Automobile Comparisons Within Countries

Armando Marquez am8245

Introduction

The data that was used was in found in the ggplot2 library of R. The information that was used was from Consumer Reports from 1990. The data compared different variables of cars that included type, country, tank size, height of car, price, length, rim, tires, turning, wheel base, and width. I just purchased a new SUV, but this is my first SUV. I am curious on gas mileage since I work in North Austin but live in Downtown Austin. I expect to see the higher the price, the bigger the tank on a vehicle.

```
#Load dataset from the package
library(dplyr)
library(tidyverse)
library(ggplot2)
library(rpart)
car_info <- car.test.frame
car_info2 <- car90
car_info3 <- cu.summary
```

Tidy

```
# Tidy the datset
car_info2 <- car_info2 %>%
 select(-Disp2, - Eng.Rev, -Frt.Leg.Room, -Frt.Shld,
 -Gear.Ratio, -Gear2, -HP.revs, - Luggage, -Mileage, -Model2,
 -Rear.Hd, -Rear.Seating, -RearShld, -Sratio.p, - Sratio.m, -Trans1,
 -Trans2)
car_info2 <- car_info2 %>%
 select(-Disp)
car_info <- car_info %>%
 select(-Reliability, -Disp.)
car_info2 <- car_info2 %>%
 select(-Reliability)
car_info2 <- car_info2 %>%
 select(-Country, -Price)
car_info <- car_info %>%
 select(-Mileage)
car_info <- car_info %>%
 select(-Weight)
car_info2 <- car_info2 %>%
 select(-Weight)

na.omit(car_info)
```

```
##                         Price   Country    Type  HP
## Eagle Summit 4           8895       USA   Small 113
## Ford Escort   4          7402       USA   Small  90
## Ford Festiva 4           6319     Korea   Small  63
## Honda Civic 4            6635 Japan/USA   Small  92
## Mazda Protege 4          6599     Japan   Small 103
## Mercury Tracer 4         8672    Mexico   Small  82
## Nissan Sentra 4          7399 Japan/USA   Small  90
## Pontiac LeMans 4         7254     Korea   Small  74
## Subaru Loyale 4          9599     Japan   Small  90
## Subaru Justy 3           5866     Japan   Small  73
## Toyota Corolla 4         8748 Japan/USA   Small 102
## Toyota Tercel 4          6488     Japan   Small  78
## Volkswagen Jetta 4       9995   Germany   Small 100
## Chevrolet Camaro V8     11545       USA  Sporty 170
## Dodge Daytona            9745       USA  Sporty 100
## Ford Mustang V8         12164       USA  Sporty 225
## Ford Probe              11470       USA  Sporty 110
## Honda Civic CRX Si 4     9410     Japan  Sporty 108
## Honda Prelude Si 4WS 4  13945     Japan  Sporty 140
## Nissan 240SX 4          13249     Japan  Sporty 140
## Plymouth Laser          10855       USA  Sporty  92
## Subaru XT 4             13071     Japan  Sporty  97
## Audi 80 4               18900   Germany Compact 108
## Buick Skylark 4         10565       USA Compact 110
## Chevrolet Beretta 4     10320       USA Compact  95
##  [ reached 'max' / getOption("max.print") -- omitted 35 rows ]
```

```
na.omit(car_info2)
```

```
##                Front.Hd  HP Height Length Rim Steering Tank  Tires Turning
## Acura Integra       3.5 130   47.5    177 R14    power 13.2 195/60      37
## Acura Legend        2.0 160   50.0    191 R15    power 18.0 205/60      42
## Audi 100            2.5 130   51.5    193 R15    power 21.1 205/60      39
## Audi 80             4.0 108   50.5    176 R14    power 15.9 175/70      35
## BMW 325i            2.0 168   49.5    175 R14    power 16.4 195/65      35
## BMW 535i            3.0 208   51.0    186 R15    power 21.1 225/60      39
## Buick Century       4.0 110   49.5    189 R14    power 15.7 185/75      41
## Buick Electra       6.0 165   50.5    197 R14    power 18.0 205/75      43
##                    Type Wheel.base Width
## Acura Integra     Small        102    67
## Acura Legend     Medium        109    69
## Audi 100         Medium        106    71
## Audi 80         Compact        100    67
## BMW 325i        Compact        101    65
## BMW 535i         Medium        109    69
## Buick Century    Medium        105    69
## Buick Electra     Large        111    72
##  [ reached 'max' / getOption("max.print") -- omitted 97 rows ]
```

```
car_info
```

```
##                          Price   Country    Type  HP
## Eagle Summit 4            8895       USA   Small 113
## Ford Escort   4           7402       USA   Small  90
## Ford Festiva 4            6319     Korea   Small  63
## Honda Civic 4             6635 Japan/USA   Small  92
## Mazda Protege 4           6599     Japan   Small 103
## Mercury Tracer 4          8672    Mexico   Small  82
## Nissan Sentra 4           7399 Japan/USA   Small  90
## Pontiac LeMans 4          7254     Korea   Small  74
## Subaru Loyale 4           9599     Japan   Small  90
## Subaru Justy 3            5866     Japan   Small  73
## Toyota Corolla 4          8748 Japan/USA   Small 102
## Toyota Tercel 4           6488     Japan   Small  78
## Volkswagen Jetta 4        9995   Germany   Small 100
## Chevrolet Camaro V8      11545       USA  Sporty 170
## Dodge Daytona             9745       USA  Sporty 100
## Ford Mustang V8          12164       USA  Sporty 225
## Ford Probe               11470       USA  Sporty 110
## Honda Civic CRX Si 4      9410     Japan  Sporty 108
## Honda Prelude Si 4WS 4   13945     Japan  Sporty 140
## Nissan 240SX 4           13249     Japan  Sporty 140
## Plymouth Laser           10855       USA  Sporty  92
## Subaru XT 4              13071     Japan  Sporty  97
## Audi 80 4                18900   Germany Compact 108
## Buick Skylark 4          10565       USA Compact 110
## Chevrolet Beretta 4      10320       USA Compact  95
##  [ reached 'max' / getOption("max.print") -- omitted 35 rows ]
```

```
car_info2
```

```
##               Front.Hd  HP Height Length Rim Steering Tank  Tires Turning
## Acura Integra      3.5 130   47.5    177 R14    power 13.2 195/60      37
## Acura Legend       2.0 160   50.0    191 R15    power 18.0 205/60      42
## Audi 100           2.5 130   51.5    193 R15    power 21.1 205/60      39
## Audi 80            4.0 108   50.5    176 R14    power 15.9 175/70      35
## BMW 325i           2.0 168   49.5    175 R14    power 16.4 195/65      35
## BMW 535i           3.0 208   51.0    186 R15    power 21.1 225/60      39
## Buick Century      4.0 110   49.5    189 R14    power 15.7 185/75      41
## Buick Electra      6.0 165   50.5    197 R14    power 18.0 205/75      43
##                   Type Wheel.base Width
## Acura Integra    Small        102    67
## Acura Legend    Medium        109    69
## Audi 100        Medium        106    71
## Audi 80        Compact        100    67
## BMW 325i       Compact        101    65
## BMW 535i        Medium        109    69
## Buick Century  Medium        105    69
## Buick Electra    Large        111    72
##  [ reached 'max' / getOption("max.print") -- omitted 103 rows ]
```

Information that was not helpful to the project was deleted. All of the information is tidy by having each specific car divided into rows and then using categorical and numerical variabes that describe aspects of the car as columns

Join and Merge

```
#Join the two datasets through Type
car_info_join <- car_info %>%
 inner_join(car_info2, by = "Type")
na.omit(car_info_join)
```

```
##   Price Country  Type HP.x Front.Hd HP.y Height Length Rim Steering Tank  Tires
## 1  8895     USA Small  113      3.5  130   47.5    177 R14    power 13.2 195/60
## 2  8895     USA Small  113      3.0   81   50.0    159 R13     both 13.2 155/80
## 3  8895     USA Small  113      3.0   93   51.0    163 R13     both 13.0 165/80
## 4  8895     USA Small  113      4.5  113   49.5    170 R14     both 13.2 195/60
## 5  8895     USA Small  113      3.5   90   51.5    169 R14     both 13.0 175/70
## 6  8895     USA Small  113      4.0   63   52.0    141 R12     both 10.0    145
##   Turning Wheel.base Width
## 1      37        102    67
## 2      32         94    66
## 3      40         99    67
## 4      36         97    66
## 5      37         94    66
## 6      33         90    63
##  [ reached 'max' / getOption("max.print") -- omitted 1183 rows ]
```

```
car_info_join
```

```
##   Price Country  Type HP.x Front.Hd HP.y Height Length Rim Steering Tank  Tires
## 1  8895     USA Small  113      3.5  130   47.5    177 R14    power 13.2 195/60
## 2  8895     USA Small  113      3.0   81   50.0    159 R13     both 13.2 155/80
## 3  8895     USA Small  113      3.0   93   51.0    163 R13     both 13.0 165/80
## 4  8895     USA Small  113      4.5  113   49.5    170 R14     both 13.2 195/60
## 5  8895     USA Small  113      3.5   90   51.5    169 R14     both 13.0 175/70
## 6  8895     USA Small  113      4.0   63   52.0    141 R12     both 10.0    145
##   Turning Wheel.base Width
## 1      37        102    67
## 2      32         94    66
## 3      40         99    67
## 4      36         97    66
## 5      37         94    66
## 6      33         90    63
##  [ reached 'max' / getOption("max.print") -- omitted 1183 rows ]
```

```
#Filter only the USA and Large vehicles
#car_info_join %>%
# filter(Country.x == "USA & Type == "Large")
```

```
#Arranged the cars by height
car_info_join %>%
 arrange(Height)
```

```
##   Price Country   Type HP.x Front.Hd HP.y Height Length Rim Steering Tank
## 1 11545     USA Sporty  170      2.0  250   43.5    177 R17    power 20.0
## 2 11545     USA Sporty  170      3.5  116   43.5    155 R14     both 11.9
## 3  9745     USA Sporty  100      2.0  250   43.5    177 R17    power 20.0
## 4  9745     USA Sporty  100      3.5  116   43.5    155 R14     both 11.9
## 5 12164     USA Sporty  225      2.0  250   43.5    177 R17    power 20.0
## 6 12164     USA Sporty  225      3.5  116   43.5    155 R14     both 11.9
##    Tires Turning Wheel.base Width
## 1 275/40      42         96    71
## 2 185/60      33         89    66
## 3 275/40      42         96    71
## 4 185/60      33         89    66
## 5 275/40      42         96    71
## 6 185/60      33         89    66
##  [ reached 'max' / getOption("max.print") -- omitted 1183 rows ]
```

```
#Increased the price by 380% to understand pricing for the current market
#usa_PriceIncrease <- car_info_join %>%
# filter(Country == "USA" & Type == "Large")
# mutate(Price.increase = Price * 3.8)
```

Summary Statistics

```
#Found the mean for all numeric variables grouped by country
#car_info_mean <- car_info2 %>%
# group_by(Country) %>%
# summarise(mean_price = mean(Price, na.rm=T), mean_height = mean(Height,
# na.rm=T), mean_length = mean(Length, na.rm=T), mean_Tank = mean(Tank,
# na.rm=T), mean_Turning = mean(Turning, na.rm=T), mean_Wheel.base =
# mean(Wheel.base, na.rm=T), mean_Width = mean(Width, na.rm=T))
```

```
#Found the minimum of all numeric variables grouped by country
car_info_join %>%
 group_by(Country) %>%
 summarise(min_price = min(Price, na.rm=T), min_height = min(Height,
 na.rm=T), min_length = min(Length, na.rm=T), min_Tank = min(Tank,
 na.rm=T), min_Turning = min(Turning, na.rm=T), min_Wheel.base =
 min(Wheel.base, na.rm=T), min_Width = min(Width, na.rm=T))
```

```
## # A tibble: 8 x 8
##   Country   min_price min_height min_length min_Tank min_Turning min_Wheel.base
## * <fct>         <int>      <dbl>      <dbl>    <dbl>       <dbl>           <dbl>
## 1 France        15930       48.5        172     13.6          34              97
## 2 Germany        9995       47.5        141      9.2          32              90
## 3 Japan          5866       43.5        141      9.2          32              88
## 4 Japan/USA      6635       47.5        141      9.2          32              90
## 5 Korea          6319       47.5        141      9.2          32              90
## 6 Mexico         8672       47.5        141      9.2          32              90
## 7 Sweden        18450       48.5        172     13.6          34              97
## 8 USA            7402       43.5        141      9.2          32              88
## # ... with 1 more variable: min_Width <dbl>
```

```
#Found the maximum of all numeric variables grouped by country
car_info_join %>%
 group_by(Country) %>%
 summarise(max_price = max(Price, na.rm=T), max_height = max(Height,
 na.rm=T), max_length = max(Length, na.rm=T), max_Tank = max(Tank,
 na.rm=T), max_Turning = max(Turning, na.rm=T), max_Wheel.base =
 max(Wheel.base, na.rm=T), max_Width = max(Width, na.rm=T))
```

```
## # A tibble: 8 x 8
##   Country   max_price max_height max_length max_Tank max_Turning max_Wheel.base
## * <fct>         <int>      <dbl>      <dbl>    <dbl>       <dbl>          <dbl>
## 1 France        15930       52.5        190     18            41            107
## 2 Germany       18900       52.5        190     18            41            107
## 3 Japan         24760       70.5        205     27            47            119
## 4 Japan/USA     12459       52.5        190     18            41            107
## 5 Korea          9999       52          205     22.5          43            113
## 6 Mexico         8672       52          177     15.9          40            102
## 7 Sweden        18450       52.5        190     18            41            107
## 8 USA           17257       70.5        221     27            47            122
## # ... with 1 more variable: max_Width <dbl>
```

```
#Found the IQR for all numeric variables grouped by country
car_info_join %>% group_by(Country) %>%
 summarise(IQR_price = IQR(Price, na.rm=T), IQR_height = IQR(Height,
 na.rm=T), IQR_length = IQR(Length, na.rm=T), IQR_Tank = IQR(Tank,
 na.rm=T), IQR_Turning = IQR(Turning, na.rm=T), IQR_Wheel.base =
 IQR(Wheel.base, na.rm=T), IQR_Width = IQR(Width, na.rm=T))
```

```
## # A tibble: 8 x 8
##   Country   IQR_price IQR_height IQR_length IQR_Tank IQR_Turning IQR_Wheel.base
## * <fct>         <dbl>      <dbl>      <dbl>    <dbl>       <dbl>          <dbl>
## 1 France            0       1           7.5     1.20        2              2.5
## 2 Germany        8905       1          11       2.9         3              6
## 3 Japan          8300       2.5        18       4.9         4              9
## 4 Japan/USA      4746       1          11       2.7         4              6
## 5 Korea          3680       2          24       4.45        4.75          11.8
## 6 Mexico            0       1.38       10.5     1.30        1.75           3
## 7 Sweden            0       1           7.5     1.20        2              2.5
## 8 USA            3930       2          15.5     3.10        5             10
## # ... with 1 more variable: IQR_Width <dbl>
```

```
#Found the MAD for all numeric variables grouped by country
car_info_join %>%
 group_by(Country) %>%
 summarise(mad_price = mad(Price, na.rm=T), mad_height = mad(Height,
 na.rm=T), mad_length = mad(Length, na.rm=T), mad_Tank = mad(Tank,
 na.rm=T), mad_Turning = mad(Turning, na.rm=T), mad_Wheel.base =
 mad(Wheel.base, na.rm=T), mad_Width = mad(Width, na.rm=T))
```

```
## # A tibble: 8 x 8
##   Country   mad_price mad_height mad_length mad_Tank mad_Turning mad_Wheel.base
## * <fct>         <dbl>      <dbl>      <dbl>    <dbl>       <dbl>          <dbl>
## 1 France            0      0.741       5.93    0.741        1.48           1.48
## 2 Germany           0      0.741      10.4     2.67         2.97           4.45
## 3 Japan          6864.     1.48       11.9     3.41         2.97           7.41
## 4 Japan/USA      1423.     0.741       8.15    2.08         2.97           3.71
## 5 Korea          1386.     1.48       17.0     1.93         3.71           5.93
## 6 Mexico            0      0.741       6.67    1.78         1.48           2.97
## 7 Sweden            0      0.741       5.93    0.741        1.48           1.48
## 8 USA            2887.     1.48       11.9     2.97         2.97           7.41
## # ... with 1 more variable: mad_Width <dbl>
```

The minimum price for a car is 7402 dollars in the USA. The maximum price for a car is 17257 dollars. The IQR for USA is 3930 dollars. The MAD for a USA car is 2887 dollars.

Visualization

```

```

```
##           sysname           release           version          nodename
##         "Windows"          "10 x64"     "build 19041" "DESKTOP-1SMB59F"
##           machine             login              user    effective_user
##          "x86-64"           "Admin"           "Admin"           "Admin"
```