Taeyoung's Motor Show

Used_Car Korean dealer, Kevin Shin

May 14, 2018

## From this picture, how many car brands can you notice?

1. The picture below captures top 40 car makes

2. Which car make or car brand do you specficially prefer?

3. Which type of factor usually influences you the most when you determine to purchase used_car?

***

This density plot shows the overall relationship between used_car's price and general mileage of them. It seems that most cars that are traded within 10K to 20K in used_car market have relatively high mileage distribution.

```
# Simple Scatterplot
#attach(mtcars)
theme_set(theme_classic())

# Plot
g <- ggplot(used_car, aes(Price/1000))
g + geom_density(aes(fill=Mileage), alpha=0.8) +
    labs(title="Density plot",
         subtitle="Price distribution Grouped by scale of Mileage",
         caption="Data: used_car",
         x="Car Price",
         fill="Yellow") + scale_color_manual(values = c("red", "blue"))
```

Density plot

Price distribution Grouped by scale of Mileage

Data: used_car
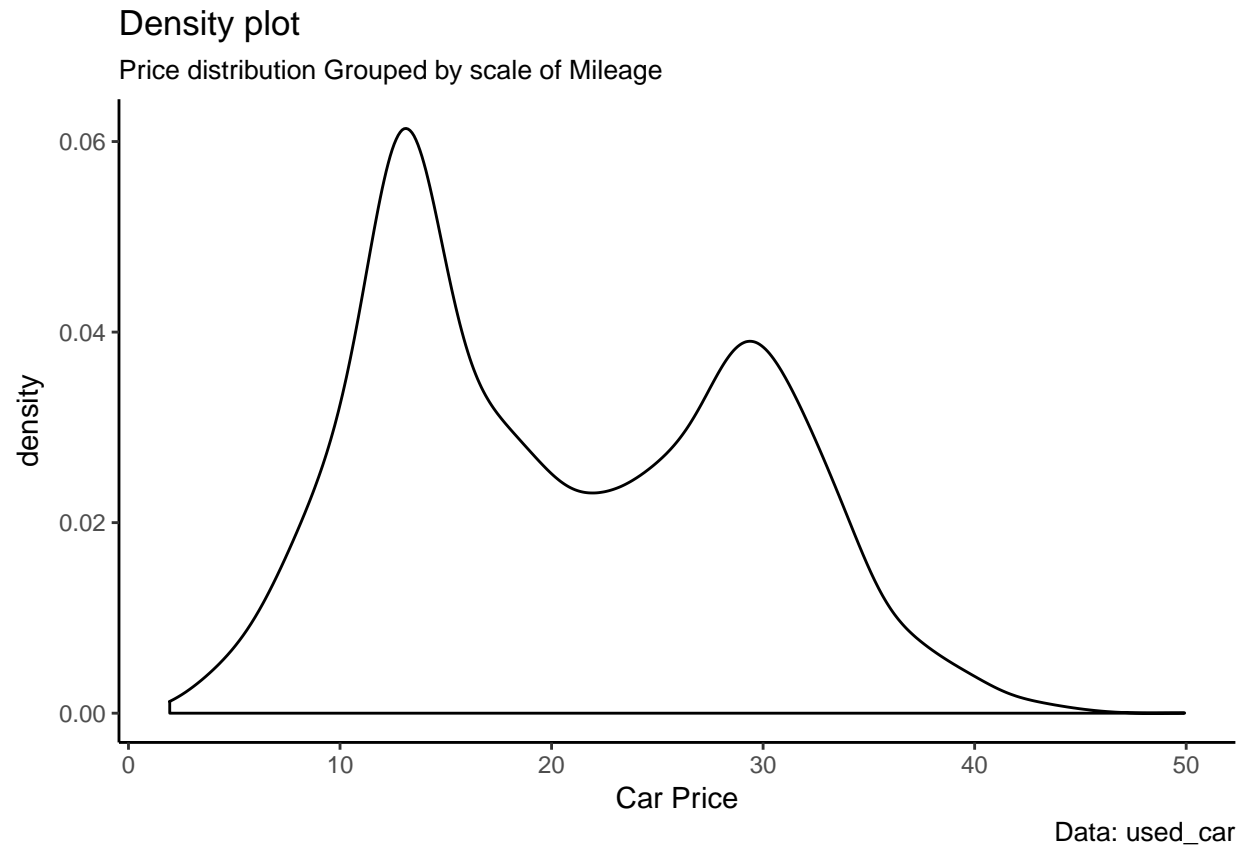
The Boxplot Demonstration of Cars Make and Cars' Price. As we can see from this boxplot chart, GMC and Audi Make seem to have relatively slow depreciation as compared to other car makes.

```
par(bg = "snow1")
plot(used_car$Make, used_car$Price, xlab = "Car Brands", ylab = "Price of Car Brands", col = "Yellow2"
```
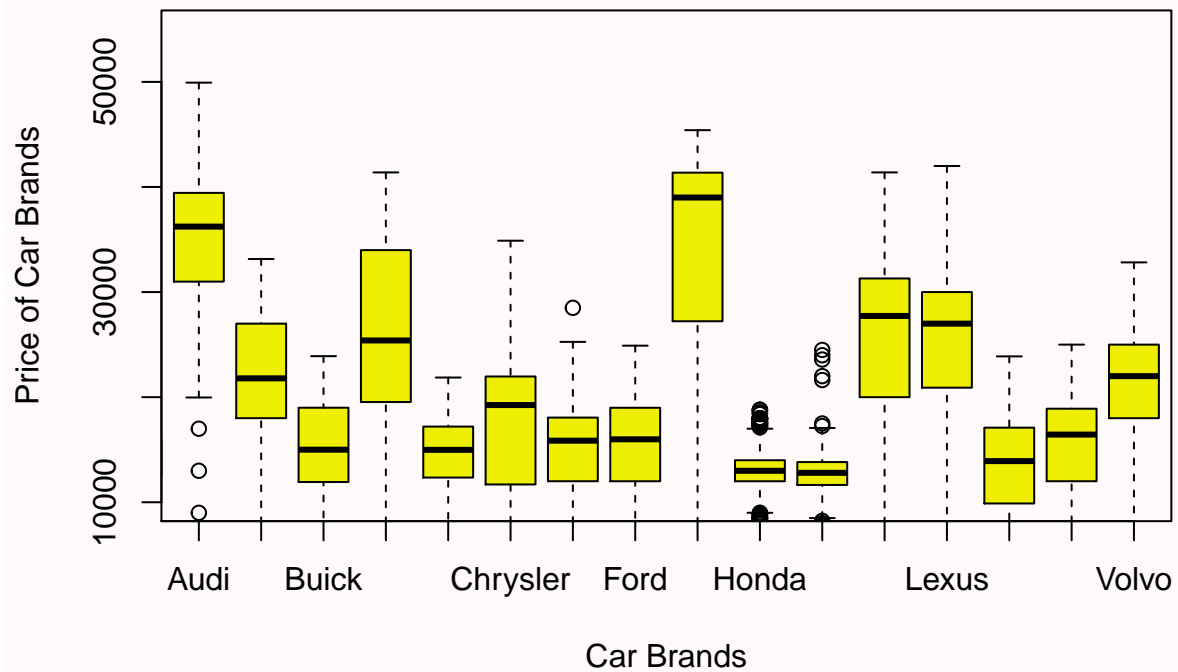
This is my personal dream car in the future

***

This scatterplot approximately shows the inverse relationship between used_cars' mileage and used_car's price. As the mileage of used_car goes up, the price of used_car goes down, which looks comprehensible.

```r
ggplot(data = used_car, aes(x = used_car$Mileage, y = used_car$Price)) +
  geom_point(colour = "red", size = 3) + xlim(10000,90000) + ylim(10000,50000)
```

Warning: Removed 2278 rows containing missing values (geom_point).
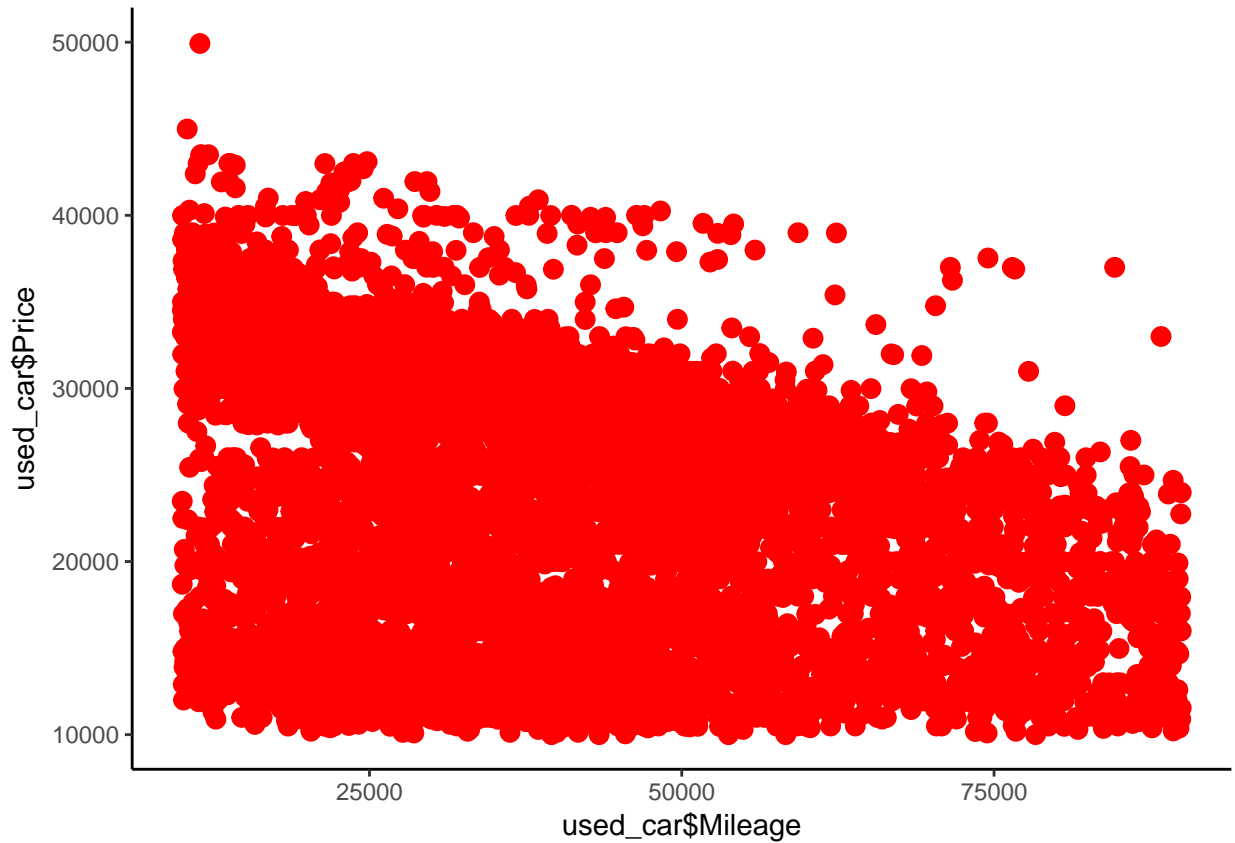
This following Piechart relatively shows the comparative price distribution among Car Make. As expected, German cars tend to be priced higher than other brands such as American, Japanese, and Korean ones as time goes on.

```r
slices <- c(15500, 35000, 16700, 12400, 10500,18500, 38050, 29020 )
lbls = c('Acura', 'BMW', 'Ford', 'Honda', 'Hyundai','Lexus', 'Mercedes', 'Porshe' )
pie3D(slices, labels = lbls, explode = 0.1, main = "Average mean price of used_cars", col = rainbow(len
```

Average mean price of used_cars
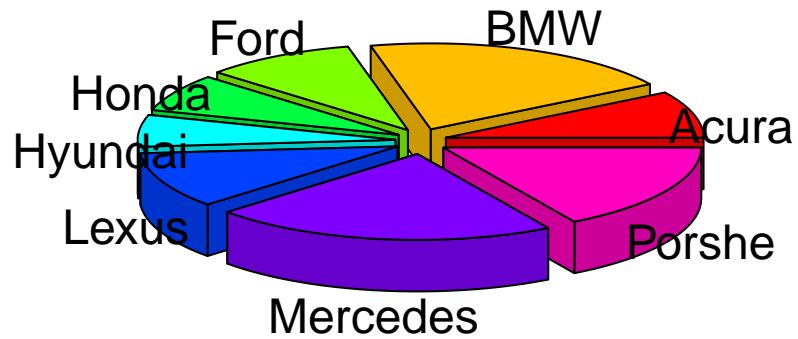
This following geom_scatter point plot shows the positive correlation between the Car Year and Car Price. In other words, the newer the car is, the more expensive the price is

```
ggplot(used_car, aes(used_car$Year,used_car$Price)) + geom_point(color= "deeppink3", alpha = 0.2) + lab
```

Correlation between Car Year and Car Price
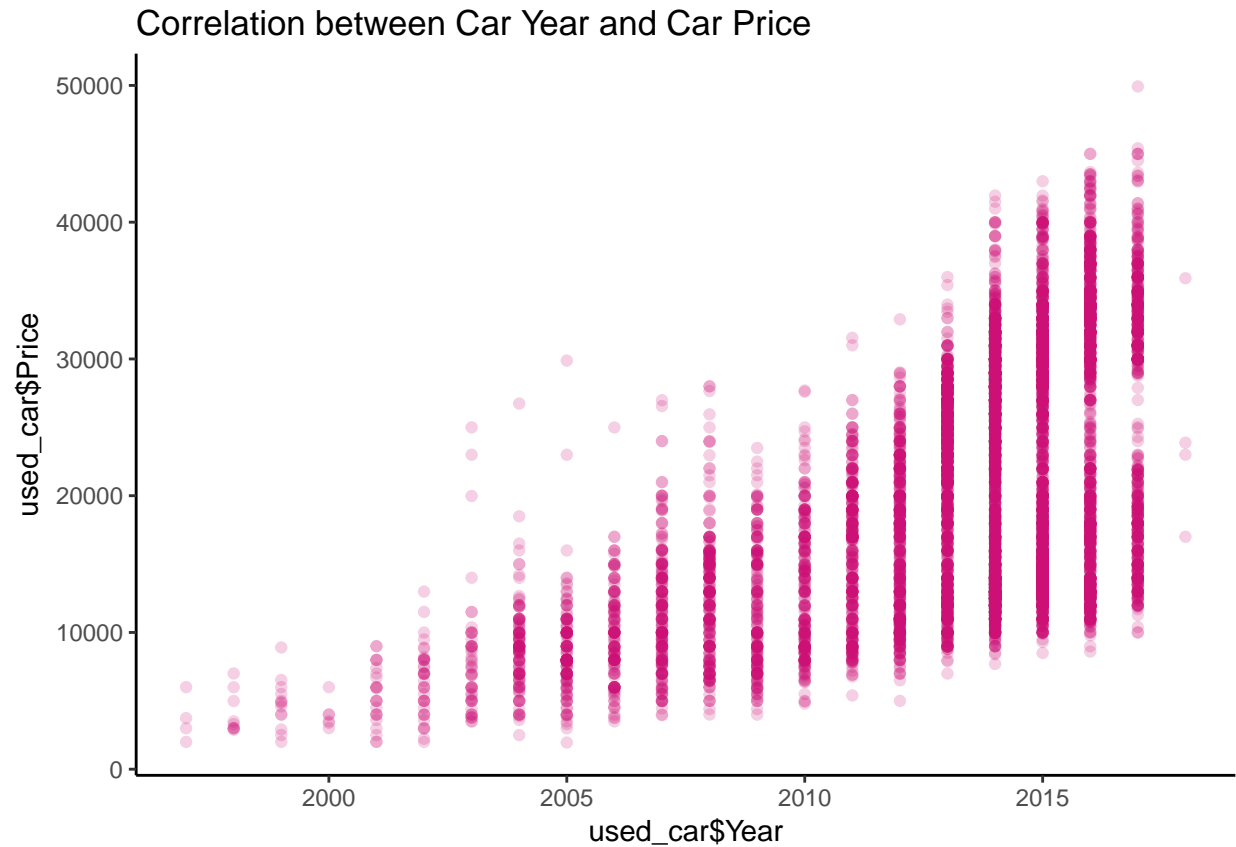
This following graph shows all states of USA and their car price variations, which do not seem quite relevant to one another since there is not specific and accurate index placeholder to be compared. However, the main property of this graph is that except few states, most of the used_car price variations of states seem to vary within decent size of range.

```
used_car %>% filter(!is.na(State)) %>% group_by(State) %>% top_n(n=10)%>%
ungroup() -> data
```

```
## Selecting by Model
```

```
ggplot(data = data, aes(x=reorder(State,Price), y=Price/1000)) + geom_col(fill="Purple", colour = "Orang
```

Price Variation by States
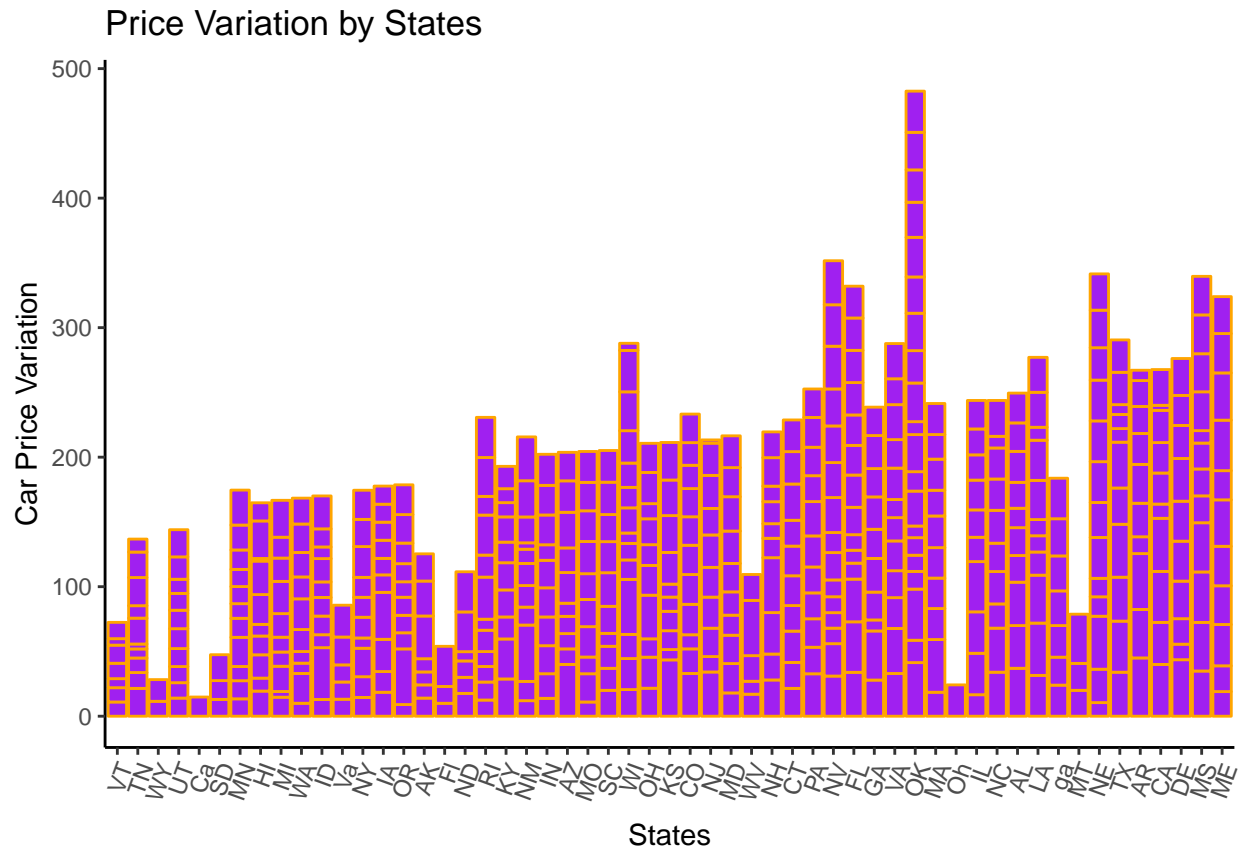
This tree map is showing how much mileage of a certain make of used_cars is in trade. In other words, the higher their mileages are when they were traded, the bigger area they will account for. In this example, it seems that Infiniti, Lexus, Honda are ranked high, meaning that used_cars from those three makes tend to be traded after a long usage.

```
treemap(used_car, #Your data frame object
        index=c("Make"),  #A list of your categorical variables
        vSize = "Mileage",  #This is your quantitative variable
        type="index", #Type sets the organization and color scheme of your treemap
        palette = "Reds",  #Select your color palette from the RColorBrewer presets or make your own.
        title="Correlation between Mileage and Make of Used_Cars", #Customize your title
        fontsize.title = 14 #Change the font size of the title
        )
```
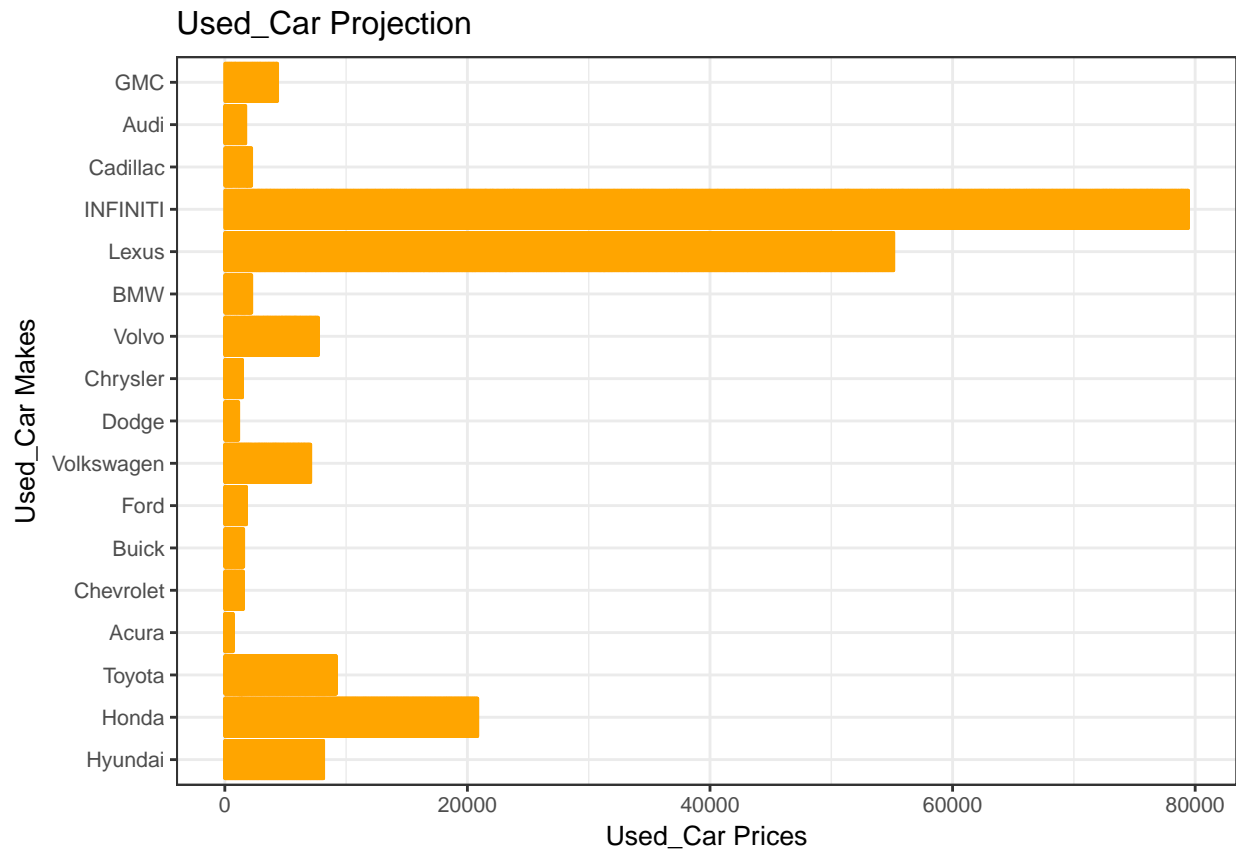
Correlation between Mileage and Make of Used_Cars
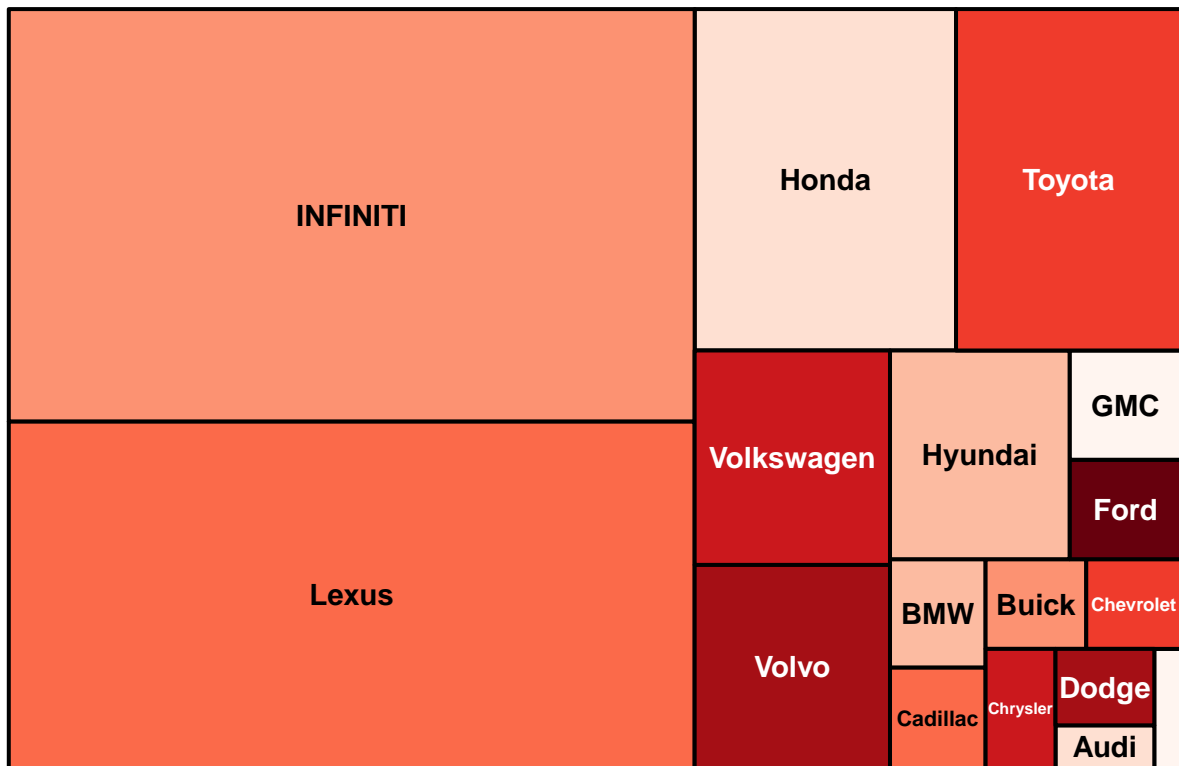
This horizontal barplot shows the relation between types of used_car makes and their prices.

```
used_car %>% filter(!is.na(Make)) %>% group_by(Make) %>%
  ungroup() -> data

ggplot(data = data, aes(x=reorder(Make,Price), y=Price/1000)) + geom_col(fill="green", colour = "Orange
```

Used_Car Projection

This line graph also demonstrates the types of Car Makes and their price distribution.

```
ggplot(data=used_car, aes(x=Year, y=Price, color = factor(Make), alpha = 0.8 )) +
  geom_line() + theme(plot.background = element_rect(fill = 'lightpink', colour = 'red'))
```
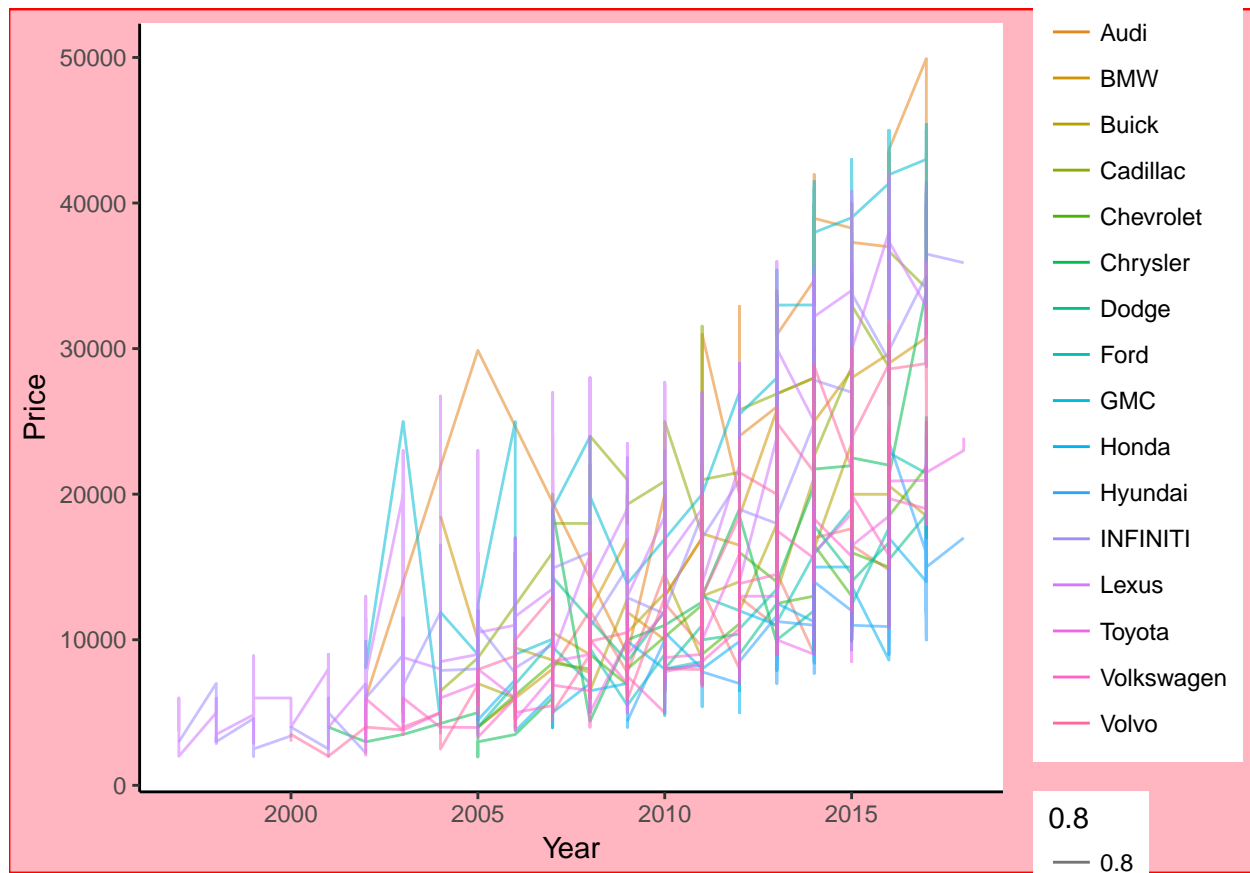
Thank you for teaching us this wonderful course this semester Dr. Christ! I really enjoy studying R with you! I hope everyone will have a great summer!