DATA 608 Final Project

Mario Pena

May 23, 2021

Contents

Introduction

Everyone shopping for a new car wants to choose something safe for themselves and their loved ones, but it's not always easy to figure out which models make the safest choices. In my final project I am interested in finding out which vehicle or set of vehicles are the safest from a body injury perspective? I believe this is a relevant topic in today's economy from the consumer's point of view as some of the vehicles getting a 5-star safety rating could be cost prohibitive for some or provide a limited variety of choices. In this case, more detailed analyses of the injury criteria and attributes are warranted.

Data

The data will be acquired from the National Highway Traffic Safety Administration (https://www-nrd.nhtsa.dot.gov/database/veh/veh.htm). The data to be used in this project provides information that refers to the performance and response of vehicles and other structures in impact. Some of the fields in the database include crash tests speed, conditions, tests' barriers, airbag, car make and model, and injury values among others. NHTSA was established through the "Highway Safety Act' and its mission has been to reduce deaths, injury and economic losses resulting from motor vehicle crashes. It started using the 5-Star safety ratings system in 1993 to help consumers make informed safety choices when buying vehicles.

Objective

Specifically, Which vehicle or set of vehicles are the safest from a head and leg injury perspective.

Injury Criteria

Head Injury Criterion, Left Femur Load, Right Femur Load

Analyses and Attributes

National Highway Traffic Safety Administration

Based on the recommendations following this section, provided by the National Highway Traffic Safety Administration, we want to perform analyses to determine which vehicles have the best and worst injury severity ratings based upon the criteria listed below and using the following attributes.

Attributes

vehicle make, Vehicle year, Vehicle body type, and occupant location (Left-front-seat, Right-front-seat).

Motiviation

The National Highway Traffic Safety Administration makes their recommendation using a star rating system. A greater number of Stars mean Safer Cars. 5-Star Safety Ratings measure the crashworthiness and rollover safety of vehicles. Five stars is the highest rating, one is the lowest. However, some of the vehicles getting a 5-star rating could be cost prohibitive for some consumers and provide a small variety of choices. In this instance, a more detailed analysis of the injury criteria and attributes are warranted.

Table ES.6: Summary of Recommended Injury Criteria for the SNPRM

Recommended Criteria	Large§ Male	Mid-Sized Male	Small Female	6 YO Child	3 YO Child	1 YO Infant
Head Criteria: HIC (15 msec)	700	700	700	700	570	390
Neck Criteria: SNPRM Nij	1.0	1.0	1.0	1.0	1.0	1.0
Critical Intercept Values						
Tension and Compression (N)	5440	4500	3370	2800	2120	1465
Flexion (Nm)	415	310	155	93	68	43
Extension (Nm)	166	125	62	39	27	17
Thoracic Criteria						
1. Chest Acceleration (g)	55	60	60	60	55	50
2. Chest Deflection (mm)	70 (2.8 in)	63 (2.5 in)	52 (2.0 in)	40 (1.6 in)	34 (1.4 in)	30* (1.2 in)
Lower Ext. Criteria: Femur Load (kN)	12.7	10.0	6.8	NA	NA	NA

Figure 1: National Highway Traffic Safety Administration Recommendations

National Highway Traffic Safety Administration Recommendations

The National Highway Traffic Safety Administration Recommendations provide a basis for vehicle safety for crashworthiness and rollover. The recommendation image provides a baseline for head, leg and chest injury. In summary, the closer the value is to absolute 0, the less likely the occupant will have a severe injury.

Preparing Environment

```r
#Loading Libraries
library(DBI)
library("knitr")
library("tidyverse")
```

```
## -- Attaching packages --------------------------------------------------- tidyverse 1.2.1 --
```

```
## v ggplot2 3.2.1     v purrr   0.3.2
## v tibble  2.1.3     v dplyr   0.8.3
## v tidyr   0.8.3     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.4.0
```

```
## -- Conflicts ------------------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library("stringr")
library("plotly")
```

```
##
## Attaching package: 'plotly'

## The following object is masked from 'package:ggplot2':
##
##     last_plot

## The following object is masked from 'package:stats':
##
##     filter

## The following object is masked from 'package:graphics':
##
##     layout
```

```r
library("htmlTable")
library("stringr")
library("stats")
library("scales")
```

```
##
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':
##
##     discard

## The following object is masked from 'package:readr':
##
##     col_factor
```

```r
library("viridis")
```

```
## Loading required package: viridisLite

##
## Attaching package: 'viridis'

## The following object is masked from 'package:scales':
##
##     viridis_pal
```

```r
library("wordcloud")
```

```
## Loading required package: RColorBrewer
```

Data Preparation for Summarized Analyses

Retrieve data from MySQL hosted in AWS

```r
cn <- dbConnect(drv      = RMySQL::MySQL(),
                username = "data622",
                password = "group4622",
                host     = "data622.c3alv39zxnhk.us-east-2.rds.amazonaws.com",
                port     = 3306,
                dbname   = "")
data <- dbGetQuery(cn, "SELECT
RD.` MAKED`
,MAX(`Head Injury Criterion`) AS HIC
,MAX(`Left Femur Load`) AS LFL
,MAX(`Right Femur Load`) AS RFL
,MIN(`Head Injury Criterion`) AS min_HIC
,MIN(`Left Femur Load`) AS min_LFL
,MIN(`Right Femur Load`) AS min_RFL
FROM
(SELECT v.` MAKED`
        , v.` YEAR`
        , AVG(convert(o.` HIC`, SIGNED INTEGER)) AS `Head Injury Criterion`
        , AVG(convert(o.` LFEM` , SIGNED INTEGER)) AS `Left Femur Load`
        , AVG(convert(o.` RFEM` , SIGNED INTEGER)) AS `Right Femur Load`
  FROM NHSA.test t
  inner join NHSA.veh v
  on(v.TSTNO = t.TSTNO)
  inner join NHSA.occ o
  on(o.TSTNO = v.TSTNO)
  inner join NHSA.rest r
  on(r.TSTNO = v.TSTNO)
  AND (r.` VEHNO` = v.` VEHNO`)
WHERE v.` MAKED` != 'NHTSA'
    AND v.` MAKED` NOT IN ('MCI')
    AND  v.` YEAR` != ''
    AND v.` YEAR` != 0
    AND  (o.` LFEM` != '' AND o.` LFEM` != 0)
    AND (o.` RFEM` != '' AND o.` RFEM` != 0)
    AND r.` DEPLOYD` = 'DEPLOYED PROPERLY'
GROUP BY
    v.` MAKED`
    ,v.` YEAR`
) AS RD
GROUP BY
RD.` MAKED`
ORDER BY
RD.` MAKED`")
```

```
## Warning in .local(conn, statement, ...): Decimal MySQL column 1 imported as
## numeric

## Warning in .local(conn, statement, ...): Decimal MySQL column 2 imported as
## numeric

## Warning in .local(conn, statement, ...): Decimal MySQL column 3 imported as
## numeric
```

```
## Warning in .local(conn, statement, ...): Decimal MySQL column 4 imported as
## numeric

## Warning in .local(conn, statement, ...): Decimal MySQL column 5 imported as
## numeric

## Warning in .local(conn, statement, ...): Decimal MySQL column 6 imported as
## numeric
```

```
head(data)
```

```
##        MAKED       HIC       LFL       RFL min_HIC min_LFL   min_RFL
## 1      ACURA  723.1429   53.5000 -329.5000  0.0000 -6788.0 -7036.750
## 2       AUDI  627.5000   44.5000 -195.6667  0.0000 -6409.5 -5085.500
## 3        BMW 1471.7059 1000.0000  999.0000  0.0000 -6962.5 -7828.000
## 4      BUICK 1273.5000   46.5000 -712.1000  0.0000 -7157.5 -7034.750
## 5   CADILLAC  666.5000   47.5000 -764.7500  0.0000 -5409.5 -5183.250
## 6 CHEVROLET 1116.2679 -625.0769 -833.3846 19.3636 -4621.5 -6234.546
```

Data Cleaning

Convert to numerics and truncate

```
data$HIC <- as.numeric(gsub(",","", data$HIC))
data$LFL <- as.numeric(gsub(",","", data$LFL))
data$RFL <- as.numeric(gsub(",","", data$RFL))
data$min_HIC <- as.numeric(gsub(",","", data$min_HIC))
data$min_LFL <- as.numeric(gsub(",","", data$min_LFL))
data$min_RFL <- as.numeric(gsub(",","", data$min_RFL))
data$HIC <- trunc(data$HIC)
data$LFL <- trunc(data$LFL)
data$RFL <- trunc(data$RFL)
data$min_HIC <- trunc(data$min_HIC)
data$min_LFL <- trunc(data$min_LFL)
data$min_RFL <- trunc(data$min_RFL)
data$MAKED <- as.factor(data$' MAKED')
```

```
str(data)
```

```
## 'data.frame':    49 obs. of  8 variables:
##  $  MAKED : chr  "ACURA" "AUDI" "BMW" "BUICK" ...
##  $ HIC    : num  723 627 1471 1273 666 ...
##  $ LFL    : num  53 44 1000 46 47 ...
##  $ RFL    : num  -329 -195 999 -712 -764 ...
##  $ min_HIC: num  0 0 0 0 0 19 0 62 0 0 ...
##  $ min_LFL: num  -6788 -6409 -6962 -7157 -5409 ...
##  $ min_RFL: num  -7036 -5085 -7828 -7034 -5183 ...
##  $ MAKED  : Factor w/ 49 levels "ACURA","AUDI",..: 1 2 3 4 5 6 7 8 9 10 ...
```

Challenges

One of the challenges encountered was to find a way to simplify the analyses, given the large amount of data, in order to present it in a way that made the most sense to the reader. I Focused on Head Injury Criterion as I believe this to be one of the most important injuries of concern in the unfortunate event someone is involved in a car accident.

HIC Discussion

The Head Injury Criterion (HIC) is a measure of the likelihood of head injury arising from an impact. The summarized data for HCI was generated by taking the maximum of HCI value for each vehicle make.

The HIC can be used to assess safety related to vehicles, personal protective gear, and sport equipment.

Normally the variable is derived from the measurements of an accelerometer mounted at the center of mass of a crash test dummy's head, when the dummy is exposed to crash forces. This means that the HIC includes the effects of head acceleration and the duration of the acceleration. Large accelerations may be tolerated for very short times.

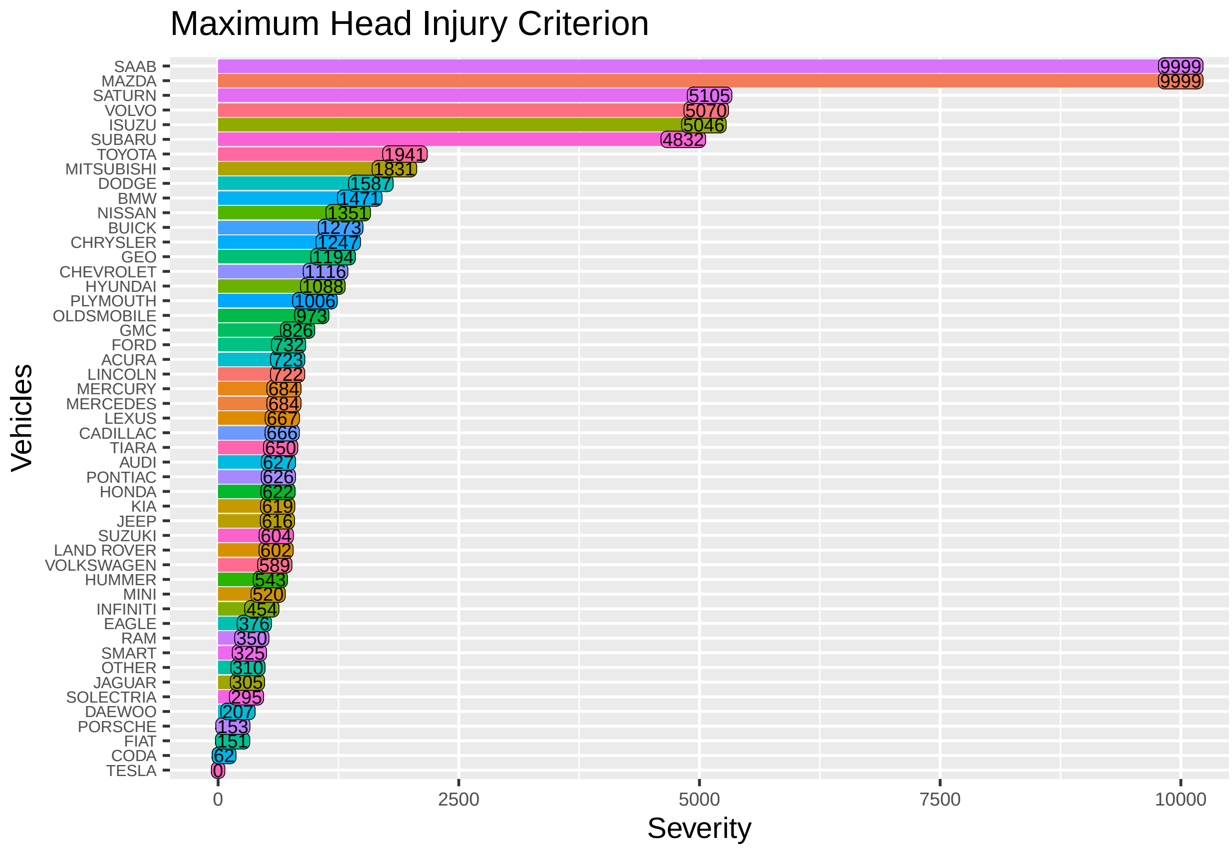
At a HIC of 1000, there is an 18% probability of a severe head injury, a 55% probability of a serious injury and a 90% probability of a moderate head injury to the average adult.

Summarized Analyses

What is the maximum Head Injury Criterion per vehicle?

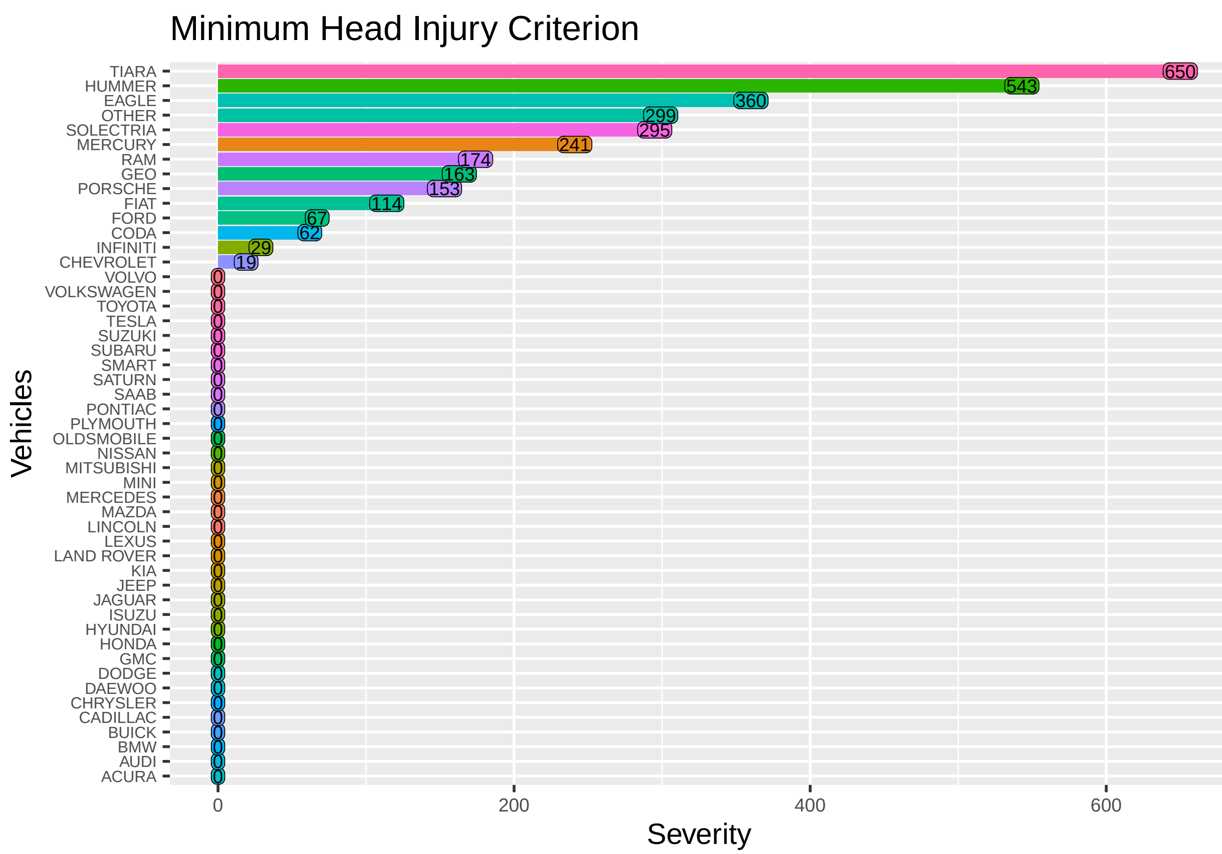
Measured in integers, 0 to 9,999, HIC is the computed value of the head injury criterion, based on the resultant acceleration pulse for the head center of gravity. (https://www.intmath.com/applications-integration/hic-part2.php) Generally, experts agree that Head Injury Criterion (HIC) values above 1000 are life threatening.

```
grid1 <- ggplot(data = data,aes(x=reorder(data$MAKED, data$HIC), y=data$HIC, fill = viridis(49), )) +
  theme(legend.position = "none", axis.text.y = element_text(size=6), axis.text.x = element_text(size=7
  geom_bar(stat = "identity") +
  geom_label(aes(label=data$HIC), position = position_dodge(width = 0.5), size = 2.4,   label.padding =
  labs(title = "Maximum Head Injury Criterion", x = "Vehicles", y = "Severity")+
  coord_flip()
grid1
```



What is minimum Head Injury Criterion per vehicle?

```
grid4 <- ggplot(data = data,aes(x=reorder(data$MAKED, data$min_HIC), y=data$min_HIC, fill = viridis(49)
  theme(legend.position = "none", axis.text.y = element_text(size=6), axis.text.x = element_text(size=7
  geom_bar(stat = "identity") +
  geom_label(aes(label=data$min_HIC), position = position_dodge(width = 0.5), size = 2.4,   label.paddi
  labs(title = "Minimum Head Injury Criterion", x = "Vehicles", y = "Severity")+
  coord_flip()
grid4
```



The summarized view of the data for HCI does not reveal the details necessary for making a more informed decision about a vehicle. For example, suppose a potential vehicle buyer wanted to know these values based on the year, type of vehicle, location of the occputant, etc. The summarized views do not provide this level of detail, so it becomes necessary to retrieve data with year and type of vehicle, among other attributes.

Data Preparation with Additional Attributes

Retrieve data from MySQL hosted in AWS

```
cn <- dbConnect(drv      = RMySQL::MySQL(),
                username = "data622",
                password = "group4622",
                host     = "data622.c3alv39zxnhk.us-east-2.rds.amazonaws.com",
                port     = 3306,
                dbname   = "")
data2 <- dbGetQuery(cn, "SELECT v.` MAKED`
        , v.` YEAR`
        , v.` BODYD`
        , o.` OCCLOCD`
        , o.` HIC` AS HIC
        , o.` LFEM` AS LFL
        , o.` RFEM` AS RFL
  FROM NHSA.test t
  inner join NHSA.veh v
```

```
  on(v.TSTNO = t.TSTNO)
  inner join NHSA.occ o
  on(o.TSTNO = v.TSTNO)
  inner join NHSA.rest r
  on(r.TSTNO = t.TSTNO)
WHERE v.` MAKED` != 'NHTSA'
    AND v.` MAKED` NOT IN ('MCI', 'OTHER')
    AND  (o.` LFEM` != '' AND o.` LFEM` != 0)
    AND (o.` RFEM` != '' AND o.` RFEM` != 0)
    AND (r.` DEPLOYD` = N'DEPLOYED PROPERLY')
    AND (o.` HIC` != '' AND o.` HIC` != 0)
    AND (v.` YEAR` != 0 and v.` YEAR` != '')
GROUP BY
v.` MAKED`
        , v.` YEAR`
        , v.` BODYD`
        , o.` OCCLOCD`
        , o.` HIC`
        , o.` LFEM`
        , o.` RFEM`
ORDER BY v.` MAKED` ")
```

```
head(data2)
```

```
##    MAKED  YEAR           BODYD          OCCLOCD HIC   LFL   RFL
## 1  ACURA  1988 FOUR DOOR SEDAN  LEFT FRONT SEAT 284 -6886 -7344
## 2  ACURA  1988 FOUR DOOR SEDAN RIGHT FRONT SEAT 387 -2931 -5494
## 3  ACURA  1992 FOUR DOOR SEDAN  LEFT FRONT SEAT 601 -9230 -8229
## 4  ACURA  1992 FOUR DOOR SEDAN  LEFT FRONT SEAT 897 -1326 -5534
## 5  ACURA  1992 FOUR DOOR SEDAN  LEFT FRONT SEAT 914 -3007 -7371
## 6  ACURA  1992 FOUR DOOR SEDAN RIGHT FRONT SEAT 433 -2140 -1401
```

Data Cleaning

Convert to numerics, factor and truncate

```
data2$HIC <- as.numeric(gsub(",","", data2$HIC))
data2$LFL <- as.numeric(gsub(",","", data2$LFL))
data2$RFL <- as.numeric(gsub(",","", data2$RFL))
data2$YEAR <- as.numeric(data2$' YEAR')
data2$BODYD <- as.factor(data2$' BODYD')
data2$MAKED <- as.factor(data2$' MAKED')
data2$OCCLOCD <- as.factor(data2$' OCCLOCD')
data2$HIC <- trunc(data2$HIC)
data2$LFL <- trunc(data2$LFL)
data2$RFL <- trunc(data2$RFL)
```

```
str(data2)
```

```
## 'data.frame':    4049 obs. of  11 variables:
```

```
##  $  MAKED  : chr  "ACURA" "ACURA" "ACURA" "ACURA" ...
##  $  YEAR   : chr  "1988" "1988" "1992" "1992" ...
##  $  BODYD  : chr  "FOUR DOOR SEDAN" "FOUR DOOR SEDAN" "FOUR DOOR SEDAN" "FOUR DOOR SEDAN" ...
##  $  OCCLOCD: chr  "LEFT FRONT SEAT" "RIGHT FRONT SEAT" "LEFT FRONT SEAT" "LEFT FRONT SEAT" ...
##  $ HIC     : num  284 387 601 897 914 433 660 443 521 334 ...
##  $ LFL     : num  -6886 -2931 -9230 -1326 -3007 ...
##  $ RFL     : num  -7344 -5494 -8229 -5534 -7371 ...
##  $ YEAR    : num  1988 1988 1992 1992 1992 ...
##  $ BODYD   : Factor w/ 16 levels "4 DOOR PICKUP",..: 5 5 5 5 5 5 5 5 5 11 ...
##  $ MAKED   : Factor w/ 48 levels "ACURA","AUDI",..: 1 1 1 1 1 1 1 1 1 1 ...
##  $ OCCLOCD : Factor w/ 8 levels "CENTER REAR SEAT",..: 3 6 3 3 3 6 6 3 6 3 ...
```

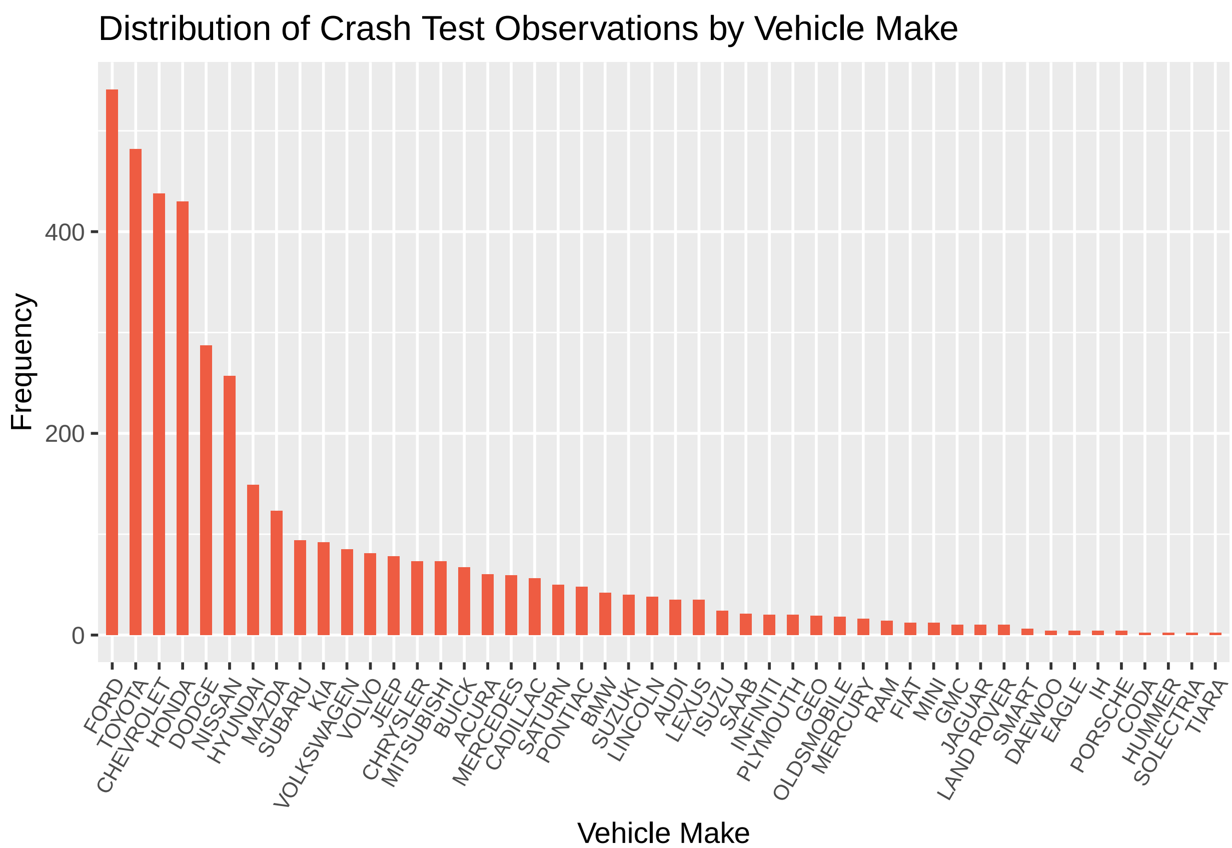
Data Analyses with Additional Attributes

In order to simplify the analyses I have decided to separete head injury criterion and compare them against the attributes I thought were most relevant. I will add a column that averages the HCI by make, year, body type of the vehicle and occupant location in the vehicle.

Crash Test Distribution

The vehicle make that has the highest frequency of crash test observations in the data is Ford, and is then followed by Toyota, Chevrolet, Honda, Dodge, and Nissan to name a few.

```
attrMake <- data2 %>% group_by(MAKED) %>% summarise("Average HIC" = mean(HIC), Count = n())
ggplot(attrMake, aes(x=reorder(MAKED, -Count), y=Count)) + geom_bar(stat="identity", width = 0.5, fill =
```



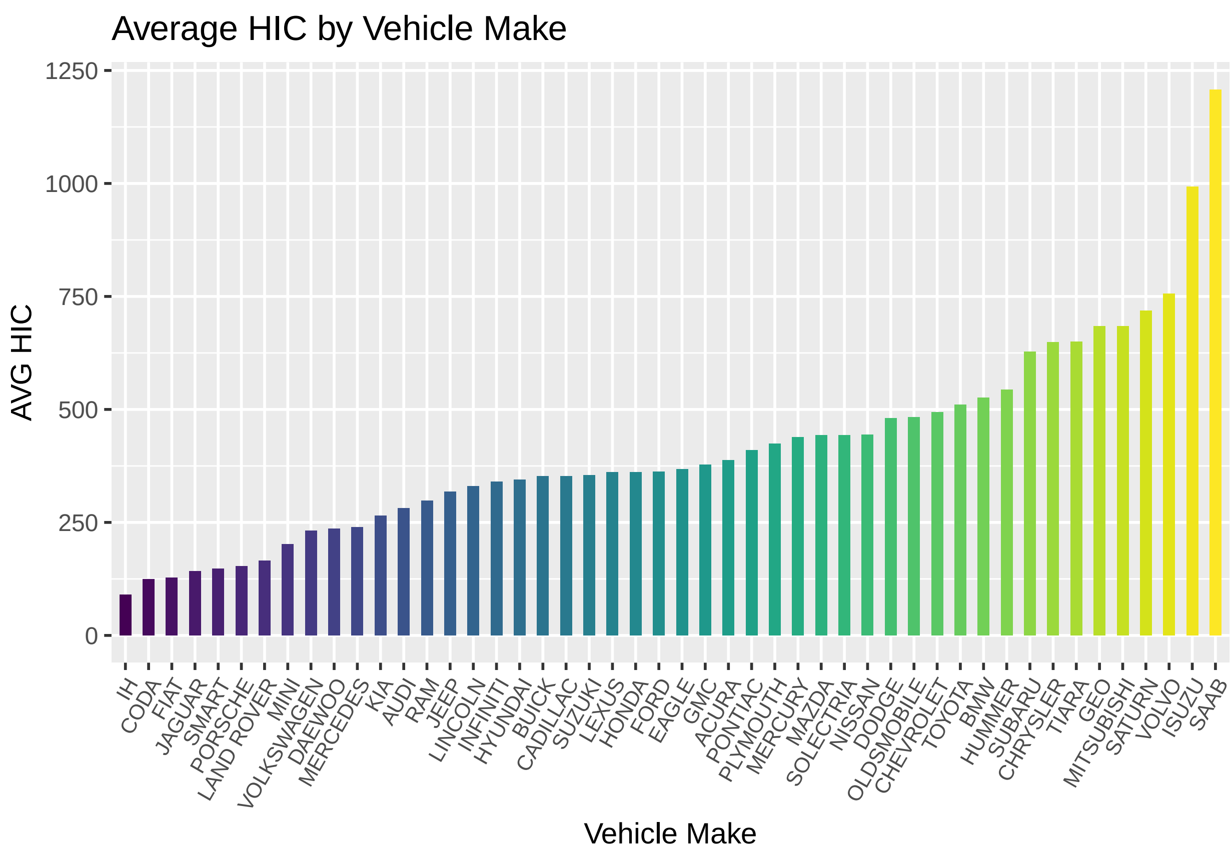
Data Sample by Vehicle Make

```r
head(attrMake)
```

```
## # A tibble: 6 x 3
##   MAKED     `Average HIC` Count
##   <fct>             <dbl> <int>
## 1 ACURA              388.    60
## 2 AUDI               282.    35
## 3 BMW                527.    42
## 4 BUICK              352.    67
## 5 CADILLAC           353.    56
## 6 CHEVROLET          494.   438
```

In the graphs below we can see from the crash tests that among the safest vehicle makes from a head injury perspective we find IH, Coda, Fiat, Jaguar, Smart Car, and Porsche among others.

```r
ggplot(attrMake, aes(x=reorder(MAKED, `Average HIC`), y=`Average HIC`)) + geom_bar(stat="identity", wid
```



Injury by Year

```
attrYear <- data2 %>% group_by(YEAR) %>% summarise("Average HIC" = mean(HIC), Count = n())
```

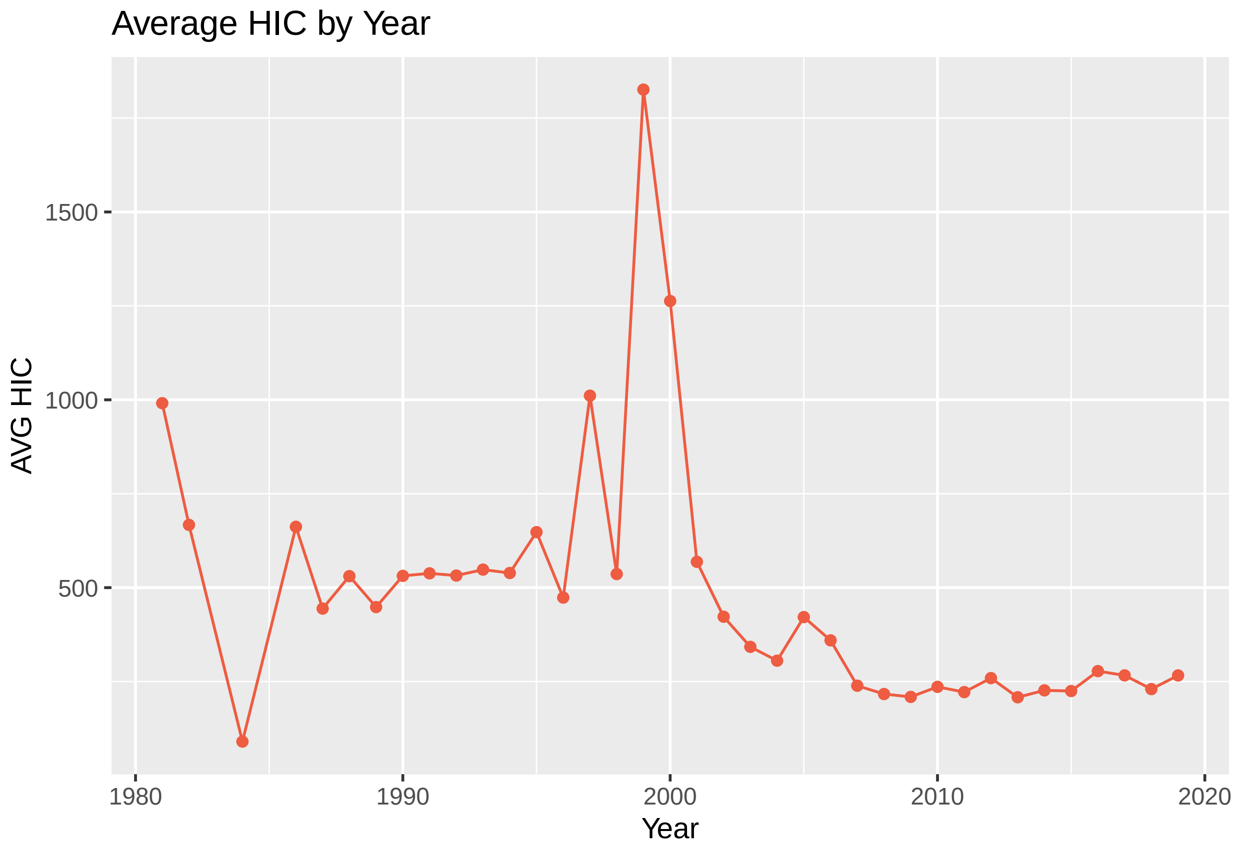
Data Sample by Year

```
head(attrYear)
```

```
## # A tibble: 6 x 3
##    YEAR `Average HIC` Count
##   <dbl>         <dbl> <int>
## 1  1981         991       2
## 2  1982         667       2
## 3  1984          90.2     4
## 4  1986         662       2
## 5  1987         444.      4
## 6  1988         531.     15
```

During 1999 The National Highway Traffic Safety Administration planned for upgrading the Federal Motor Vehicle Safety Standard (FMVSS). They added new crash specifications that required the use of additional dummies of various sizes as well as additional performance criteria that appropriately represent head injury thresholds.

```r
ggplot(attrYear, aes(x=YEAR, y=`Average HIC`)) + geom_line(color = "tomato2") + labs(x = "Year", y = "A
```



Injury by Vehicle Body Type

```r
attrBody <- data2 %>% group_by(BODYD) %>% filter (BODYD != "OTHER") %>% summarise("Average HIC" = mean(
```

Data Sample by Vehicle Body Type

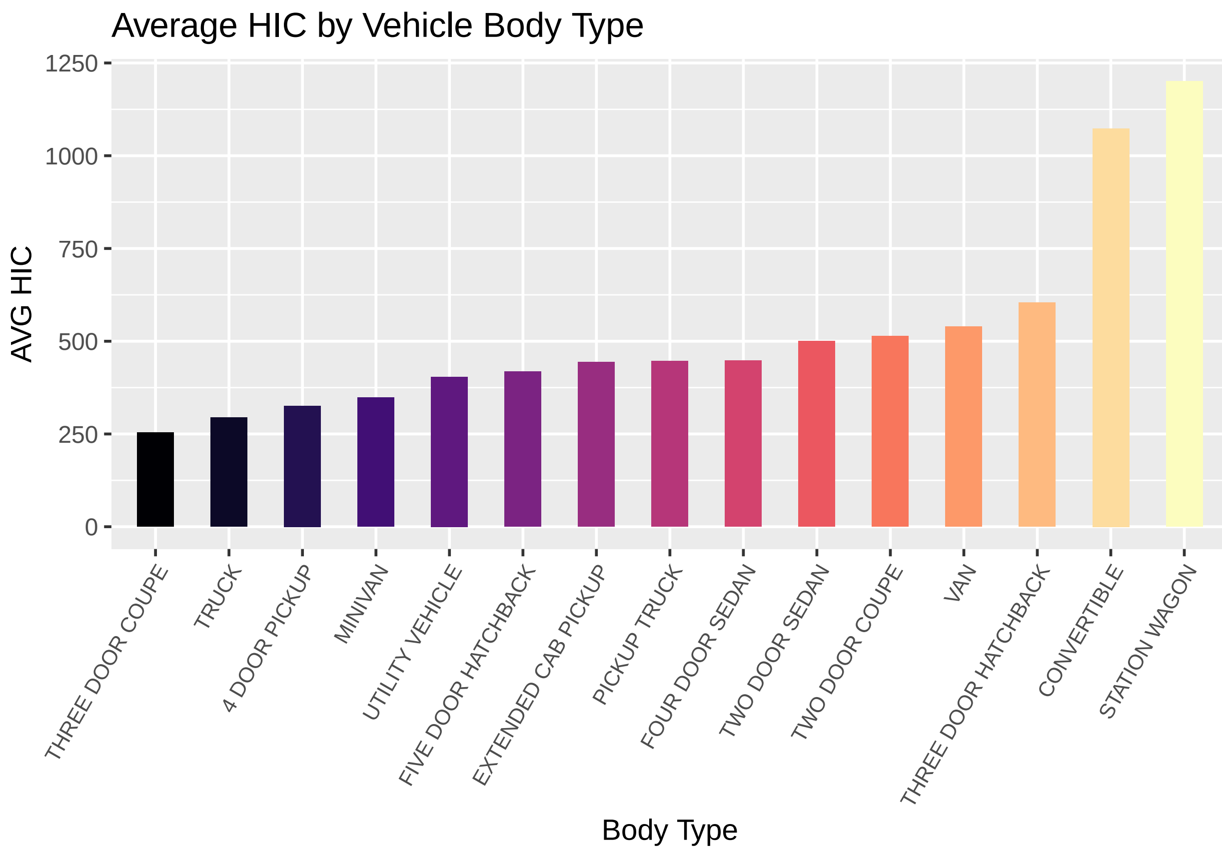
```r
head(attrBody)
```

```
## # A tibble: 6 x 3
##   BODYD                `Average HIC` Count
##   <fct>                        <dbl> <int>
## 1 4 DOOR PICKUP                 326.    97
## 2 CONVERTIBLE                  1074.    52
## 3 EXTENDED CAB PICKUP           444.    85
## 4 FIVE DOOR HATCHBACK           418.   184
## 5 FOUR DOOR SEDAN               448.  1920
## 6 MINIVAN                       349.    65
```

According to the graphs below, among the safest vehicle body types from a head injury perspective we find that three door coupes, Trucks, 4 door pickups, and Minivans have the lowest injury averages.

Please Note: The category "OTHER" for body type has been removed as it is not clear what type of vehicles are included

```
ggplot(attrBody, aes(x=reorder(BODYD, `Average HIC`), y=`Average HIC`)) + geom_bar(stat="identity", wid
```



Injury by Occupant Location

```
attrOcc <- data2 %>% group_by(OCCLOCD) %>% summarise("Average HIC" = mean(HIC), Count = n())
```

Data Sample by Occupant Location

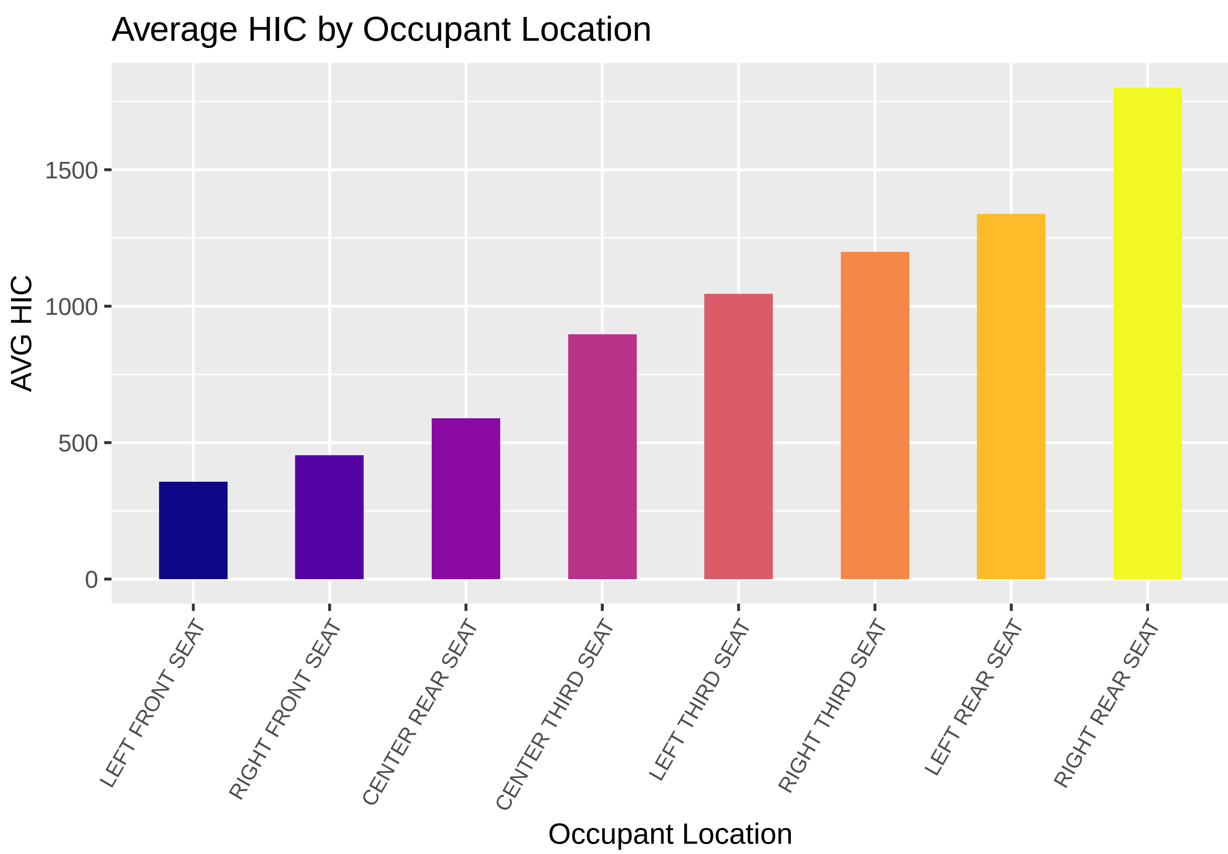
```
head(attrOcc)
```

```
## # A tibble: 6 x 3
##   OCCLOCD            `Average HIC` Count
##   <fct>                      <dbl> <int>
## 1 CENTER REAR SEAT            588.    15
## 2 CENTER THIRD SEAT           896.     2
## 3 LEFT FRONT SEAT             356.  2188
```

```
## 4 LEFT REAR SEAT             1338.    109
## 5 LEFT THIRD SEAT            1045.      4
## 6 RIGHT FRONT SEAT            453.   1685
```

According to the graphs below, one of the safest places to sit in a vehicle in the unfortunate event of a crash is the left front seat. In other words, you are the most safe when you are the driver. The second safest place to sit is in the front passenger seat.

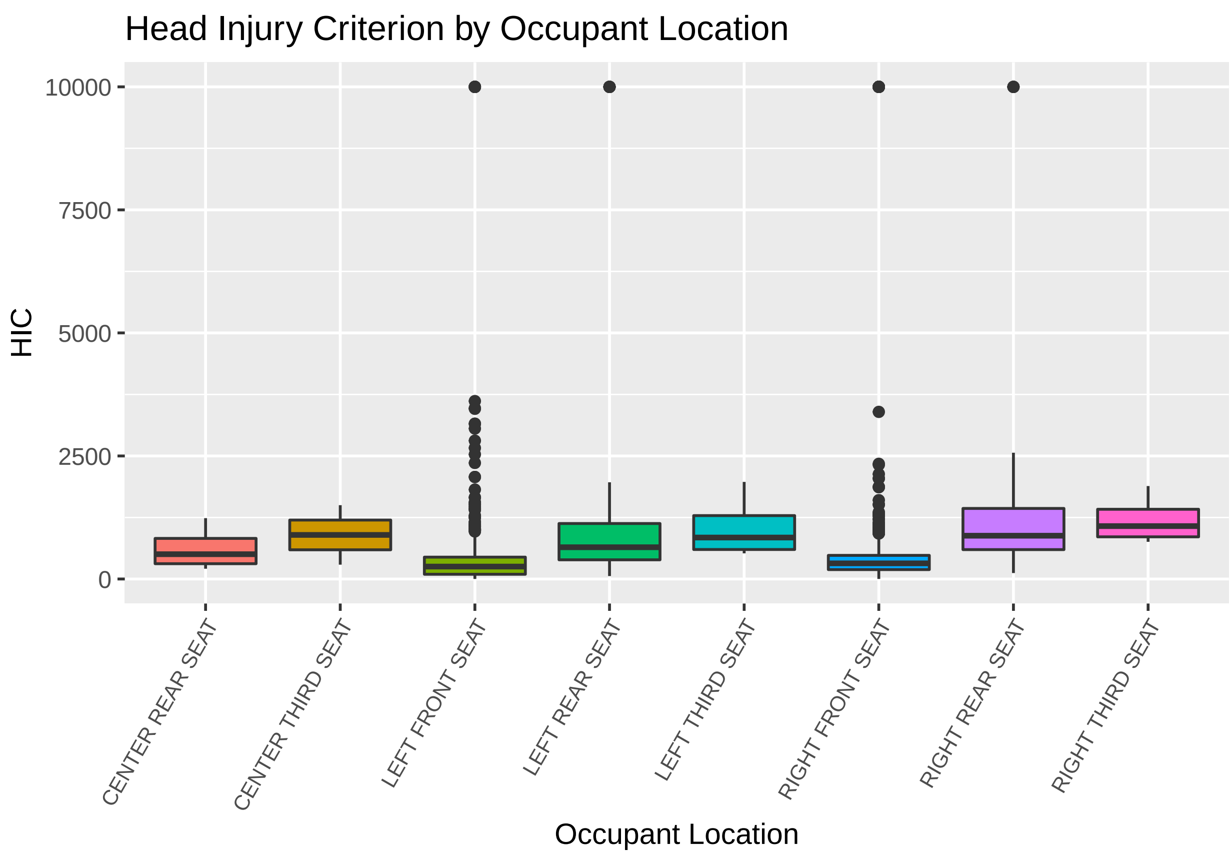
```
ggplot(attrOcc, aes(x=reorder(OCCLOCD, `Average HIC`), y=`Average HIC`)) + geom_bar(stat="identity", wi
```



Additional Analyses

Below we have another perspective for head injury criterion by occupant location looking at the data through boxplots. We can see that our boxplots agree with our analyses above, the two safest places to sit in a vehicle are in the driver seat and front passenger seat.

```
qplot(OCCLOCD, HIC, data = data2, geom= "boxplot", fill = OCCLOCD) + labs(x = "Occupant Location", y =
```

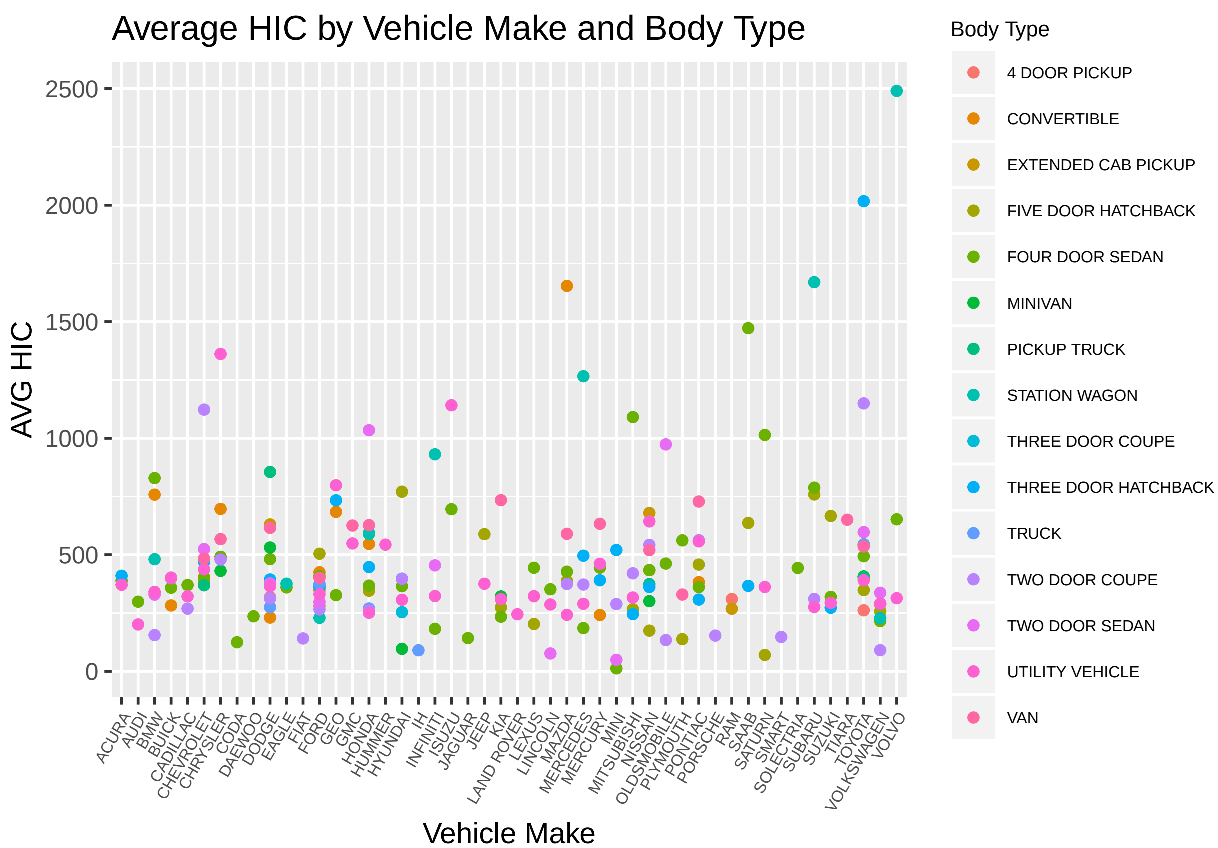


We can also look at our data by comparing vehicle make, body type and head injury criterion in the same graph in order to have a better visualization of our aggregate data.

Please Note: I have removed an observation (Chevrolet convertible) that was very far from the rest of the data in order to make the graph more readable. Additionally, the category "OTHER" for body type has been removed as it is not clear what type of vehicles are included

```
attrMakeBody <- data2 %>% group_by(MAKED, BODYD) %>% filter (MAKED != "CHEVROLET" | BODYD != "CONVERTIBL
```

```
ggplot(attrMakeBody, aes(MAKED, `Average HIC`)) + geom_point(aes(color = BODYD)) + labs(x = "Vehicle Mal
```



Transformation

The data obtained from the more granular query has many more rows with repeating fields of data. To get a better understanding of this data, we will transform into a easier visual model.

```
library(data.tree)

data2$pathString <- paste("Vehicle_Make",
                          data2$YEAR,
                          data2$MAKED,
                          data2$BODYD,
                          data2$OCCLOCD,
                          sep = "/")
data2tree <- as.Node(data2)
```

Reasons for transformation: It is increasingly becoming difficult to view the results because of the number of records in the resulset. Every time another variable of interest is added it exponentially increases the output of records. The difficulty becomes in "visually" consuming the data without removing any of the results. The following data has been transformed into a tree structure and is still too large to consume visually, thus, I decided to answer the question using subsets of the data as depicted below.

```
print(data2tree, "HIC", "LFL", "RFL")
```

```
##                           levelName HIC    LFL    RFL
```

```
## 1   Vehicle_Make                      NA    NA    NA
## 2    ¦--1988                          NA    NA    NA
## 3    ¦   ¦--ACURA                     NA    NA    NA
## 4    ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 5    ¦   ¦       ¦--LEFT FRONT SEAT  284 -6886 -7344
## 6    ¦   ¦       °--RIGHT FRONT SEAT 387 -2931 -5494
## 7    ¦   ¦--DODGE                     NA    NA    NA
## 8    ¦   ¦   °--THREE DOOR HATCHBACK  NA    NA    NA
## 9    ¦   ¦       ¦--LEFT FRONT SEAT  194 -3011 -6397
## 10   ¦   ¦       °--RIGHT FRONT SEAT 319 -4408 -3670
## 11   ¦   ¦--FORD                      NA    NA    NA
## 12   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 13   ¦   ¦       ¦--LEFT FRONT SEAT  456 -5089 -5378
## 14   ¦   ¦       °--RIGHT FRONT SEAT 561 -2971 -2322
## 15   ¦   ¦--OLDSMOBILE                NA    NA    NA
## 16   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 17   ¦   ¦       ¦--LEFT FRONT SEAT  709 -6010 -9399
## 18   ¦   ¦       °--RIGHT FRONT SEAT 539 -2318 -2620
## 19   ¦   °--VOLVO                     NA    NA    NA
## 20   ¦       °--FOUR DOOR SEDAN       NA    NA    NA
## 21   ¦           ¦--LEFT FRONT SEAT  519 -5084 -7976
## 22   ¦           °--RIGHT FRONT SEAT 445 -3216 -2820
## 23   ¦--1992                          NA    NA    NA
## 24   ¦   ¦--ACURA                     NA    NA    NA
## 25   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 26   ¦   ¦       ¦--LEFT FRONT SEAT  914 -3007 -7371
## 27   ¦   ¦       °--RIGHT FRONT SEAT 660 -2940 -1517
## 28   ¦   ¦--BMW                       NA    NA    NA
## 29   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 30   ¦   ¦       ¦--LEFT FRONT SEAT  705 -5418 -5196
## 31   ¦   ¦       °--RIGHT FRONT SEAT 698 -3127 -2077
## 32   ¦   ¦--CHEVROLET                 NA    NA    NA
## 33   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 34   ¦   ¦       °--LEFT FRONT SEAT  960 -9190 -5480
## 35   ¦   ¦--DODGE                     NA    NA    NA
## 36   ¦   ¦   °--VAN                   NA    NA    NA
## 37   ¦   ¦       ¦--LEFT FRONT SEAT  407 -2046 -6944
## 38   ¦   ¦       °--RIGHT FRONT SEAT 427 -6210 -2607
## 39   ¦   ¦--FORD                      NA    NA    NA
## 40   ¦   ¦   ¦--CONVERTIBLE           NA    NA    NA
## 41   ¦   ¦   ¦   ¦--LEFT FRONT SEAT  811 -5556 -5298
## 42   ¦   ¦   ¦   °--RIGHT FRONT SEAT 128 -1076  -649
## 43   ¦   ¦   ¦--FOUR DOOR SEDAN       NA    NA    NA
## 44   ¦   ¦   ¦   ¦--LEFT FRONT SEAT  907 -6401 -3763
## 45   ¦   ¦   ¦   °--RIGHT FRONT SEAT 331 -6067 -3759
## 46   ¦   ¦   °--VAN                   NA    NA    NA
## 47   ¦   ¦       ¦--LEFT FRONT SEAT  698 -6797 -3545
## 48   ¦   ¦       °--RIGHT FRONT SEAT 723 -4212 -1521
## 49   ¦   ¦--GEO                       NA    NA    NA
## 50   ¦   ¦   °--THREE DOOR HATCHBACK  NA    NA    NA
## 51   ¦   ¦       ¦--LEFT FRONT SEAT   75 -6121 -6005
## 52   ¦   ¦       °--RIGHT FRONT SEAT 613 -1517 -1459
## 53   ¦   ¦--HONDA                     NA    NA    NA
## 54   ¦   ¦   ¦--FOUR DOOR SEDAN       NA    NA    NA
```

```
## 55   ¦   ¦   ¦   ¦--LEFT FRONT SEAT  612 -2574 -4664
## 56   ¦   ¦   ¦   °--RIGHT FRONT SEAT 712 -3630   -63
## 57   ¦   ¦   °--THREE DOOR HATCHBACK  NA    NA    NA
## 58   ¦   ¦       ¦--LEFT FRONT SEAT  302 -7908 -5954
## 59   ¦   ¦       °--RIGHT FRONT SEAT 119 -1214  -819
## 60   ¦   ¦--MITSUBISHI                NA    NA    NA
## 61   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 62   ¦   ¦       ¦--LEFT FRONT SEAT  679 -6788 -3487
## 63   ¦   ¦       °--RIGHT FRONT SEAT 472 -1993 -3692
## 64   ¦   ¦--NISSAN                    NA    NA    NA
## 65   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 66   ¦   ¦       ¦--LEFT FRONT SEAT  818 -6219 -3043
## 67   ¦   ¦       °--RIGHT FRONT SEAT 907 -4284 -4559
## 68   ¦   ¦--OLDSMOBILE                NA    NA    NA
## 69   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 70   ¦   ¦       ¦--LEFT FRONT SEAT  473 -4777 -5849
## 71   ¦   ¦       °--RIGHT FRONT SEAT 829 -5035 -5058
## 72   ¦   ¦--PLYMOUTH                  NA    NA    NA
## 73   ¦   ¦   °--VAN                   NA    NA    NA
## 74   ¦   ¦       ¦--LEFT FRONT SEAT  426 -1793 -3625
## 75   ¦   ¦       °--RIGHT FRONT SEAT 175  -231 -2202
## 76   ¦   ¦--PONTIAC                   NA    NA    NA
## 77   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 78   ¦   ¦       ¦--LEFT FRONT SEAT  360 -6122 -6291
## 79   ¦   ¦       °--RIGHT FRONT SEAT 768 -6668 -3134
## 80   ¦   ¦--SAAB                      NA    NA    NA
## 81   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 82   ¦   ¦       °--LEFT FRONT SEAT  361 -4786 -5534
## 83   ¦   ¦--TOYOTA                    NA    NA    NA
## 84   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 85   ¦   ¦       ¦--LEFT FRONT SEAT  428 -6183 -7139
## 86   ¦   ¦       °--RIGHT FRONT SEAT 649 -1361 -1210
## 87   ¦   °--VOLVO                     NA    NA    NA
## 88   ¦       °--FOUR DOOR SEDAN       NA    NA    NA
## 89   ¦           ¦--LEFT FRONT SEAT  282 -5774 -3376
## 90   ¦           °--RIGHT FRONT SEAT 835 -4333  -205
## 91   ¦--1993                          NA    NA    NA
## 92   ¦   ¦--ACURA                     NA    NA    NA
## 93   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 94   ¦   ¦       ¦--LEFT FRONT SEAT  443 -1500 -1950
## 95   ¦   ¦       °--RIGHT FRONT SEAT 521 -1329 -1745
## 96   ¦   ¦--BUICK                     NA    NA    NA
## 97   ¦   ¦   °--FOUR DOOR SEDAN       NA    NA    NA
## 98   ¦   ¦       ¦--LEFT FRONT SEAT  442 -4324 -3416
## 99   ¦   ¦       °--RIGHT FRONT SEAT 931 -2703 -2047
## 100  ¦   °--... 20 nodes w/ 90 sub    NA    NA    NA
## 101  °--... 34 nodes w/ 3715 sub      NA    NA    NA
```

Conclusion

In the first output, I selected a couple of vehicles with the least amount of injury to head. In addition, two additional ranges of vehicles are provided, where minimal head injuries were reported. These display results with vehicle year, vehicle body type, and occupant location in the vehicle. The occupant location, is the test dummy used to measure different crash impact forces.

Difficulty Encountered: As mentioned above, visualizing the large number of rows is difficult. This presented challenges in displaying the results so a consumer could easily find their vehicle of choice. So, I decided to present the results in subsets of the overall results.

References:

https://www.intmath.com/applications-integration/hic-part2.php

https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/vehdb-v4.pdf

https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/rev_criteria_0.pdf

https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/rev_criteria_0.pdf

https://en.wikipedia.org > wiki > Head_injury_criterion

https://www.safercar.gov/Vehicle-Shoppers