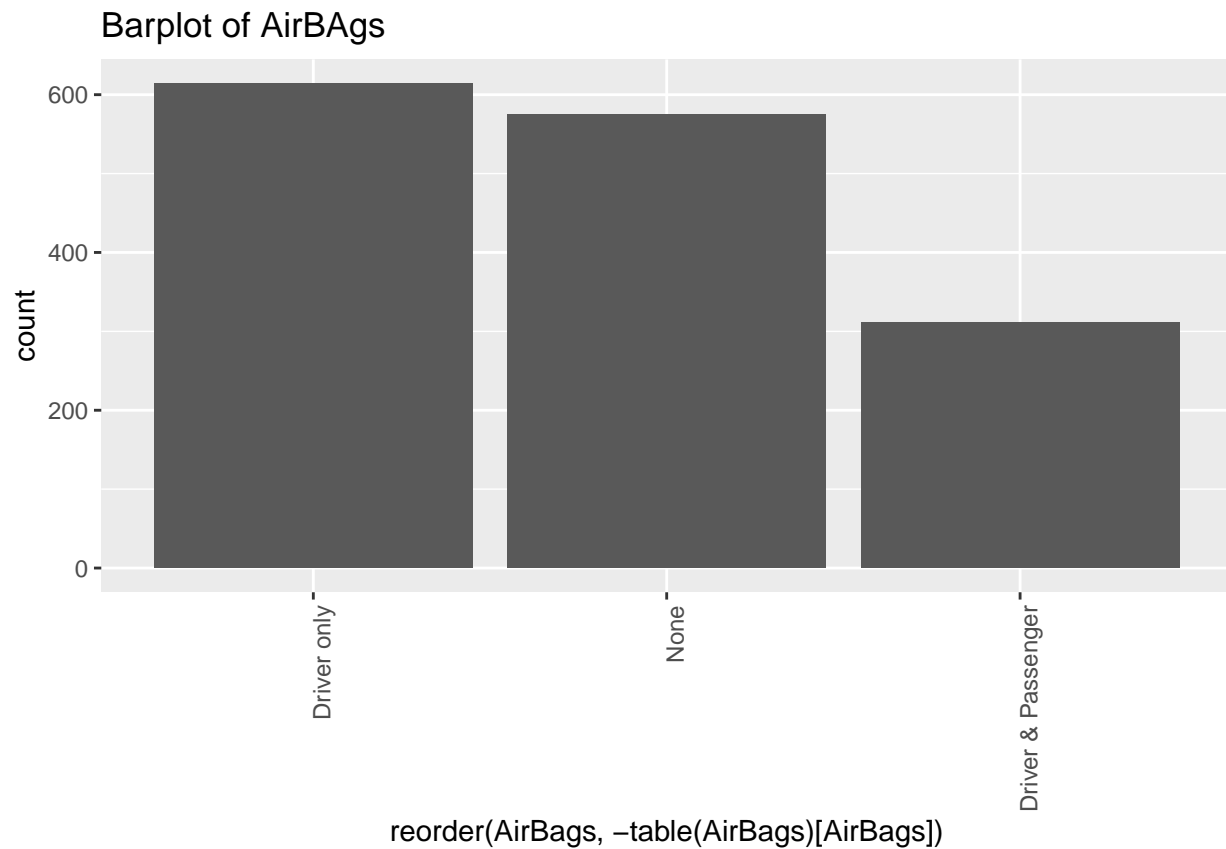
```
ggplot(data=cars.train, aes(x=reorder(Type, -table(Type)[Type]))) + geom_bar() + theme(axis.text.x = el
```

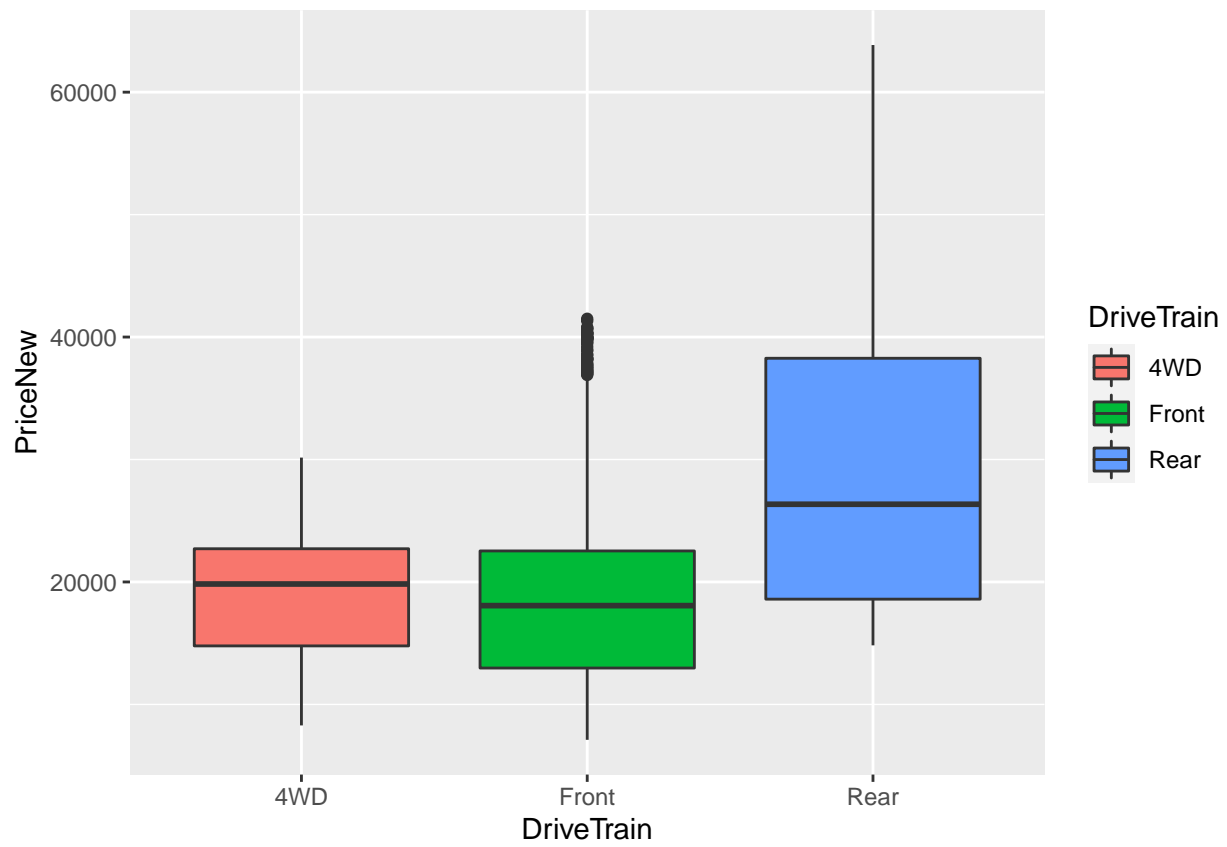
```
# AirBags  
ggplot(cars.train, aes(x=AirBags, y=PriceNew)) + geom_boxplot(aes(fill= AirBags, shape = AirBags))
```



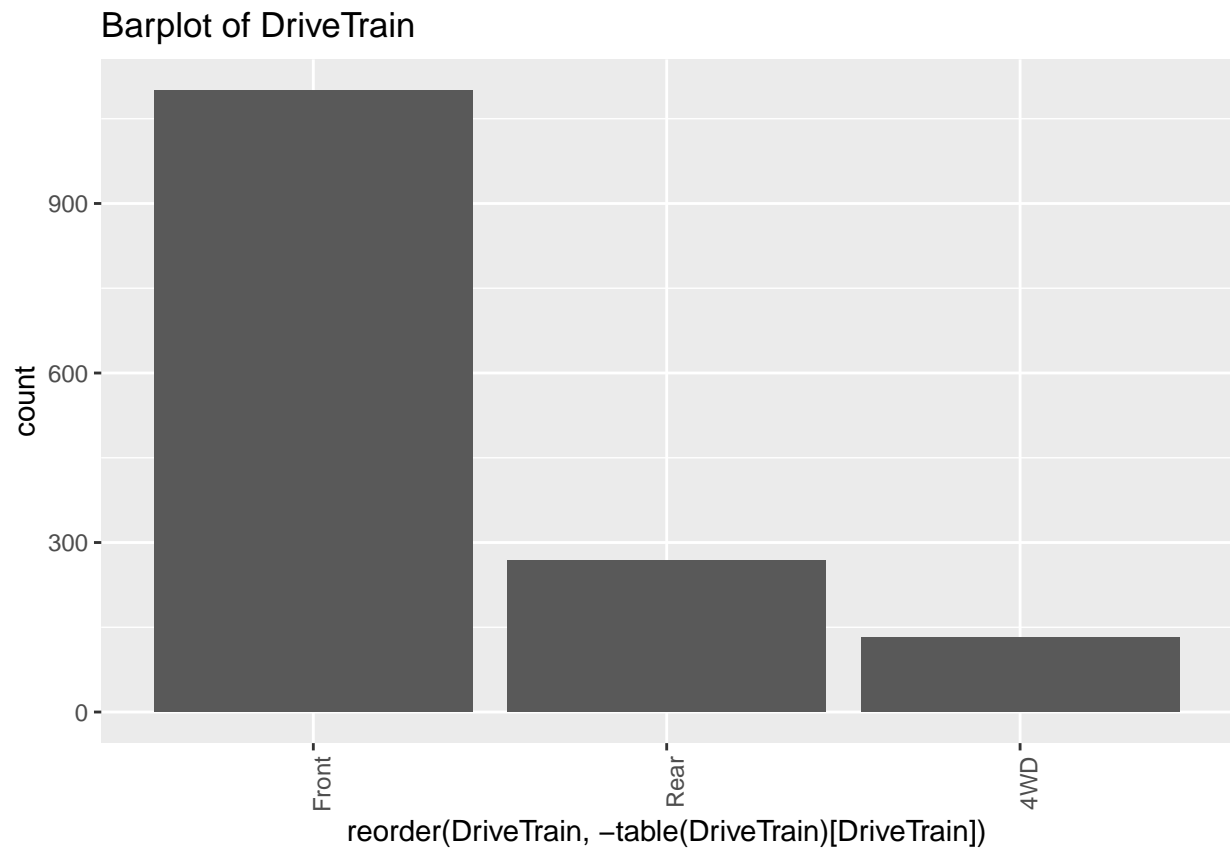
```
ggplot(data=cars.train, aes(x=reorder(AirBags, -table(AirBags)[AirBags]))) + geom_bar() + theme(axis.te
```



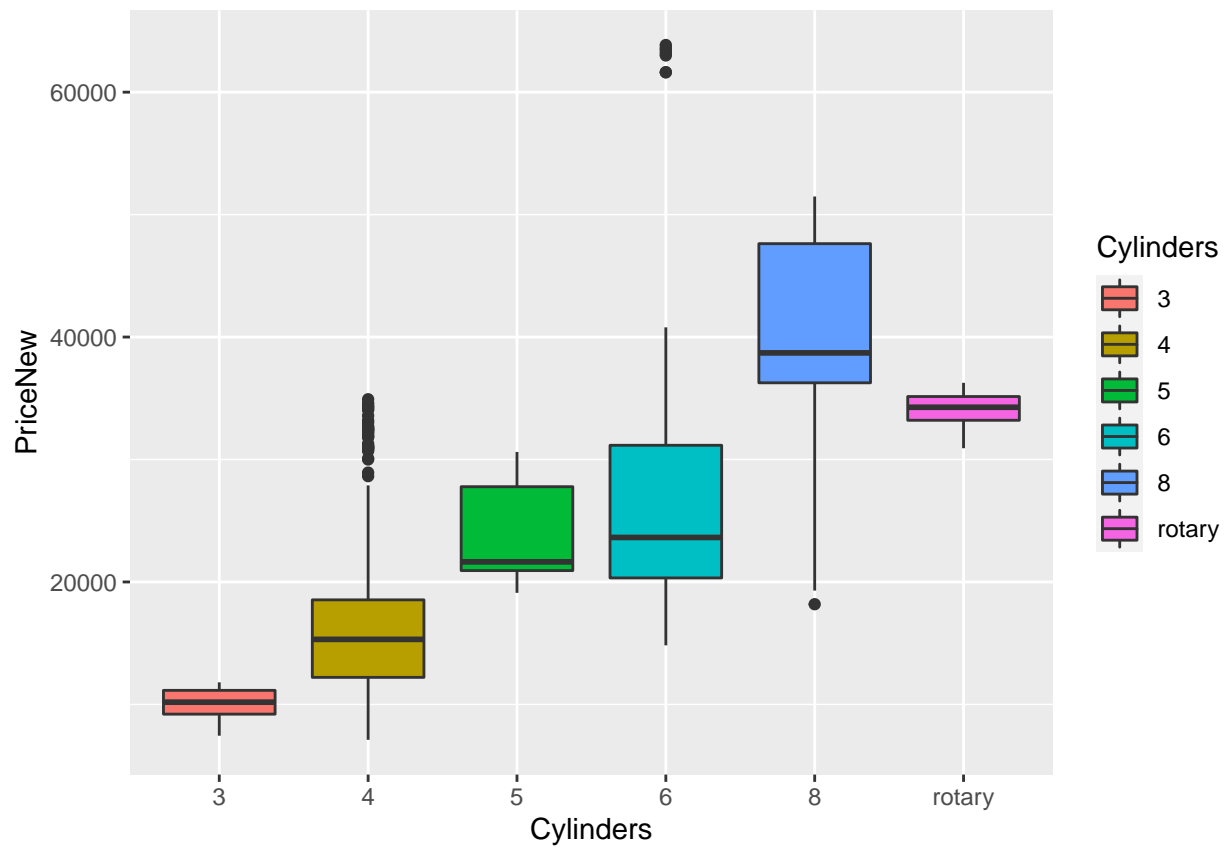
```
# DriveTrain  
ggplot(cars.train, aes(x=DriveTrain, y=PriceNew)) + geom_boxplot(aes(fill= DriveTrain, shape = DriveTrain))
```



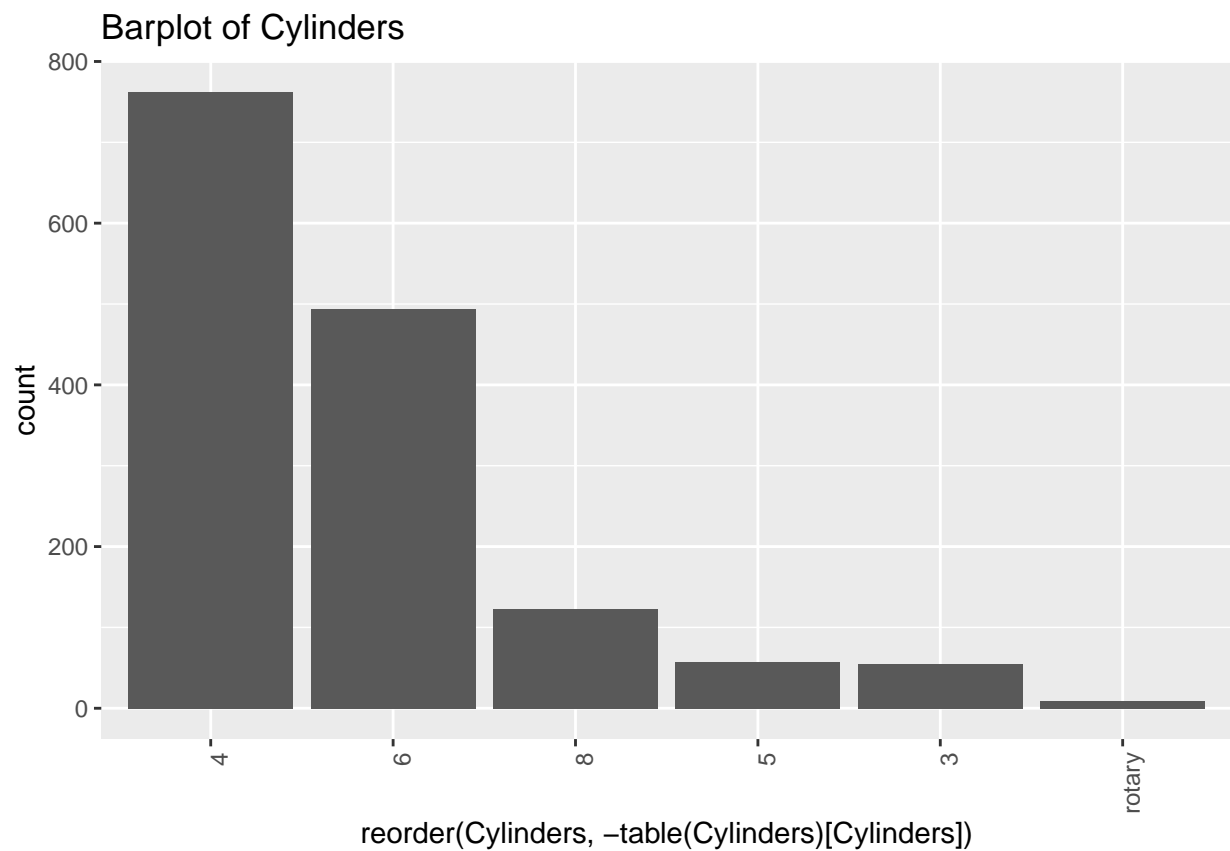
```
ggplot(data=cars.train, aes(x=reorder(DriveTrain, -table(DriveTrain)[DriveTrain]))) + geom_bar() + theme
```



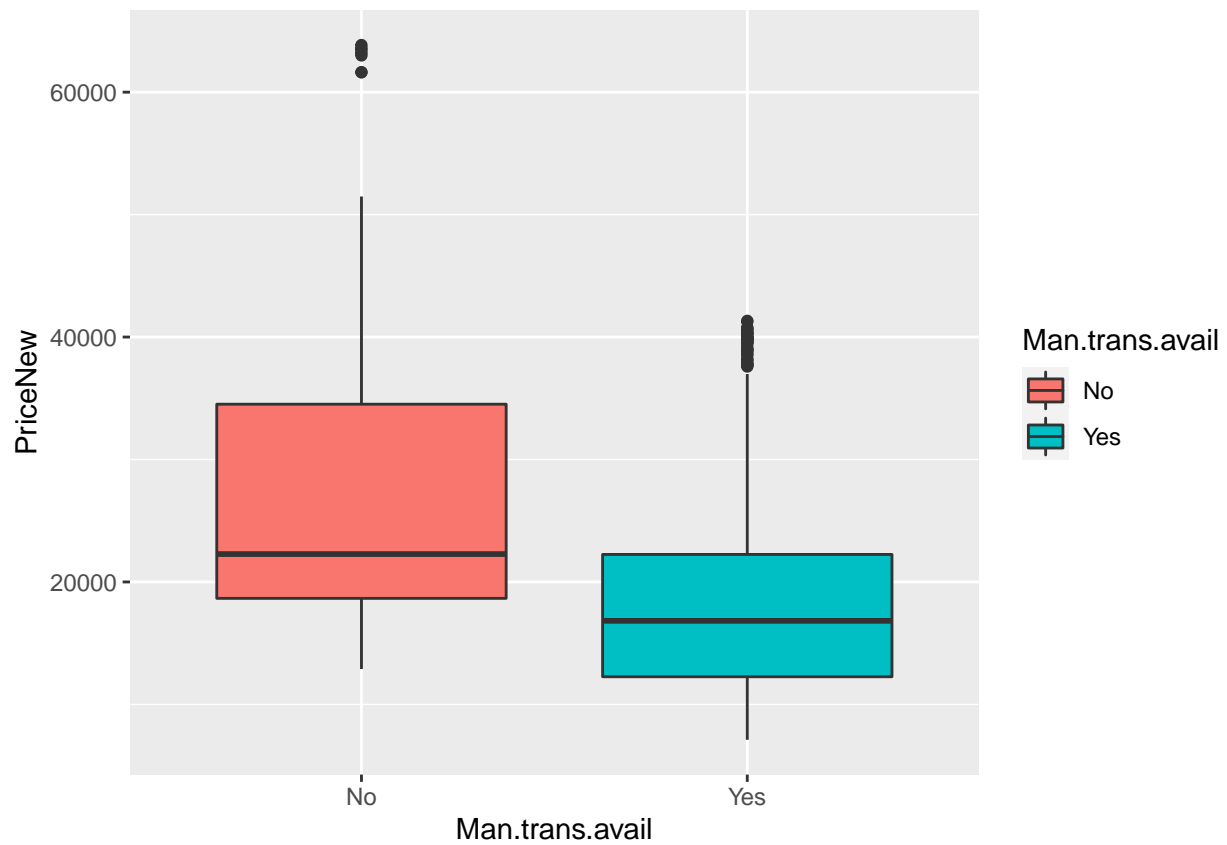
```
# Cylinders  
ggplot(cars.train, aes(x=Cylinders, y=PriceNew)) +geom_boxplot(aes(fill= Cylinders,shape = Cylinders))
```



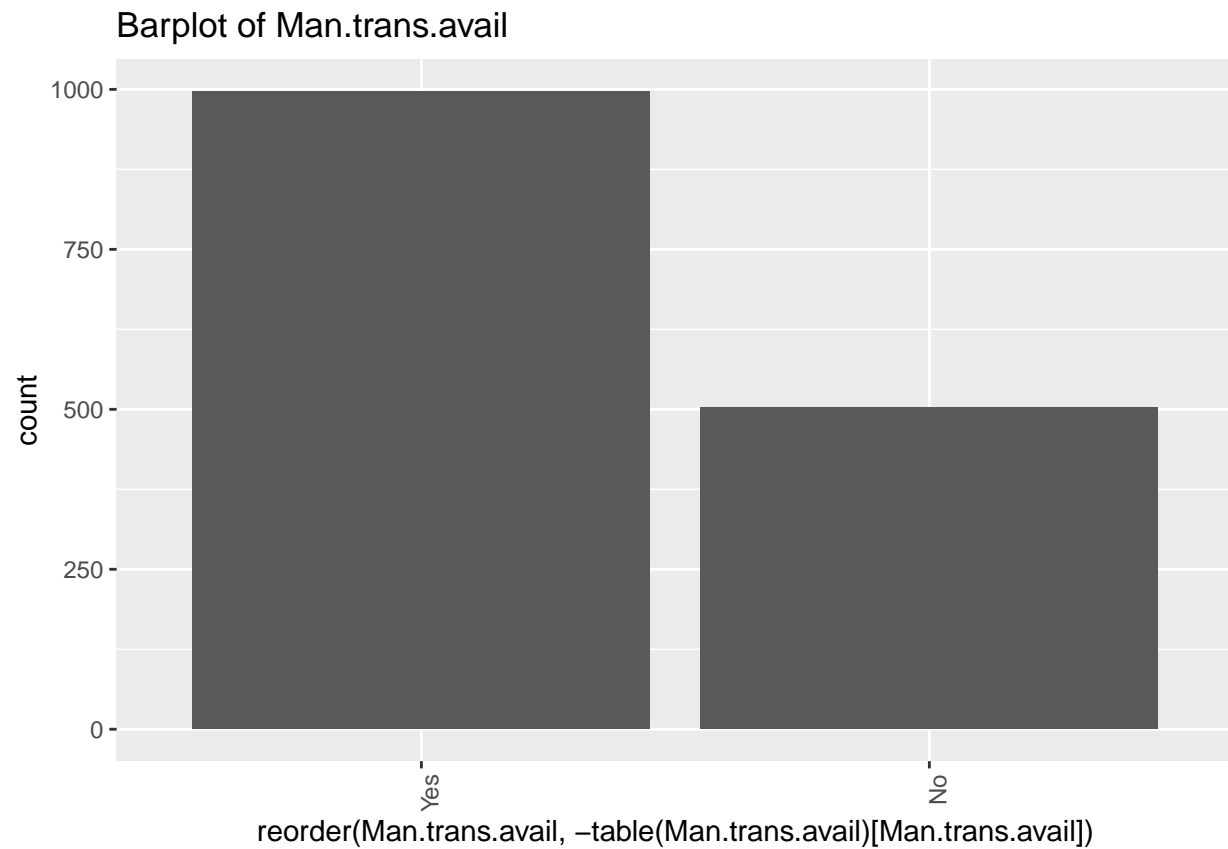
```
ggplot(data=cars.train, aes(x=reorder(Cylinders, -table(Cylinders)[Cylinders]))) + geom_bar() + theme(a
```



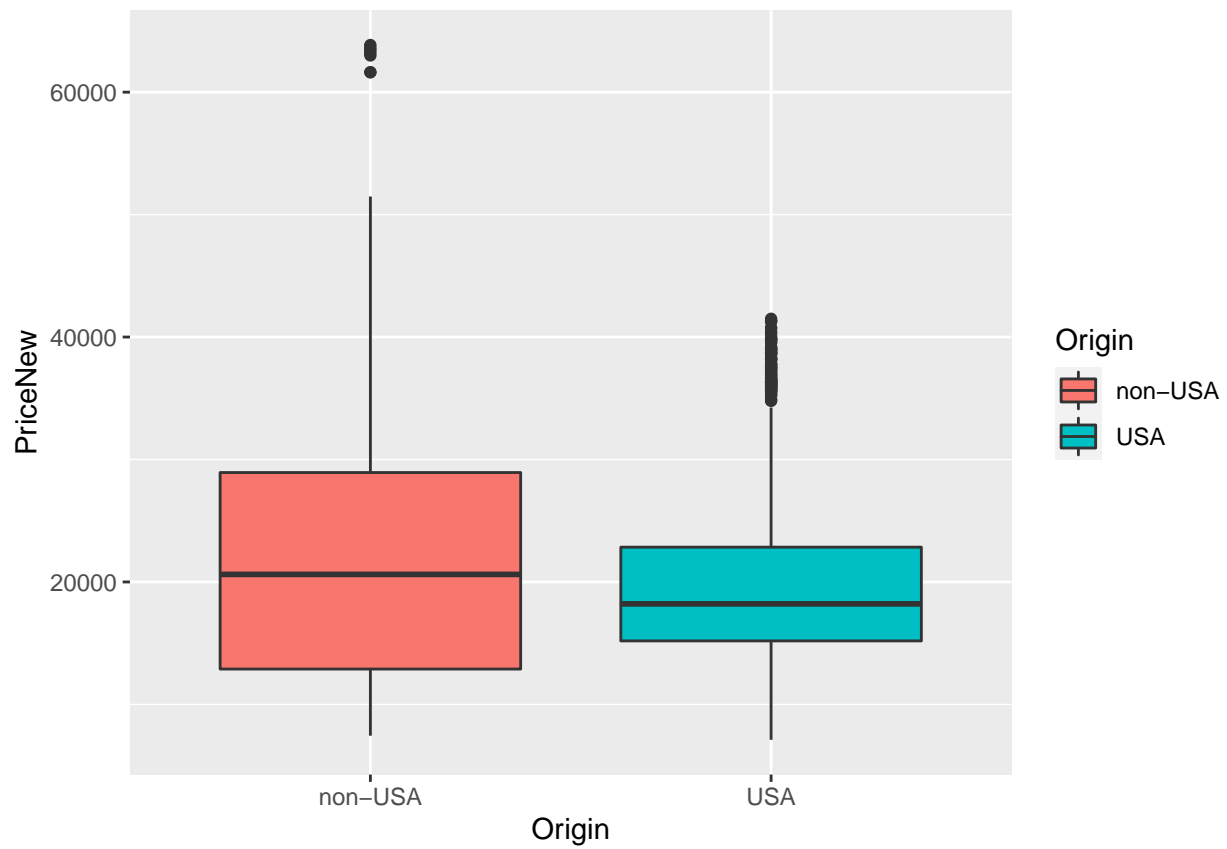
```
# Man.trans.avail  
ggplot(cars.train, aes(x=Man.trans.avail, y=PriceNew)) + geom_boxplot(aes(fill= Man.trans.avail, shape =
```



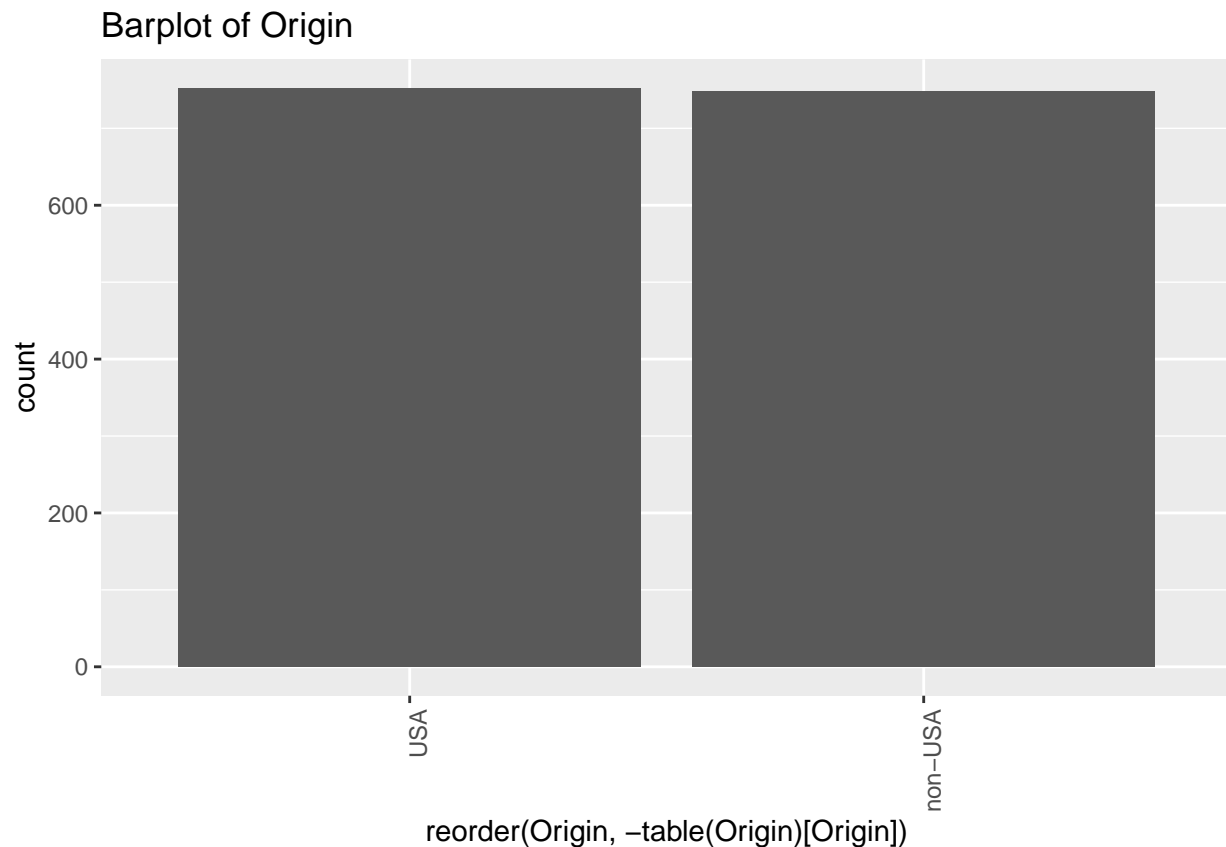
```
ggplot(data=cars.train, aes(x=reorder(Man.trans.avail, -table(Man.trans.avail)[Man.trans.avail]))) + ge
```

```
# Origin  
ggplot(cars.train, aes(x=Origin, y=PriceNew)) +geom_boxplot(aes(fill= Origin,shape = Origin))
```

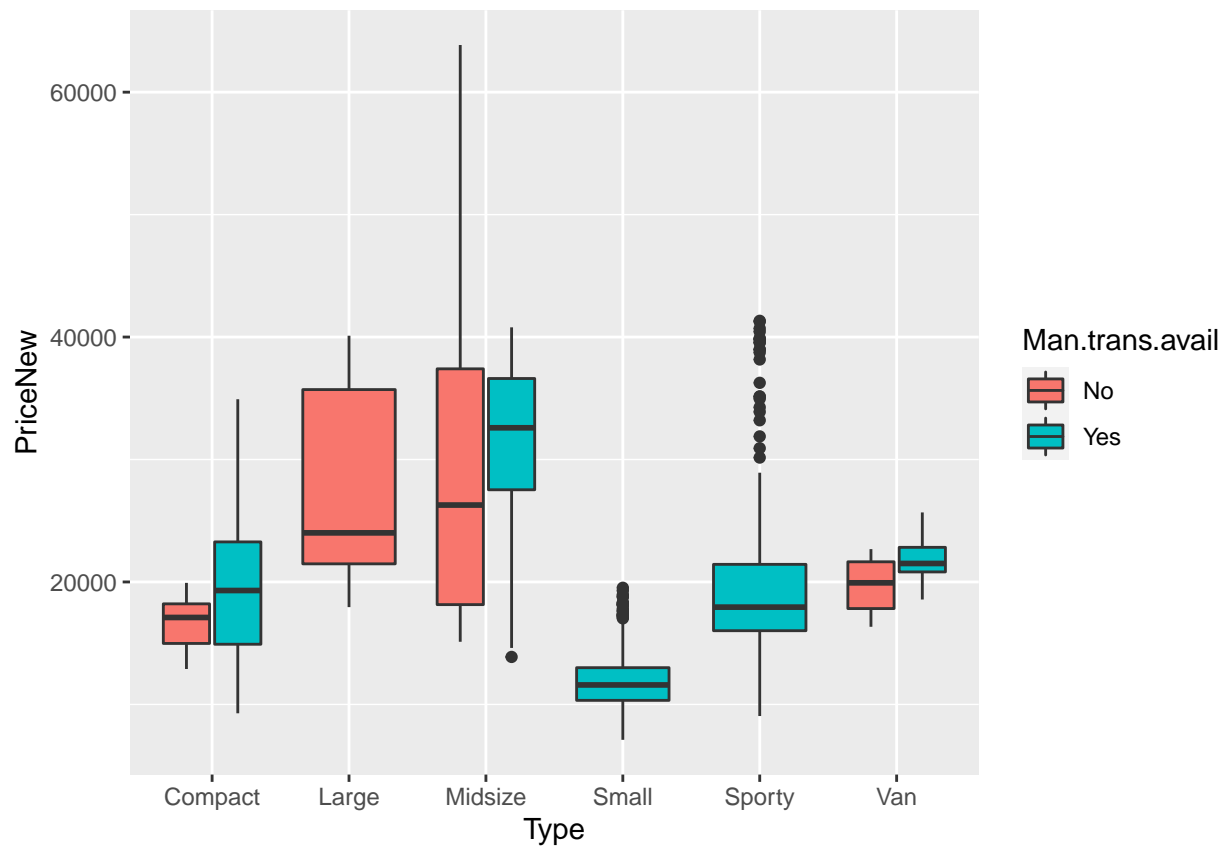


```
ggplot(data=cars.train, aes(x=reorder(Origin, -table(Origin)[Origin]))) + geom_bar() + theme(axis.text..
```

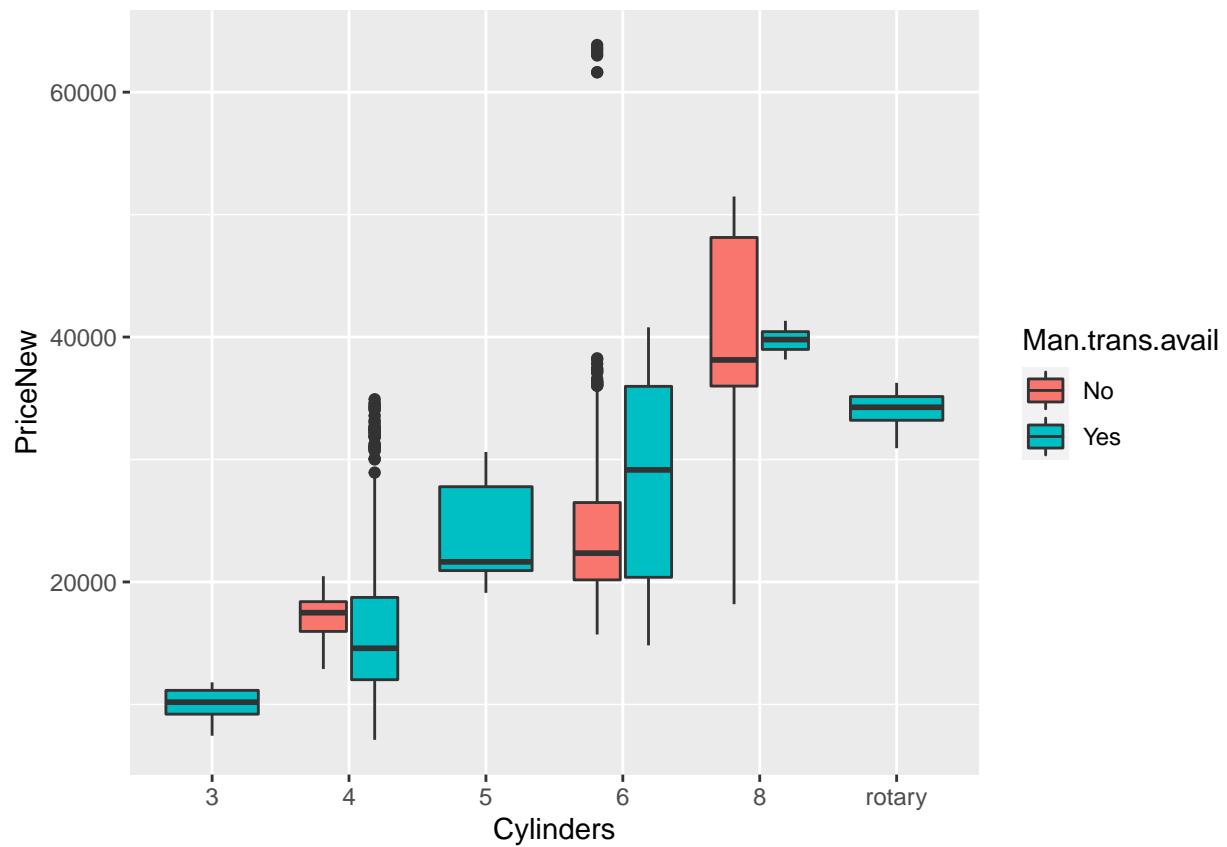


Two Categorical Variables We have seen how each categorical predictor interacts with the price of cars. Let's see some of these variables further broken down by categorical predictors with two outcomes, such as ManTrans and Origin.

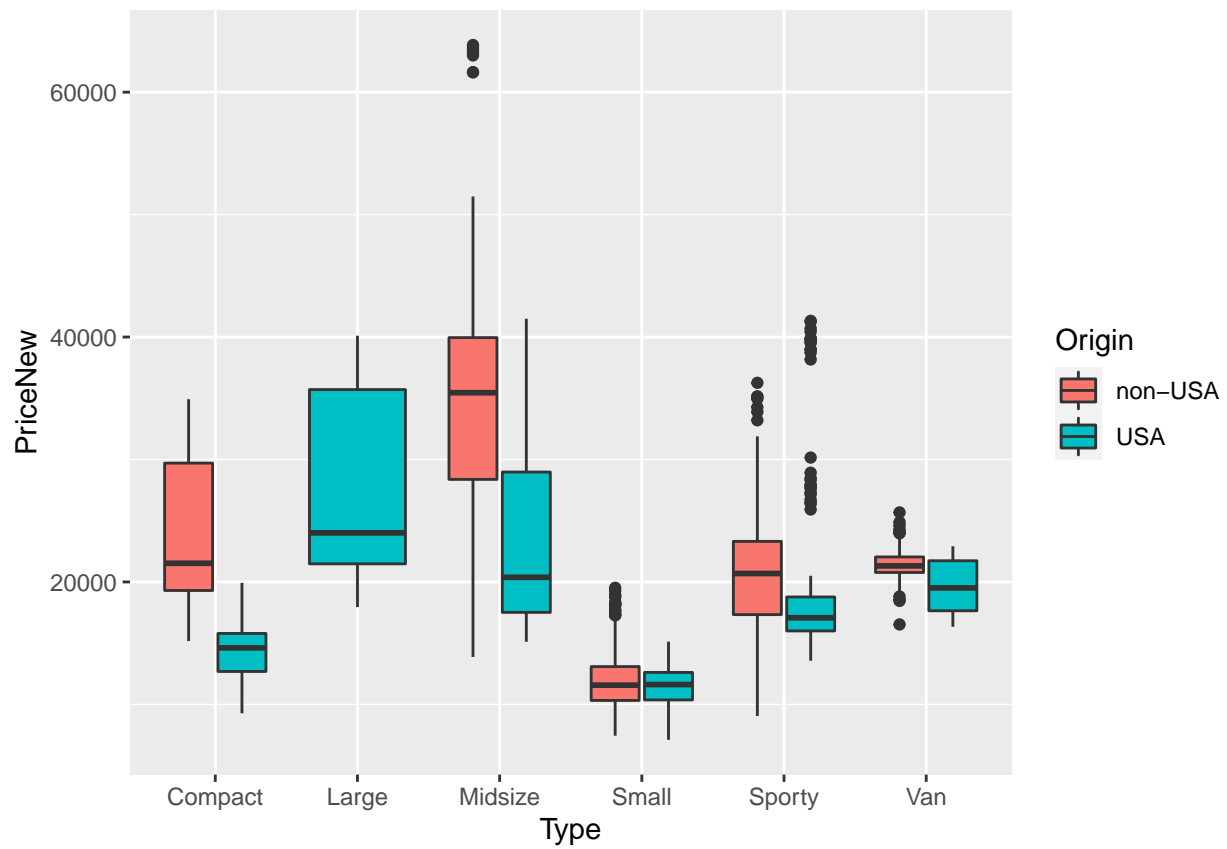
```
# Type by Mantrans
ggplot(cars.train, aes(x=Type, y=PriceNew)) +geom_boxplot(aes(fill= Man.trans.avail,shape = Man.trans.a
```



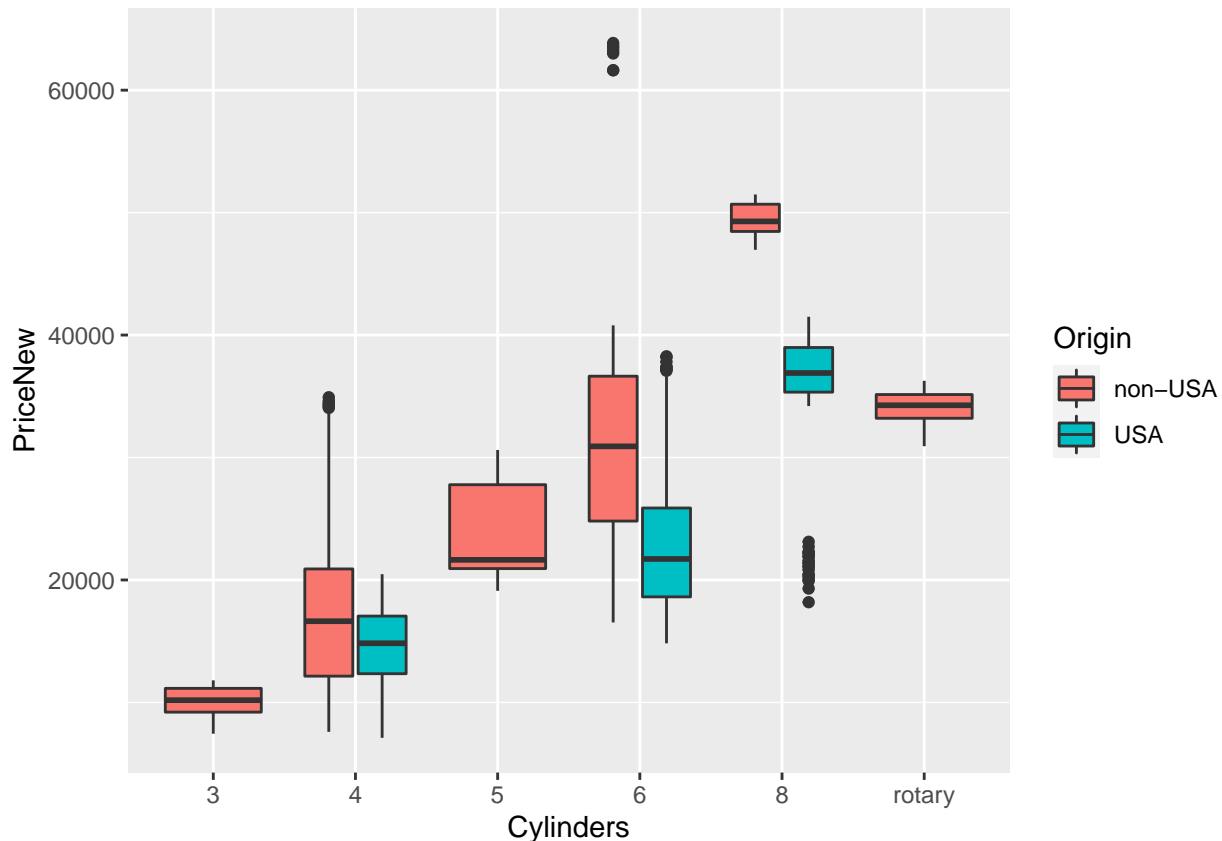
```
# Cylinders by Manstrans
ggplot(cars.train, aes(x=Cylinders, y=PriceNew)) +geom_boxplot(aes(fill= Man.trans.avail,shape = Man.tr
```



```
# Type by Origin
ggplot(cars.train, aes(x=Type, y=PriceNew)) +geom_boxplot(aes(fill= Origin,shape = Origin))
```



```
# Cylinders by Origin (Combine)
ggplot(cars.train, aes(x=Cylinders, y=PriceNew)) +geom_boxplot(aes(fill= Origin,shape = Origin))
```



From these visualizations we can make several initial observations.

First, Model and Make provide the same information, each with 93 unique categories. It would be inefficient for our regression model to predict the prices of cars using their names since this is all the information we would need. Furthermore, a model using this information would have many betas and lead to overfitting. Instead we can summarize this information by creating a new categorical variable that synthesizes this data into different bins. We will create four bins from the model data: Luxury, High, Medium, and Low.

Next, we can group the categories 4WD and Front together since they appear to have similar average prices.

Subsetting cylinders by origin appears to have a significant difference in car prices.

We want to further explore the significance of the rest of the categorical predictors so we will leave them as they are.

Numerical Variables

Next, let's visualize the numerical variables.

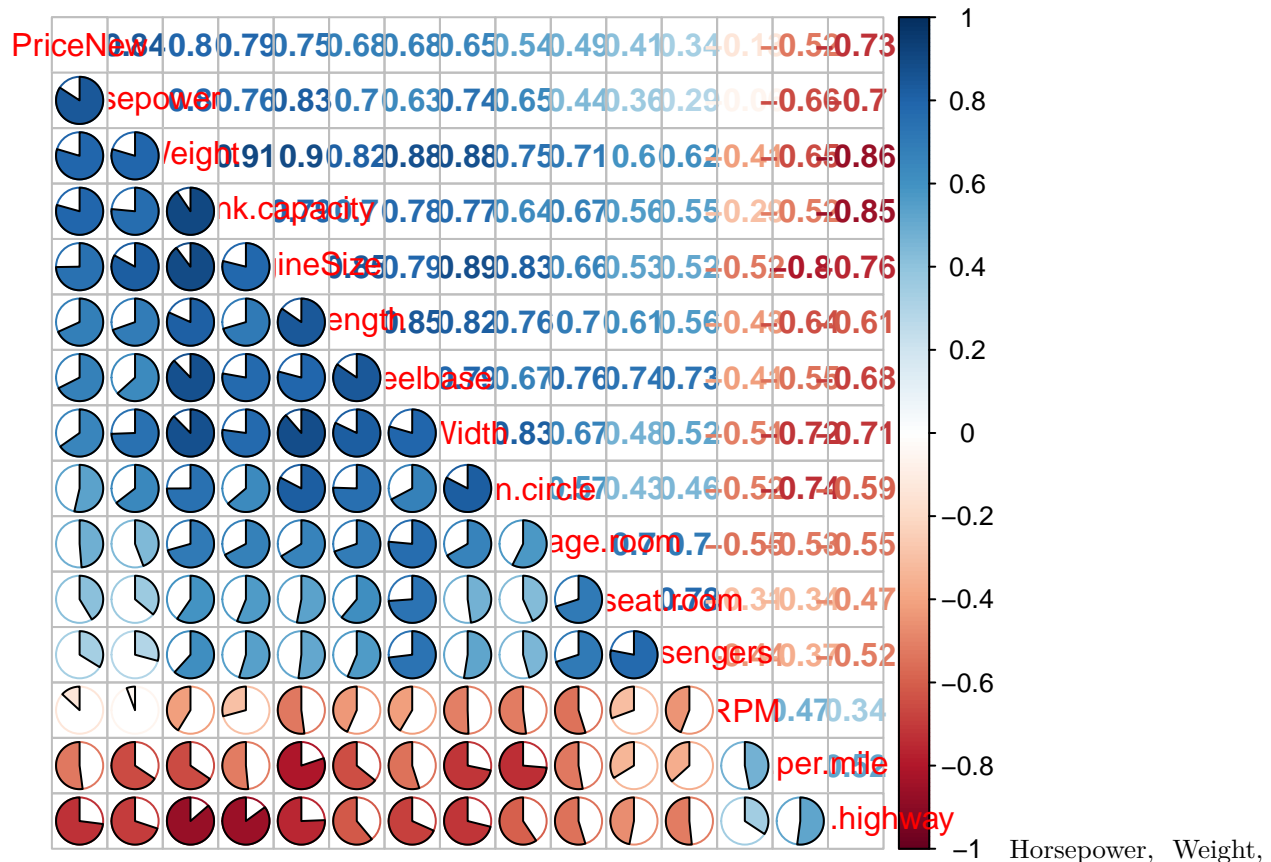
```
numeric <- cars.train[c("PriceNew", "Horsepower", "Weight", "Fuel.tank.capacity", "EngineSize", "Length")]
numeric.cor <- cor(numeric, method = c("spearman"))
# Correlation Matrix
numeric.cor
```

	PriceNew	Horsepower	Weight	Fuel.tank.capacity
## PriceNew	1.0000000	0.84157345	0.7960047	0.7932735
## Horsepower	0.8415734	1.0000000	0.7955683	0.7647021
## Weight	0.7960047	0.79556835	1.0000000	0.9055242
## Fuel.tank.capacity	0.7932735	0.76470206	0.9055242	1.0000000
## EngineSize	0.7474788	0.82993174	0.8977826	0.7889379
## Length	0.6824272	0.69638260	0.8177324	0.7046024

## Wheelbase	0.6757729	0.63289001	0.8778402		0.7781147
## Width	0.6529375	0.74432589	0.8757911		0.7717914
## Turn.circle	0.5362739	0.64526960	0.7493602		0.6383263
## Luggage.room	0.4869604	0.44247987	0.7062168		0.6732930
## Rear.seat.room	0.4133382	0.36110052	0.5957408		0.5613753
## Passengers	0.3382573	0.29188670	0.6173508		0.5464237
## RPM	-0.1317892	-0.05791055	-0.4126433		-0.2918557
## Rev.per.mile	-0.5205092	-0.65642506	-0.6543657		-0.5155725
## MPG.highway	-0.7298849	-0.69888953	-0.8633088		-0.8505750
##	EngineSize	Length	Wheelbase	Width	Turn.circle
## PriceNew	0.7474788	0.6824272	0.6757729	0.6529375	0.5362739
## Horsepower	0.8299317	0.6963826	0.6328900	0.7443259	0.6452696
## Weight	0.8977826	0.8177324	0.8778402	0.8757911	0.7493602
## Fuel.tank.capacity	0.7889379	0.7046024	0.7781147	0.7717914	0.6383263
## EngineSize	1.0000000	0.8470256	0.7912021	0.8862335	0.8261247
## Length	0.8470256	1.0000000	0.8451968	0.8209058	0.7551547
## Wheelbase	0.7912021	0.8451968	1.0000000	0.7946504	0.6730567
## Width	0.8862335	0.8209058	0.7946504	1.0000000	0.8260158
## Turn.circle	0.8261247	0.7551547	0.6730567	0.8260158	1.0000000
## Luggage.room	0.6616642	0.6975178	0.7646091	0.6676682	0.5742701
## Rear.seat.room	0.5302730	0.6109570	0.7387207	0.4778486	0.4339578
## Passengers	0.5193854	0.5632721	0.7309220	0.5249309	0.4580523
## RPM	-0.5208886	-0.4341898	-0.4149213	-0.5056740	-0.5198411
## Rev.per.mile	-0.8013462	-0.6420070	-0.5488991	-0.7177080	-0.7363931
## MPG.highway	-0.7562031	-0.6105728	-0.6833100	-0.7111176	-0.5936773
##	Luggage.room	Rear.seat.room	Passengers		RPM
## PriceNew	0.4869604	0.4133382	0.3382573	-0.13178917	
## Horsepower	0.4424799	0.3611005	0.2918867	-0.05791055	
## Weight	0.7062168	0.5957408	0.6173508	-0.41264326	
## Fuel.tank.capacity	0.6732930	0.5613753	0.5464237	-0.29185570	
## EngineSize	0.6616642	0.5302730	0.5193854	-0.52088859	
## Length	0.6975178	0.6109570	0.5632721	-0.43418977	
## Wheelbase	0.7646091	0.7387207	0.7309220	-0.41492129	
## Width	0.6676682	0.4778486	0.5249309	-0.50567403	
## Turn.circle	0.5742701	0.4339578	0.4580523	-0.51984113	
## Luggage.room	1.0000000	0.7008921	0.7006482	-0.54814222	
## Rear.seat.room	0.7008921	1.0000000	0.7799654	-0.30542290	
## Passengers	0.7006482	0.7799654	1.0000000	-0.44110472	
## RPM	-0.5481422	-0.3054229	-0.4411047	1.00000000	
## Rev.per.mile	-0.5284816	-0.3373567	-0.3670689	0.47004964	
## MPG.highway	-0.5469450	-0.4690272	-0.5166257	0.34305958	
##	Rev.per.mile	MPG.highway			
## PriceNew	-0.5205092	-0.7298849			
## Horsepower	-0.6564251	-0.6988895			
## Weight	-0.6543657	-0.8633088			
## Fuel.tank.capacity	-0.5155725	-0.8505750			
## EngineSize	-0.8013462	-0.7562031			
## Length	-0.6420070	-0.6105728			
## Wheelbase	-0.5488991	-0.6833100			
## Width	-0.7177080	-0.7111176			
## Turn.circle	-0.7363931	-0.5936773			
## Luggage.room	-0.5284816	-0.5469450			
## Rear.seat.room	-0.3373567	-0.4690272			
## Passengers	-0.3670689	-0.5166257			


```
## RPM                0.4700496    0.3430596
## Rev.per.mile       1.0000000    0.5208963
## MPG.highway        0.5208963    1.0000000
```

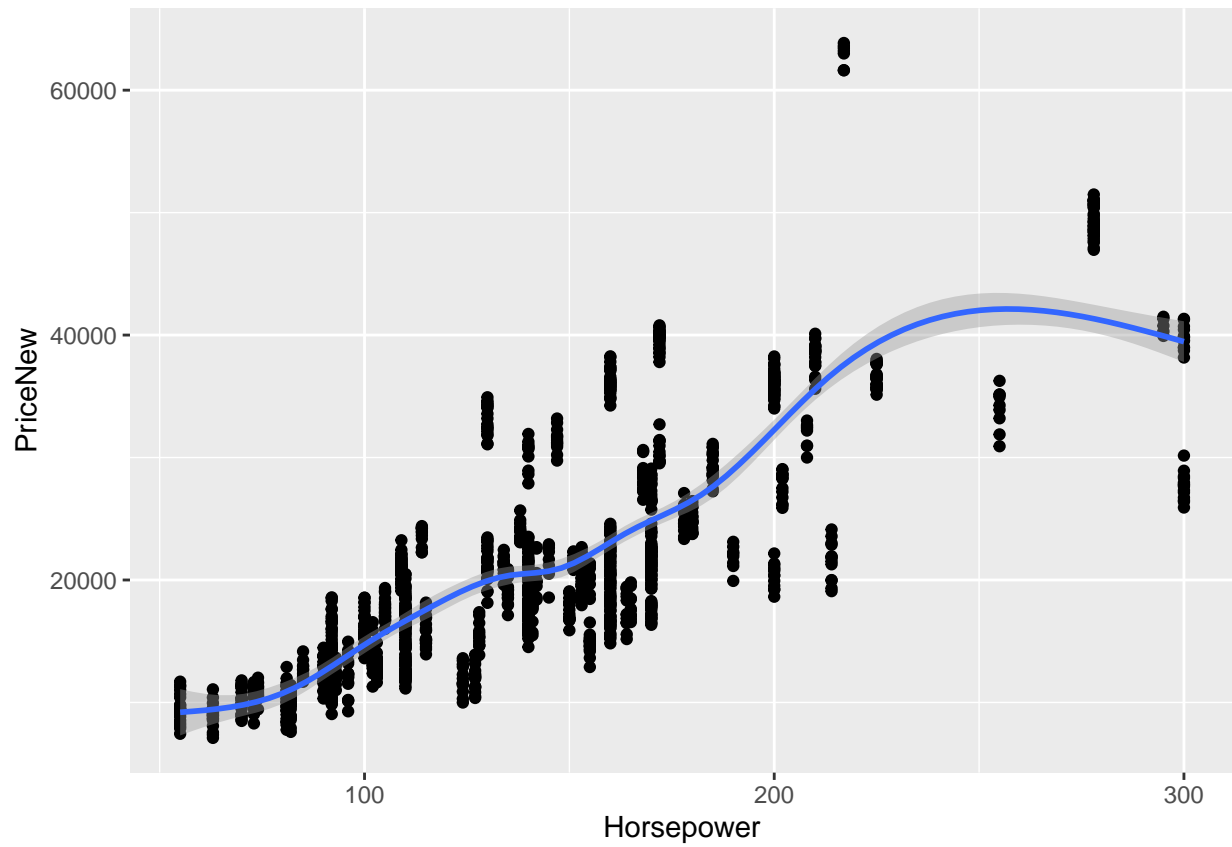
```
corrplot.mixed(numeric.cor, upper="number", lower="pie")
```



and Fuel tank capacity have the strongest positive correlation to price, while MPG, rev per mile, and RPM have the strongest negative correlation. Let's visualize each of our independent continuous variables along with Price. ## Visualizing numerical variables

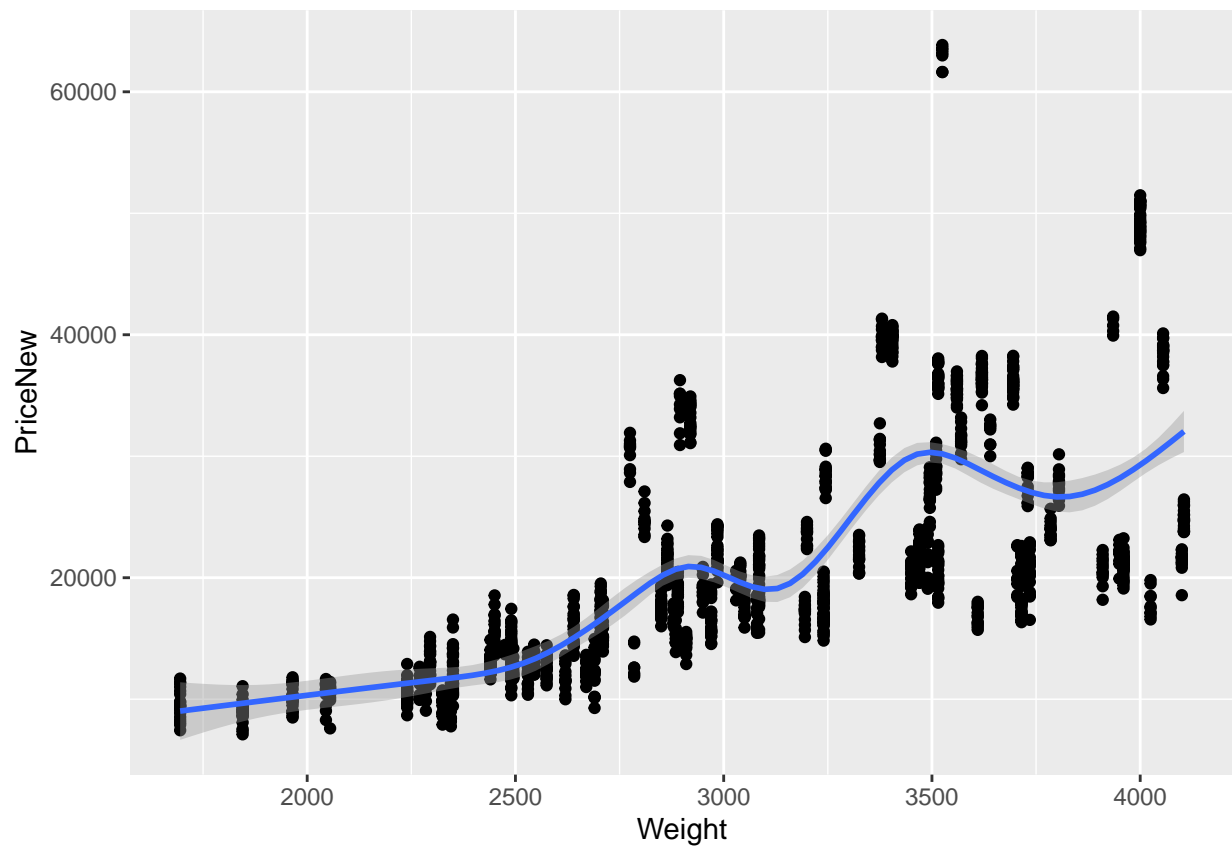
```
ggplot(cars.train, aes(x=Horsepower, y=PriceNew)) + geom_point() + geom_smooth()
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

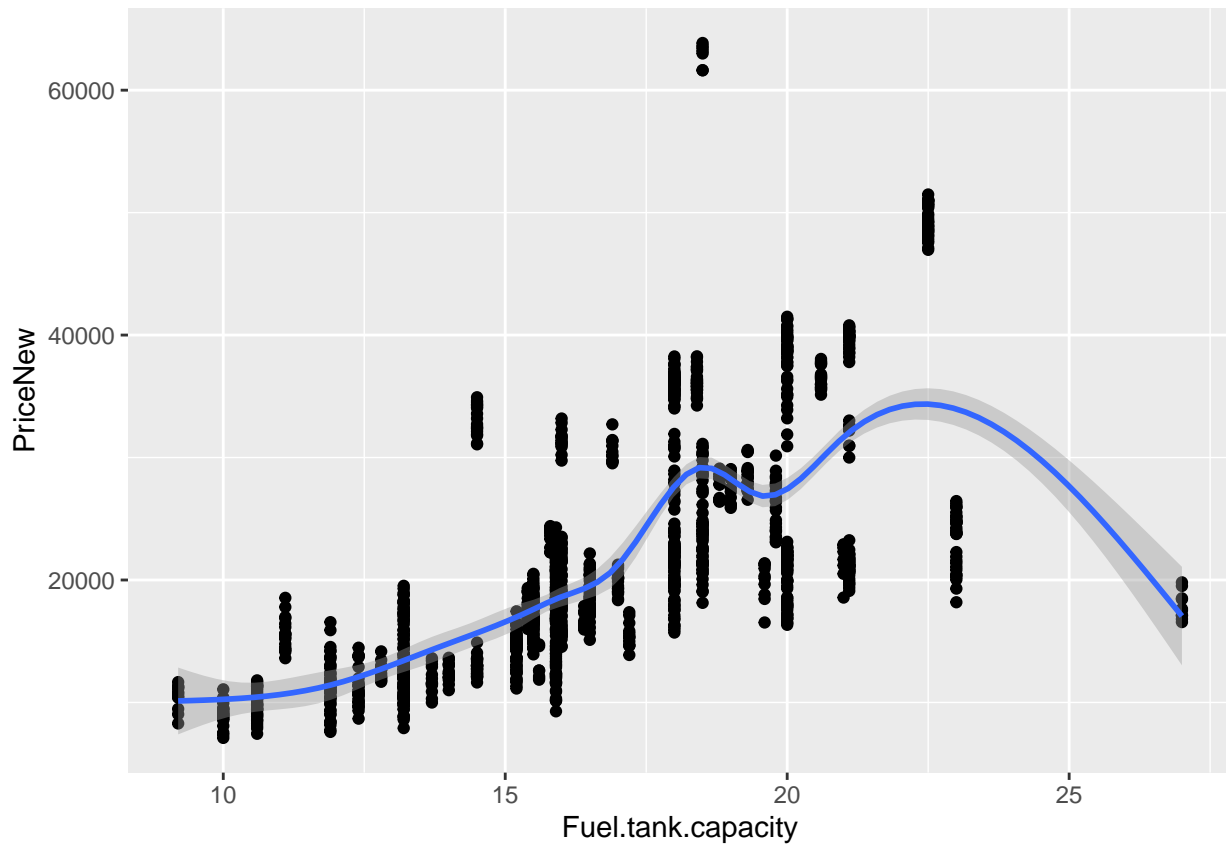


```
ggplot(cars.train, aes(x=Horsepower, y=PriceNew)) + geom_point() + geom_smooth()
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

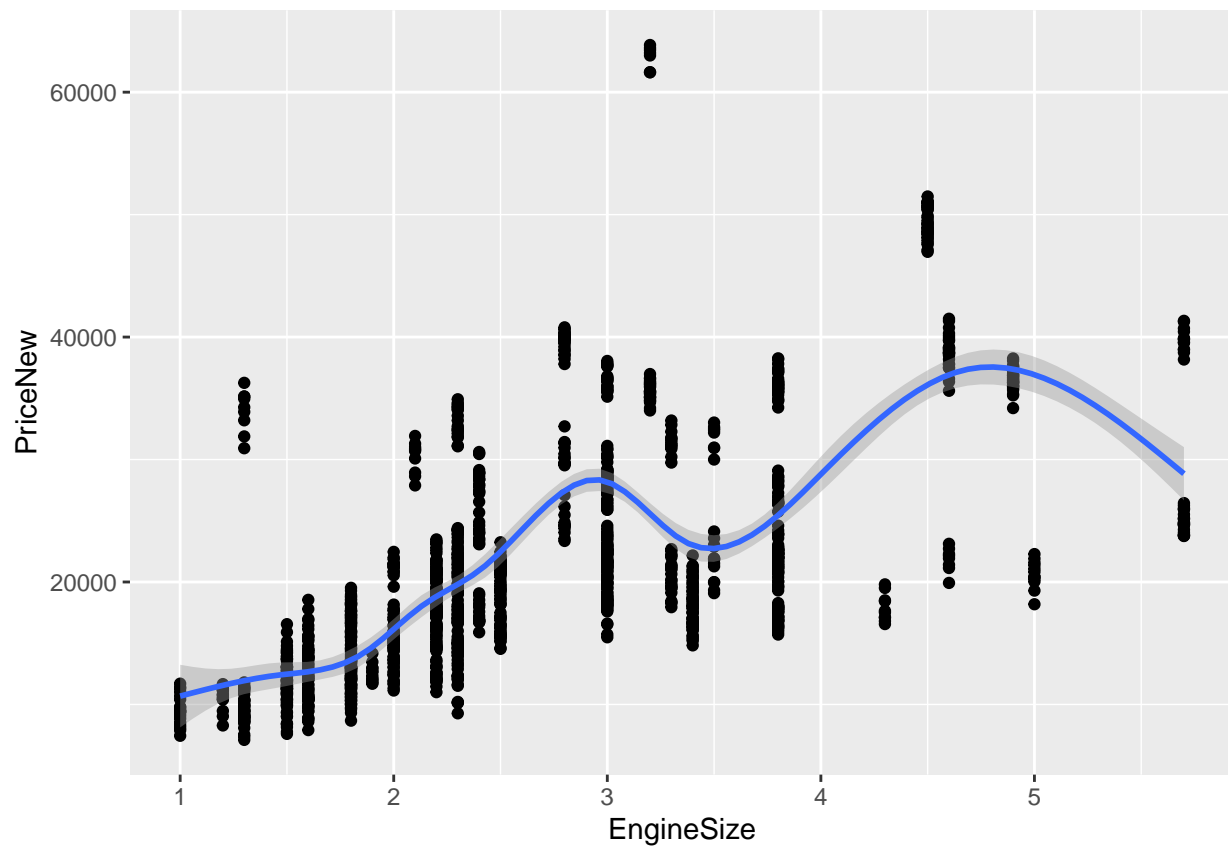


```
ggplot(cars.train, aes(x=Fuel.tank.capacity, y=PriceNew)) + geom_point() + geom_smooth()  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



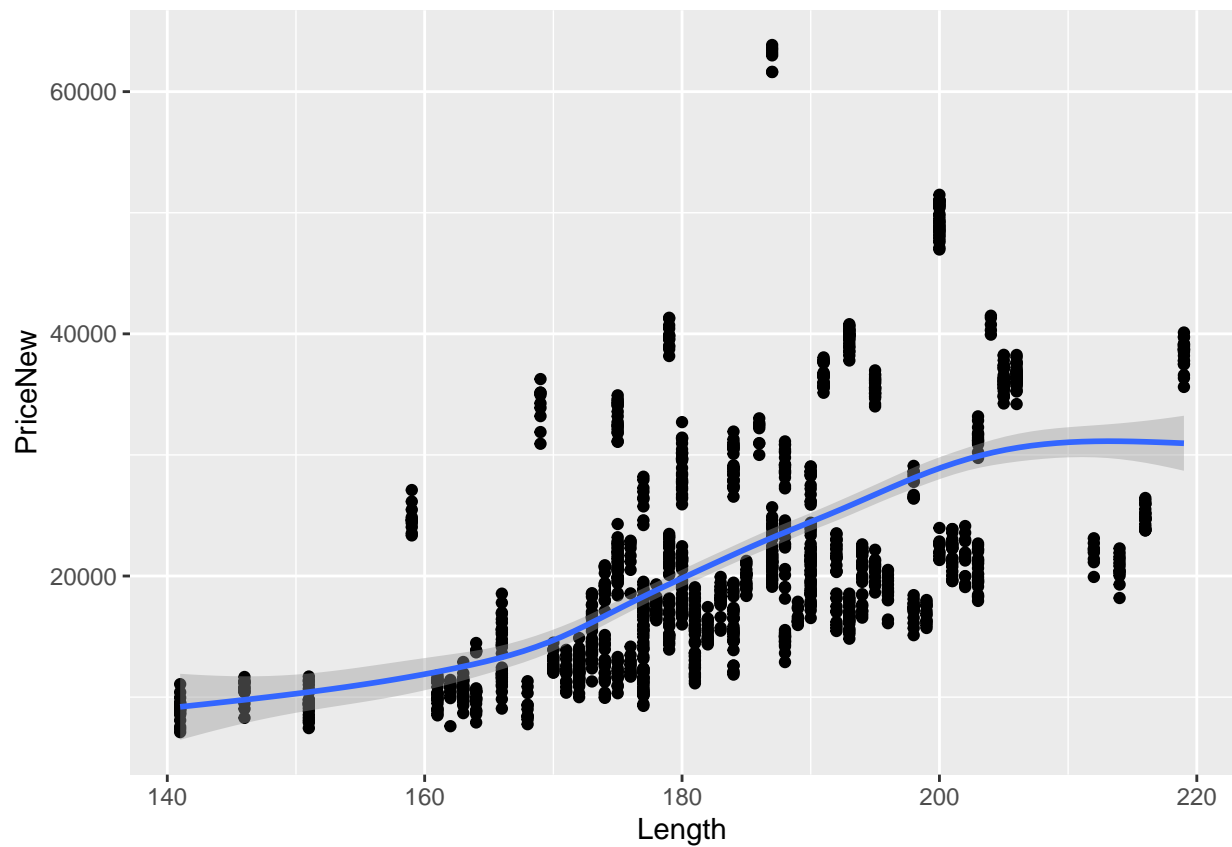
```
ggplot(cars.train, aes(x=EngineSize, y=PriceNew)) + geom_point() + geom_smooth()
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



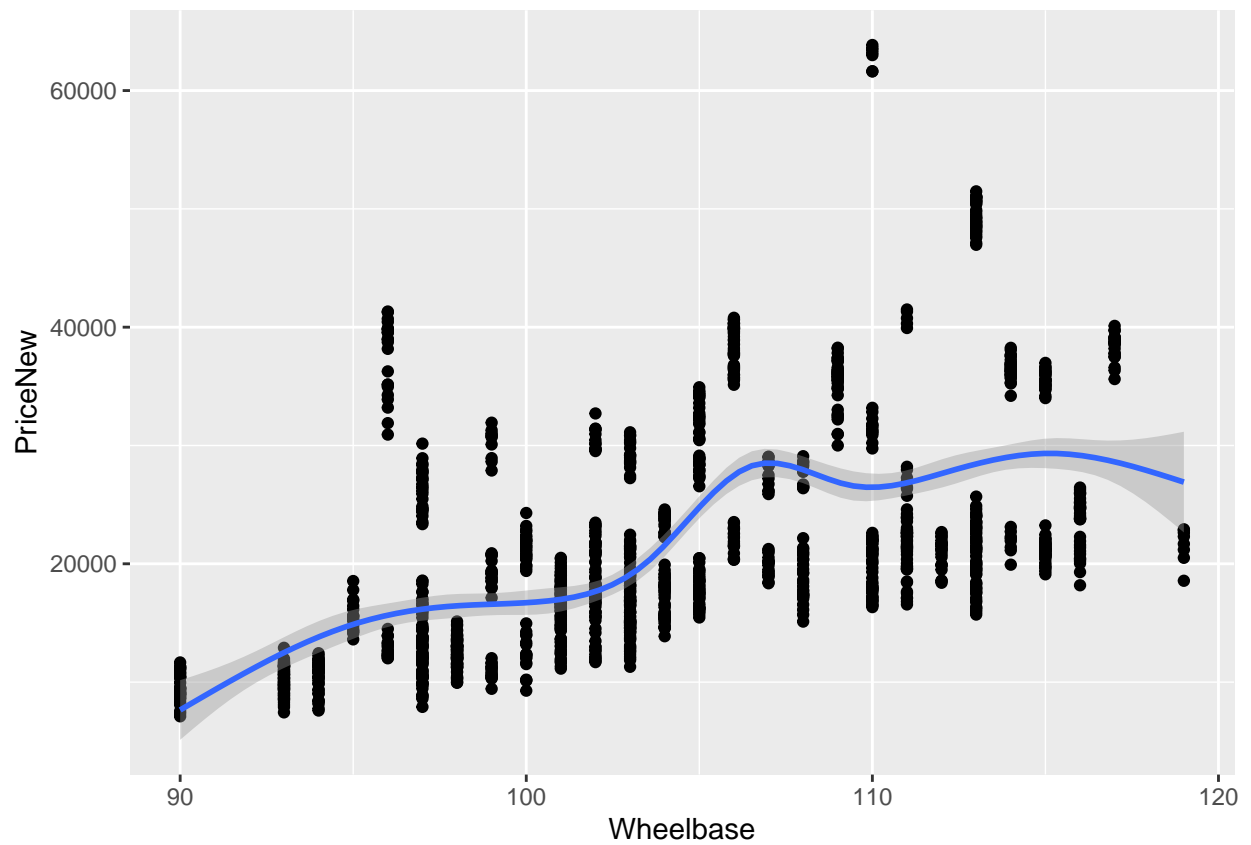
```
ggplot(cars.train, aes(x=Length, y=PriceNew)) + geom_point() + geom_smooth()
```

```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

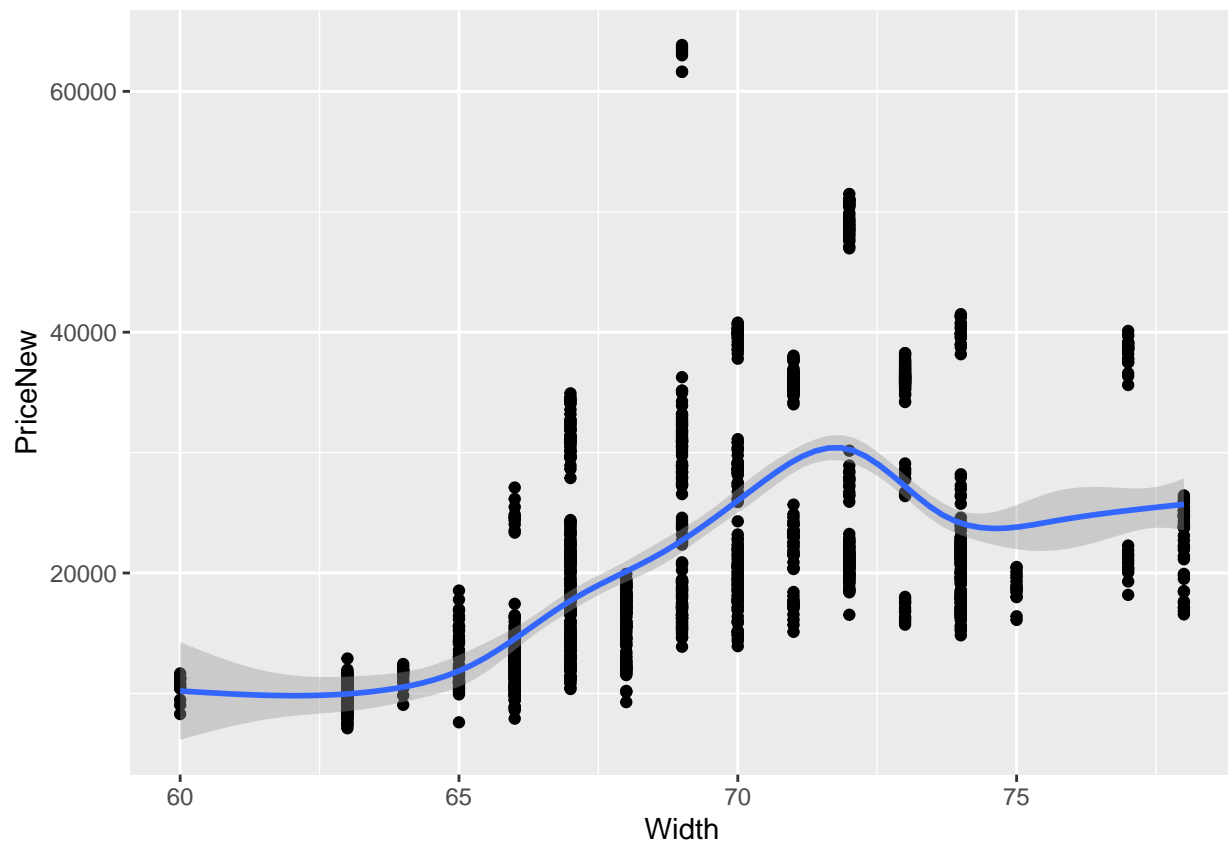


```
ggplot(cars.train, aes(x=Wheelbase, y=PriceNew)) + geom_point() + geom_smooth()
```

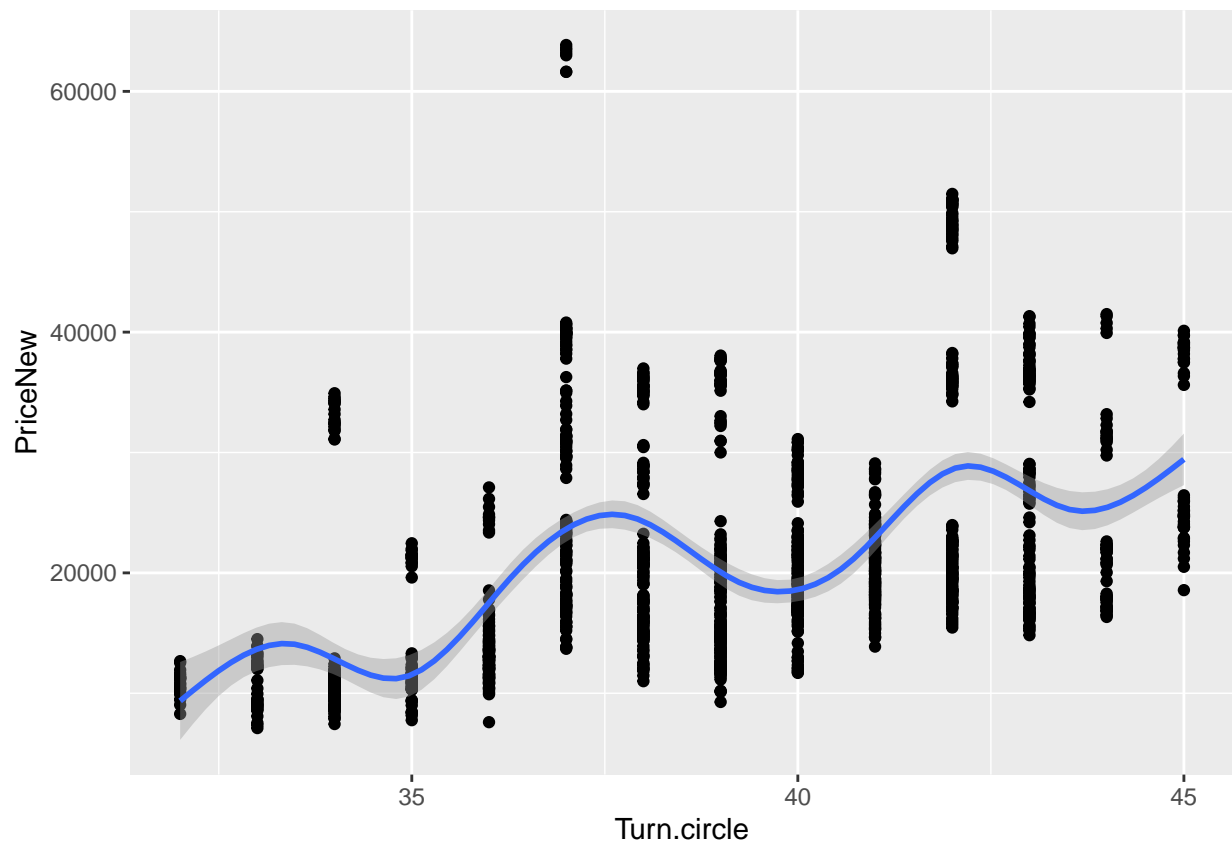
```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



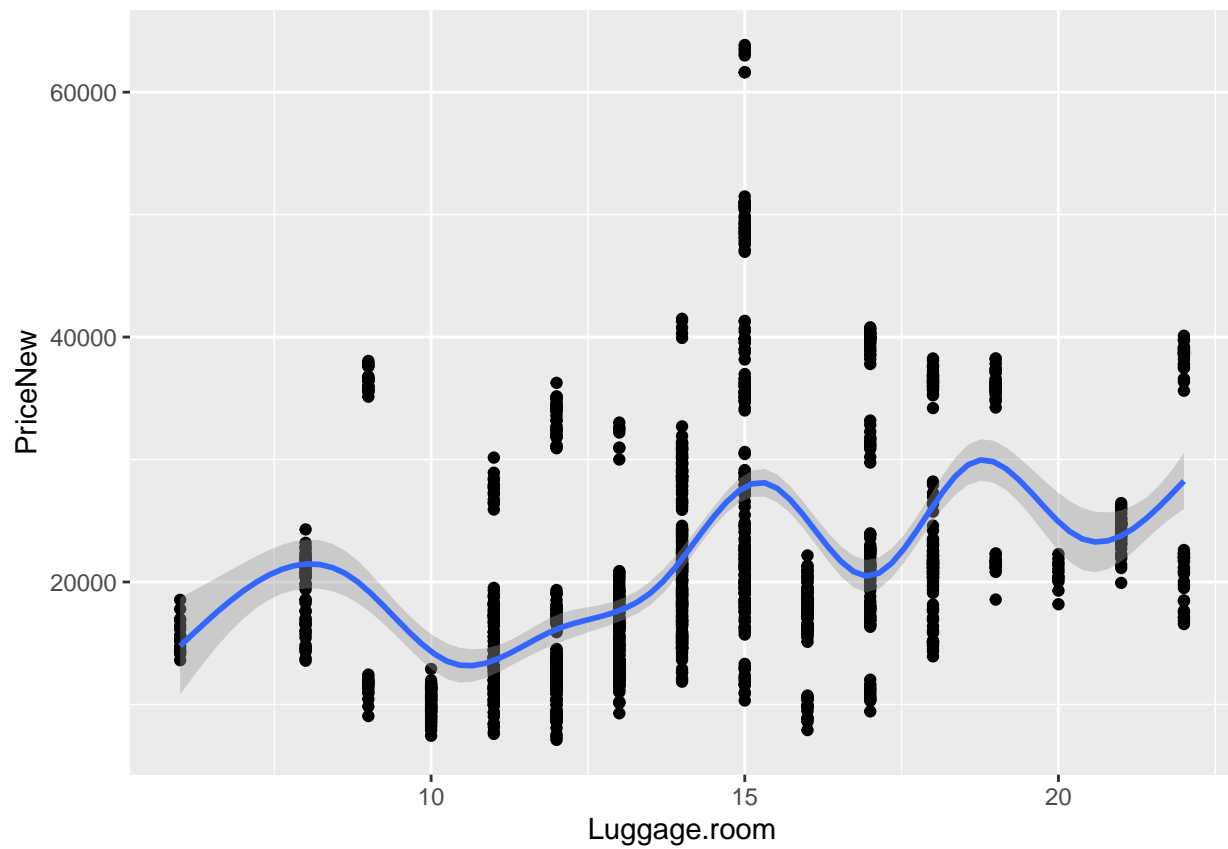
```
ggplot(cars.train, aes(x=Wheelbase, y=PriceNew)) + geom_point() + geom_smooth()  
  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



```
ggplot(cars.train, aes(x=Turn.circle, y=PriceNew)) + geom_point() + geom_smooth()  
  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

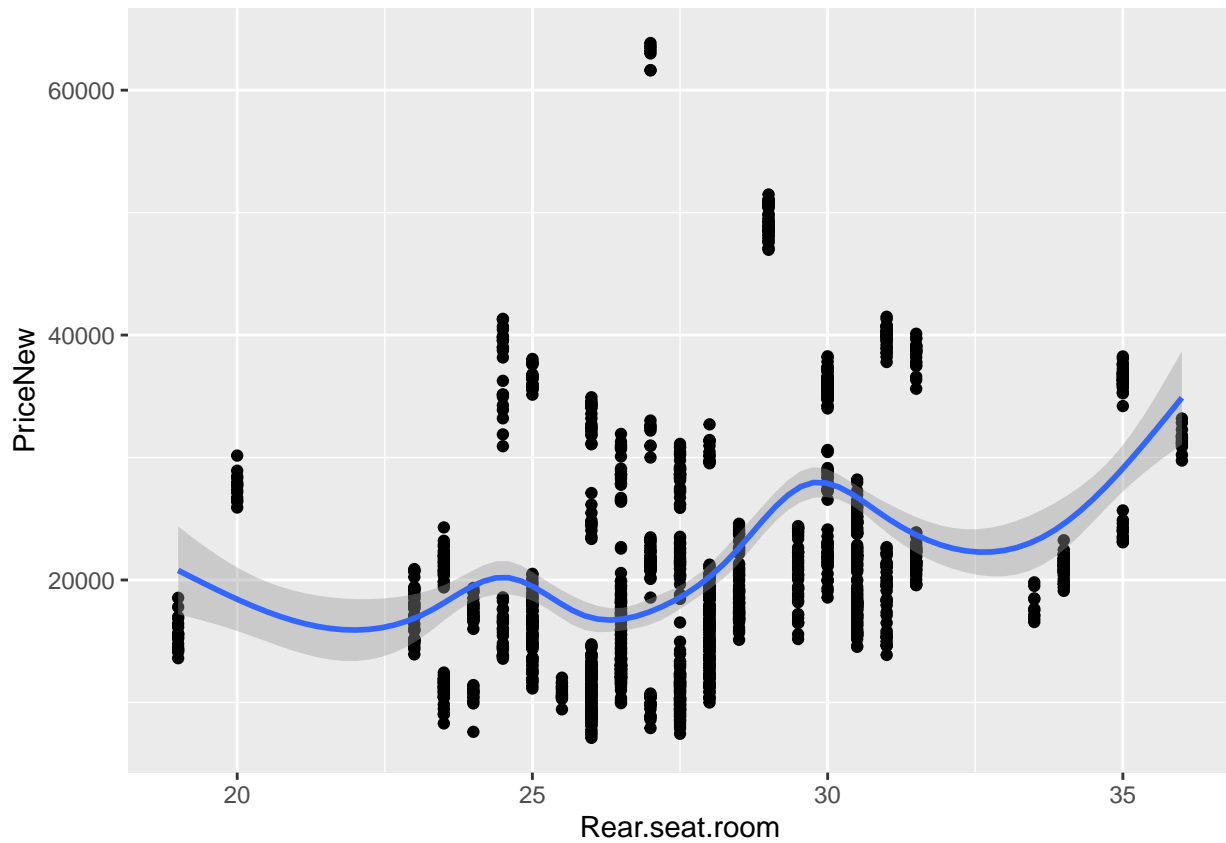



```
ggplot(cars.train, aes(x=Luggage.room, y=PriceNew)) + geom_point() + geom_smooth()  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



```
ggplot(cars.train, aes(x=Rear.seat.room, y=PriceNew)) + geom_point() + geom_smooth()

## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

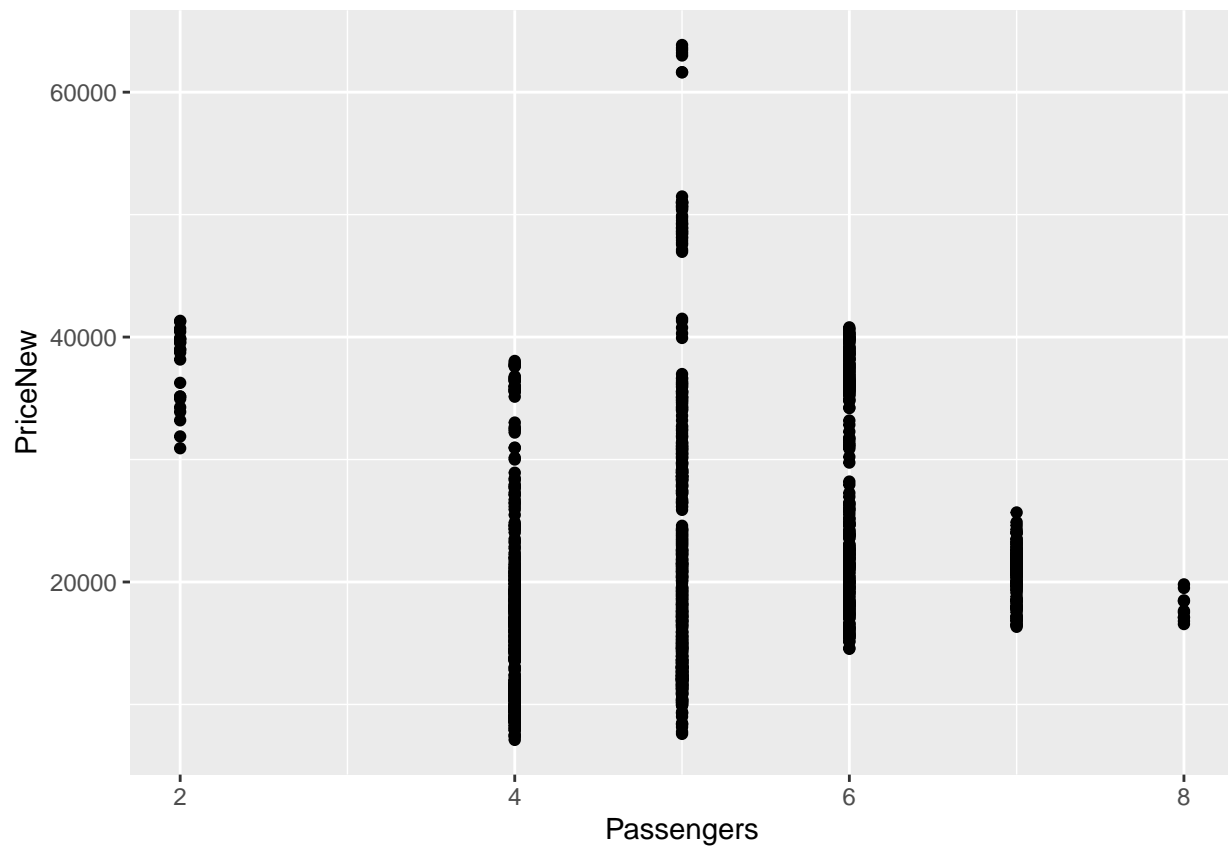


```
ggplot(cars.train, aes(x=Rear.seat.room, y=PriceNew)) + geom_point() + geom_smooth()
```

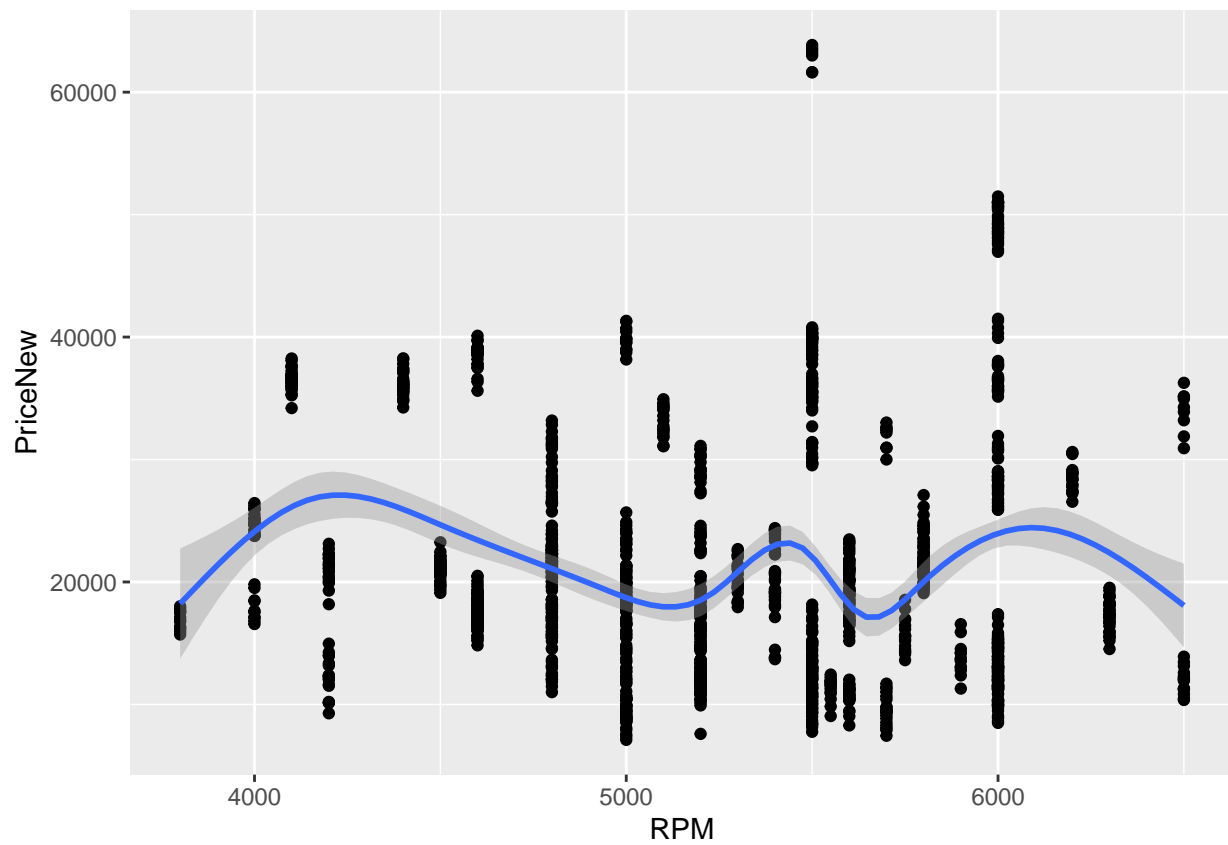
```
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```

```
## Warning: Computation failed in `stat_smooth()`:
```

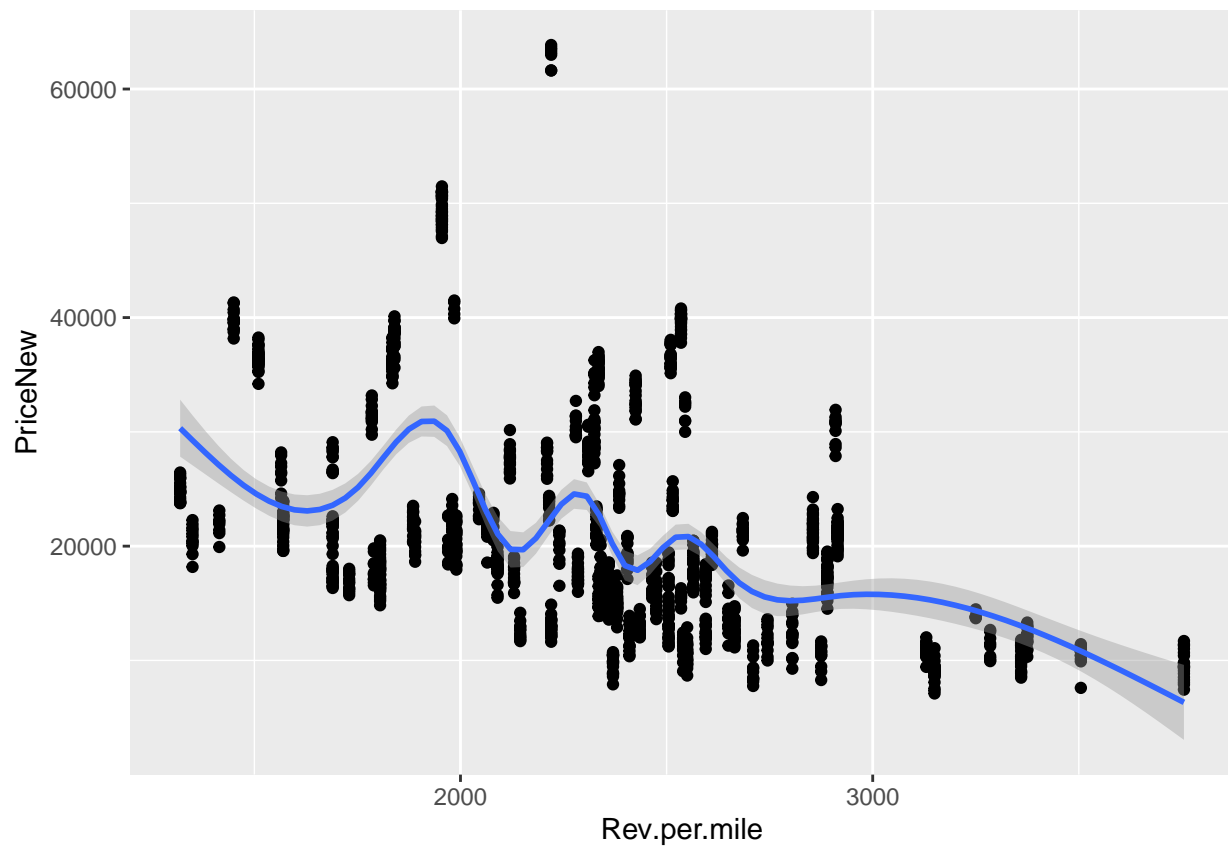
```
## x has insufficient unique values to support 10 knots: reduce k.
```



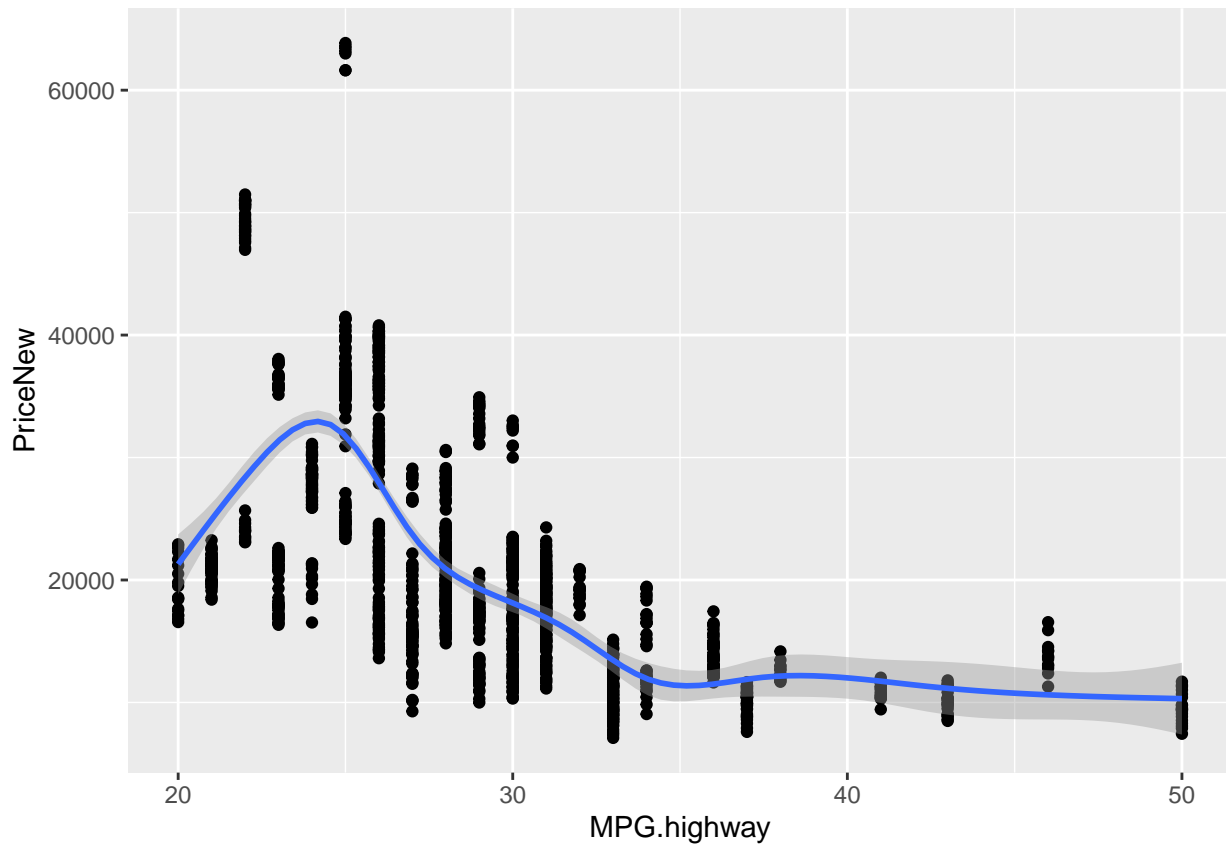
```
ggplot(cars.train, aes(x=RPM, y=PriceNew)) + geom_point() + geom_smooth()  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



```
ggplot(cars.train, aes(x=Rev.per.mile, y=PriceNew)) + geom_point() + geom_smooth()  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



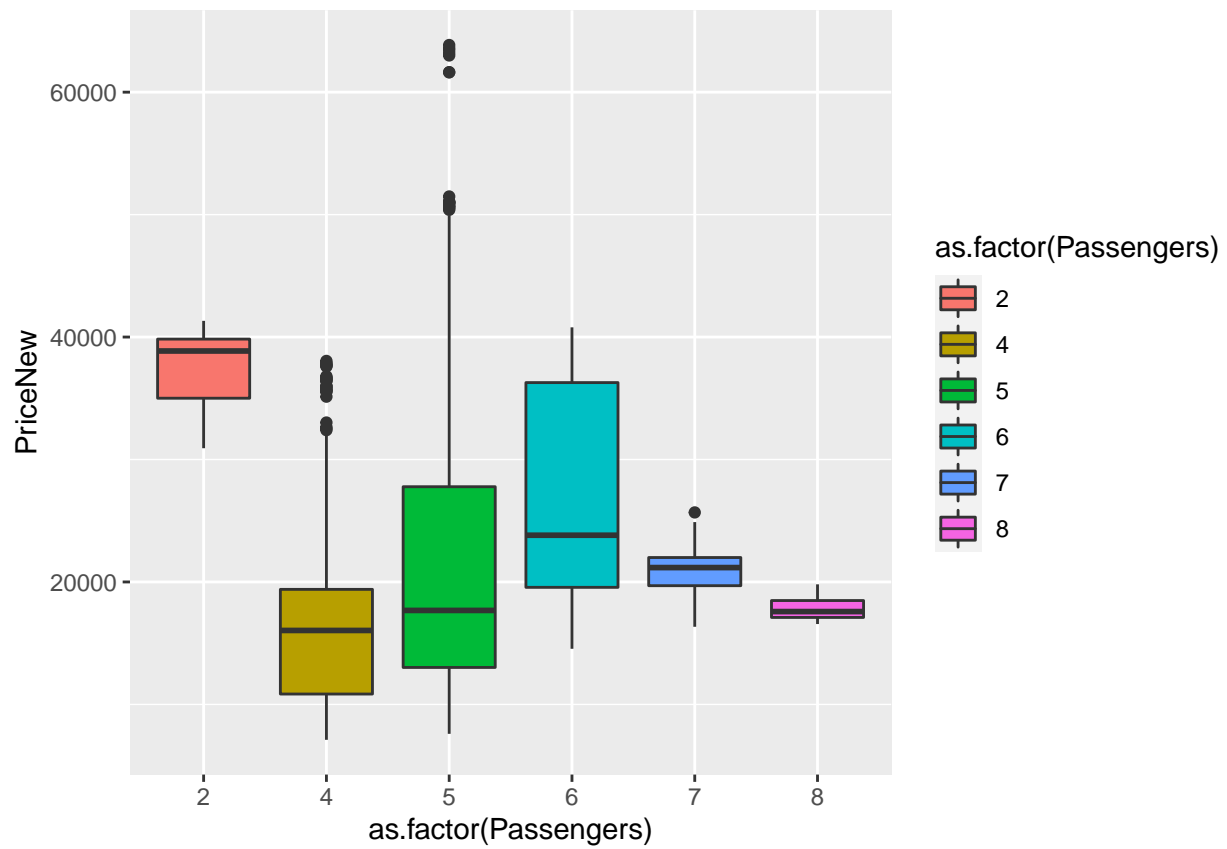
```
ggplot(cars.train, aes(x=MPG.highway, y=PriceNew)) + geom_point() + geom_smooth()  
## `geom_smooth()` using method = 'gam' and formula 'y ~ s(x, bs = "cs")'
```



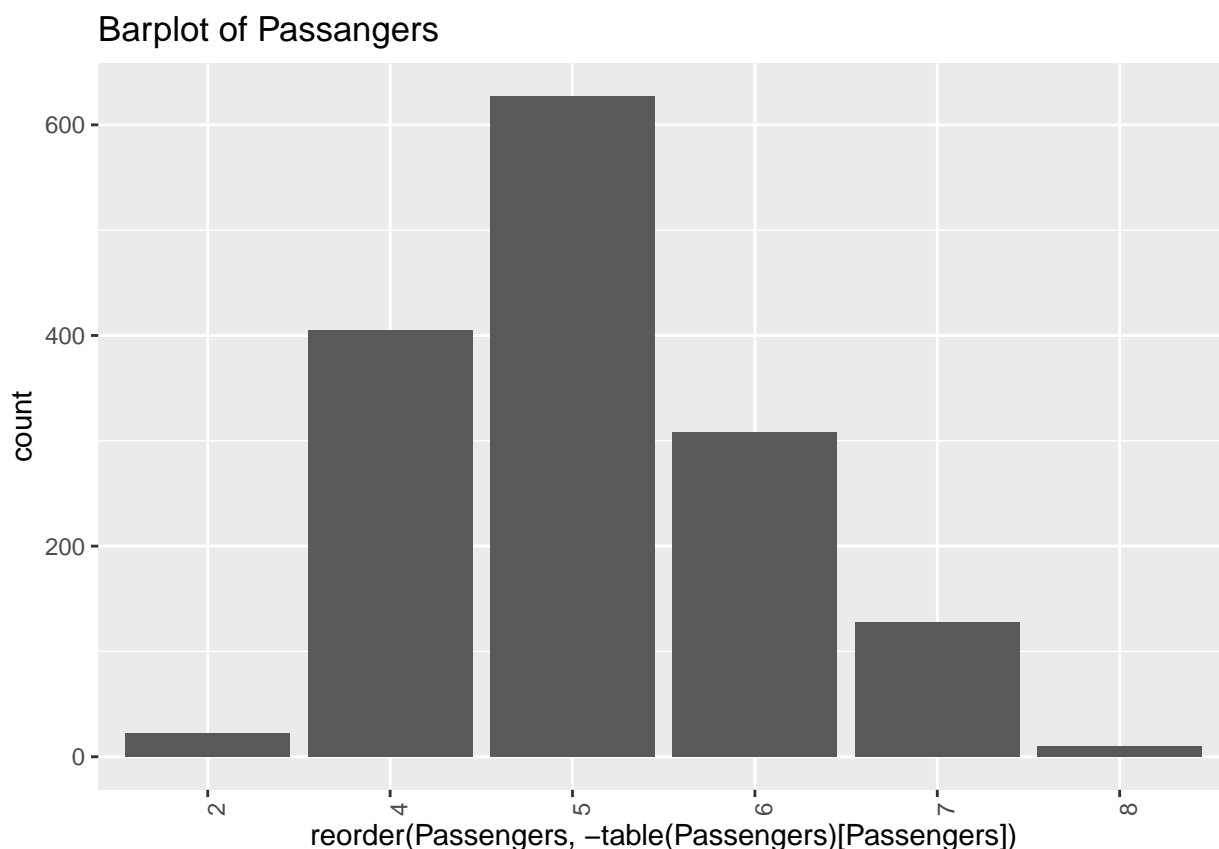
From these plots of the numerical predictors we can see that passengers would be better off as a categorical predictor. Let's treat it as a factor and plot it again.

Passengers as a categorical predictor

```
ggplot(cars.train, aes(x=as.factor(Passengers), y=PriceNew)) +geom_boxplot(aes(fill= as.factor(Passenger
```



```
ggplot(data=cars.train, aes(x=reorder(Passengers, -table(Passengers)[Passengers]))) + geom_bar() + theme
```

Creating new variables

We can now create new variables from the various observations we made above.

```
# New categorical variable for the models based on price
cars.train$mod_category <- ifelse(mean_prices_model[cars.train$Model] >= 40000, "Luxury", ifelse(mean_
# Combine 4WD and Front together
cars.train$DriveTrain <- ifelse(cars.train$DriveTrain == "Rear", "Rear", "Other")
# Combining Cylinders and Origin
cars.train$Cylinders-Origin <- paste(cars.train$Cylinders, cars.train$Origin)
```

Create a new training data frame

```
# Drop the Ob column and original categorical predictors for which we created new variables
# 1-Ob, 2-Manufacturer, 3-Model, 8-Cylinders, 23-Origin, 24-Make
cars.train.new <- cars.train[-c(1,2,3,8,23,24)]
dmy <- dummyVars( ~. ,data=cars.train.new,fullRank =T)
cars.train.new <- data.frame(predict(dmy,newdata = cars.train.new))
dim(cars.train.new)

## [1] 1500 35

head(cars.train.new)

##   TypeLarge TypeMidsize TypeSmall TypeSporty TypeVan MPG.highway
## 1         0         0         0         1         0          25
```

## 2	0	1	0	0	0	27
## 3	0	1	0	0	0	22
## 4	0	0	0	0	0	34
## 5	0	0	0	1	0	25
## 6	0	1	0	0	0	26
##	AirBagsDriver.only	AirBagsNone	DriveTrainRear	EngineSize	Horsepower	RPM
## 1	0	1	0	2.8	178	5800
## 2	1	0	0	3.8	170	4800
## 3	1	0	1	4.5	278	6000
## 4	1	0	0	2.5	164	5600
## 5	1	0	1	5.7	300	5000
## 6	0	0	0	3.8	160	4400
##	Rev.per.mile	Man.trans.availYes	Fuel.tank.capacity	Passengers	Length	
## 1	2385	1	18.5	4	159	
## 2	1690	0	18.8	5	198	
## 3	1955	0	22.5	5	200	
## 4	2505	1	15.5	5	184	
## 5	1450	1	20.0	2	179	
## 6	1835	0	18.4	6	205	
##	Wheelbase	Width	Turn.circle	Rear.seat.room	Luggage.room	Weight
## 1	97	66	36	26.0	15	2810
## 2	108	73	41	26.5	14	3495
## 3	113	72	42	29.0	15	4000
## 4	103	69	40	29.5	14	2970
## 5	96	74	43	24.5	15	3380
## 6	109	73	42	30.0	19	3695
##	mod_categoryLow	mod_categoryLuxury	mod_categoryMedium			
## 1	0	0	1			
## 2	0	0	1			
## 3	0	1	0			
## 4	0	0	1			
## 5	0	0	0			
## 6	0	0	0			
##	Cylinders_Origin4.non.USA	Cylinders_Origin4.USA	Cylinders_Origin5.non.USA			
## 1	0	0	0			
## 2	0	0	0			
## 3	0	0	0			
## 4	1	0	0			
## 5	0	0	0			
## 6	0	0	0			
##	Cylinders_Origin6.non.USA	Cylinders_Origin6.USA	Cylinders_Origin8.non.USA			
## 1	1	0	0			
## 2	0	1	0			
## 3	0	0	1			
## 4	0	0	0			
## 5	0	0	0			
## 6	0	1	0			
##	Cylinders_Origin8.USA	Cylinders_Originrotary.non.USA				
## 1	0	0				
## 2	0	0				
## 3	0	0				
## 4	0	0				
## 5	1	0				
## 6	0	0				

Full model using all variables

We now create our first regression model using all predictors in our new training data frame.

```
# Full model
full_model <- lm(PriceNew~., data=cars.train.new)
summary(full_model)

##
## Call:
## lm(formula = PriceNew ~ ., data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -12344.0  -1389.8   -21.6   1366.5   9165.5
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.887e+04  4.312e+03  11.334 < 2e-16 ***
## TypeLarge       1.644e+03  4.797e+02   3.427 0.000627 ***
## TypeMidsize     1.570e+03  3.153e+02   4.978 7.18e-07 ***
## TypeSmall       2.175e+02  3.352e+02   0.649 0.516479
## TypeSporty      1.209e+03  3.661e+02   3.301 0.000987 ***
## TypeVan        -1.449e+03  6.737e+02  -2.150 0.031687 *
## MPG.highway     1.001e+01  3.181e+01   0.315 0.753063
## AirBagsDriver.only -3.141e+02  2.076e+02  -1.513 0.130577
## AirBagsNone     -2.115e+03  2.554e+02  -8.282 2.70e-16 ***
## DriveTrainRear   6.020e+02  2.631e+02   2.288 0.022263 *
## EngineSize     -2.377e+02  3.179e+02  -0.748 0.454770
## Horsepower      5.006e+01  5.736e+00   8.726 < 2e-16 ***
## RPM            -3.884e-01  2.752e-01  -1.411 0.158407
## Rev.per.mile    -3.409e-01  2.825e-01  -1.207 0.227648
## Man.trans.availYes -3.708e+02  2.840e+02  -1.305 0.191951
## Fuel.tank.capacity 3.306e+02  5.909e+01   5.594 2.64e-08 ***
## Passengers      6.040e+02  1.973e+02   3.062 0.002241 **
## Length         -1.924e+01  1.516e+01  -1.270 0.204464
## Wheelbase       3.055e+01  3.414e+01   0.895 0.371068
## Width          -4.159e+02  7.289e+01  -5.706 1.40e-08 ***
## Turn.circle     -6.188e+01  4.732e+01  -1.308 0.191111
## Rear.seat.room  -9.464e+01  4.408e+01  -2.147 0.031962 *
## Luggage.room    2.460e+02  3.632e+01   6.773 1.83e-11 ***
## Weight         4.519e-01  6.749e-01   0.670 0.503248
## mod_categoryLow -1.506e+04  3.849e+02 -39.126 < 2e-16 ***
## mod_categoryLuxury 1.391e+04  7.366e+02  18.878 < 2e-16 ***
## mod_categoryMedium -1.146e+04  2.729e+02 -41.985 < 2e-16 ***
## Cylinders_Origin4.non.USA -4.325e+02  5.138e+02  -0.842 0.400045
## Cylinders_Origin4.USA -1.438e+03  5.667e+02  -2.537 0.011282 *
## Cylinders_Origin5.non.USA 3.860e+03  6.970e+02   5.537 3.64e-08 ***
## Cylinders_Origin6.non.USA 1.640e+03  6.669e+02   2.459 0.014040 *
## Cylinders_Origin6.USA -6.550e+02  7.417e+02  -0.883 0.377283
## Cylinders_Origin8.non.USA -5.541e+03  1.111e+03  -4.985 6.93e-07 ***
## Cylinders_Origin8.USA -8.727e+02  9.284e+02  -0.940 0.347355
## Cylinders_Originrotary.non.USA -3.873e+03  1.409e+03  -2.749 0.006058 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 2263 on 1465 degrees of freedom
## Multiple R-squared:  0.9477, Adjusted R-squared:  0.9465
## F-statistic: 781.5 on 34 and 1465 DF,  p-value: < 2.2e-16
```

```
anova(full_model)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: PriceNew
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
TypeLarge	1	6.8849e+09	6.8849e+09	1344.7426	< 2.2e-16
TypeMidsize	1	4.5920e+10	4.5920e+10	8968.9572	< 2.2e-16
TypeSmall	1	1.4674e+10	1.4674e+10	2866.0212	< 2.2e-16
TypeSporty	1	1.8126e+04	1.8126e+04	0.0035	0.9525615
TypeVan	1	7.0571e+07	7.0571e+07	13.7838	0.0002128
MPG.highway	1	1.1482e+10	1.1482e+10	2242.6219	< 2.2e-16
AirBagsDriver.only	1	4.9270e+08	4.9270e+08	96.2333	< 2.2e-16
AirBagsNone	1	7.5473e+09	7.5473e+09	1474.1146	< 2.2e-16
DriveTrainRear	1	5.2173e+09	5.2173e+09	1019.0252	< 2.2e-16
EngineSize	1	3.0072e+09	3.0072e+09	587.3603	< 2.2e-16
Horsepower	1	1.3761e+10	1.3761e+10	2687.6843	< 2.2e-16
RPM	1	4.4791e+06	4.4791e+06	0.8749	0.3497701
Rev.per.mile	1	2.2191e+09	2.2191e+09	433.4385	< 2.2e-16
Man.trans.availYes	1	5.8485e+07	5.8485e+07	11.4231	0.0007444
Fuel.tank.capacity	1	1.3339e+05	1.3339e+05	0.0261	0.8717933
Passengers	1	1.0783e+08	1.0783e+08	21.0603	4.831e-06
Length	1	4.5748e+08	4.5748e+08	89.3531	< 2.2e-16
Wheelbase	1	6.3645e+08	6.3645e+08	124.3107	< 2.2e-16
Width	1	7.8934e+09	7.8934e+09	1541.7272	< 2.2e-16
Turn.circle	1	1.3403e+09	1.3403e+09	261.7869	< 2.2e-16
Rear.seat.room	1	6.2002e+08	6.2002e+08	121.1005	< 2.2e-16
Luggage.room	1	5.9044e+08	5.9044e+08	115.3240	< 2.2e-16
Weight	1	4.4225e+07	4.4225e+07	8.6380	0.0033437
mod_categoryLow	1	7.0566e+08	7.0566e+08	137.8281	< 2.2e-16
mod_categoryLuxury	1	1.8467e+09	1.8467e+09	360.6961	< 2.2e-16
mod_categoryMedium	1	9.5358e+09	9.5358e+09	1862.5194	< 2.2e-16
Cylinders_Origin4.non.USA	1	2.2081e+06	2.2081e+06	0.4313	0.5114636
Cylinders_Origin4.USA	1	9.4590e+07	9.4590e+07	18.4752	1.835e-05
Cylinders_Origin5.non.USA	1	1.6673e+08	1.6673e+08	32.5651	1.393e-08
Cylinders_Origin6.non.USA	1	4.6299e+08	4.6299e+08	90.4296	< 2.2e-16
Cylinders_Origin6.USA	1	3.9483e+06	3.9483e+06	0.7712	0.3799970
Cylinders_Origin8.non.USA	1	1.5628e+08	1.5628e+08	30.5236	3.898e-08
Cylinders_Origin8.USA	1	9.0647e+05	9.0647e+05	0.1770	0.6739827
Cylinders_Originrotary.non.USA	1	3.8680e+07	3.8680e+07	7.5549	0.0060582
Residuals	1465	7.5006e+09	5.1199e+06		

```
##
```

TypeLarge	***
TypeMidsize	***
TypeSmall	***
TypeSporty	
TypeVan	***
MPG.highway	***
AirBagsDriver.only	***
AirBagsNone	***

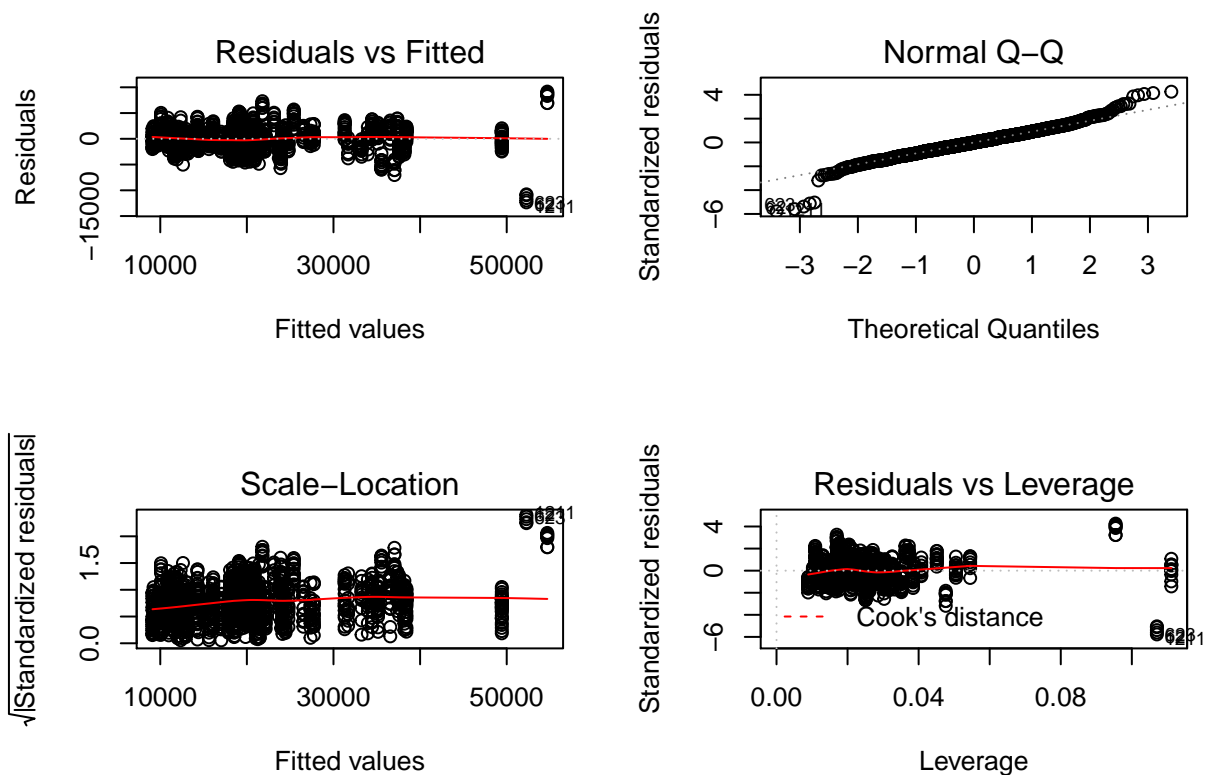
```
## DriveTrainRear          ***
## EngineSize              ***
## Horsepower              ***
## RPM                     ***
## Rev.per.mile            ***
## Man.trans.availYes      ***
## Fuel.tank.capacity      ***
## Passengers              ***
## Length                  ***
## Wheelbase               ***
## Width                   ***
## Turn.circle             ***
## Rear.seat.room          ***
## Luggage.room            ***
## Weight                  **
## mod_categoryLow         ***
## mod_categoryLuxury      ***
## mod_categoryMedium      ***
## Cylinders_Origin4.non.USA
## Cylinders_Origin4.USA    ***
## Cylinders_Origin5.non.USA ***
## Cylinders_Origin6.non.USA ***
## Cylinders_Origin6.USA
## Cylinders_Origin8.non.USA ***
## Cylinders_Origin8.USA
## Cylinders_Originrotary.non.USA **
## Residuals
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
vif(full_model)
```

```
##              TypeLarge              TypeMidsize
##              7.254404              5.209939
##              TypeSmall              TypeSporty
##              5.863621              5.188381
##              TypeVan                MPG.highway
##              11.109131              8.595856
##              AirBagsDriver.only      AirBagsNone
##              3.053550              4.518034
##              DriveTrainRear          EngineSize
##              2.975708              32.611714
##              Horsepower              RPM
##              25.277147              8.231508
##              Rev.per.mile            Man.trans.availYes
##              5.759132              5.267537
##              Fuel.tank.capacity      Passengers
##              10.886931              11.424500
##              Length                  Wheelbase
##              15.036337              16.251097
##              Width                   Turn.circle
##              21.878331              6.698631
##              Rear.seat.room          Luggage.room
##              5.478407              4.307976
##              Weight                  mod_categoryLow
```

```
##          47.756441          8.421639
##      mod_categoryLuxury      mod_categoryMedium
##          4.725783          5.408235
##      Cylinders_Origin4.non.USA      Cylinders_Origin4.USA
##          15.615684          16.528679
##      Cylinders_Origin5.non.USA      Cylinders_Origin6.non.USA
##          5.203763          13.227319
##      Cylinders_Origin6.USA      Cylinders_Origin8.non.USA
##          27.170698          8.017778
##      Cylinders_Origin8.USA Cylinders_Originrotary.non.USA
##          14.092834          3.469608
```

```
par(mfrow=c(2,2))
plot(full_model)
```

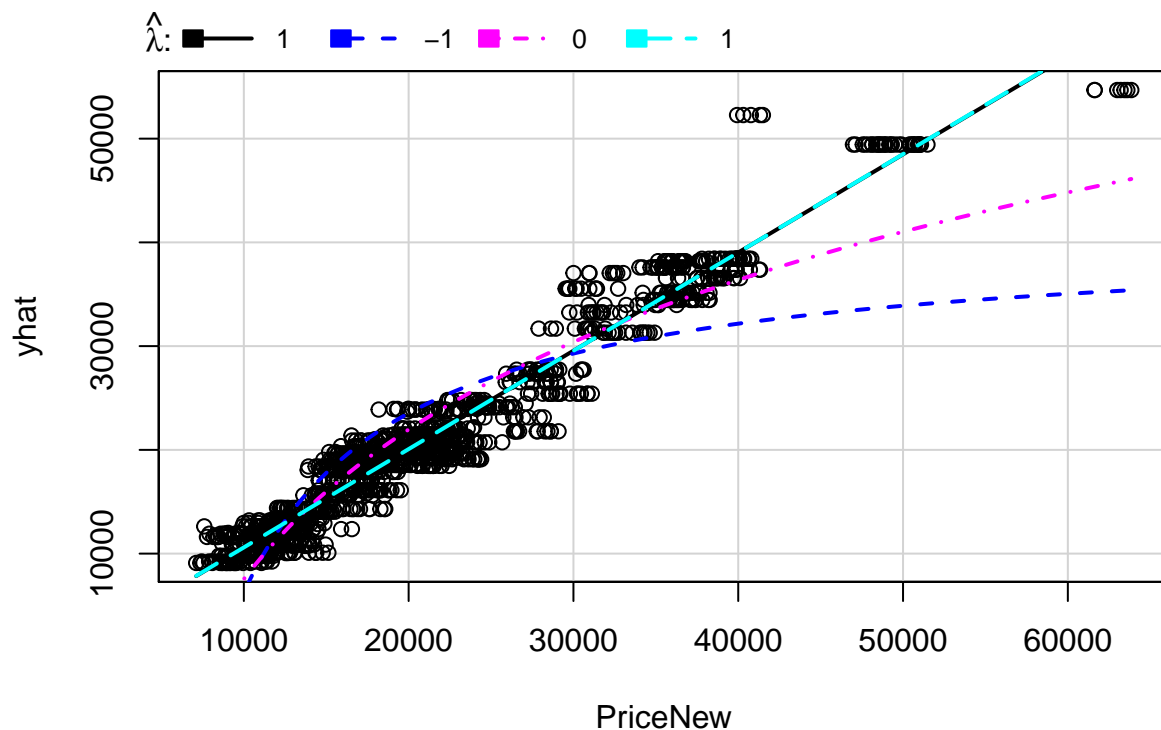


Initial Transformations

```
summary(powerTransform(cbind(PriceNew, MPG.highway, EngineSize, Horsepower, RPM, Rev.per.mile, Fuel.tan
```

```
## bcPower Transformations to Multinormality
##      Est Power Rounded Pwr Wald Lwr Bnd Wald Up Bnd
## PriceNew      -0.3431      -0.33      -0.4271      -0.2591
## MPG.highway    -1.1597     -1.00     -1.3287     -0.9907
## EngineSize     -0.1315     -0.13     -0.1912     -0.0718
## Horsepower      0.0660      0.07      0.0019      0.1301
## RPM             1.6476      1.65      1.3454      1.9498
## Rev.per.mile   -0.1007      0.00     -0.2522      0.0509
## Fuel.tank.capacity 0.4328      0.50      0.3024      0.5632
## Passengers      1.4747      1.47      1.3131      1.6363
## Length          1.1197      1.00      0.8134      1.4261
## Wheelbase     -1.4069     -1.00     -1.8757     -0.9381
```

```
## Width          -1.4400      -1.00      -1.9604      -0.9196
## Turn.circle    1.4497       1.00       0.9536       1.9457
## Rear.seat.room 0.6661       0.50       0.4047       0.9276
## Luggage.room   0.9331       1.00       0.8000       1.0663
## Weight         0.4946       0.50       0.3938       0.5954
##
## Likelihood ratio test that transformation parameters are equal to 0
## (all log transformations)
##                                     LRT df      pval
## LR test, lambda = (0 0 0 0 0 0 0 0 0 0 0 0 0 0 0) 1308.748 15 < 2.22e-16
##
## Likelihood ratio test that no transformations are needed
##                                     LRT df      pval
## LR test, lambda = (1 1 1 1 1 1 1 1 1 1 1 1 1 1 1) 3693.051 15 < 2.22e-16
inverseResponsePlot(full_model, lam=c(-1,0,1))
```



```
##      lambda      RSS
## 1  0.9965508  7108552318
## 2 -1.0000000  37290468091
## 3  0.0000000  15652685593
## 4  1.0000000  7108656717
```

Let's see the effect of the power transformation.

```
summary(lm(I(PriceNew^(-0.33)) ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty + TypeVan + I(MPG.highway^(-1)) +
##
## Call:
## lm(formula = I(PriceNew^(-0.33)) ~ TypeLarge + TypeMidsize +
##     TypeSmall + TypeSporty + TypeVan + I(MPG.highway^(-1)) +
##     AirBagsDriver.only + AirBagsNone + DriveTrainRear + I(EngineSize^(-0.13)) +
##     I(Horsepower^(0.07)) + I(RPM^(1.65)) + log(Rev.per.mile) +
```

```

##      Man.trans.availYes + log(Passengers) + I(Fuel.tank.capacity^(0.5)) +
##      Length + I(Wheelbase^(-1)) + I(Width^(-1)) + Turn.circle +
##      I(Rear.seat.room^(0.5)) + Luggage.room + I(Weight^(0.5)) +
##      mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.non.USA +
##      Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA, data = cars.train.new)
##
## Residuals:
##      Min      1Q    Median      3Q      Max
## -0.0042674 -0.0008684 -0.0000881  0.0008517  0.0065290
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.383e-01  1.443e-02   9.589 < 2e-16 ***
## TypeLarge      4.708e-04  2.876e-04   1.637 0.101805
## TypeMidsize    1.196e-04  1.934e-04   0.618 0.536482
## TypeSmall      2.005e-05  2.123e-04   0.094 0.924767
## TypeSporty     -6.556e-04  2.278e-04  -2.878 0.004063 **
## TypeVan        1.464e-03  4.118e-04   3.555 0.000391 ***
## I(MPG.highway^(-1)) 8.313e-03  2.049e-02   0.406 0.685015
## AirBagsDriver.only -5.111e-04  1.279e-04  -3.996 6.76e-05 ***
## AirBagsNone     8.766e-04  1.587e-04   5.523 3.93e-08 ***
## DriveTrainRear  1.233e-04  1.579e-04   0.781 0.435054
## I(EngineSize^(-0.13)) -2.242e-02  6.258e-03  -3.582 0.000352 ***
## I(Horsepower^(0.07)) -5.267e-02  6.444e-03  -8.173 6.44e-16 ***
## I(RPM^(1.65))    4.893e-10  4.627e-10   1.058 0.290407
## log(Rev.per.mile)  1.058e-03  4.495e-04   2.355 0.018670 *
## Man.trans.availYes  6.156e-04  1.706e-04   3.608 0.000319 ***
## log(Passengers)  -5.981e-04  4.932e-04  -1.213 0.225436
## I(Fuel.tank.capacity^(0.5)) -2.346e-03  2.976e-04  -7.882 6.22e-15 ***
## Length          -6.752e-06  9.552e-06  -0.707 0.479761
## I(Wheelbase^(-1))  8.326e-01  2.253e-01   3.696 0.000227 ***
## I(Width^(-1))     -8.430e-01  2.077e-01  -4.059 5.19e-05 ***
## Turn.circle      2.113e-05  2.919e-05   0.724 0.469200
## I(Rear.seat.room^(0.5)) 3.220e-05  2.873e-04   0.112 0.910774
## Luggage.room     -2.034e-05  2.281e-05  -0.892 0.372778
## I(Weight^(0.5))   -9.258e-05  4.333e-05  -2.137 0.032792 *
## mod_categoryLow   8.243e-03  2.354e-04  35.019 < 2e-16 ***
## mod_categoryLuxury -1.831e-03  4.484e-04  -4.082 4.70e-05 ***
## mod_categoryMedium 4.888e-03  1.641e-04  29.783 < 2e-16 ***
## Cylinders_Origin4.non.USA 3.656e-04  2.839e-04   1.288 0.198034
## Cylinders_Origin4.USA   7.598e-04  3.100e-04   2.451 0.014378 *
## Cylinders_Origin5.non.USA -2.363e-03  4.223e-04  -5.596 2.62e-08 ***
## Cylinders_Origin6.non.USA -5.402e-04  4.054e-04  -1.333 0.182861
## Cylinders_Origin6.USA   2.687e-04  4.510e-04   0.596 0.551405
## Cylinders_Origin8.non.USA 2.180e-03  6.888e-04   3.164 0.001585 **
## Cylinders_Origin8.USA   6.680e-04  5.622e-04   1.188 0.234950
## Cylinders_Originrotary.non.USA 3.742e-03  9.302e-04   4.023 6.04e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.001395 on 1465 degrees of freedom
## Multiple R-squared:  0.9359, Adjusted R-squared:  0.9344

```


F-statistic: 628.7 on 34 and 1465 DF, p-value: < 2.2e-16

```
summary(lm(log(PriceNew) ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty + TypeVan + log(MPG.highway)
```

##

Call:

```
## lm(formula = log(PriceNew) ~ TypeLarge + TypeMidsize + TypeSmall +  
##     TypeSporty + TypeVan + log(MPG.highway) + AirBagsDriver.only +  
##     AirBagsNone + DriveTrainRear + log(EngineSize) + log(Horsepower) +  
##     log(RPM) + log(Rev.per.mile) + Man.trans.availYes + log(Passengers) +  
##     log(Fuel.tank.capacity) + log(Length) + log(Wheelbase) +  
##     log(Width) + log(Turn.circle) + log(Rear.seat.room) + log(Luggage.room) +  
##     log(Weight) + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +  
##     Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +  
##     Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.non.USA +  
##     Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA, data = cars.train.new)
```

##

Residuals:

```
##      Min       1Q   Median       3Q      Max  
## -0.40840 -0.06605  0.00683  0.06702  0.30230
```

##

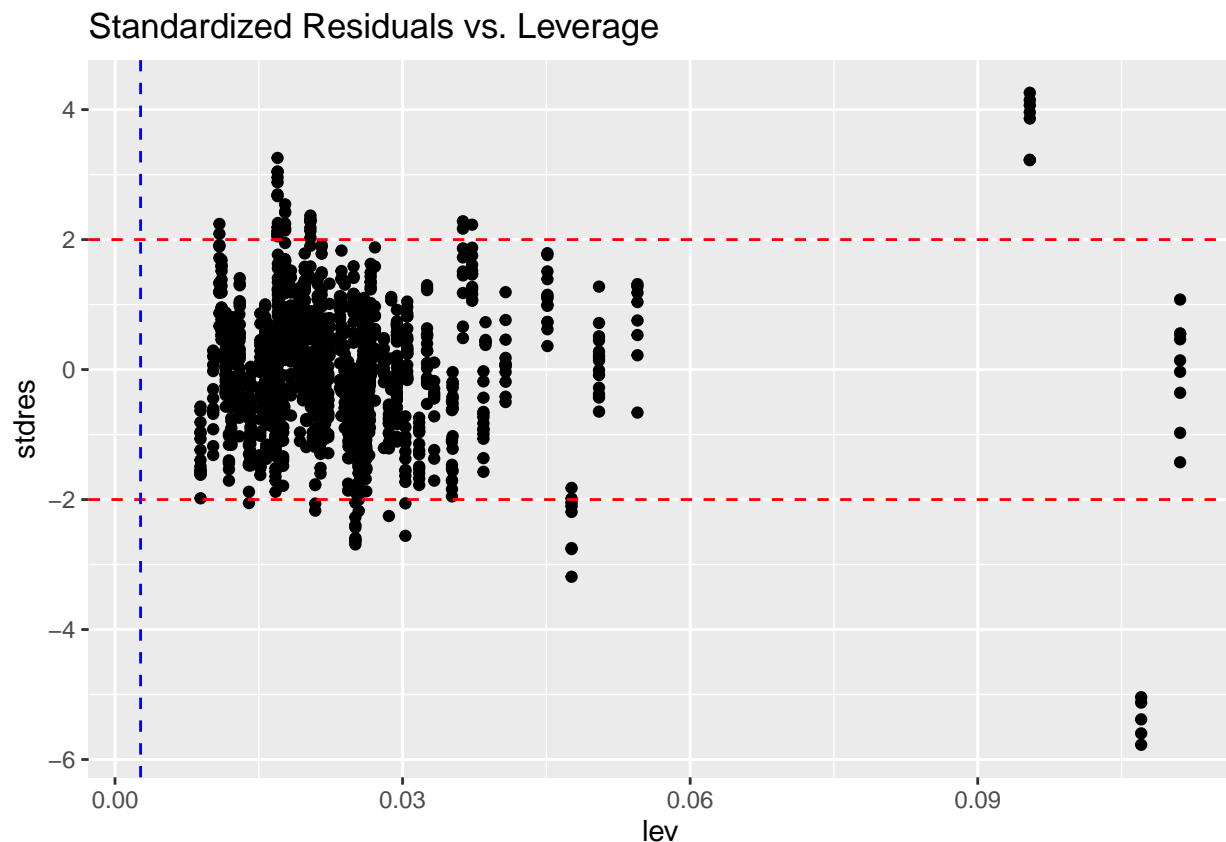
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	10.626478	1.132607	9.382	< 2e-16 ***
## TypeLarge	0.016835	0.022898	0.735	0.462325
## TypeMidsize	0.017512	0.014800	1.183	0.236879
## TypeSmall	0.007833	0.015665	0.500	0.617156
## TypeSporty	0.033084	0.017262	1.917	0.055487 .
## TypeVan	-0.071774	0.030921	-2.321	0.020412 *
## log(MPG.highway)	0.038967	0.050851	0.766	0.443631
## AirBagsDriver.only	0.025567	0.009834	2.600	0.009420 **
## AirBagsNone	-0.082342	0.012247	-6.724	2.53e-11 ***
## DriveTrainRear	-0.000193	0.011853	-0.016	0.987011
## log(EngineSize)	-0.227491	0.054750	-4.155	3.44e-05 ***
## log(Horsepower)	0.400807	0.046296	8.657	< 2e-16 ***
## log(RPM)	-0.143768	0.076325	-1.884	0.059812 .
## log(Rev.per.mile)	-0.076605	0.033744	-2.270	0.023341 *
## Man.trans.availYes	-0.046064	0.012760	-3.610	0.000316 ***
## log(Passengers)	0.017552	0.037921	0.463	0.643528
## log(Fuel.tank.capacity)	0.414112	0.045032	9.196	< 2e-16 ***
## log(Length)	-0.029285	0.125481	-0.233	0.815498
## log(Wheelbase)	0.285639	0.164922	1.732	0.083490 .
## log(Width)	-0.975264	0.226671	-4.303	1.80e-05 ***
## log(Turn.circle)	-0.114990	0.084931	-1.354	0.175970
## log(Rear.seat.room)	-0.005546	0.057333	-0.097	0.922948
## log(Luggage.room)	0.019268	0.022334	0.863	0.388413
## log(Weight)	0.229690	0.089046	2.579	0.009992 **
## mod_categoryLow	-0.671928	0.017680	-38.005	< 2e-16 ***
## mod_categoryLuxury	0.243497	0.033735	7.218	8.44e-13 ***
## mod_categoryMedium	-0.430162	0.012357	-34.810	< 2e-16 ***
## Cylinders_Origin4.non.USA	-0.057261	0.021796	-2.627	0.008700 **
## Cylinders_Origin4.USA	-0.080894	0.023650	-3.421	0.000642 ***
## Cylinders_Origin5.non.USA	0.154076	0.032078	4.803	1.72e-06 ***
## Cylinders_Origin6.non.USA	0.028556	0.030343	0.941	0.346805
## Cylinders_Origin6.USA	-0.032051	0.033401	-0.960	0.337419

```
## Cylinders_Origin8.non.USA      -0.158117    0.051795   -3.053 0.002309 **
## Cylinders_Origin8.USA         -0.041043    0.041676   -0.985 0.324887
## Cylinders_Originrotary.non.USA -0.360224    0.072666   -4.957 7.98e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.105 on 1465 degrees of freedom
## Multiple R-squared:  0.9419, Adjusted R-squared:  0.9406
## F-statistic: 698.8 on 34 and 1465 DF,  p-value: < 2.2e-16
```

It appears that the untransformed full model is better based on the r-squared and adjusted r-squared score. Let's check the full model's outliers and leverage points. # Outliers and leverage values.

```
# Plot of leverage vs standardized residuals
cars.train.new$lev <- hatvalues(full_model)
cars.train.new$stdres <- rstandard(full_model)
rvl_plot <- qplot(lev, stdres, data=cars.train.new) + geom_hline(yintercept=c(-2,2), linetype="dashed",
# Add boundary lines
rvl_plot <- rvl_plot + geom_vline(xintercept=4/dim(cars.train.new)[1], linetype="dashed", col="blue")
rvl_plot <- rvl_plot + labs(title="Standardized Residuals vs. Leverage")
rvl_plot
```



```
# Bad Leverage points
c(which((cars.train.new$lev > 4/dim(cars.train.new)[1]) & (abs(cars.train.new$stdres) > 2)))

## [1] 2 37 62 73 95 141 173 176 203 204 230 231 270 272 309
## [16] 333 343 385 390 409 445 465 468 592 623 653 688 731 736 739
## [31] 776 785 795 836 845 850 860 861 903 904 954 959 972 1030 1064
## [46] 1116 1191 1211 1224 1254 1263 1266 1326 1333 1354 1360 1369 1373 1381 1388
```

```
## [61] 1395 1404 1416 1430 1432 1447 1462
```

Let's try deleting the bad leverage values.

```
# Delete bad leverage points
```

```
cars.train.new <- cars.train.new[-c(which((cars.train.new$lev > 4/dim(cars.train.new)[1]) & (abs(cars.train.new$lev - 4/dim(cars.train.new)[1]) > 2)))
```

```
## [1] 1433 37
```

Finally, let's update and revisit our full model.

```
full_model <- update(full_model, ~.-lev-stdres)
summary(full_model)
```

```
##
```

```
## Call:
```

```
## lm(formula = PriceNew ~ TypeLarge + TypeMidsize + TypeSmall +
##     TypeSporty + TypeVan + MPG.highway + AirBagsDriver.only +
##     AirBagsNone + DriveTrainRear + EngineSize + Horsepower +
##     RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##     Passengers + Length + Wheelbase + Width + Turn.circle + Rear.seat.room +
##     Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##     mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
##     Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##     Cylinders_Origin8.non.USA + Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA,
##     data = cars.train.new)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max
## -6079.5 -1190.8    -2.5   1165.7   5692.9
```

```
##
```

```
## Coefficients: (1 not defined because of singularities)
```

	Estimate	Std. Error	t value	Pr(> t)
## (Intercept)	3.564e+04	3.424e+03	10.409	< 2e-16 ***
## TypeLarge	1.544e+03	3.993e+02	3.868	0.000115 ***
## TypeMidsize	1.509e+03	2.686e+02	5.618	2.32e-08 ***
## TypeSmall	-1.649e+02	2.644e+02	-0.624	0.532977
## TypeSporty	1.144e+03	2.917e+02	3.922	9.20e-05 ***
## TypeVan	-2.191e+03	5.461e+02	-4.012	6.34e-05 ***
## MPG.highway	1.421e+02	2.580e+01	5.508	4.32e-08 ***
## AirBagsDriver.only	-2.530e+02	1.666e+02	-1.519	0.129074
## AirBagsNone	-1.604e+03	2.025e+02	-7.923	4.71e-15 ***
## DriveTrainRear	-1.787e+02	2.239e+02	-0.798	0.424921
## EngineSize	6.753e+02	2.549e+02	2.649	0.008156 **
## Horsepower	2.776e+01	4.690e+00	5.920	4.05e-09 ***
## RPM	8.202e-01	2.218e-01	3.698	0.000226 ***
## Rev.per.mile	3.351e-01	2.232e-01	1.502	0.133445
## Man.trans.availYes	4.255e+02	2.282e+02	1.864	0.062516 .
## Fuel.tank.capacity	2.996e+02	4.677e+01	6.404	2.06e-10 ***
## Passengers	6.732e+02	1.556e+02	4.327	1.62e-05 ***
## Length	-6.298e+01	1.243e+01	-5.065	4.62e-07 ***
## Wheelbase	-4.136e+01	2.716e+01	-1.523	0.127922
## Width	-3.813e+02	5.827e+01	-6.543	8.41e-11 ***
## Turn.circle	2.824e+01	3.760e+01	0.751	0.452712
## Rear.seat.room	-8.483e+01	3.499e+01	-2.424	0.015465 *
## Luggage.room	1.802e+02	2.951e+01	6.106	1.32e-09 ***

```
## Weight          4.197e+00  5.598e-01   7.497 1.16e-13 ***
## mod_categoryLow -1.592e+04  3.087e+02 -51.572 < 2e-16 ***
## mod_categoryLuxury 9.664e+03  7.562e+02  12.780 < 2e-16 ***
## mod_categoryMedium -1.247e+04  2.242e+02 -55.616 < 2e-16 ***
## Cylinders_Origin4.non.USA 7.606e+02  4.067e+02   1.870 0.061689 .
## Cylinders_Origin4.USA 3.556e+02  4.506e+02   0.789 0.430178
## Cylinders_Origin5.non.USA 4.411e+03  5.512e+02   8.003 2.54e-15 ***
## Cylinders_Origin6.non.USA 2.475e+03  5.330e+02   4.644 3.73e-06 ***
## Cylinders_Origin6.USA 9.519e+02  5.855e+02   1.626 0.104229
## Cylinders_Origin8.non.USA NA NA NA NA
## Cylinders_Origin8.USA 2.054e+03  7.436e+02   2.763 0.005810 **
## Cylinders_Originrotary.non.USA -1.187e+02  1.116e+03 -0.106 0.915308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1759 on 1399 degrees of freedom
## Multiple R-squared:  0.9658, Adjusted R-squared:  0.965
## F-statistic: 1196 on 33 and 1399 DF, p-value: < 2.2e-16
```

Variable selection

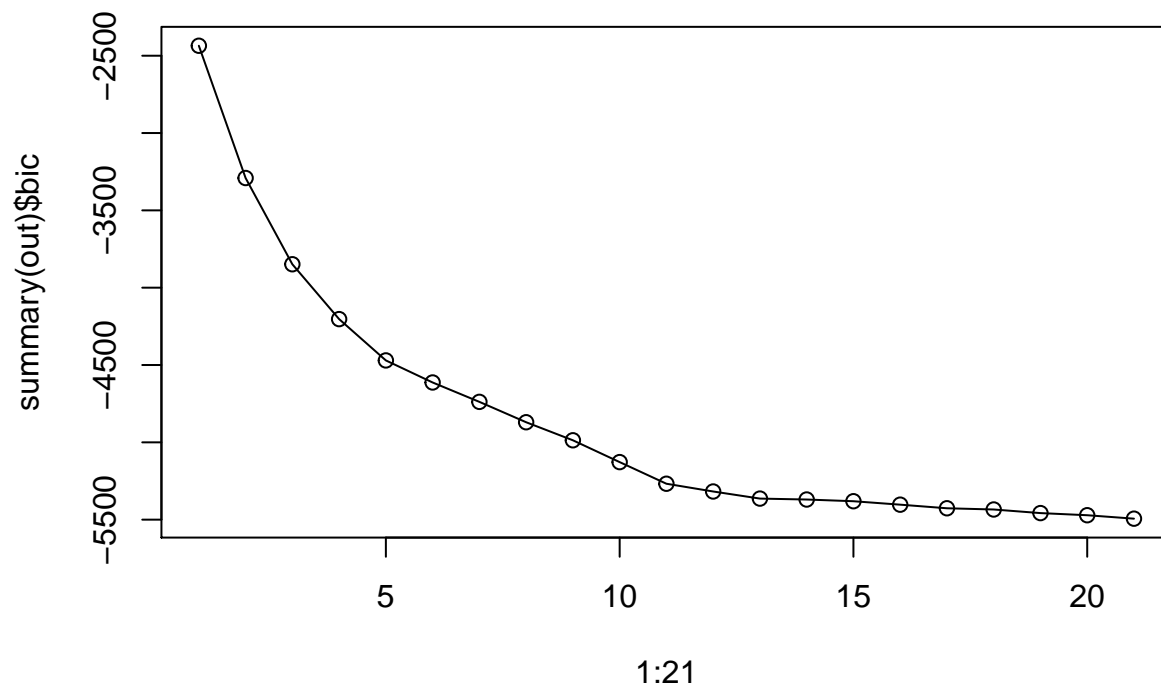
Now we use the `regsubsets` function to determine the optimal number of variables in our regression model.

```
out <- regsubsets(as.matrix(cars.train.new[, -23]), cars.train.new[, 23], nvmax = 20)
```

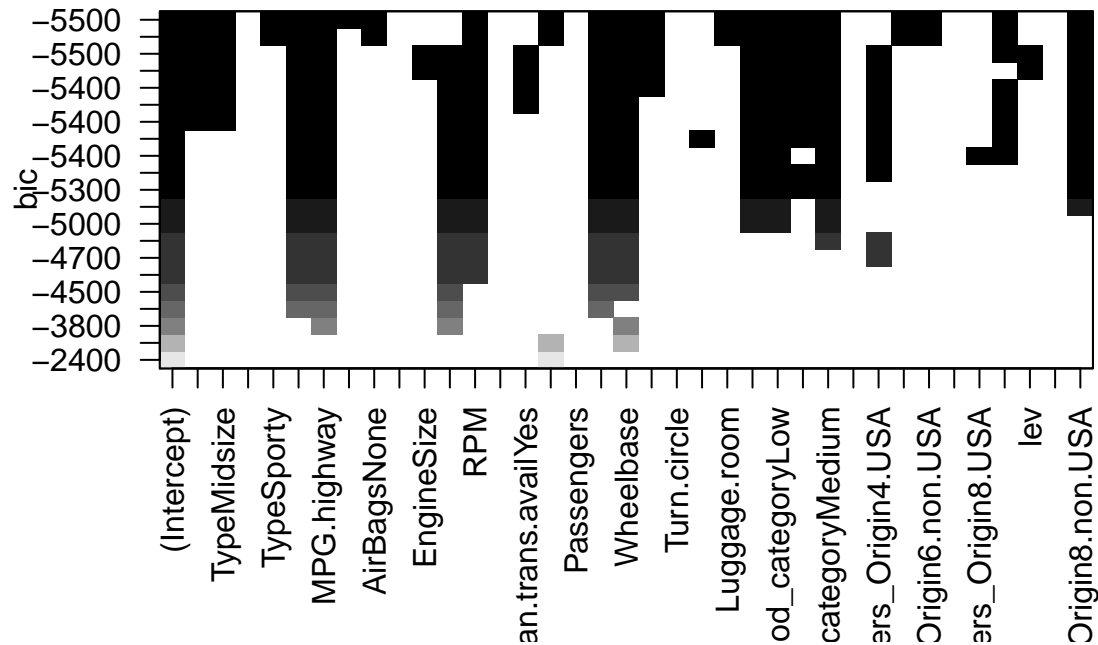
```
## Warning in leaps.setup(x, y, wt = weights, nbest = nbest, nvmax = nvmax, : 1
## linear dependencies found
```

```
## Reordering variables and trying again:
```

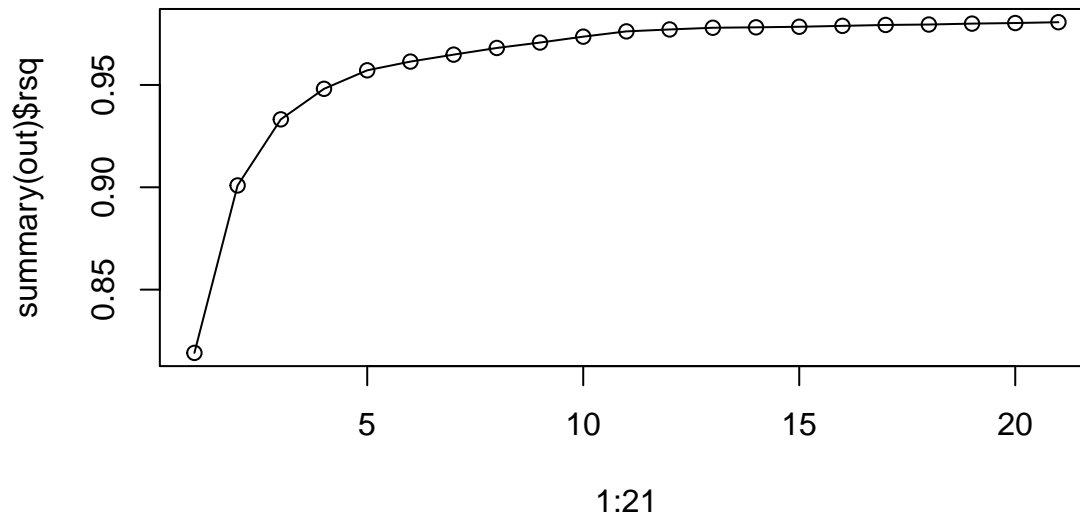
```
plot(1:21, summary(out)$bic)
lines(1:21, summary(out)$bic)
```



```
plot(out, scale="bic")
```



```
plot(1:21,summary(out)$rsq)
lines(1:21,summary(out)$rsq)
```



We also use the step functions for similar analysis.

```
backAIC <- step(full_model, direction='backward')
```

```
## Start: AIC=21450.38
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
## TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
## DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
## Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
## Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
## Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
## Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
## Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.non.USA +
```

```

##      Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA
##
##
## Step:  AIC=21450.38
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
##      TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
##      DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
##      Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
##      Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA +
##      Cylinders_Originrotary.non.USA
##
##
##      Df Sum of Sq      RSS   AIC
## - Cylinders_Originrotary.non.USA  1      35019 4.3305e+09 21448
## - TypeSmall                      1      1203807 4.3316e+09 21449
## - Turn.circle                    1      1746317 4.3322e+09 21449
## - Cylinders_Origin4.USA          1      1927509 4.3324e+09 21449
## - DriveTrainRear                 1      1971877 4.3324e+09 21449
## <none>                           1      4.3304e+09 21450
## - Rev.per.mile                   1      6978784 4.3374e+09 21451
## - AirBagsDriver.only             1      7138965 4.3376e+09 21451
## - Wheelbase                     1      7182204 4.3376e+09 21451
## - Cylinders_Origin6.USA          1      8181233 4.3386e+09 21451
## - Man.trans.availYes             1     10756086 4.3412e+09 21452
## - Cylinders_Origin4.non.USA      1     10824518 4.3413e+09 21452
## - Rear.seat.room                 1     18191812 4.3486e+09 21454
## - EngineSize                     1     21726557 4.3522e+09 21456
## - Cylinders_Origin8.USA          1     23623380 4.3541e+09 21456
## - RPM                            1     42330342 4.3728e+09 21462
## - TypeLarge                      1     46302005 4.3767e+09 21464
## - TypeSporty                     1     47620734 4.3781e+09 21464
## - TypeVan                        1     49820314 4.3803e+09 21465
## - Passengers                     1     57967050 4.3884e+09 21467
## - Cylinders_Origin6.non.USA      1     66767873 4.3972e+09 21470
## - Length                         1     79417574 4.4099e+09 21474
## - MPG.highway                    1     93892818 4.4243e+09 21479
## - TypeMidsize                    1     97707405 4.4281e+09 21480
## - Horsepower                     1    108470926 4.4389e+09 21484
## - Luggage.room                   1    115397552 4.4458e+09 21486
## - Fuel.tank.capacity             1    126958775 4.4574e+09 21490
## - Width                          1    132534796 4.4630e+09 21492
## - Weight                         1    173963678 4.5044e+09 21505
## - AirBagsNone                    1    194290244 4.5247e+09 21511
## - Cylinders_Origin5.non.USA      1    198229883 4.5287e+09 21512
## - mod_categoryLuxury             1    505578655 4.8360e+09 21607
## - mod_categoryLow                1    8232819583 1.2563e+10 22975
## - mod_categoryMedium             1    9574469828 1.3905e+10 23120
##
## Step:  AIC=21448.39
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
##      TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
##      DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +

```

```

##      Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
##      Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##              Df    Sum of Sq      RSS    AIC
## - TypeSmall          1      1329014 4.3318e+09 21447
## - Turn.circle         1      1775527 4.3322e+09 21447
## - DriveTrainRear       1      2020097 4.3325e+09 21447
## - Cylinders_Origin4.USA 1      3086389 4.3336e+09 21447
## <none>                                4.3305e+09 21448
## - AirBagsDriver.only   1       7346053 4.3378e+09 21449
## - Wheelbase            1       7475456 4.3379e+09 21449
## - Rev.per.mile          1       8398802 4.3389e+09 21449
## - Cylinders_Origin6.USA 1     11622982 4.3421e+09 21450
## - Man.trans.availYes    1     12959135 4.3434e+09 21451
## - Cylinders_Origin4.non.USA 1    16634240 4.3471e+09 21452
## - Rear.seat.room        1     18635822 4.3491e+09 21452
## - Cylinders_Origin8.USA 1     32222385 4.3627e+09 21457
## - EngineSize            1     36192277 4.3667e+09 21458
## - TypeLarge            1     46295142 4.3768e+09 21462
## - TypeSporty           1     47660370 4.3781e+09 21462
## - TypeVan              1     49942423 4.3804e+09 21463
## - RPM                  1     50868301 4.3813e+09 21463
## - Passengers            1     63724982 4.3942e+09 21467
## - Length                1     81998284 4.4125e+09 21473
## - Cylinders_Origin6.non.USA 1    93602154 4.4241e+09 21477
## - TypeMidsize           1    102065355 4.4325e+09 21480
## - MPG.highway           1    115820567 4.4463e+09 21484
## - Luggage.room          1    116538782 4.4470e+09 21484
## - Horsepower            1    134446666 4.4649e+09 21490
## - Width                 1    138331303 4.4688e+09 21492
## - Fuel.tank.capacity     1    138705571 4.4692e+09 21492
## - AirBagsNone           1    194262805 4.5247e+09 21509
## - Weight                1    203112826 4.5336e+09 21512
## - Cylinders_Origin5.non.USA 1   231368909 4.5618e+09 21521
## - mod_categoryLuxury     1    658718950 4.9892e+09 21649
## - mod_categoryLow        1   8307799647 1.2638e+10 22981
## - mod_categoryMedium     1   9904841970 1.4235e+10 23152
##
## Step:  AIC=21446.83
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##      AirBagsDriver.only + AirBagsNone + DriveTrainRear + EngineSize +
##      Horsepower + RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##      Passengers + Length + Wheelbase + Width + Turn.circle + Rear.seat.room +
##      Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##      Cylinders_Origin8.USA
##
##              Df    Sum of Sq      RSS    AIC
## - Turn.circle          1      1987069 4.3338e+09 21446
## - DriveTrainRear        1      2285517 4.3341e+09 21446

```

```

## - Cylinders_Origin4.USA      1      3054518 4.3349e+09 21446
## <none>                        4.3318e+09 21447
## - Wheelbase                  1      6466535 4.3383e+09 21447
## - AirBagsDriver.only         1      8029881 4.3398e+09 21448
## - Rev.per.mile                1      8269018 4.3401e+09 21448
## - Cylinders_Origin6.USA      1     10998201 4.3428e+09 21448
## - Man.trans.availYes         1     12216158 4.3440e+09 21449
## - Cylinders_Origin4.non.USA  1     16014051 4.3478e+09 21450
## - Rear.seat.room             1     19254324 4.3511e+09 21451
## - Cylinders_Origin8.USA      1     31048855 4.3629e+09 21455
## - EngineSize                 1     35660513 4.3675e+09 21457
## - TypeLarge                  1     46563818 4.3784e+09 21460
## - TypeVan                    1     49139453 4.3809e+09 21461
## - RPM                        1     50455608 4.3823e+09 21461
## - TypeSporty                 1     61729362 4.3935e+09 21465
## - Passengers                 1     63628997 4.3954e+09 21466
## - Length                     1     81222062 4.4130e+09 21472
## - Cylinders_Origin6.non.USA  1     92273260 4.4241e+09 21475
## - TypeMidsize                1    104541425 4.4363e+09 21479
## - MPG.highway                1    115194318 4.4470e+09 21482
## - Luggage.room               1    115567785 4.4474e+09 21483
## - Horsepower                 1    147195566 4.4790e+09 21493
## - Fuel.tank.capacity         1    159476950 4.4913e+09 21497
## - Width                      1    174425491 4.5062e+09 21501
## - Weight                     1    202004716 4.5338e+09 21510
## - AirBagsNone                1    202416110 4.5342e+09 21510
## - Cylinders_Origin5.non.USA  1    230110633 4.5619e+09 21519
## - mod_categoryLuxury         1    675332189 5.0071e+09 21652
## - mod_categoryLow            1   8659051535 1.2991e+10 23019
## - mod_categoryMedium         1   9970500109 1.4302e+10 23156
##
## Step:  AIC=21445.49
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsDriver.only + AirBagsNone + DriveTrainRear + EngineSize +
##   Horsepower + RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##   Passengers + Length + Wheelbase + Width + Rear.seat.room +
##   Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
##   Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##   Cylinders_Origin8.USA
##
##           Df  Sum of Sq      RSS    AIC
## - DriveTrainRear      1 2.8204e+06 4.3366e+09 21444
## - Cylinders_Origin4.USA 1 3.5603e+06 4.3373e+09 21445
## <none>                  4.3338e+09 21446
## - Wheelbase           1 7.1894e+06 4.3410e+09 21446
## - Rev.per.mile         1 7.5900e+06 4.3414e+09 21446
## - AirBagsDriver.only   1 9.0802e+06 4.3429e+09 21446
## - Cylinders_Origin6.USA 1 1.0514e+07 4.3443e+09 21447
## - Man.trans.availYes   1 1.1590e+07 4.3454e+09 21447
## - Cylinders_Origin4.non.USA 1 1.5626e+07 4.3494e+09 21449
## - Rear.seat.room       1 1.7440e+07 4.3512e+09 21449
## - Cylinders_Origin8.USA 1 3.0649e+07 4.3644e+09 21454
## - EngineSize           1 3.9191e+07 4.3730e+09 21456

```



```

## - TypeLarge          1 4.6544e+07 4.3803e+09 21459
## - TypeVan            1 4.7924e+07 4.3817e+09 21459
## - RPM                1 4.8789e+07 4.3826e+09 21460
## - Passengers         1 6.1659e+07 4.3954e+09 21464
## - TypeSporty         1 6.2999e+07 4.3968e+09 21464
## - Length             1 8.1090e+07 4.4149e+09 21470
## - Cylinders_Origin6.non.USA 1 9.0290e+07 4.4241e+09 21473
## - TypeMidsize        1 1.0957e+08 4.4434e+09 21479
## - Luggage.room       1 1.1360e+08 4.4474e+09 21481
## - MPG.highway        1 1.1627e+08 4.4501e+09 21481
## - Horsepower         1 1.5453e+08 4.4883e+09 21494
## - Fuel.tank.capacity 1 1.5824e+08 4.4920e+09 21495
## - Width              1 1.7327e+08 4.5071e+09 21500
## - Weight             1 2.0218e+08 4.5360e+09 21509
## - Cylinders_Origin5.non.USA 1 2.3344e+08 4.5672e+09 21519
## - AirBagsNone        1 2.3664e+08 4.5704e+09 21520
## - mod_categoryLuxury 1 6.7376e+08 5.0076e+09 21651
## - mod_categoryLow     1 8.7985e+09 1.3132e+10 23032
## - mod_categoryMedium 1 1.0840e+10 1.5174e+10 23239
##
## Step: AIC=21444.42
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsDriver.only + AirBagsNone + EngineSize + Horsepower +
##   RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##   Passengers + Length + Wheelbase + Width + Rear.seat.room +
##   Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
##   Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##   Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS   AIC
## <none>                4.3366e+09 21444
## - Cylinders_Origin4.USA 1 6.2652e+06 4.3429e+09 21444
## - Wheelbase            1 7.2221e+06 4.3438e+09 21445
## - Rev.per.mile         1 8.8383e+06 4.3454e+09 21445
## - AirBagsDriver.only   1 9.0210e+06 4.3456e+09 21445
## - Man.trans.availYes   1 1.1040e+07 4.3476e+09 21446
## - Cylinders_Origin6.USA 1 1.4915e+07 4.3515e+09 21447
## - Rear.seat.room       1 1.5227e+07 4.3518e+09 21447
## - Cylinders_Origin4.non.USA 1 2.1536e+07 4.3581e+09 21450
## - Cylinders_Origin8.USA 1 3.5451e+07 4.3721e+09 21454
## - EngineSize           1 3.7434e+07 4.3740e+09 21455
## - TypeVan              1 4.6739e+07 4.3833e+09 21458
## - TypeLarge            1 4.9426e+07 4.3860e+09 21459
## - RPM                  1 5.0721e+07 4.3873e+09 21459
## - TypeSporty           1 7.7164e+07 4.4138e+09 21468
## - Length               1 8.6180e+07 4.4228e+09 21471
## - Passengers           1 8.6940e+07 4.4235e+09 21471
## - Cylinders_Origin6.non.USA 1 1.0530e+08 4.4419e+09 21477
## - TypeMidsize          1 1.2301e+08 4.4596e+09 21482
## - Luggage.room         1 1.2346e+08 4.4601e+09 21483
## - MPG.highway          1 1.2900e+08 4.4656e+09 21484
## - Fuel.tank.capacity    1 1.5632e+08 4.4929e+09 21493
## - Horsepower           1 1.7678e+08 4.5134e+09 21500

```

```
## - Weight 1 2.0060e+08 4.5372e+09 21507
## - Width 1 2.1119e+08 4.5478e+09 21511
## - AirBagsNone 1 2.3641e+08 4.5730e+09 21518
## - Cylinders_Origin5.non.USA 1 2.5266e+08 4.5893e+09 21524
## - mod_categoryLuxury 1 6.7096e+08 5.0076e+09 21649
## - mod_categoryLow 1 8.9354e+09 1.3272e+10 23045
## - mod_categoryMedium 1 1.1250e+10 1.5587e+10 23276

forwardAIC <- step(full_model, direction='forward')

## Start: AIC=21450.38
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
## TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
## DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
## Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
## Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
## Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
## Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
## Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.non.USA +
## Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA

backBIC <- step(full_model, direction='backward', k=length(cars.train.new$PriceNew))

## Start: AIC=70104.38
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
## TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
## DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
## Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
## Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
## Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
## Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
## Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.non.USA +
## Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA
##
##
## Step: AIC=70104.38
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
## TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
## DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
## Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
## Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
## Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
## Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
## Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA +
## Cylinders_Originrotary.non.USA
##
##
## Df Sum of Sq RSS AIC
## - Cylinders_Originrotary.non.USA 1 35019 4.3305e+09 68671
## - TypeSmall 1 1203807 4.3316e+09 68672
## - Turn.circle 1 1746317 4.3322e+09 68672
## - Cylinders_Origin4.USA 1 1927509 4.3324e+09 68672
## - DriveTrainRear 1 1971877 4.3324e+09 68672
## - Rev.per.mile 1 6978784 4.3374e+09 68674
## - AirBagsDriver.only 1 7138965 4.3376e+09 68674
## - Wheelbase 1 7182204 4.3376e+09 68674
```

```

## - Cylinders_Origin6.USA      1      8181233 4.3386e+09 68674
## - Man.trans.availYes        1      10756086 4.3412e+09 68675
## - Cylinders_Origin4.non.USA  1      10824518 4.3413e+09 68675
## - Rear.seat.room            1      18191812 4.3486e+09 68677
## - EngineSize                 1      21726557 4.3522e+09 68679
## - Cylinders_Origin8.USA     1      23623380 4.3541e+09 68679
## - RPM                        1      42330342 4.3728e+09 68685
## - TypeLarge                  1      46302005 4.3767e+09 68687
## - TypeSporty                 1      47620734 4.3781e+09 68687
## - TypeVan                    1      49820314 4.3803e+09 68688
## - Passengers                 1      57967050 4.3884e+09 68690
## - Cylinders_Origin6.non.USA  1      66767873 4.3972e+09 68693
## - Length                     1      79417574 4.4099e+09 68697
## - MPG.highway                1      93892818 4.4243e+09 68702
## - TypeMidsize                1      97707405 4.4281e+09 68703
## - Horsepower                 1     108470926 4.4389e+09 68707
## - Luggage.room               1     115397552 4.4458e+09 68709
## - Fuel.tank.capacity         1     126958775 4.4574e+09 68713
## - Width                      1     132534796 4.4630e+09 68715
## - Weight                     1     173963678 4.5044e+09 68728
## - AirBagsNone                1     194290244 4.5247e+09 68734
## - Cylinders_Origin5.non.USA  1     198229883 4.5287e+09 68736
## - mod_categoryLuxury         1     505578655 4.8360e+09 68830
## <none>                       1              4.3304e+09 70104
## - mod_categoryLow            1     8232819583 1.2563e+10 70198
## - mod_categoryMedium         1     9574469828 1.3905e+10 70343
##
## Step: AIC=68671.39
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
##   TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
##   DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
##   Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
##   Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
##   Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##   Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
##   Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##              Df  Sum of Sq      RSS    AIC
## - TypeSmall      1    1329014 4.3318e+09 67239
## - Turn.circle     1    1775527 4.3322e+09 67239
## - DriveTrainRear  1    2020097 4.3325e+09 67239
## - Cylinders_Origin4.USA  1    3086389 4.3336e+09 67239
## - AirBagsDriver.only  1    7346053 4.3378e+09 67241
## - Wheelbase       1    7475456 4.3379e+09 67241
## - Rev.per.mile     1    8398802 4.3389e+09 67241
## - Cylinders_Origin6.USA  1   11622982 4.3421e+09 67242
## - Man.trans.availYes  1   12959135 4.3434e+09 67243
## - Cylinders_Origin4.non.USA  1   16634240 4.3471e+09 67244
## - Rear.seat.room   1   18635822 4.3491e+09 67245
## - Cylinders_Origin8.USA  1   32222385 4.3627e+09 67249
## - EngineSize       1   36192277 4.3667e+09 67250
## - TypeLarge       1   46295142 4.3768e+09 67254
## - TypeSporty       1   47660370 4.3781e+09 67254
## - TypeVan          1   49942423 4.3804e+09 67255

```

```

## - RPM 1 50868301 4.3813e+09 67255
## - Passengers 1 63724982 4.3942e+09 67259
## - Length 1 81998284 4.4125e+09 67265
## - Cylinders_Origin6.non.USA 1 93602154 4.4241e+09 67269
## - TypeMidsize 1 102065355 4.4325e+09 67272
## - MPG.highway 1 115820567 4.4463e+09 67276
## - Luggage.room 1 116538782 4.4470e+09 67276
## - Horsepower 1 134446666 4.4649e+09 67282
## - Width 1 138331303 4.4688e+09 67283
## - Fuel.tank.capacity 1 138705571 4.4692e+09 67284
## - AirBagsNone 1 194262805 4.5247e+09 67301
## - Weight 1 203112826 4.5336e+09 67304
## - Cylinders_Origin5.non.USA 1 231368909 4.5618e+09 67313
## - mod_categoryLuxury 1 658718950 4.9892e+09 67441
## <none> 4.3305e+09 68671
## - mod_categoryLow 1 8307799647 1.2638e+10 68773
## - mod_categoryMedium 1 9904841970 1.4235e+10 68944
##
## Step: AIC=67238.83
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
## AirBagsDriver.only + AirBagsNone + DriveTrainRear + EngineSize +
## Horsepower + RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
## Passengers + Length + Wheelbase + Width + Turn.circle + Rear.seat.room +
## Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
## mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
## Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
## Cylinders_Origin8.USA
##
## Df Sum of Sq RSS AIC
## - Turn.circle 1 1987069 4.3338e+09 65806
## - DriveTrainRear 1 2285517 4.3341e+09 65807
## - Cylinders_Origin4.USA 1 3054518 4.3349e+09 65807
## - Wheelbase 1 6466535 4.3383e+09 65808
## - AirBagsDriver.only 1 8029881 4.3398e+09 65808
## - Rev.per.mile 1 8269018 4.3401e+09 65809
## - Cylinders_Origin6.USA 1 10998201 4.3428e+09 65809
## - Man.trans.availYes 1 12216158 4.3440e+09 65810
## - Cylinders_Origin4.non.USA 1 16014051 4.3478e+09 65811
## - Rear.seat.room 1 19254324 4.3511e+09 65812
## - Cylinders_Origin8.USA 1 31048855 4.3629e+09 65816
## - EngineSize 1 35660513 4.3675e+09 65818
## - TypeLarge 1 46563818 4.3784e+09 65821
## - TypeVan 1 49139453 4.3809e+09 65822
## - RPM 1 50455608 4.3823e+09 65822
## - TypeSporty 1 61729362 4.3935e+09 65826
## - Passengers 1 63628997 4.3954e+09 65827
## - Length 1 81222062 4.4130e+09 65832
## - Cylinders_Origin6.non.USA 1 92273260 4.4241e+09 65836
## - TypeMidsize 1 104541425 4.4363e+09 65840
## - MPG.highway 1 115194318 4.4470e+09 65843
## - Luggage.room 1 115567785 4.4474e+09 65844
## - Horsepower 1 147195566 4.4790e+09 65854
## - Fuel.tank.capacity 1 159476950 4.4913e+09 65858
## - Width 1 174425491 4.5062e+09 65862

```

```

## - Weight 1 202004716 4.5338e+09 65871
## - AirBagsNone 1 202416110 4.5342e+09 65871
## - Cylinders_Origin5.non.USA 1 230110633 4.5619e+09 65880
## - mod_categoryLuxury 1 675332189 5.0071e+09 66013
## <none> 4.3318e+09 67239
## - mod_categoryLow 1 8659051535 1.2991e+10 67380
## - mod_categoryMedium 1 9970500109 1.4302e+10 67517
##
## Step: AIC=65806.49
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
## AirBagsDriver.only + AirBagsNone + DriveTrainRear + EngineSize +
## Horsepower + RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
## Passengers + Length + Wheelbase + Width + Rear.seat.room +
## Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
## mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
## Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
## Cylinders_Origin8.USA
##
## Df Sum of Sq RSS AIC
## - DriveTrainRear 1 2.8204e+06 4.3366e+09 64374
## - Cylinders_Origin4.USA 1 3.5603e+06 4.3373e+09 64375
## - Wheelbase 1 7.1894e+06 4.3410e+09 64376
## - Rev.per.mile 1 7.5900e+06 4.3414e+09 64376
## - AirBagsDriver.only 1 9.0802e+06 4.3429e+09 64376
## - Cylinders_Origin6.USA 1 1.0514e+07 4.3443e+09 64377
## - Man.trans.availYes 1 1.1590e+07 4.3454e+09 64377
## - Cylinders_Origin4.non.USA 1 1.5626e+07 4.3494e+09 64379
## - Rear.seat.room 1 1.7440e+07 4.3512e+09 64379
## - Cylinders_Origin8.USA 1 3.0649e+07 4.3644e+09 64384
## - EngineSize 1 3.9191e+07 4.3730e+09 64386
## - TypeLarge 1 4.6544e+07 4.3803e+09 64389
## - TypeVan 1 4.7924e+07 4.3817e+09 64389
## - RPM 1 4.8789e+07 4.3826e+09 64390
## - Passengers 1 6.1659e+07 4.3954e+09 64394
## - TypeSporty 1 6.2999e+07 4.3968e+09 64394
## - Length 1 8.1090e+07 4.4149e+09 64400
## - Cylinders_Origin6.non.USA 1 9.0290e+07 4.4241e+09 64403
## - TypeMidsize 1 1.0957e+08 4.4434e+09 64409
## - Luggage.room 1 1.1360e+08 4.4474e+09 64411
## - MPG.highway 1 1.1627e+08 4.4501e+09 64411
## - Horsepower 1 1.5453e+08 4.4883e+09 64424
## - Fuel.tank.capacity 1 1.5824e+08 4.4920e+09 64425
## - Width 1 1.7327e+08 4.5071e+09 64430
## - Weight 1 2.0218e+08 4.5360e+09 64439
## - Cylinders_Origin5.non.USA 1 2.3344e+08 4.5672e+09 64449
## - AirBagsNone 1 2.3664e+08 4.5704e+09 64450
## - mod_categoryLuxury 1 6.7376e+08 5.0076e+09 64581
## <none> 4.3338e+09 65806
## - mod_categoryLow 1 8.7985e+09 1.3132e+10 65962
## - mod_categoryMedium 1 1.0840e+10 1.5174e+10 66169
##
## Step: AIC=64374.42
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
## AirBagsDriver.only + AirBagsNone + EngineSize + Horsepower +

```

```

##      RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##      Passengers + Length + Wheelbase + Width + Rear.seat.room +
##      Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##      Cylinders_Origin8.USA
##
##
##              Df  Sum of Sq      RSS    AIC
## - Cylinders_Origin4.USA      1 6.2652e+06 4.3429e+09 62943
## - Wheelbase                  1 7.2221e+06 4.3438e+09 62944
## - Rev.per.mile                1 8.8383e+06 4.3454e+09 62944
## - AirBagsDriver.only         1 9.0210e+06 4.3456e+09 62944
## - Man.trans.availYes         1 1.1040e+07 4.3476e+09 62945
## - Cylinders_Origin6.USA      1 1.4915e+07 4.3515e+09 62946
## - Rear.seat.room             1 1.5227e+07 4.3518e+09 62946
## - Cylinders_Origin4.non.USA  1 2.1536e+07 4.3581e+09 62949
## - Cylinders_Origin8.USA      1 3.5451e+07 4.3721e+09 62953
## - EngineSize                 1 3.7434e+07 4.3740e+09 62954
## - TypeVan                    1 4.6739e+07 4.3833e+09 62957
## - TypeLarge                  1 4.9426e+07 4.3860e+09 62958
## - RPM                        1 5.0721e+07 4.3873e+09 62958
## - TypeSporty                 1 7.7164e+07 4.4138e+09 62967
## - Length                     1 8.6180e+07 4.4228e+09 62970
## - Passengers                 1 8.6940e+07 4.4235e+09 62970
## - Cylinders_Origin6.non.USA  1 1.0530e+08 4.4419e+09 62976
## - TypeMidsize                1 1.2301e+08 4.4596e+09 62982
## - Luggage.room               1 1.2346e+08 4.4601e+09 62982
## - MPG.highway                1 1.2900e+08 4.4656e+09 62983
## - Fuel.tank.capacity         1 1.5632e+08 4.4929e+09 62992
## - Horsepower                 1 1.7678e+08 4.5134e+09 62999
## - Weight                     1 2.0060e+08 4.5372e+09 63006
## - Width                      1 2.1119e+08 4.5478e+09 63010
## - AirBagsNone                1 2.3641e+08 4.5730e+09 63017
## - Cylinders_Origin5.non.USA  1 2.5266e+08 4.5893e+09 63023
## - mod_categoryLuxury         1 6.7096e+08 5.0076e+09 63148
## <none>                      4.3366e+09 64374
## - mod_categoryLow            1 8.9354e+09 1.3272e+10 64544
## - mod_categoryMedium         1 1.1250e+10 1.5587e+10 64775
##
## Step:  AIC=62943.49
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##      AirBagsDriver.only + AirBagsNone + EngineSize + Horsepower +
##      RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##      Passengers + Length + Wheelbase + Width + Rear.seat.room +
##      Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##
##              Df  Sum of Sq      RSS    AIC
## - Wheelbase                  1 4.6267e+06 4.3475e+09 61512
## - Rev.per.mile                1 6.5638e+06 4.3494e+09 61513
## - AirBagsDriver.only         1 8.9567e+06 4.3518e+09 61513
## - Cylinders_Origin6.USA      1 8.9772e+06 4.3519e+09 61513
## - Man.trans.availYes         1 1.0940e+07 4.3538e+09 61514

```

```

## - Rear.seat.room          1 2.0282e+07 4.3632e+09 61517
## - Cylinders_Origin4.non.USA 1 2.8289e+07 4.3712e+09 61520
## - Cylinders_Origin8.USA    1 3.0420e+07 4.3733e+09 61520
## - EngineSize              1 3.8610e+07 4.3815e+09 61523
## - TypeLarge               1 4.4091e+07 4.3870e+09 61525
## - RPM                     1 4.8329e+07 4.3912e+09 61526
## - TypeVan                 1 5.7893e+07 4.4008e+09 61529
## - Length                  1 8.1380e+07 4.4243e+09 61537
## - TypeSporty              1 8.1604e+07 4.4245e+09 61537
## - Passengers              1 1.0423e+08 4.4471e+09 61544
## - TypeMidsize             1 1.1852e+08 4.4614e+09 61549
## - Luggage.room            1 1.3158e+08 4.4745e+09 61553
## - MPG.highway             1 1.3440e+08 4.4773e+09 61554
## - Cylinders_Origin6.non.USA 1 1.4573e+08 4.4886e+09 61558
## - Fuel.tank.capacity      1 1.5214e+08 4.4950e+09 61560
## - Horsepower              1 1.7956e+08 4.5224e+09 61569
## - Weight                  1 1.9561e+08 4.5385e+09 61574
## - Width                   1 2.0678e+08 4.5497e+09 61577
## - AirBagsNone             1 2.3988e+08 4.5828e+09 61588
## - Cylinders_Origin5.non.USA 1 3.0423e+08 4.6471e+09 61608
## - mod_categoryLuxury      1 8.2605e+08 5.1689e+09 61760
## <none>                    4.3429e+09 62943
## - mod_categoryLow         1 9.3301e+09 1.3673e+10 63154
## - mod_categoryMedium      1 1.1498e+10 1.5841e+10 63365
##
## Step: AIC=61512.01
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsDriver.only + AirBagsNone + EngineSize + Horsepower +
##   RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##   Passengers + Length + Width + Rear.seat.room + Luggage.room +
##   Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##   Cylinders_Origin4.non.USA + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##   Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##           Df Sum of Sq      RSS   AIC
## - AirBagsDriver.only      1 7.5762e+06 4.3551e+09 60082
## - Rev.per.mile             1 7.6835e+06 4.3552e+09 60082
## - Man.trans.availYes       1 8.8335e+06 4.3563e+09 60082
## - Cylinders_Origin6.USA     1 1.3947e+07 4.3614e+09 60084
## - Rear.seat.room          1 2.3064e+07 4.3706e+09 60087
## - Cylinders_Origin4.non.USA 1 3.1120e+07 4.3786e+09 60089
## - Cylinders_Origin8.USA     1 3.5999e+07 4.3835e+09 60091
## - EngineSize              1 3.9364e+07 4.3869e+09 60092
## - TypeLarge               1 3.9607e+07 4.3871e+09 60092
## - RPM                     1 4.8353e+07 4.3959e+09 60095
## - TypeVan                 1 7.6279e+07 4.4238e+09 60104
## - Length                  1 9.6811e+07 4.4443e+09 60111
## - TypeSporty              1 1.0059e+08 4.4481e+09 60112
## - Passengers              1 1.0521e+08 4.4527e+09 60113
## - TypeMidsize             1 1.1390e+08 4.4614e+09 60116
## - MPG.highway             1 1.2982e+08 4.4773e+09 60121
## - Luggage.room            1 1.3333e+08 4.4808e+09 60122
## - Cylinders_Origin6.non.USA 1 1.5717e+08 4.5047e+09 60130
## - Fuel.tank.capacity       1 1.6254e+08 4.5100e+09 60132

```

```

## - Horsepower          1 1.8822e+08 4.5357e+09 60140
## - Weight              1 1.9859e+08 4.5461e+09 60143
## - Width               1 2.2430e+08 4.5718e+09 60151
## - AirBagsNone         1 2.3563e+08 4.5831e+09 60155
## - Cylinders_Origin5.non.USA 1 3.2143e+08 4.6689e+09 60181
## - mod_categoryLuxury   1 8.2350e+08 5.1710e+09 60328
## <none>                4.3475e+09 61512
## - mod_categoryLow      1 9.7422e+09 1.4090e+10 61764
## - mod_categoryMedium   1 1.1541e+10 1.5888e+10 61936
##
## Step: AIC=60081.51
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsNone + EngineSize + Horsepower + RPM + Rev.per.mile +
##   Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
##   Width + Rear.seat.room + Luggage.room + Weight + mod_categoryLow +
##   mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin4.non.USA +
##   Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##   Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS   AIC
## - Man.trans.availYes      1 7.5722e+06 4.3626e+09 58651
## - Rev.per.mile            1 1.0090e+07 4.3652e+09 58652
## - Rear.seat.room         1 1.9526e+07 4.3746e+09 58655
## - Cylinders_Origin6.USA    1 1.9892e+07 4.3750e+09 58655
## - Cylinders_Origin4.non.USA 1 3.1729e+07 4.3868e+09 58659
## - TypeLarge               1 3.3286e+07 4.3884e+09 58659
## - Cylinders_Origin8.USA    1 3.8192e+07 4.3933e+09 58661
## - EngineSize              1 3.9673e+07 4.3947e+09 58662
## - RPM                     1 5.5355e+07 4.4104e+09 58667
## - Length                  1 9.4758e+07 4.4498e+09 58679
## - TypeVan                 1 9.5094e+07 4.4502e+09 58679
## - TypeSporty              1 1.0427e+08 4.4593e+09 58682
## - TypeMidsize             1 1.0695e+08 4.4620e+09 58683
## - Passengers              1 1.0886e+08 4.4639e+09 58684
## - Luggage.room            1 1.2632e+08 4.4814e+09 58689
## - MPG.highway             1 1.3123e+08 4.4863e+09 58691
## - Fuel.tank.capacity       1 1.6194e+08 4.5170e+09 58701
## - Horsepower              1 1.8235e+08 4.5374e+09 58707
## - Cylinders_Origin6.non.USA 1 1.8568e+08 4.5408e+09 58708
## - Weight                  1 1.9733e+08 4.5524e+09 58712
## - Width                   1 2.1824e+08 4.5733e+09 58719
## - AirBagsNone             1 3.4750e+08 4.7026e+09 58759
## - Cylinders_Origin5.non.USA 1 3.6604e+08 4.7211e+09 58764
## - mod_categoryLuxury      1 8.1608e+08 5.1712e+09 58895
## <none>                    4.3551e+09 60082
## - mod_categoryLow         1 1.0186e+10 1.4541e+10 60376
## - mod_categoryMedium      1 1.2553e+10 1.6908e+10 60592
##
## Step: AIC=58651
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsNone + EngineSize + Horsepower + RPM + Rev.per.mile +
##   Fuel.tank.capacity + Passengers + Length + Width + Rear.seat.room +
##   Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin5.non.USA +

```



```

##      Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##      Df Sum of Sq      RSS      AIC
## - Rev.per.mile      1 1.2815e+07 4.3755e+09 57222
## - Rear.seat.room    1 1.7383e+07 4.3800e+09 57224
## - Cylinders_Origin6.USA 1 2.3899e+07 4.3865e+09 57226
## - TypeLarge         1 2.6229e+07 4.3889e+09 57227
## - EngineSize        1 3.4966e+07 4.3976e+09 57229
## - Cylinders_Origin8.USA 1 4.1668e+07 4.4043e+09 57232
## - Cylinders_Origin4.non.USA 1 4.6934e+07 4.4096e+09 57233
## - RPM              1 5.2790e+07 4.4154e+09 57235
## - Length           1 9.5148e+07 4.4578e+09 57249
## - Passengers       1 1.0131e+08 4.4640e+09 57251
## - TypeSporty       1 1.0141e+08 4.4641e+09 57251
## - TypeMidsize      1 1.0230e+08 4.4649e+09 57251
## - TypeVan          1 1.0886e+08 4.4715e+09 57253
## - Luggage.room     1 1.2134e+08 4.4840e+09 57257
## - MPG.highway      1 1.2453e+08 4.4872e+09 57258
## - Fuel.tank.capacity 1 1.7510e+08 4.5377e+09 57274
## - Horsepower       1 1.8144e+08 4.5441e+09 57276
## - Weight           1 1.9095e+08 4.5536e+09 57279
## - Width            1 2.1089e+08 4.5735e+09 57286
## - Cylinders_Origin6.non.USA 1 2.2428e+08 4.5869e+09 57290
## - AirBagsNone      1 3.5544e+08 4.7181e+09 57330
## - Cylinders_Origin5.non.USA 1 4.8493e+08 4.8476e+09 57369
## - mod_categoryLuxury 1 8.1416e+08 5.1768e+09 57463
## <none>                                4.3626e+09 58651
## - mod_categoryLow   1 1.0304e+10 1.4667e+10 58956
## - mod_categoryMedium 1 1.2876e+10 1.7238e+10 59187
##
## Step: AIC=57222.2
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##      AirBagsNone + EngineSize + Horsepower + RPM + Fuel.tank.capacity +
##      Passengers + Length + Width + Rear.seat.room + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin4.non.USA + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##      Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##      Df Sum of Sq      RSS      AIC
## - Rear.seat.room    1 1.6024e+07 4.3915e+09 55794
## - Cylinders_Origin6.USA 1 1.9025e+07 4.3945e+09 55795
## - TypeLarge         1 2.4902e+07 4.4004e+09 55797
## - EngineSize        1 2.5761e+07 4.4012e+09 55798
## - Cylinders_Origin8.USA 1 3.9665e+07 4.4151e+09 55802
## - Cylinders_Origin4.non.USA 1 4.2940e+07 4.4184e+09 55803
## - RPM              1 5.2639e+07 4.4281e+09 55806
## - TypeSporty       1 9.5416e+07 4.4709e+09 55820
## - TypeMidsize      1 9.9843e+07 4.4753e+09 55822
## - Passengers       1 1.0106e+08 4.4765e+09 55822
## - Length           1 1.0227e+08 4.4777e+09 55822
## - TypeVan          1 1.0822e+08 4.4837e+09 55824
## - Luggage.room     1 1.1223e+08 4.4877e+09 55825
## - MPG.highway      1 1.4522e+08 4.5207e+09 55836
## - Horsepower       1 1.7693e+08 4.5524e+09 55846

```

```

## - Fuel.tank.capacity      1 1.8954e+08 4.5650e+09 55850
## - Weight                  1 1.8964e+08 4.5651e+09 55850
## - Width                   1 2.0190e+08 4.5774e+09 55854
## - Cylinders_Origin6.non.USA 1 2.2782e+08 4.6033e+09 55862
## - AirBagsNone             1 3.4740e+08 4.7229e+09 55899
## - Cylinders_Origin5.non.USA 1 5.1553e+08 4.8910e+09 55949
## - mod_categoryLuxury      1 8.2523e+08 5.2007e+09 56037
## <none>                    4.3755e+09 57222
## - mod_categoryLow         1 1.0632e+10 1.5008e+10 57555
## - mod_categoryMedium      1 1.3629e+10 1.8005e+10 57816
##
## Step: AIC=55794.44
## PriceNew ~ TypeLarge + TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsNone + EngineSize + Horsepower + RPM + Fuel.tank.capacity +
##   Passengers + Length + Width + Luggage.room + Weight + mod_categoryLow +
##   mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin4.non.USA +
##   Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##   Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS   AIC
## - TypeLarge      1 1.5964e+07 4.4075e+09 54367
## - Cylinders_Origin6.USA 1 2.0374e+07 4.4119e+09 54368
## - EngineSize      1 2.3319e+07 4.4148e+09 54369
## - Cylinders_Origin4.non.USA 1 3.7293e+07 4.4288e+09 54374
## - Cylinders_Origin8.USA 1 4.2715e+07 4.4342e+09 54375
## - RPM             1 5.7873e+07 4.4494e+09 54380
## - Passengers      1 8.7743e+07 4.4792e+09 54390
## - TypeMidsize     1 9.2482e+07 4.4840e+09 54391
## - Luggage.room    1 9.8717e+07 4.4902e+09 54393
## - Length          1 1.0918e+08 4.5007e+09 54397
## - TypeVan         1 1.1765e+08 4.5091e+09 54399
## - TypeSporty      1 1.2407e+08 4.5156e+09 54401
## - MPG.highway     1 1.3752e+08 4.5290e+09 54406
## - Horsepower      1 1.7032e+08 4.5618e+09 54416
## - Fuel.tank.capacity 1 1.8479e+08 4.5763e+09 54421
## - Width           1 1.8921e+08 4.5807e+09 54422
## - Weight          1 1.9878e+08 4.5903e+09 54425
## - Cylinders_Origin6.non.USA 1 2.3602e+08 4.6275e+09 54436
## - AirBagsNone     1 3.6558e+08 4.7571e+09 54476
## - Cylinders_Origin5.non.USA 1 5.0755e+08 4.8990e+09 54518
## - mod_categoryLuxury 1 8.3538e+08 5.2269e+09 54611
## <none>            4.3915e+09 55794
## - mod_categoryLow  1 1.1160e+10 1.5551e+10 56173
## - mod_categoryMedium 1 1.4075e+10 1.8467e+10 56420
##
## Step: AIC=54366.64
## PriceNew ~ TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsNone + EngineSize + Horsepower + RPM + Fuel.tank.capacity +
##   Passengers + Length + Width + Luggage.room + Weight + mod_categoryLow +
##   mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin4.non.USA +
##   Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##   Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS   AIC

```

```

## - EngineSize          1 2.4081e+07 4.4315e+09 52941
## - Cylinders_Origin4.non.USA 1 3.2228e+07 4.4397e+09 52944
## - Cylinders_Origin6.USA 1 4.4432e+07 4.4519e+09 52948
## - RPM                 1 7.4976e+07 4.4824e+09 52958
## - TypeMidsize         1 8.1776e+07 4.4892e+09 52960
## - Cylinders_Origin8.USA 1 8.6123e+07 4.4936e+09 52961
## - Luggage.room        1 9.7850e+07 4.5053e+09 52965
## - Length              1 1.0833e+08 4.5158e+09 52968
## - Passengers          1 1.0915e+08 4.5166e+09 52969
## - TypeSporty          1 1.1178e+08 4.5192e+09 52970
## - MPG.highway         1 1.3949e+08 4.5469e+09 52978
## - Horsepower          1 1.5473e+08 4.5622e+09 52983
## - Width               1 1.7925e+08 4.5867e+09 52991
## - Fuel.tank.capacity  1 2.0062e+08 4.6081e+09 52997
## - Weight              1 2.1047e+08 4.6179e+09 53000
## - TypeVan             1 2.2564e+08 4.6331e+09 53005
## - Cylinders_Origin6.non.USA 1 2.8885e+08 4.6963e+09 53025
## - AirBagsNone         1 3.8446e+08 4.7919e+09 53053
## - Cylinders_Origin5.non.USA 1 5.5346e+08 4.9609e+09 53103
## - mod_categoryLuxury  1 9.0446e+08 5.3119e+09 53201
## <none>                 4.4075e+09 54367
## - mod_categoryLow     1 1.1246e+10 1.5653e+10 54750
## - mod_categoryMedium  1 1.4067e+10 1.8474e+10 54987
##
## Step: AIC=52941.45
## PriceNew ~ TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsNone + Horsepower + RPM + Fuel.tank.capacity + Passengers +
##   Length + Width + Luggage.room + Weight + mod_categoryLow +
##   mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin4.non.USA +
##   Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##   Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS    AIC
## - Cylinders_Origin4.non.USA 1 3.3084e+07 4.4646e+09 51519
## - RPM                       1 5.0976e+07 4.4825e+09 51525
## - TypeMidsize               1 6.6805e+07 4.4983e+09 51530
## - Cylinders_Origin6.USA     1 8.5918e+07 4.5175e+09 51536
## - TypeSporty                1 9.5885e+07 4.5274e+09 51539
## - Length                    1 1.0100e+08 4.5325e+09 51541
## - Luggage.room              1 1.0210e+08 4.5336e+09 51541
## - Passengers                1 1.1184e+08 4.5434e+09 51544
## - MPG.highway               1 1.3623e+08 4.5678e+09 51552
## - Width                     1 1.6474e+08 4.5963e+09 51561
## - Cylinders_Origin8.USA     1 1.6548e+08 4.5970e+09 51561
## - Weight                    1 2.0526e+08 4.6368e+09 51573
## - Fuel.tank.capacity        1 2.3046e+08 4.6620e+09 51581
## - Horsepower                1 2.3567e+08 4.6672e+09 51583
## - TypeVan                   1 2.6327e+08 4.6948e+09 51591
## - Cylinders_Origin6.non.USA 1 3.5531e+08 4.7868e+09 51619
## - AirBagsNone               1 3.8896e+08 4.8205e+09 51629
## - Cylinders_Origin5.non.USA 1 5.5585e+08 4.9874e+09 51678
## - mod_categoryLuxury        1 1.3179e+09 5.7494e+09 51882
## <none>                       4.4315e+09 52941
## - mod_categoryLow           1 1.1226e+10 1.5658e+10 53317

```

```

## - mod_categoryMedium          1 1.4150e+10 1.8582e+10 53563
##
## Step: AIC=51519.11
## PriceNew ~ TypeMidsize + TypeSporty + TypeVan + MPG.highway +
##   AirBagsNone + Horsepower + RPM + Fuel.tank.capacity + Passengers +
##   Length + Width + Luggage.room + Weight + mod_categoryLow +
##   mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##   Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##
##          Df Sum of Sq      RSS      AIC
## - TypeMidsize          1 5.6023e+07 4.5206e+09 50104
## - Cylinders_Origin6.USA  1 6.3876e+07 4.5285e+09 50106
## - RPM                  1 7.4631e+07 4.5392e+09 50110
## - TypeSporty           1 8.5317e+07 4.5499e+09 50113
## - Luggage.room         1 9.8540e+07 4.5632e+09 50117
## - Passengers           1 1.0357e+08 4.5682e+09 50119
## - Length               1 1.1641e+08 4.5810e+09 50123
## - Cylinders_Origin8.USA 1 1.4567e+08 4.6103e+09 50132
## - MPG.highway          1 1.4605e+08 4.6107e+09 50132
## - Width                1 1.6706e+08 4.6317e+09 50139
## - Horsepower           1 2.2124e+08 4.6859e+09 50155
## - Fuel.tank.capacity    1 2.3653e+08 4.7011e+09 50160
## - Weight               1 2.5700e+08 4.7216e+09 50166
## - TypeVan              1 2.7678e+08 4.7414e+09 50172
## - Cylinders_Origin6.non.USA 1 3.2385e+08 4.7885e+09 50186
## - AirBagsNone          1 3.8884e+08 4.8535e+09 50206
## - Cylinders_Origin5.non.USA 1 5.2566e+08 4.9903e+09 50246
## - mod_categoryLuxury    1 1.3015e+09 5.7661e+09 50453
## <none>                  4.4646e+09 51519
## - mod_categoryLow      1 1.1297e+10 1.5762e+10 51894
## - mod_categoryMedium    1 1.4320e+10 1.8784e+10 52145
##
## Step: AIC=50103.98
## PriceNew ~ TypeSporty + TypeVan + MPG.highway + AirBagsNone +
##   Horsepower + RPM + Fuel.tank.capacity + Passengers + Length +
##   Width + Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##   Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##
##          Df Sum of Sq      RSS      AIC
## - TypeSporty          1 5.3428e+07 4.5741e+09 48688
## - Cylinders_Origin6.USA 1 7.6658e+07 4.5973e+09 48695
## - Passengers          1 7.9253e+07 4.5999e+09 48696
## - RPM                 1 7.9435e+07 4.6001e+09 48696
## - Length              1 1.0465e+08 4.6253e+09 48704
## - Luggage.room        1 1.1271e+08 4.6334e+09 48706
## - Cylinders_Origin8.USA 1 1.2411e+08 4.6447e+09 48710
## - MPG.highway         1 1.2452e+08 4.6452e+09 48710
## - Width               1 1.5648e+08 4.6771e+09 48720
## - Horsepower          1 1.9999e+08 4.7206e+09 48733
## - Fuel.tank.capacity   1 2.4220e+08 4.7628e+09 48746
## - Weight              1 2.4520e+08 4.7658e+09 48747
## - TypeVan             1 3.3556e+08 4.8562e+09 48774
## - AirBagsNone         1 3.8135e+08 4.9020e+09 48787

```

```

## - Cylinders_Origin6.non.USA 1 5.0734e+08 5.0280e+09 48823
## - Cylinders_Origin5.non.USA 1 6.2565e+08 5.1463e+09 48857
## - mod_categoryLuxury 1 1.5610e+09 6.0816e+09 49096
## <none> 4.5206e+09 50104
## - mod_categoryLow 1 1.2535e+10 1.7056e+10 50574
## - mod_categoryMedium 1 1.4731e+10 1.9251e+10 50747
##
## Step: AIC=48687.81
## PriceNew ~ TypeVan + MPG.highway + AirBagsNone + Horsepower +
## RPM + Fuel.tank.capacity + Passengers + Length + Width +
## Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
## mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
## Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
## Df Sum of Sq RSS AIC
## - Passengers 1 3.2776e+07 4.6068e+09 47265
## - RPM 1 6.8372e+07 4.6424e+09 47276
## - Cylinders_Origin6.USA 1 7.3795e+07 4.6479e+09 47278
## - Luggage.room 1 8.7688e+07 4.6618e+09 47282
## - MPG.highway 1 9.0662e+07 4.6647e+09 47283
## - Length 1 1.1054e+08 4.6846e+09 47289
## - Width 1 1.1915e+08 4.6932e+09 47292
## - Cylinders_Origin8.USA 1 1.2406e+08 4.6981e+09 47293
## - Horsepower 1 1.8767e+08 4.7617e+09 47312
## - Weight 1 2.2423e+08 4.7983e+09 47323
## - Fuel.tank.capacity 1 2.3328e+08 4.8073e+09 47326
## - TypeVan 1 2.9413e+08 4.8682e+09 47344
## - AirBagsNone 1 4.2434e+08 4.9984e+09 47382
## - Cylinders_Origin6.non.USA 1 4.8424e+08 5.0583e+09 47399
## - Cylinders_Origin5.non.USA 1 6.0878e+08 5.1829e+09 47434
## - mod_categoryLuxury 1 1.5550e+09 6.1291e+09 47674
## <none> 4.5741e+09 48688
## - mod_categoryLow 1 1.2875e+10 1.7449e+10 49173
## - mod_categoryMedium 1 1.4724e+10 1.9298e+10 49318
##
## Step: AIC=47265.05
## PriceNew ~ TypeVan + MPG.highway + AirBagsNone + Horsepower +
## RPM + Fuel.tank.capacity + Length + Width + Luggage.room +
## Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
## Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
## Cylinders_Origin8.USA
##
## Df Sum of Sq RSS AIC
## - MPG.highway 1 7.6166e+07 4.6830e+09 45856
## - Length 1 8.6862e+07 4.6937e+09 45859
## - Cylinders_Origin6.USA 1 9.4963e+07 4.7018e+09 45861
## - RPM 1 1.0179e+08 4.7086e+09 45863
## - Luggage.room 1 1.2060e+08 4.7274e+09 45869
## - Cylinders_Origin8.USA 1 1.4452e+08 4.7514e+09 45876
## - Width 1 1.4688e+08 4.7537e+09 45877
## - Horsepower 1 1.5504e+08 4.7619e+09 45879
## - Fuel.tank.capacity 1 2.4066e+08 4.8475e+09 45905
## - TypeVan 1 2.6143e+08 4.8683e+09 45911
## - Weight 1 2.7586e+08 4.8827e+09 45915

```

```

## - AirBagsNone          1 4.2550e+08 5.0323e+09 45959
## - Cylinders_Origin6.non.USA 1 4.9097e+08 5.0978e+09 45977
## - Cylinders_Origin5.non.USA 1 5.8162e+08 5.1885e+09 46002
## - mod_categoryLuxury    1 1.5838e+09 6.1907e+09 46255
## <none>                  4.6068e+09 47265
## - mod_categoryLow       1 1.2885e+10 1.7492e+10 47744
## - mod_categoryMedium    1 1.4709e+10 1.9316e+10 47886
##
## Step: AIC=45855.54
## PriceNew ~ TypeVan + AirBagsNone + Horsepower + RPM + Fuel.tank.capacity +
##   Length + Width + Luggage.room + Weight + mod_categoryLow +
##   mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##   Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##           Df Sum of Sq      RSS   AIC
## - Length          1 7.0890e+07 4.7539e+09 44444
## - Cylinders_Origin6.USA 1 1.1737e+08 4.8004e+09 44458
## - Width           1 1.2657e+08 4.8096e+09 44461
## - Horsepower      1 1.3934e+08 4.8224e+09 44465
## - RPM             1 1.4074e+08 4.8237e+09 44465
## - Cylinders_Origin8.USA 1 1.7048e+08 4.8535e+09 44474
## - Luggage.room    1 1.7492e+08 4.8579e+09 44475
## - Fuel.tank.capacity 1 2.0635e+08 4.8894e+09 44484
## - Weight          1 2.0923e+08 4.8922e+09 44485
## - TypeVan         1 2.7259e+08 4.9556e+09 44504
## - AirBagsNone     1 4.7617e+08 5.1592e+09 44561
## - Cylinders_Origin6.non.USA 1 5.2468e+08 5.2077e+09 44575
## - Cylinders_Origin5.non.USA 1 6.1964e+08 5.3027e+09 44601
## - mod_categoryLuxury 1 1.6984e+09 6.3814e+09 44866
## <none>            4.6830e+09 45856
## - mod_categoryLow 1 1.2895e+10 1.7578e+10 46318
## - mod_categoryMedium 1 1.4637e+10 1.9320e+10 46453
##
## Step: AIC=44444.07
## PriceNew ~ TypeVan + AirBagsNone + Horsepower + RPM + Fuel.tank.capacity +
##   Width + Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##   Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##           Df Sum of Sq      RSS   AIC
## - RPM          1 9.6712e+07 4.8506e+09 43040
## - Cylinders_Origin6.USA 1 1.3002e+08 4.8839e+09 43050
## - Luggage.room  1 1.3835e+08 4.8922e+09 43052
## - Weight        1 1.3983e+08 4.8937e+09 43053
## - Cylinders_Origin8.USA 1 1.6303e+08 4.9169e+09 43059
## - Fuel.tank.capacity 1 2.0936e+08 4.9633e+09 43073
## - TypeVan       1 2.2706e+08 4.9810e+09 43078
## - Width         1 2.8727e+08 5.0412e+09 43095
## - Horsepower    1 3.2372e+08 5.0776e+09 43105
## - AirBagsNone   1 5.6022e+08 5.3141e+09 43171
## - Cylinders_Origin6.non.USA 1 5.9675e+08 5.3506e+09 43181
## - Cylinders_Origin5.non.USA 1 7.4900e+08 5.5029e+09 43221
## - mod_categoryLuxury 1 1.7251e+09 6.4790e+09 43455
## <none>          4.7539e+09 44444

```

```

## - mod_categoryLow          1 1.2892e+10 1.7646e+10 44891
## - mod_categoryMedium       1 1.5057e+10 1.9811e+10 45056
##
## Step: AIC=43039.93
## PriceNew ~ TypeVan + AirBagsNone + Horsepower + Fuel.tank.capacity +
##   Width + Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##   Cylinders_Origin6.USA + Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS   AIC
## - Cylinders_Origin6.USA      1 6.9495e+07 4.9201e+09 41627
## - Cylinders_Origin8.USA      1 8.7913e+07 4.9385e+09 41633
## - Weight                     1 9.4480e+07 4.9451e+09 41635
## - Luggage.room               1 1.4462e+08 4.9952e+09 41649
## - Fuel.tank.capacity         1 1.7146e+08 5.0221e+09 41657
## - TypeVan                    1 2.0780e+08 5.0584e+09 41667
## - Width                      1 3.1834e+08 5.1689e+09 41698
## - AirBagsNone                1 5.4042e+08 5.3910e+09 41758
## - Cylinders_Origin6.non.USA  1 5.7330e+08 5.4239e+09 41767
## - Cylinders_Origin5.non.USA  1 8.1958e+08 5.6702e+09 41831
## - Horsepower                 1 1.0028e+09 5.8534e+09 41876
## - mod_categoryLuxury         1 1.6747e+09 6.5253e+09 42032
## <none>                      4.8506e+09 43040
## - mod_categoryLow           1 1.2801e+10 1.7652e+10 43458
## - mod_categoryMedium         1 1.4965e+10 1.9816e+10 43624
##
## Step: AIC=41627.32
## PriceNew ~ TypeVan + AirBagsNone + Horsepower + Fuel.tank.capacity +
##   Width + Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##   Cylinders_Origin8.USA
##
##              Df Sum of Sq      RSS   AIC
## - Cylinders_Origin8.USA      1 2.6381e+07 4.9465e+09 40202
## - Fuel.tank.capacity         1 1.2208e+08 5.0422e+09 40229
## - Weight                     1 1.3437e+08 5.0545e+09 40233
## - Luggage.room               1 1.6343e+08 5.0835e+09 40241
## - TypeVan                    1 2.0499e+08 5.1251e+09 40253
## - Width                      1 2.4884e+08 5.1690e+09 40265
## - AirBagsNone                1 5.0072e+08 5.4208e+09 40333
## - Cylinders_Origin6.non.USA  1 5.1869e+08 5.4388e+09 40338
## - Cylinders_Origin5.non.USA  1 7.5015e+08 5.6703e+09 40398
## - Horsepower                 1 1.2203e+09 6.1404e+09 40512
## - mod_categoryLuxury         1 1.6992e+09 6.6193e+09 40619
## <none>                      4.9201e+09 41627
## - mod_categoryLow           1 1.2758e+10 1.7678e+10 42027
## - mod_categoryMedium         1 1.5467e+10 2.0387e+10 42231
##
## Step: AIC=40201.98
## PriceNew ~ TypeVan + AirBagsNone + Horsepower + Fuel.tank.capacity +
##   Width + Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##
##              Df Sum of Sq      RSS   AIC

```

```

## - Fuel.tank.capacity      1 1.1241e+08 5.0589e+09 38801
## - Weight                  1 1.2694e+08 5.0734e+09 38805
## - Luggage.room           1 2.0011e+08 5.1466e+09 38826
## - TypeVan                 1 2.1830e+08 5.1648e+09 38831
## - Width                   1 2.2867e+08 5.1752e+09 38834
## - AirBagsNone             1 4.8890e+08 5.4354e+09 38904
## - Cylinders_Origin6.non.USA 1 4.9239e+08 5.4389e+09 38905
## - Cylinders_Origin5.non.USA 1 7.6668e+08 5.7132e+09 38975
## - Horsepower              1 1.3783e+09 6.3248e+09 39121
## - mod_categoryLuxury      1 1.7058e+09 6.6523e+09 39194
## <none>                    4.9465e+09 40202
## - mod_categoryLow         1 1.3283e+10 1.8229e+10 40638
## - mod_categoryMedium      1 1.7992e+10 2.2938e+10 40967
##
## Step: AIC=38801.18
## PriceNew ~ TypeVan + AirBagsNone + Horsepower + Width + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##
##              Df Sum of Sq      RSS    AIC
## - TypeVan      1 1.8158e+08 5.2405e+09 37419
## - Width        1 1.9860e+08 5.2575e+09 37423
## - Weight       1 2.1909e+08 5.2780e+09 37429
## - Luggage.room 1 2.9672e+08 5.3556e+09 37450
## - AirBagsNone  1 4.4883e+08 5.5077e+09 37490
## - Cylinders_Origin6.non.USA 1 6.4933e+08 5.7082e+09 37541
## - Cylinders_Origin5.non.USA 1 9.3335e+08 5.9922e+09 37611
## - Horsepower   1 1.6839e+09 6.7428e+09 37780
## - mod_categoryLuxury 1 1.9575e+09 7.0164e+09 37837
## <none>          5.0589e+09 38801
## - mod_categoryLow 1 1.3603e+10 1.8661e+10 39239
## - mod_categoryMedium 1 1.7952e+10 2.3011e+10 39539
##
## Step: AIC=37418.72
## PriceNew ~ AirBagsNone + Horsepower + Width + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##
##              Df Sum of Sq      RSS    AIC
## - Weight      1 1.0476e+08 5.3452e+09 36014
## - Width       1 1.7009e+08 5.4106e+09 36031
## - Luggage.room 1 2.9505e+08 5.5355e+09 36064
## - Cylinders_Origin6.non.USA 1 6.2493e+08 5.8654e+09 36147
## - AirBagsNone  1 8.2752e+08 6.0680e+09 36196
## - Cylinders_Origin5.non.USA 1 8.3843e+08 6.0789e+09 36198
## - mod_categoryLuxury 1 1.9938e+09 7.2343e+09 36448
## - Horsepower   1 2.1400e+09 7.3805e+09 36476
## <none>          5.2405e+09 37419
## - mod_categoryLow 1 1.4164e+10 1.9404e+10 37862
## - mod_categoryMedium 1 1.8914e+10 2.4154e+10 38175
##
## Step: AIC=36014.08
## PriceNew ~ AirBagsNone + Horsepower + Width + Luggage.room +
##      mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +

```



```

##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##
##              Df  Sum of Sq      RSS   AIC
## - Width      1 7.4600e+07 5.4198e+09 34601
## - Luggage.room 1 5.6307e+08 5.9083e+09 34725
## - AirBagsNone 1 8.0807e+08 6.1533e+09 34783
## - Cylinders_Origin6.non.USA 1 9.3898e+08 6.2842e+09 34813
## - Cylinders_Origin5.non.USA 1 1.3409e+09 6.6861e+09 34902
## - mod_categoryLuxury      1 2.4379e+09 7.7831e+09 35120
## - Horsepower              1 2.7026e+09 8.0479e+09 35167
## <none>                    5.3452e+09 36014
## - mod_categoryLow         1 1.6256e+10 2.1601e+10 36582
## - mod_categoryMedium      1 1.8974e+10 2.4319e+10 36752
##
## Step:  AIC=34600.94
## PriceNew ~ AirBagsNone + Horsepower + Luggage.room + mod_categoryLow +
##      mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA
##
##              Df  Sum of Sq      RSS   AIC
## - Luggage.room      1 5.8371e+08 6.0035e+09 33315
## - AirBagsNone       1 7.6954e+08 6.1894e+09 33358
## - Cylinders_Origin6.non.USA 1 1.0611e+09 6.4810e+09 33424
## - Cylinders_Origin5.non.USA 1 1.3569e+09 6.7767e+09 33488
## - mod_categoryLuxury 1 2.6787e+09 8.0985e+09 33743
## - Horsepower        1 3.1914e+09 8.6112e+09 33831
## <none>              5.4198e+09 34601
## - mod_categoryLow   1 1.6186e+10 2.1606e+10 35150
## - mod_categoryMedium 1 2.0624e+10 2.6044e+10 35417
##
## Step:  AIC=33314.51
## PriceNew ~ AirBagsNone + Horsepower + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##
##              Df  Sum of Sq      RSS   AIC
## - AirBagsNone       1 6.7204e+08 6.6756e+09 32034
## - Cylinders_Origin6.non.USA 1 9.0277e+08 6.9063e+09 32082
## - Cylinders_Origin5.non.USA 1 1.6286e+09 7.6321e+09 32225
## - mod_categoryLuxury 1 2.4367e+09 8.4403e+09 32370
## - Horsepower        1 3.6897e+09 9.6932e+09 32568
## <none>              6.0035e+09 33315
## - mod_categoryLow   1 1.8515e+10 2.4518e+10 33898
## - mod_categoryMedium 1 2.1612e+10 2.7616e+10 34068
##
## Step:  AIC=32033.56
## PriceNew ~ Horsepower + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##
##              Df  Sum of Sq      RSS   AIC
## - Cylinders_Origin6.non.USA 1 7.6631e+08 7.4419e+09 30756
## - Cylinders_Origin5.non.USA 1 1.3521e+09 8.0277e+09 30865
## - mod_categoryLuxury      1 2.2847e+09 8.9602e+09 31022
## - Horsepower              1 4.2111e+09 1.0887e+10 31301
## <none>                    6.6756e+09 32034

```

```

## - mod_categoryMedium          1 2.4197e+10 3.0872e+10 32795
## - mod_categoryLow             1 2.4354e+10 3.1030e+10 32802
##
## Step: AIC=30756.29
## PriceNew ~ Horsepower + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium + Cylinders_Origin5.non.USA
##
##              Df Sum of Sq      RSS   AIC
## - Cylinders_Origin5.non.USA  1 1.2564e+09 8.6983e+09 29547
## - mod_categoryLuxury         1 1.8313e+09 9.2732e+09 29638
## - Horsepower                 1 4.9438e+09 1.2386e+10 30053
## <none>                       7.4419e+09 30756
## - mod_categoryLow            1 2.6522e+10 3.3964e+10 31499
## - mod_categoryMedium         1 2.7114e+10 3.4556e+10 31524
##
## Step: AIC=29546.84
## PriceNew ~ Horsepower + mod_categoryLow + mod_categoryLuxury +
##   mod_categoryMedium
##
##              Df Sum of Sq      RSS   AIC
## - mod_categoryLuxury         1 1.9559e+09 1.0654e+10 28404
## - Horsepower                 1 4.4637e+09 1.3162e+10 28707
## <none>                       8.6983e+09 29547
## - mod_categoryMedium         1 2.6403e+10 3.5101e+10 30113
## - mod_categoryLow            1 2.7566e+10 3.6264e+10 30160
##
## Step: AIC=28404.49
## PriceNew ~ Horsepower + mod_categoryLow + mod_categoryMedium
##
##              Df Sum of Sq      RSS   AIC
## - Horsepower                 1 7.8021e+09 1.8456e+10 27759
## <none>                       1.0654e+10 28404
## - mod_categoryLow            1 2.6929e+10 3.7583e+10 28778
## - mod_categoryMedium         1 2.8607e+10 3.9261e+10 28840
##
## Step: AIC=27758.86
## PriceNew ~ mod_categoryLow + mod_categoryMedium
##
##              Df Sum of Sq      RSS   AIC
## <none>                       1.8456e+10 27759
## - mod_categoryMedium         1 6.0670e+10 7.9126e+10 28412
## - mod_categoryLow            1 1.0677e+11 1.2523e+11 29070

```

```
forwardBIC <- step(full_model, direction='forward', k=length(cars.train.new$PriceNew))
```

```

## Start: AIC=70104.38
## PriceNew ~ TypeLarge + TypeMidsize + TypeSmall + TypeSporty +
##   TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
##   DriveTrainRear + EngineSize + Horsepower + RPM + Rev.per.mile +
##   Man.trans.availYes + Fuel.tank.capacity + Passengers + Length +
##   Wheelbase + Width + Turn.circle + Rear.seat.room + Luggage.room +
##   Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##   Cylinders_Origin4.non.USA + Cylinders_Origin4.USA + Cylinders_Origin5.non.USA +
##   Cylinders_Origin6.non.USA + Cylinders_Origin6.USA + Cylinders_Origin8.non.USA +
##   Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA

```

```
summary(backAIC)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeLarge + TypeMidsize + TypeSporty +
##     TypeVan + MPG.highway + AirBagsDriver.only + AirBagsNone +
##     EngineSize + Horsepower + RPM + Rev.per.mile + Man.trans.availYes +
##     Fuel.tank.capacity + Passengers + Length + Wheelbase + Width +
##     Rear.seat.room + Luggage.room + Weight + mod_categoryLow +
##     mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin4.non.USA +
##     Cylinders_Origin4.USA + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##     Cylinders_Origin6.USA + Cylinders_Origin8.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6240.0 -1214.8    -8.6   1162.5   5524.5
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.704e+04  3.217e+03  11.514 < 2e-16 ***
## TypeLarge         1.587e+03  3.969e+02   3.999 6.70e-05 ***
## TypeMidsize       1.607e+03  2.547e+02   6.309 3.77e-10 ***
## TypeSporty        1.299e+03  2.600e+02   4.996 6.57e-07 ***
## TypeVan          -2.003e+03  5.150e+02  -3.889 0.000106 ***
## MPG.highway       1.479e+02  2.290e+01   6.460 1.44e-10 ***
## AirBagsDriver.only -2.794e+02  1.636e+02  -1.708 0.087789 .
## AirBagsNone       -1.670e+03  1.910e+02  -8.746 < 2e-16 ***
## EngineSize        6.909e+02  1.985e+02   3.480 0.000517 ***
## Horsepower        2.955e+01  3.907e+00   7.563 7.11e-14 ***
## RPM               7.849e-01  1.938e-01   4.051 5.38e-05 ***
## Rev.per.mile      3.483e-01  2.060e-01   1.691 0.091063 .
## Man.trans.availYes 3.970e+02  2.101e+02   1.890 0.058979 .
## Fuel.tank.capacity 3.025e+02  4.254e+01   7.112 1.82e-12 ***
## Passengers        7.149e+02  1.348e+02   5.303 1.32e-07 ***
## Length           -5.998e+01  1.136e+01  -5.280 1.49e-07 ***
## Wheelbase        -3.989e+01  2.610e+01  -1.529 0.126595
## Width            -4.119e+02  4.983e+01  -8.266 3.18e-16 ***
## Rear.seat.room    -7.307e+01  3.292e+01  -2.220 0.026611 *
## Luggage.room       1.802e+02  2.851e+01   6.320 3.51e-10 ***
## Weight            4.184e+00  5.194e-01   8.056 1.67e-15 ***
## mod_categoryLow    -1.588e+04  2.954e+02 -53.766 < 2e-16 ***
## mod_categoryLuxury  9.554e+03  6.485e+02  14.733 < 2e-16 ***
## mod_categoryMedium -1.236e+04  2.048e+02 -60.330 < 2e-16 ***
## Cylinders_Origin4.non.USA 8.511e+02  3.225e+02   2.640 0.008392 **
## Cylinders_Origin4.USA  5.159e+02  3.623e+02   1.424 0.154754
## Cylinders_Origin5.non.USA 4.405e+03  4.872e+02   9.041 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.526e+03  4.328e+02   5.837 6.60e-09 ***
## Cylinders_Origin6.USA  1.063e+03  4.840e+02   2.197 0.028208 *
## Cylinders_Origin8.USA  2.136e+03  6.308e+02   3.387 0.000727 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1758 on 1403 degrees of freedom
## Multiple R-squared:  0.9657, Adjusted R-squared:  0.965
```

```
## F-statistic: 1363 on 29 and 1403 DF, p-value: < 2.2e-16
```

```
summary(forwardAIC)
```

```
##
```

```
## Call:
```

```
## lm(formula = PriceNew ~ TypeLarge + TypeMidsize + TypeSmall +  
##     TypeSporty + TypeVan + MPG.highway + AirBagsDriver.only +  
##     AirBagsNone + DriveTrainRear + EngineSize + Horsepower +  
##     RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +  
##     Passengers + Length + Wheelbase + Width + Turn.circle + Rear.seat.room +  
##     Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +  
##     mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +  
##     Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +  
##     Cylinders_Origin8.non.USA + Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA,  
##     data = cars.train.new)
```

```
##
```

```
## Residuals:
```

```
##      Min       1Q   Median       3Q      Max  
## -6079.5 -1190.8    -2.5   1165.7   5692.9
```

```
##
```

```
## Coefficients: (1 not defined because of singularities)
```

	Estimate	Std. Error	t value	Pr(> t)	
## (Intercept)	3.564e+04	3.424e+03	10.409	< 2e-16	***
## TypeLarge	1.544e+03	3.993e+02	3.868	0.000115	***
## TypeMidsize	1.509e+03	2.686e+02	5.618	2.32e-08	***
## TypeSmall	-1.649e+02	2.644e+02	-0.624	0.532977	
## TypeSporty	1.144e+03	2.917e+02	3.922	9.20e-05	***
## TypeVan	-2.191e+03	5.461e+02	-4.012	6.34e-05	***
## MPG.highway	1.421e+02	2.580e+01	5.508	4.32e-08	***
## AirBagsDriver.only	-2.530e+02	1.666e+02	-1.519	0.129074	
## AirBagsNone	-1.604e+03	2.025e+02	-7.923	4.71e-15	***
## DriveTrainRear	-1.787e+02	2.239e+02	-0.798	0.424921	
## EngineSize	6.753e+02	2.549e+02	2.649	0.008156	**
## Horsepower	2.776e+01	4.690e+00	5.920	4.05e-09	***
## RPM	8.202e-01	2.218e-01	3.698	0.000226	***
## Rev.per.mile	3.351e-01	2.232e-01	1.502	0.133445	
## Man.trans.availYes	4.255e+02	2.282e+02	1.864	0.062516	.
## Fuel.tank.capacity	2.996e+02	4.677e+01	6.404	2.06e-10	***
## Passengers	6.732e+02	1.556e+02	4.327	1.62e-05	***
## Length	-6.298e+01	1.243e+01	-5.065	4.62e-07	***
## Wheelbase	-4.136e+01	2.716e+01	-1.523	0.127922	
## Width	-3.813e+02	5.827e+01	-6.543	8.41e-11	***
## Turn.circle	2.824e+01	3.760e+01	0.751	0.452712	
## Rear.seat.room	-8.483e+01	3.499e+01	-2.424	0.015465	*
## Luggage.room	1.802e+02	2.951e+01	6.106	1.32e-09	***
## Weight	4.197e+00	5.598e-01	7.497	1.16e-13	***
## mod_categoryLow	-1.592e+04	3.087e+02	-51.572	< 2e-16	***
## mod_categoryLuxury	9.664e+03	7.562e+02	12.780	< 2e-16	***
## mod_categoryMedium	-1.247e+04	2.242e+02	-55.616	< 2e-16	***
## Cylinders_Origin4.non.USA	7.606e+02	4.067e+02	1.870	0.061689	.
## Cylinders_Origin4.USA	3.556e+02	4.506e+02	0.789	0.430178	
## Cylinders_Origin5.non.USA	4.411e+03	5.512e+02	8.003	2.54e-15	***
## Cylinders_Origin6.non.USA	2.475e+03	5.330e+02	4.644	3.73e-06	***
## Cylinders_Origin6.USA	9.519e+02	5.855e+02	1.626	0.104229	

```
## Cylinders_Origin8.non.USA      NA      NA      NA      NA
## Cylinders_Origin8.USA          2.054e+03  7.436e+02  2.763 0.005810 **
## Cylinders_Originrotary.non.USA -1.187e+02  1.116e+03 -0.106 0.915308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1759 on 1399 degrees of freedom
## Multiple R-squared:  0.9658, Adjusted R-squared:  0.965
## F-statistic: 1196 on 33 and 1399 DF, p-value: < 2.2e-16
```

```
summary(backBIC)
```

```
##
## Call:
## lm(formula = PriceNew ~ mod_categoryLow + mod_categoryMedium,
##     data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9918.9 -2294.1  -267.7   1640.4  13675.3
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    37805.6     223.7   169.03  <2e-16 ***
## mod_categoryLow -26195.6     288.0   -90.95  <2e-16 ***
## mod_categoryMedium -17681.6     257.9   -68.56  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3593 on 1430 degrees of freedom
## Multiple R-squared:  0.8541, Adjusted R-squared:  0.8539
## F-statistic: 4185 on 2 and 1430 DF, p-value: < 2.2e-16
```

```
summary(forwardBIC)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeLarge + TypeMidsize + TypeSmall +
##     TypeSporty + TypeVan + MPG.highway + AirBagsDriver.only +
##     AirBagsNone + DriveTrainRear + EngineSize + Horsepower +
##     RPM + Rev.per.mile + Man.trans.availYes + Fuel.tank.capacity +
##     Passengers + Length + Wheelbase + Width + Turn.circle + Rear.seat.room +
##     Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##     mod_categoryMedium + Cylinders_Origin4.non.USA + Cylinders_Origin4.USA +
##     Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin6.USA +
##     Cylinders_Origin8.non.USA + Cylinders_Origin8.USA + Cylinders_Originrotary.non.USA,
##     data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6079.5 -1190.8    -2.5   1165.7   5692.9
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.564e+04  3.424e+03  10.409  < 2e-16 ***
```

```
## TypeLarge          1.544e+03  3.993e+02   3.868 0.000115 ***
## TypeMidsize        1.509e+03  2.686e+02   5.618 2.32e-08 ***
## TypeSmall          -1.649e+02  2.644e+02  -0.624 0.532977
## TypeSporty         1.144e+03  2.917e+02   3.922 9.20e-05 ***
## TypeVan            -2.191e+03  5.461e+02  -4.012 6.34e-05 ***
## MPG.highway        1.421e+02  2.580e+01   5.508 4.32e-08 ***
## AirBagsDriver.only -2.530e+02  1.666e+02  -1.519 0.129074
## AirBagsNone        -1.604e+03  2.025e+02  -7.923 4.71e-15 ***
## DriveTrainRear     -1.787e+02  2.239e+02  -0.798 0.424921
## EngineSize         6.753e+02  2.549e+02   2.649 0.008156 **
## Horsepower         2.776e+01  4.690e+00   5.920 4.05e-09 ***
## RPM               8.202e-01  2.218e-01   3.698 0.000226 ***
## Rev.per.mile       3.351e-01  2.232e-01   1.502 0.133445
## Man.trans.availYes  4.255e+02  2.282e+02   1.864 0.062516 .
## Fuel.tank.capacity  2.996e+02  4.677e+01   6.404 2.06e-10 ***
## Passengers         6.732e+02  1.556e+02   4.327 1.62e-05 ***
## Length            -6.298e+01  1.243e+01  -5.065 4.62e-07 ***
## Wheelbase         -4.136e+01  2.716e+01  -1.523 0.127922
## Width             -3.813e+02  5.827e+01  -6.543 8.41e-11 ***
## Turn.circle        2.824e+01  3.760e+01   0.751 0.452712
## Rear.seat.room     -8.483e+01  3.499e+01  -2.424 0.015465 *
## Luggage.room       1.802e+02  2.951e+01   6.106 1.32e-09 ***
## Weight            4.197e+00  5.598e-01   7.497 1.16e-13 ***
## mod_categoryLow    -1.592e+04  3.087e+02 -51.572 < 2e-16 ***
## mod_categoryLuxury  9.664e+03  7.562e+02  12.780 < 2e-16 ***
## mod_categoryMedium -1.247e+04  2.242e+02 -55.616 < 2e-16 ***
## Cylinders_Origin4.non.USA 7.606e+02  4.067e+02   1.870 0.061689 .
## Cylinders_Origin4.USA   3.556e+02  4.506e+02   0.789 0.430178
## Cylinders_Origin5.non.USA 4.411e+03  5.512e+02   8.003 2.54e-15 ***
## Cylinders_Origin6.non.USA 2.475e+03  5.330e+02   4.644 3.73e-06 ***
## Cylinders_Origin6.USA   9.519e+02  5.855e+02   1.626 0.104229
## Cylinders_Origin8.non.USA NA NA NA NA
## Cylinders_Origin8.USA   2.054e+03  7.436e+02   2.763 0.005810 **
## Cylinders_Originrotary.non.USA -1.187e+02  1.116e+03  -0.106 0.915308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1759 on 1399 degrees of freedom
## Multiple R-squared:  0.9658, Adjusted R-squared:  0.965
## F-statistic: 1196 on 33 and 1399 DF, p-value: < 2.2e-16
```

From our analysis it appears that having more variables is optimal. However, our goal is to optimize a model with the least amount of predictors so let's take out the variables which have low p-values and update our full model.

```
prediction_model <- update(full_model, ~. - TypeSmall - AirBagsDriver.only - DriveTrainRear - Rev.per.mile)
sort(vif(prediction_model),decreasing =T)
```

```
##          Weight          Horsepower          EngineSize
##      34.761422          16.202453          15.688418
##          Width          Length          Passengers
##     13.704057          12.026198          7.620457
##          TypeVan      Fuel.tank.capacity      mod_categoryLow
##          7.544289          7.401176          6.768036
##          RPM          MPG.highway          Rear.seat.room
```

```
##           5.930751           4.975065           4.439148
##           TypeLarge           Luggage.room           mod_categoryMedium
##           4.162307           3.986524           3.983024
##           TypeSporty           TypeMidsize           Cylinders_Origin8.USA
##           3.749284           3.214398           2.265200
##           mod_categoryLuxury Cylinders_Origin6.non.USA           AirBagsNone
##           2.097058           2.004839           2.000805
## Cylinders_Origin5.non.USA
##           1.883240
```

```
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeLarge + TypeMidsize + TypeSporty +
##     TypeVan + MPG.highway + AirBagsNone + EngineSize + Horsepower +
##     RPM + Fuel.tank.capacity + Passengers + Length + Width +
##     Rear.seat.room + Luggage.room + Weight + mod_categoryLow +
##     mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##     Cylinders_Origin6.non.USA + Cylinders_Origin8.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6184.9 -1178.7   -34.8   1154.5   5617.7
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.282e+04  2.708e+03  12.119 < 2e-16 ***
## TypeLarge       9.884e+02  2.896e+02   3.413 0.000660 ***
## TypeMidsize     1.267e+03  2.029e+02   6.243 5.67e-10 ***
## TypeSporty      1.352e+03  2.457e+02   5.505 4.39e-08 ***
## TypeVan        -2.683e+03  4.388e+02  -6.116 1.24e-09 ***
## MPG.highway     1.398e+02  1.912e+01   7.310 4.46e-13 ***
## AirBagsNone    -1.413e+03  1.352e+02 -10.447 < 2e-16 ***
## EngineSize      6.641e+02  1.760e+02   3.774 0.000168 ***
## Horsepower      2.920e+01  3.711e+00   7.869 7.09e-15 ***
## RPM             8.924e-01  1.851e-01   4.822 1.58e-06 ***
## Fuel.tank.capacity 3.014e+02  3.870e+01   7.789 1.31e-14 ***
## Passengers      6.945e+02  1.271e+02   5.463 5.54e-08 ***
## Length         -7.015e+01  1.076e+01  -6.522 9.62e-11 ***
## Width          -3.558e+02  4.586e+01  -7.759 1.63e-14 ***
## Rear.seat.room  -6.058e+01  3.135e+01  -1.932 0.053511 .
## Luggage.room     1.597e+02  2.757e+01   5.790 8.67e-09 ***
## Weight          4.147e+00  4.581e-01   9.053 < 2e-16 ***
## mod_categoryLow -1.591e+04  2.732e+02 -58.222 < 2e-16 ***
## mod_categoryLuxury 7.975e+03  4.455e+02  17.902 < 2e-16 ***
## mod_categoryMedium -1.257e+04  1.877e+02 -66.952 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.014e+03  3.287e+02  12.211 < 2e-16 ***
## Cylinders_Origin6.non.USA 1.799e+03  2.172e+02   8.283 2.76e-16 ***
## Cylinders_Origin8.USA  9.226e+02  3.034e+02   3.041 0.002401 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1772 on 1410 degrees of freedom
## Multiple R-squared:  0.965, Adjusted R-squared:  0.9644
```



```
## F-statistic: 1766 on 22 and 1410 DF, p-value: < 2.2e-16
```

We will proceed by removing points based on their significance and relative VIF score. The procedure is as follows: Delete predictors with high P-values/high VIFs until all VIF scores are under 5 and multicollinearity is dealt with.

```
prediction_model <- update(prediction_model, ~. -Rear.seat.room)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeLarge + TypeMidsize + TypeSporty +
##     TypeVan + MPG.highway + AirBagsNone + EngineSize + Horsepower +
##     RPM + Fuel.tank.capacity + Passengers + Length + Width +
##     Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##     mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA +
##     Cylinders_Origin8.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6083.6 -1178.5    4.5   1151.2   5514.6
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.083e+04  2.508e+03  12.296 < 2e-16 ***
## TypeLarge       8.315e+02  2.782e+02   2.988 0.002853 **
## TypeMidsize     1.240e+03  2.026e+02   6.118 1.23e-09 ***
## TypeSporty      1.484e+03  2.363e+02   6.278 4.55e-10 ***
## TypeVan        -2.744e+03  4.381e+02  -6.264 4.99e-10 ***
## MPG.highway     1.363e+02  1.905e+01   7.152 1.37e-12 ***
## AirBagsNone    -1.435e+03  1.349e+02 -10.643 < 2e-16 ***
## EngineSize      6.480e+02  1.760e+02   3.683 0.000239 ***
## Horsepower      2.890e+01  3.711e+00   7.788 1.31e-14 ***
## RPM             9.093e-01  1.850e-01   4.914 9.97e-07 ***
## Fuel.tank.capacity 2.954e+02  3.862e+01   7.651 3.68e-14 ***
## Passengers      6.377e+02  1.238e+02   5.150 2.97e-07 ***
## Length         -7.164e+01  1.074e+01  -6.671 3.63e-11 ***
## Width          -3.408e+02  4.524e+01  -7.533 8.79e-14 ***
## Luggage.room     1.476e+02  2.688e+01   5.489 4.77e-08 ***
## Weight          4.191e+00  4.580e-01   9.150 < 2e-16 ***
## mod_categoryLow -1.577e+04  2.639e+02 -59.757 < 2e-16 ***
## mod_categoryLuxury 8.008e+03  4.456e+02  17.973 < 2e-16 ***
## mod_categoryMedium -1.249e+04  1.835e+02 -68.068 < 2e-16 ***
## Cylinders_Origin5.non.USA 3.824e+03  3.140e+02  12.179 < 2e-16 ***
## Cylinders_Origin6.non.USA 1.834e+03  2.166e+02   8.465 < 2e-16 ***
## Cylinders_Origin8.USA 9.403e+02  3.035e+02   3.098 0.001987 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1774 on 1411 degrees of freedom
## Multiple R-squared:  0.9649, Adjusted R-squared:  0.9644
## F-statistic: 1847 on 21 and 1411 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##              Weight              Horsepower              EngineSize
```



```
##           34.678290           16.175041           15.653341
##           Width           Length           TypeVan
##           13.311853           11.964633           7.506053
##           Fuel.tank.capacity           Passengers           mod_categoryLow
##           7.353409           7.213195           6.301041
##           RPM           MPG.highway           TypeLarge
##           5.917532           4.930112           3.834826
##           mod_categoryMedium           Luggage.room           TypeSporty
##           3.798378           3.781615           3.462371
##           TypeMidsize           Cylinders_Origin8.USA           mod_categoryLuxury
##           3.198891           2.263124           2.093904
## Cylinders_Origin6.non.USA           AirBagsNone Cylinders_Origin5.non.USA
##           1.990953           1.985918           1.714771
```

```
prediction_model <- update(prediction_model, ~. -TypeLarge)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeMidsize + TypeSporty + TypeVan +
##     MPG.highway + AirBagsNone + EngineSize + Horsepower + RPM +
##     Fuel.tank.capacity + Passengers + Length + Width + Luggage.room +
##     Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##     Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA + Cylinders_Origin8.USA,
##     data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5967.2 -1143.5   44.1  1179.1  5367.5
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.783e+04  2.303e+03  12.081 < 2e-16 ***
## TypeMidsize     9.254e+02  1.737e+02   5.328 1.16e-07 ***
## TypeSporty     1.395e+03  2.351e+02   5.934 3.71e-09 ***
## TypeVan       -3.345e+03  3.902e+02  -8.572 < 2e-16 ***
## MPG.highway    1.404e+02  1.905e+01   7.366 2.97e-13 ***
## AirBagsNone   -1.446e+03  1.352e+02 -10.698 < 2e-16 ***
## EngineSize     7.526e+02  1.729e+02   4.352 1.44e-05 ***
## Horsepower     2.718e+01  3.676e+00   7.393 2.45e-13 ***
## RPM           9.889e-01  1.836e-01   5.385 8.48e-08 ***
## Fuel.tank.capacity 2.891e+02  3.866e+01   7.476 1.34e-13 ***
## Passengers     7.300e+02  1.202e+02   6.071 1.63e-09 ***
## Length        -7.300e+01  1.076e+01  -6.785 1.70e-11 ***
## Width         -3.098e+02  4.416e+01  -7.016 3.53e-12 ***
## Luggage.room    1.469e+02  2.696e+01   5.451 5.90e-08 ***
## Weight         4.331e+00  4.569e-01   9.479 < 2e-16 ***
## mod_categoryLow -1.580e+04  2.644e+02 -59.767 < 2e-16 ***
## mod_categoryLuxury 7.943e+03  4.463e+02  17.799 < 2e-16 ***
## mod_categoryMedium -1.251e+04  1.838e+02 -68.044 < 2e-16 ***
## Cylinders_Origin5.non.USA 3.920e+03  3.132e+02  12.515 < 2e-16 ***
## Cylinders_Origin6.non.USA 1.889e+03  2.165e+02   8.726 < 2e-16 ***
## Cylinders_Origin8.USA 1.131e+03  2.976e+02   3.799 0.000152 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 1779 on 1412 degrees of freedom
## Multiple R-squared:  0.9647, Adjusted R-squared:  0.9642
## F-statistic: 1928 on 20 and 1412 DF,  p-value: < 2.2e-16

sort(vif(prediction_model),decreasing =T)

##           Weight           Horsepower           EngineSize
##      34.314089         15.784172         15.033590
##           Width           Length      Fuel.tank.capacity
##      12.613324         11.943114         7.330980
##      Passengers      mod_categoryLow           TypeVan
##      6.764666         6.289969         5.922649
##           RPM           MPG.highway      mod_categoryMedium
##      5.795075         4.904338         3.792675
##      Luggage.room           TypeSporty           TypeMidsize
##      3.781383         3.408116         2.337710
##      Cylinders_Origin8.USA      mod_categoryLuxury      AirBagsNone
##      2.163553         2.088926         1.984454
##      Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
##      1.976583         1.696934

prediction_model <- update(prediction_model,~. -Cylinders_Origin8.USA)
summary(prediction_model)

##
## Call:
## lm(formula = PriceNew ~ TypeMidsize + TypeSporty + TypeVan +
##      MPG.highway + AirBagsNone + EngineSize + Horsepower + RPM +
##      Fuel.tank.capacity + Passengers + Length + Width + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6102.8 -1145.9    26.6  1175.1  5353.4
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.816e+04  2.313e+03  12.176 < 2e-16 ***
## TypeMidsize     7.721e+02  1.698e+02   4.548 5.87e-06 ***
## TypeSporty     1.381e+03  2.362e+02   5.847 6.23e-09 ***
## TypeVan       -3.307e+03  3.919e+02  -8.438 < 2e-16 ***
## MPG.highway     1.402e+02  1.915e+01   7.322 4.08e-13 ***
## AirBagsNone    -1.426e+03  1.357e+02 -10.506 < 2e-16 ***
## EngineSize      9.223e+02  1.678e+02   5.495 4.63e-08 ***
## Horsepower      2.885e+01  3.667e+00   7.868 7.12e-15 ***
## RPM            9.573e-01  1.843e-01   5.194 2.36e-07 ***
## Fuel.tank.capacity 2.753e+02  3.868e+01   7.117 1.75e-12 ***
## Passengers      7.072e+02  1.207e+02   5.861 5.71e-09 ***
## Length        -6.847e+01  1.074e+01  -6.373 2.50e-10 ***
## Width         -3.166e+02  4.433e+01  -7.142 1.47e-12 ***
## Luggage.room    1.624e+02  2.678e+01   6.064 1.70e-09 ***
## Weight         4.058e+00  4.533e-01   8.952 < 2e-16 ***
## mod_categoryLow -1.599e+04  2.609e+02 -61.298 < 2e-16 ***
```

```
## mod_categoryLuxury      7.595e+03  4.388e+02  17.307 < 2e-16 ***
## mod_categoryMedium     -1.271e+04  1.770e+02 -71.812 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.148e+03  3.089e+02  13.430 < 2e-16 ***
## Cylinders_Origin6.non.USA 1.862e+03  2.174e+02   8.567 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1787 on 1413 degrees of freedom
## Multiple R-squared:  0.9643, Adjusted R-squared:  0.9638
## F-statistic: 2009 on 19 and 1413 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##              Weight      Horsepower      EngineSize
##      33.466382      15.557528      14.030504
##              Width      Length      Fuel.tank.capacity
##      12.592608      11.796346      7.266250
##      Passengers      mod_categoryLow      TypeVan
##      6.747905      6.066265      5.918736
##              RPM      MPG.highway      Luggage.room
##      5.783248      4.904308      3.695621
##      mod_categoryMedium      TypeSporty      TypeMidsize
##      3.481808      3.407240      2.211547
##      mod_categoryLuxury      AirBagsNone Cylinders_Origin6.non.USA
##      2.000873      1.981312      1.974513
## Cylinders_Origin5.non.USA
##      1.634498
```

```
prediction_model <- update(prediction_model,~. -TypeMidsize)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeSporty + TypeVan + MPG.highway +
##      AirBagsNone + EngineSize + Horsepower + RPM + Fuel.tank.capacity +
##      Passengers + Length + Width + Luggage.room + Weight + mod_categoryLow +
##      mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6358.6 -1167.7    53.2  1187.7  5244.8
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.741e+04  2.323e+03  11.800 < 2e-16 ***
## TypeSporty        1.025e+03  2.244e+02   4.569 5.34e-06 ***
## TypeVan          -3.713e+03  3.843e+02  -9.662 < 2e-16 ***
## MPG.highway       1.295e+02  1.913e+01   6.771 1.87e-11 ***
## AirBagsNone      -1.402e+03  1.366e+02 -10.267 < 2e-16 ***
## EngineSize        7.371e+02  1.640e+02   4.496 7.50e-06 ***
## Horsepower        2.863e+01  3.692e+00   7.754 1.70e-14 ***
## RPM               8.745e-01  1.847e-01   4.735 2.41e-06 ***
## Fuel.tank.capacity 2.787e+02  3.894e+01   7.158 1.31e-12 ***
## Passengers        6.207e+02  1.200e+02   5.174 2.63e-07 ***
```

```
## Length -6.622e+01 1.081e+01 -6.127 1.16e-09 ***
## Width -2.849e+02 4.408e+01 -6.463 1.41e-10 ***
## Luggage.room 1.699e+02 2.691e+01 6.314 3.63e-10 ***
## Weight 4.046e+00 4.565e-01 8.864 < 2e-16 ***
## mod_categoryLow -1.624e+04 2.567e+02 -63.269 < 2e-16 ***
## mod_categoryLuxury 8.301e+03 4.133e+02 20.083 < 2e-16 ***
## mod_categoryMedium -1.275e+04 1.780e+02 -71.605 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.335e+03 3.083e+02 14.064 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.264e+03 2.000e+02 11.317 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1800 on 1414 degrees of freedom
## Multiple R-squared: 0.9638, Adjusted R-squared: 0.9633
## F-statistic: 2091 on 18 and 1414 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
## Weight Horsepower EngineSize
## 33.465301 15.554716 13.204785
## Width Length Fuel.tank.capacity
## 12.281117 11.771222 7.263464
## Passengers mod_categoryLow RPM
## 6.580249 5.793922 5.726761
## TypeVan MPG.highway Luggage.room
## 5.611917 4.831137 3.681437
## mod_categoryMedium TypeSporty AirBagsNone
## 3.474242 3.033989 1.978368
## mod_categoryLuxury Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
## 1.750657 1.648996 1.605555
```

```
prediction_model <- update(prediction_model,.-EngineSize)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeSporty + TypeVan + MPG.highway +
## AirBagsNone + Horsepower + RPM + Fuel.tank.capacity + Passengers +
## Length + Width + Luggage.room + Weight + mod_categoryLow +
## mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
## Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
## Min 1Q Median 3Q Max
## -6212.2 -1208.3 24.2 1232.1 5248.8
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.672e+04 2.333e+03 11.452 < 2e-16 ***
## TypeSporty 9.004e+02 2.242e+02 4.016 6.23e-05 ***
## TypeVan -3.964e+03 3.828e+02 -10.357 < 2e-16 ***
## MPG.highway 1.349e+02 1.923e+01 7.016 3.53e-12 ***
## AirBagsNone -1.377e+03 1.374e+02 -10.023 < 2e-16 ***
## Horsepower 3.797e+01 3.072e+00 12.360 < 2e-16 ***
## RPM 3.234e-01 1.391e-01 2.325 0.0202 *
```

```
## Fuel.tank.capacity      2.786e+02  3.920e+01  7.107 1.88e-12 ***
## Passengers              6.815e+02  1.200e+02  5.679 1.64e-08 ***
## Length                 -6.412e+01  1.087e+01 -5.899 4.56e-09 ***
## Width                  -2.346e+02  4.293e+01 -5.465 5.45e-08 ***
## Luggage.room           1.788e+02  2.702e+01  6.617 5.17e-11 ***
## Weight                  3.972e+00  4.593e-01  8.648 < 2e-16 ***
## mod_categoryLow        -1.620e+04  2.583e+02 -62.722 < 2e-16 ***
## mod_categoryLuxury      8.721e+03  4.054e+02  21.514 < 2e-16 ***
## mod_categoryMedium     -1.273e+04  1.792e+02 -71.041 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.228e+03  3.094e+02  13.664 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.232e+03  2.013e+02  11.092 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1812 on 1415 degrees of freedom
## Multiple R-squared:  0.9633, Adjusted R-squared:  0.9628
## F-statistic: 2183 on 17 and 1415 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##              Weight              Length              Width
##          33.420870              11.749381              11.490736
##          Horsepower      Fuel.tank.capacity      Passengers
##          10.625313              7.263460              6.496564
##          mod_categoryLow      TypeVan      MPG.highway
##          5.785930              5.493106              4.812541
##          Luggage.room      mod_categoryMedium      RPM
##          3.661571              3.472477              3.203963
##          TypeSporty      AirBagsNone      mod_categoryLuxury
##          2.987472              1.975021              1.661278
## Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
##          1.646983              1.595905
```

```
prediction_model <- update(prediction_model,~. -RPM)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeSporty + TypeVan + MPG.highway +
##      AirBagsNone + Horsepower + Fuel.tank.capacity + Passengers +
##      Length + Width + Luggage.room + Weight + mod_categoryLow +
##      mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6199.5 -1232.9    46.1  1222.3  5252.7
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.002e+04  1.855e+03  16.180 < 2e-16 ***
## TypeSporty      8.471e+02  2.234e+02   3.792 0.000156 ***
## TypeVan        -3.784e+03  3.754e+02 -10.079 < 2e-16 ***
## MPG.highway     1.409e+02  1.908e+01   7.385 2.59e-13 ***
## AirBagsNone    -1.400e+03  1.372e+02 -10.208 < 2e-16 ***
```

```
## Horsepower          4.230e+01  2.447e+00  17.290 < 2e-16 ***
## Fuel.tank.capacity  2.809e+02  3.925e+01   7.157 1.32e-12 ***
## Passengers          7.134e+02  1.194e+02   5.974 2.93e-09 ***
## Length              -5.860e+01  1.062e+01  -5.516 4.11e-08 ***
## Width               -2.680e+02  4.052e+01  -6.615 5.24e-11 ***
## Luggage.room        1.650e+02  2.640e+01   6.250 5.40e-10 ***
## Weight              3.585e+00  4.288e-01   8.361 < 2e-16 ***
## mod_categoryLow     -1.616e+04  2.580e+02 -62.623 < 2e-16 ***
## mod_categoryLuxury   8.869e+03  4.010e+02  22.119 < 2e-16 ***
## mod_categoryMedium  -1.261e+04  1.716e+02 -73.448 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.414e+03  2.993e+02  14.748 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.342e+03  1.960e+02  11.949 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1815 on 1416 degrees of freedom
## Multiple R-squared:  0.9631, Adjusted R-squared:  0.9627
## F-statistic: 2312 on 16 and 1416 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##              Weight          Length          Width
##          29.038390          11.188794          10.203935
##      Fuel.tank.capacity      Horsepower      Passengers
##          7.258824          6.718523          6.412006
##      mod_categoryLow          TypeVan      MPG.highway
##          5.754737          5.268253          4.725249
##          Luggage.room      mod_categoryMedium      TypeSporty
##          3.485928          3.176986          2.956185
##          AirBagsNone      mod_categoryLuxury Cylinders_Origin6.non.USA
##          1.964135          1.620358          1.556901
## Cylinders_Origin5.non.USA
##          1.488800
```

```
prediction_model <- update(prediction_model,`-`-TypeSporty)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeVan + MPG.highway + AirBagsNone +
##      Horsepower + Fuel.tank.capacity + Passengers + Length + Width +
##      Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA,
##      data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6152.4 -1240.6    21.4  1251.5  5071.4
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.993e+04  1.864e+03  16.057 < 2e-16 ***
## TypeVan        -3.497e+03  3.694e+02  -9.465 < 2e-16 ***
## MPG.highway     1.195e+02  1.831e+01   6.525 9.46e-11 ***
## AirBagsNone    -1.464e+03  1.368e+02 -10.700 < 2e-16 ***
```

```
## Horsepower          4.069e+01  2.421e+00  16.809 < 2e-16 ***
## Fuel.tank.capacity  2.748e+02  3.940e+01   6.974 4.71e-12 ***
## Passengers          4.388e+02  9.541e+01   4.599 4.62e-06 ***
## Length              -6.090e+01  1.066e+01  -5.715 1.34e-08 ***
## Width               -2.151e+02  3.821e+01  -5.628 2.19e-08 ***
## Luggage.room        1.472e+02  2.610e+01   5.638 2.07e-08 ***
## Weight              3.452e+00  4.293e-01   8.041 1.86e-15 ***
## mod_categoryLow     -1.628e+04  2.572e+02 -63.274 < 2e-16 ***
## mod_categoryLuxury   8.834e+03  4.027e+02  21.935 < 2e-16 ***
## mod_categoryMedium  -1.255e+04  1.717e+02 -73.065 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.331e+03  2.999e+02  14.440 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.259e+03  1.957e+02  11.545 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1823 on 1417 degrees of freedom
## Multiple R-squared:  0.9628, Adjusted R-squared:  0.9624
## F-statistic: 2442 on 15 and 1417 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##              Weight          Length          Width
##          28.845198          11.152394          8.990476
##      Fuel.tank.capacity      Horsepower      mod_categoryLow
##          7.246585          6.515653          5.667661
##          TypeVan          MPG.highway          Passengers
##          5.053445          4.309900          4.054195
##          Luggage.room      mod_categoryMedium      AirBagsNone
##          3.374948          3.150348          1.934885
##      mod_categoryLuxury Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
##          1.619521          1.537621          1.480718
```

```
prediction_model <- update(prediction_model,~. -Passengers)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeVan + MPG.highway + AirBagsNone +
##      Horsepower + Fuel.tank.capacity + Length + Width + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6111.3 -1252.6   -18.1  1261.3  4845.4
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.126e+04  1.855e+03  16.854 < 2e-16 ***
## TypeVan        -2.819e+03  3.412e+02  -8.263 3.23e-16 ***
## MPG.highway     1.108e+02  1.834e+01   6.040 1.97e-09 ***
## AirBagsNone    -1.458e+03  1.378e+02 -10.581 < 2e-16 ***
## Horsepower      3.719e+01  2.315e+00  16.069 < 2e-16 ***
## Fuel.tank.capacity 2.724e+02  3.968e+01   6.866 9.83e-12 ***
## Length         -4.525e+01  1.017e+01  -4.450 9.28e-06 ***
```



```
## Width -2.546e+02 3.750e+01 -6.790 1.65e-11 ***
## Luggage.room 1.723e+02 2.571e+01 6.701 2.98e-11 ***
## Weight 3.805e+00 4.254e-01 8.943 < 2e-16 ***
## mod_categoryLow -1.612e+04 2.568e+02 -62.773 < 2e-16 ***
## mod_categoryLuxury 8.848e+03 4.056e+02 21.815 < 2e-16 ***
## mod_categoryMedium -1.250e+04 1.727e+02 -72.410 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.147e+03 2.993e+02 13.854 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.243e+03 1.970e+02 11.384 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1836 on 1418 degrees of freedom
## Multiple R-squared: 0.9622, Adjusted R-squared: 0.9618
## F-statistic: 2578 on 14 and 1418 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
## Weight Length Width
## 27.926379 10.015269 8.535457
## Fuel.tank.capacity Horsepower mod_categoryLow
## 7.245360 5.873051 5.570133
## MPG.highway TypeVan Luggage.room
## 4.263955 4.250016 3.227501
## mod_categoryMedium AirBagsNone mod_categoryLuxury
## 3.140542 1.934698 1.619435
## Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
## 1.537134 1.454535
```

```
prediction_model <- update(prediction_model,~. -Length)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeVan + MPG.highway + AirBagsNone +
## Horsepower + Fuel.tank.capacity + Width + Luggage.room +
## Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
## Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
## Min 1Q Median 3Q Max
## -5852.3 -1254.9 -13.6 1300.9 5669.8
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.055e+04 1.860e+03 16.424 < 2e-16 ***
## TypeVan -1.776e+03 2.494e+02 -7.119 1.73e-12 ***
## MPG.highway 9.790e+01 1.823e+01 5.370 9.18e-08 ***
## AirBagsNone -1.564e+03 1.366e+02 -11.453 < 2e-16 ***
## Horsepower 4.181e+01 2.082e+00 20.080 < 2e-16 ***
## Fuel.tank.capacity 2.673e+02 3.992e+01 6.695 3.11e-11 ***
## Width -3.167e+02 3.503e+01 -9.039 < 2e-16 ***
## Luggage.room 1.551e+02 2.558e+01 6.061 1.73e-09 ***
## Weight 2.682e+00 3.447e-01 7.779 1.40e-14 ***
## mod_categoryLow -1.595e+04 2.554e+02 -62.422 < 2e-16 ***
## mod_categoryLuxury 8.843e+03 4.083e+02 21.660 < 2e-16 ***
```



```
## mod_categoryMedium      -1.233e+04  1.695e+02 -72.760 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.335e+03  2.983e+02  14.532 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.323e+03  1.975e+02  11.765 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1848 on 1419 degrees of freedom
## Multiple R-squared:  0.9617, Adjusted R-squared:  0.9613
## F-statistic: 2739 on 13 and 1419 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##              Weight              Width      Fuel.tank.capacity
##          18.097929          7.353567          7.239153
##      mod_categoryLow          Horsepower      MPG.highway
##          5.438883          4.691290          4.157949
##      Luggage.room      mod_categoryMedium      TypeVan
##          3.154459          2.986821          2.241391
##          AirBagsNone      mod_categoryLuxury Cylinders_Origin6.non.USA
##          1.876383          1.619425          1.524189
## Cylinders_Origin5.non.USA
##          1.425633
```

```
prediction_model <- update(prediction_model,~. -MPG.highway)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeVan + AirBagsNone + Horsepower +
##      Fuel.tank.capacity + Width + Luggage.room + Weight + mod_categoryLow +
##      mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6080.0 -1296.5    11.5  1292.7  5651.9
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.348e+04  1.795e+03  18.644 < 2e-16 ***
## TypeVan          -1.972e+03  2.491e+02  -7.916 4.89e-15 ***
## AirBagsNone      -1.628e+03  1.374e+02 -11.847 < 2e-16 ***
## Horsepower        4.182e+01  2.103e+00  19.892 < 2e-16 ***
## Fuel.tank.capacity  2.243e+02  3.949e+01   5.681 1.63e-08 ***
## Width            -2.816e+02  3.475e+01  -8.102 1.15e-15 ***
## Luggage.room       1.895e+02  2.501e+01   7.579 6.24e-14 ***
## Weight            1.910e+00  3.164e-01   6.037 2.01e-09 ***
## mod_categoryLow    -1.573e+04  2.547e+02 -61.751 < 2e-16 ***
## mod_categoryLuxury  9.073e+03  4.100e+02  22.129 < 2e-16 ***
## mod_categoryMedium -1.225e+04  1.705e+02 -71.867 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.456e+03  3.004e+02  14.836 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.369e+03  1.992e+02  11.889 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1866 on 1420 degrees of freedom
## Multiple R-squared:  0.9609, Adjusted R-squared:  0.9606
## F-statistic: 2908 on 12 and 1420 DF,  p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##           Weight           Width      Fuel.tank.capacity
##      14.955381      7.097681      6.949212
## mod_categoryLow      Horsepower mod_categoryMedium
##      5.304453      4.691285      2.963951
##      Luggage.room      TypeVan      AirBagsNone
##      2.955892      2.193082      1.862313
## mod_categoryLuxury Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
##      1.601705      1.521404      1.417492
```

```
prediction_model <- update(prediction_model, ~. -Fuel.tank.capacity)
summary(prediction_model)
```

```
##
## Call:
## lm(formula = PriceNew ~ TypeVan + AirBagsNone + Horsepower +
##      Width + Luggage.room + Weight + mod_categoryLow + mod_categoryLuxury +
##      mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA,
##      data = cars.train.new)
##
```

```
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6051.3 -1296.9   -25.4  1313.6  5512.6
##
```

```
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.324e+04  1.815e+03  18.319 < 2e-16 ***
## TypeVan        -1.782e+03  2.496e+02  -7.142 1.47e-12 ***
## AirBagsNone    -1.552e+03  1.382e+02 -11.228 < 2e-16 ***
## Horsepower      4.479e+01  2.059e+00  21.748 < 2e-16 ***
## Width          -2.610e+02  3.494e+01  -7.469 1.41e-13 ***
## Luggage.room    2.239e+02  2.453e+01   9.129 < 2e-16 ***
## Weight         2.410e+00  3.073e-01   7.845 8.46e-15 ***
## mod_categoryLow -1.586e+04  2.565e+02 -61.813 < 2e-16 ***
## mod_categoryLuxury 9.530e+03  4.064e+02  23.449 < 2e-16 ***
## mod_categoryMedium -1.224e+04  1.723e+02 -71.011 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.810e+03  2.970e+02  16.192 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.641e+03  1.955e+02  13.505 < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 1887 on 1421 degrees of freedom
## Multiple R-squared:  0.96, Adjusted R-squared:  0.9597
## F-statistic: 3101 on 11 and 1421 DF,  p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##           Weight           Width      mod_categoryLow
##      13.797225      7.020382      5.263708
##      Horsepower      mod_categoryMedium      Luggage.room
##      4.402976      2.963244      2.782616
```

```
##              TypeVan              AirBagsNone              mod_categoryLuxury
##              2.153659              1.844921              1.539962
## Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
##              1.433652              1.356568

prediction_model <- update(prediction_model, ~. -TypeVan)
summary(prediction_model)

##
## Call:
## lm(formula = PriceNew ~ AirBagsNone + Horsepower + Width + Luggage.room +
##      Weight + mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6143.7 -1348.1   -12.4  1299.0  5715.1
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.417e+04  1.842e+03  18.554 < 2e-16 ***
## AirBagsNone      -1.939e+03  1.294e+02 -14.985 < 2e-16 ***
## Horsepower       4.868e+01  2.020e+00  24.098 < 2e-16 ***
## Width           -2.407e+02  3.543e+01  -6.794 1.60e-11 ***
## Luggage.room     2.233e+02  2.495e+01   8.948 < 2e-16 ***
## Weight           1.525e+00  2.861e-01   5.332 1.13e-07 ***
## mod_categoryLow  -1.607e+04  2.592e+02 -61.994 < 2e-16 ***
## mod_categoryLuxury 9.614e+03  4.133e+02  23.260 < 2e-16 ***
## mod_categoryMedium -1.242e+04  1.734e+02 -71.639 < 2e-16 ***
## Cylinders_Origin5.non.USA 4.514e+03  2.993e+02  15.083 < 2e-16 ***
## Cylinders_Origin6.non.USA 2.589e+03  1.988e+02  13.022 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1920 on 1422 degrees of freedom
## Multiple R-squared:  0.9586, Adjusted R-squared:  0.9583
## F-statistic: 3290 on 10 and 1422 DF, p-value: < 2.2e-16

sort(vif(prediction_model),decreasing =T)

##              Weight              Width              mod_categoryLow
##              11.551317              6.974132              5.192205
##              Horsepower      mod_categoryMedium      Luggage.room
##              4.093695              2.897499              2.782580
##              AirBagsNone      mod_categoryLuxury Cylinders_Origin6.non.USA
##              1.561626              1.538676              1.431673
## Cylinders_Origin5.non.USA
##              1.330232

prediction_model <- update(prediction_model, ~. -Weight)
summary(prediction_model)

##
## Call:
## lm(formula = PriceNew ~ AirBagsNone + Horsepower + Width + Luggage.room +
##      mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
```

```
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -6638  -1342       12    1306   5535
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    29761.505    1661.487   17.913 < 2e-16 ***
## AirBagsNone    -1915.009     130.565  -14.667 < 2e-16 ***
## Horsepower         52.019        1.939   26.823 < 2e-16 ***
## Width          -127.820        28.682   -4.456 8.99e-06 ***
## Luggage.room      279.529        22.831   12.243 < 2e-16 ***
## mod_categoryLow  -16472.062    250.393  -65.785 < 2e-16 ***
## mod_categoryLuxury 10220.104    401.169   25.476 < 2e-16 ***
## mod_categoryMedium -12439.429    175.025  -71.072 < 2e-16 ***
## Cylinders_Origin5.non.USA  5182.969    274.321   18.894 < 2e-16 ***
## Cylinders_Origin6.non.USA  2965.695    187.577   15.811 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1938 on 1423 degrees of freedom
## Multiple R-squared:  0.9577, Adjusted R-squared:  0.9575
## F-statistic: 3583 on 9 and 1423 DF, p-value: < 2.2e-16
```

```
sort(vif(prediction_model),decreasing =T)
```

```
##      mod_categoryLow      Width      Horsepower
##      4.753049      4.483198      3.701008
##      mod_categoryMedium      Luggage.room      AirBagsNone
##      2.896443      2.285226      1.559720
##      mod_categoryLuxury Cylinders_Origin6.non.USA Cylinders_Origin5.non.USA
##      1.422142      1.250551      1.096490
```

This is our first model where all VIF values are under 5. Let's see if it is worth taking out more predictors.

```
anova(lm(formula = PriceNew ~ AirBagsNone + Horsepower + Luggage.room +
      mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new), prediction_model)
```

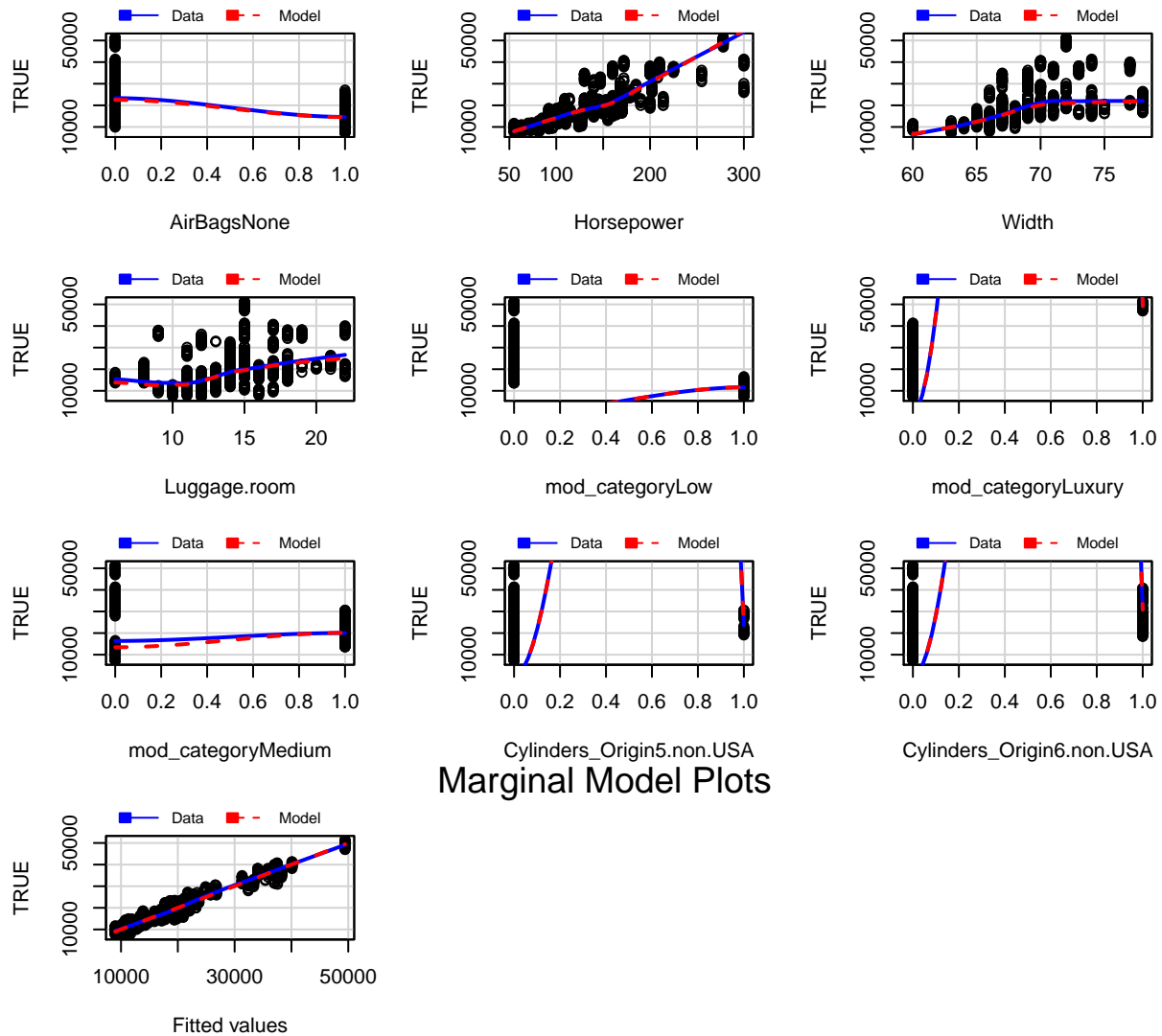
```
## Analysis of Variance Table
##
## Model 1: PriceNew ~ AirBagsNone + Horsepower + Luggage.room + mod_categoryLow +
##      mod_categoryLuxury + mod_categoryMedium + Cylinders_Origin5.non.USA +
##      Cylinders_Origin6.non.USA
## Model 2: PriceNew ~ AirBagsNone + Horsepower + Width + Luggage.room +
##      mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
##      Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA
##   Res.Df    RSS Df Sum of Sq    F    Pr(>F)
## 1    1424 5419834606
## 2    1423 5345235022  1  74599585 19.86 8.987e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the above anova output we can see that the F-score and corresponding p-value are significant. This means that a lot of information is lost by reducing the number of predictors in the model. Therefore, the

difference between having 9 predictors and 8 is significant and it is better to have 9.

mmps

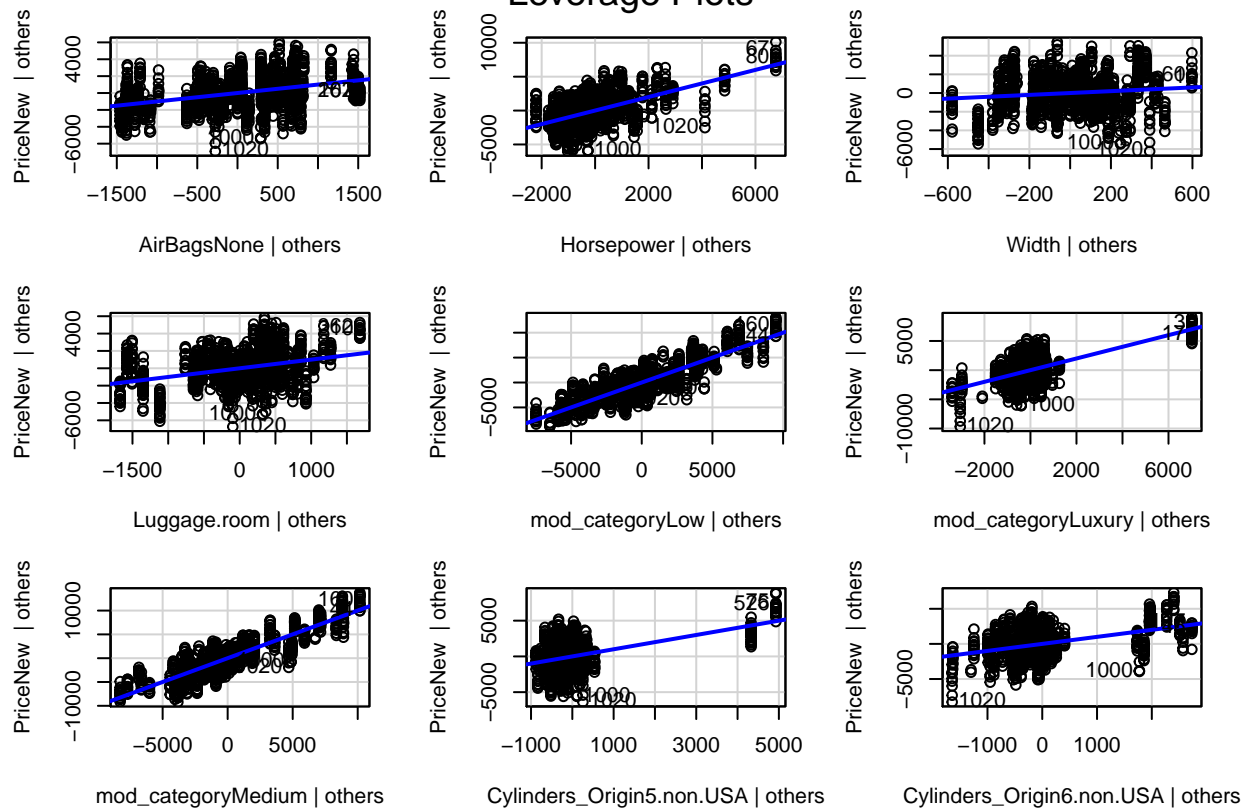
```
mmps(prediction_model)
```



Marginal Model Plots

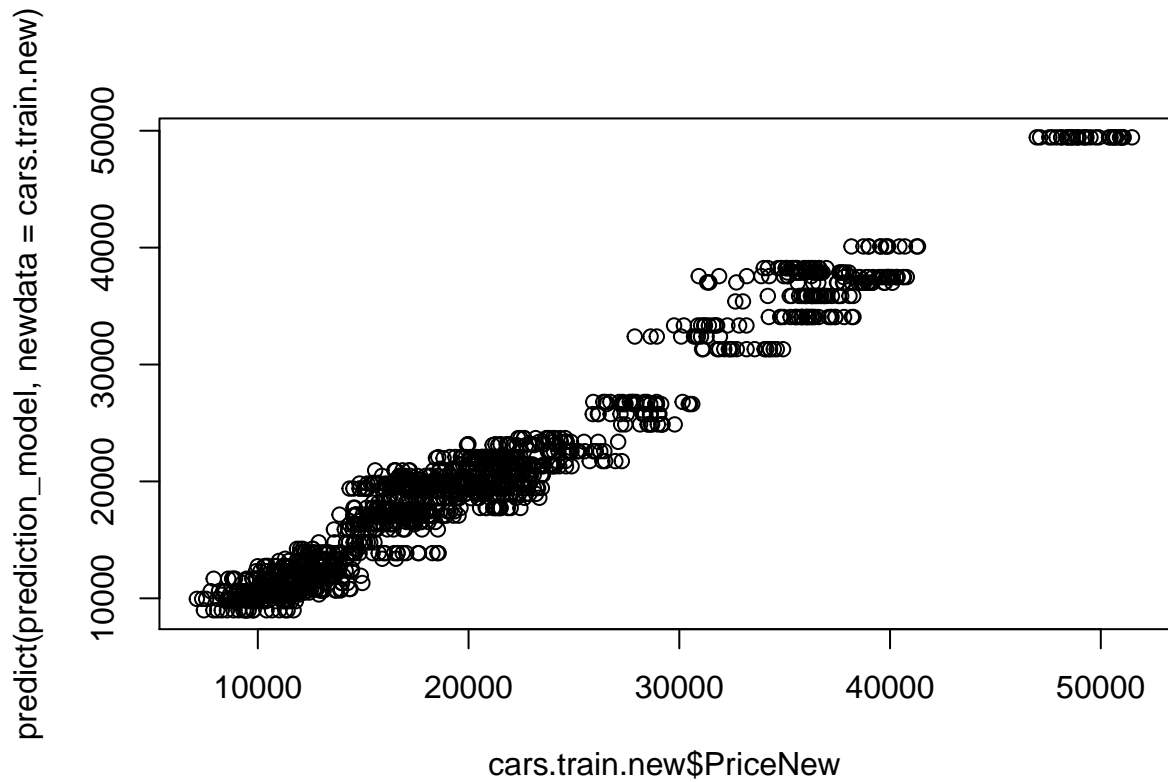
```
leveragePlots(prediction_model)
```

Leverage Plots



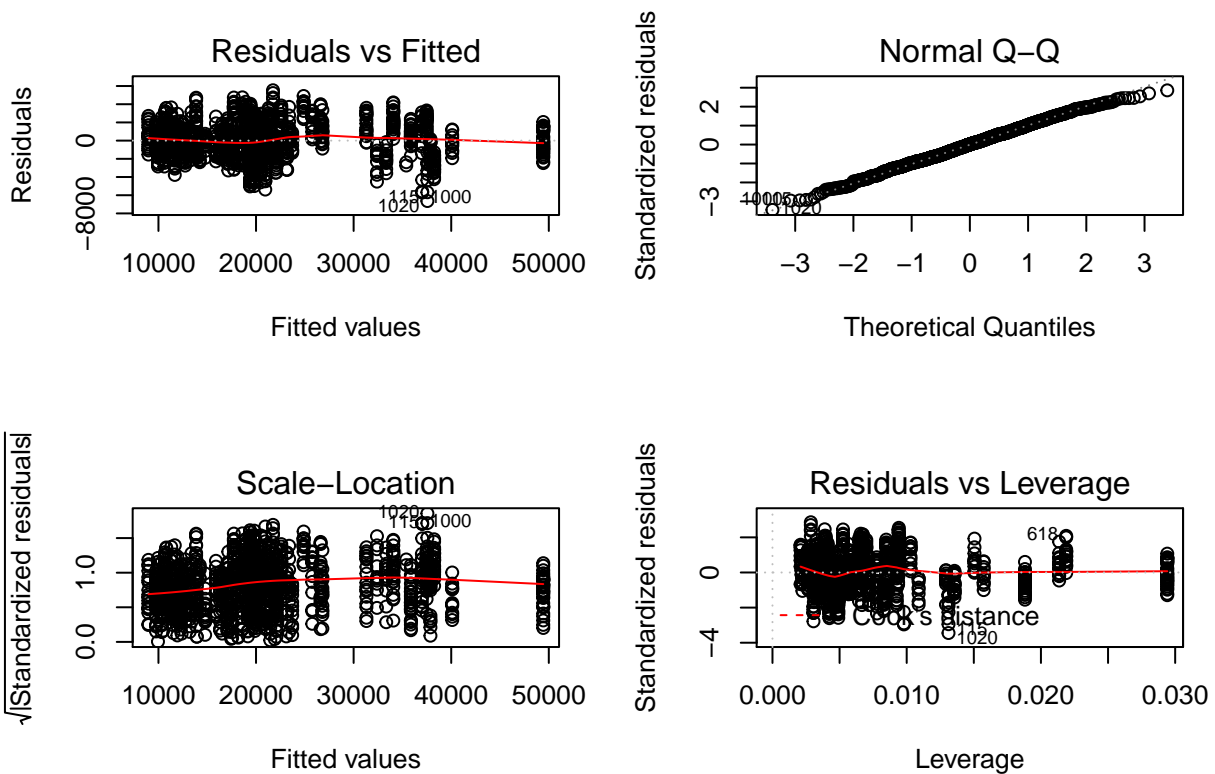
Y vs y-hat

```
plot(cars.train.new$PriceNew, predict(prediction_model, newdata = cars.train.new))
```

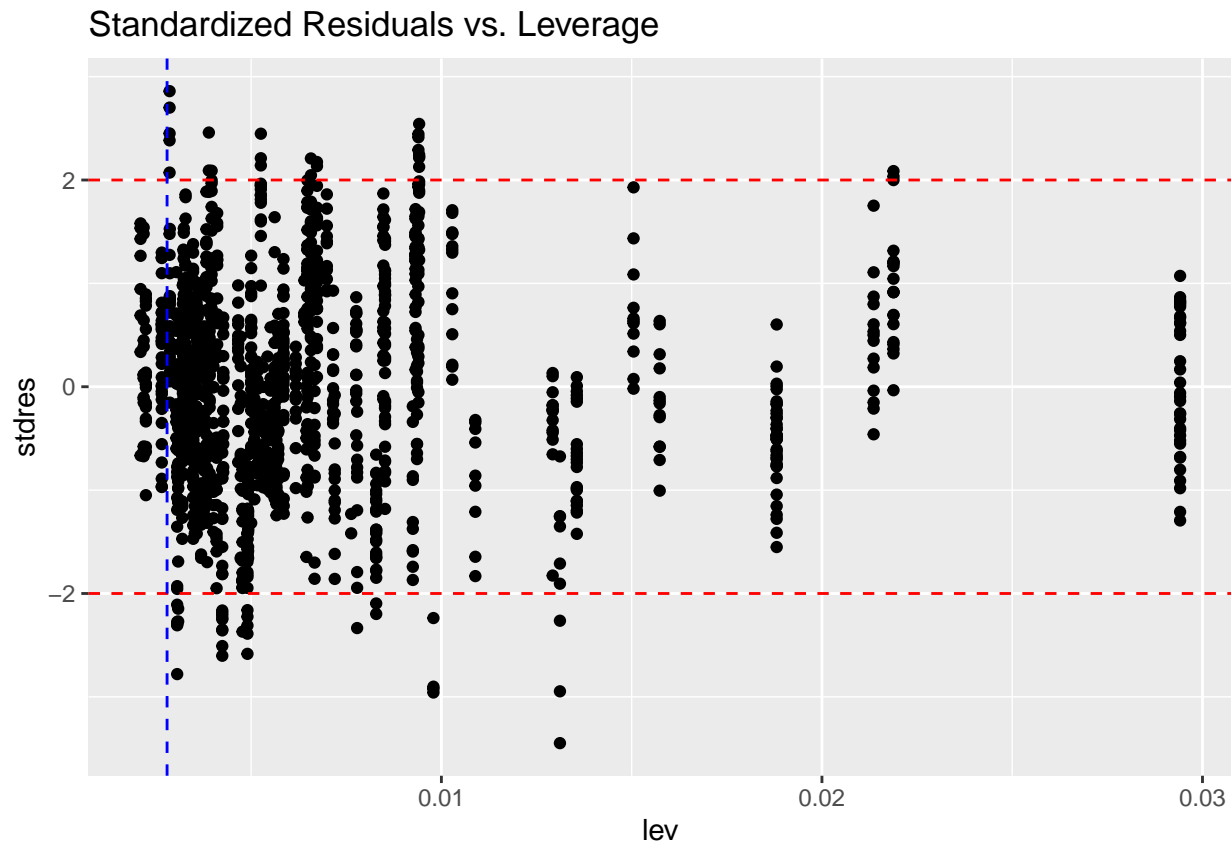


Let's look at the diagnostic plots.

```
# Diagnostic Plots
par(mfrow=c(2,2))
plot(prediction_model)
```



```
# Plot of leverage vs standardized residuals
cars.train.new$lev <- hatvalues(prediction_model)
cars.train.new$stdres <- rstandard(prediction_model)
rvl_plot <- qplot(lev, stdres, data=cars.train.new) + geom_hline(yintercept=c(-2,2), linetype="dashed",
# Add boundary lines
rvl_plot <- rvl_plot + geom_vline(xintercept=4/dim(cars.train.new)[1], linetype="dashed", col="blue")
rvl_plot <- rvl_plot + labs(title="Standardized Residuals vs. Leverage")
rvl_plot
```



The assumptions of random variance appear to hold, while normality and constant variance seem to be violated as apparent in the non-linear line in the Normal Q-Q plot as well as the upward-trending line in the Scale-Location plot. Let's see if we can make transformations to improve our model.

Box Cox Transformation

```
summary(powerTransform(cbind(PriceNew, Horsepower, Width, Luggage.room)~1,data=cars.train.new))

## bcPower Transformations to Multinormality
##           Est Power Rounded Pwr Wald Lwr Bnd Wald Up Bnd
## PriceNew   -0.4574      -0.50   -0.5510   -0.3637
## Horsepower  -0.2116      -0.21   -0.2981   -0.1250
## Width       -1.4302     -1.00   -2.0766   -0.7839
## Luggage.room  1.1903       1.19    1.0370    1.3436
##
## Likelihood ratio test that transformation parameters are equal to 0
## (all log transformations)
##           LRT df      pval
```



```
## LR test, lambda = (0 0 0 0) 380.9827 4 < 2.22e-16
##
## Likelihood ratio test that no transformations are needed
##               LRT df      pval
## LR test, lambda = (1 1 1 1) 1197.073 4 < 2.22e-16
```

Power Transformed model

```
prediction_model_pt <- lm(formula = log(PriceNew) ~ AirBagsNone + log(Horsepower) + log(Width) + log(Luggage.room) +
  mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
  Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
summary(prediction_model_pt)
```

```
##
## Call:
## lm(formula = log(PriceNew) ~ AirBagsNone + log(Horsepower) +
##     log(Width) + log(Luggage.room) + mod_categoryLow + mod_categoryLuxury +
##     mod_categoryMedium + Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA,
##     data = cars.train.new)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.34988	-0.07138	0.00708	0.06969	0.26365

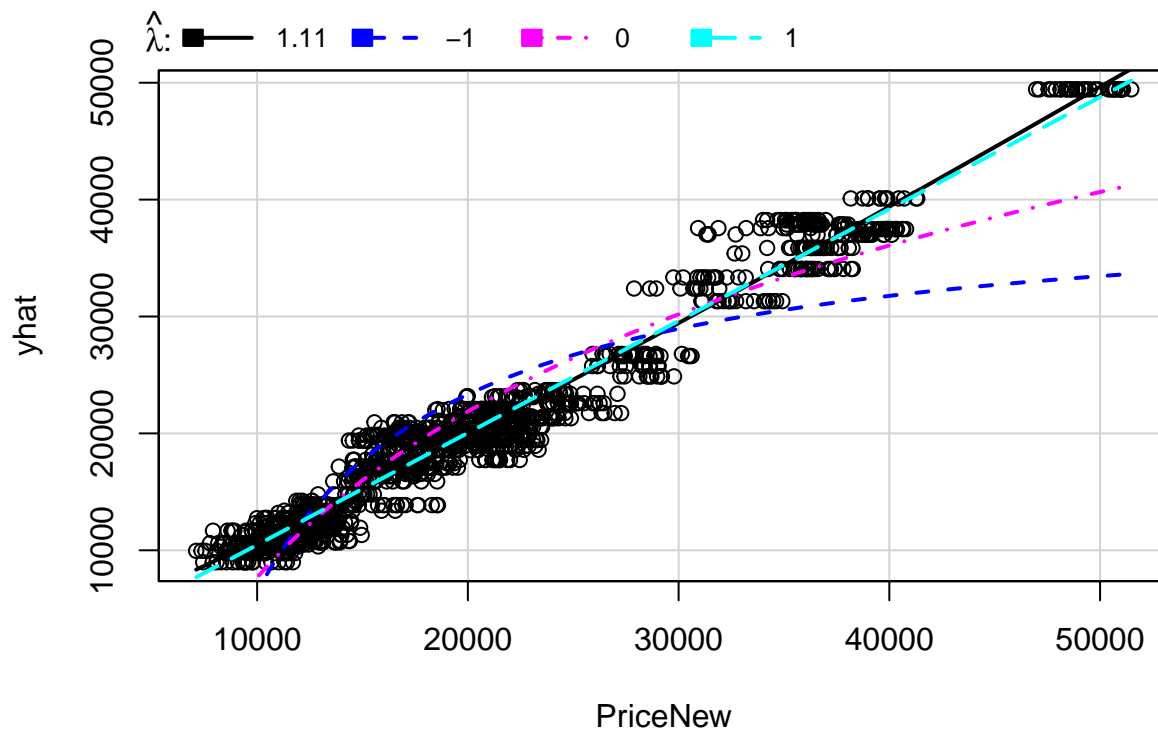
```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9.044664	0.376470	24.025	< 2e-16 ***
AirBagsNone	-0.093304	0.006948	-13.430	< 2e-16 ***
log(Horsepower)	0.397084	0.015487	25.639	< 2e-16 ***
log(Width)	-0.236301	0.102635	-2.302	0.0215 *
log(Luggage.room)	0.123100	0.015187	8.106	1.12e-15 ***
mod_categoryLow	-0.727161	0.013484	-53.928	< 2e-16 ***
mod_categoryLuxury	0.205949	0.020376	10.108	< 2e-16 ***
mod_categoryMedium	-0.443073	0.009022	-49.110	< 2e-16 ***
Cylinders_Origin5.non.USA	0.255856	0.014628	17.491	< 2e-16 ***
Cylinders_Origin6.non.USA	0.114274	0.010051	11.370	< 2e-16 ***

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1031 on 1423 degrees of freedom
## Multiple R-squared:  0.9413, Adjusted R-squared:  0.9409
## F-statistic: 2534 on 9 and 1423 DF, p-value: < 2.2e-16
```

Inverse Response Plot

```
inverseResponsePlot(prediction_model, lam=c(-1,0,1))
```



```
##      lambda      RSS
## 1  1.111753 5043541038
## 2 -1.000000 31733415729
## 3  0.000000 12995331320
## 4  1.000000 5119348278
```

```
prediction_model_inv <- lm(formula = PriceNew^1.111753 ~ AirBagsNone + Horsepower + Width + Luggage.room
  mod_categoryLow + mod_categoryLuxury + mod_categoryMedium +
  Cylinders_Origin5.non.USA + Cylinders_Origin6.non.USA, data = cars.train.new)
```

Create new test data frame based on same steps as training set

```
# New categorical variable for the models based on price
cars.test$mod_category <- ifelse(mean_prices_model[cars.test$Model] >= 40000, "Luxury", ifelse(mean_pr
# Combine 4WD and Front together
cars.test$DriveTrain <- ifelse(cars.test$DriveTrain == "Rear", "Rear", "Other")
# Combining Cylinders and Origin
cars.test$Cylinders_Origin <- paste(cars.test$Cylinders, cars.test$Origin)
```

Create a new testing data frame

```
# Drop the Ob column and original categorical predictors for which we created new variables
# 2-Manufacturer, 3-Model, 8-Cylinders, 23-Origin, 24-Make
cars.test.new <- cars.test[-c(2,3,8,23,24)]
dmy <- dummyVars( ~. ,data=cars.test.new,fullRank =T)
cars.test.new <- data.frame(predict(dmy,newdata = cars.test.new))
```

New Test data

```
write.csv(cars.test.new, 'cars.test.new.csv')
```

Fit model

```
cars.test.new$PriceNew <- predict(prediction_model, newdata = cars.test.new)
cars.test.new$PriceNew <- cars.test.new$PriceNew
write.csv(cars.test.new[c("Ob", "PriceNew")], 'car_prediction14.csv', row.names = FALSE)
```

```
extractAIC(prediction_model, k=log(dim(cars.train.new[1])))
```

```
## [1] 10.00 21756.75 21684.08
```