IST 687 – Introduction to Data Science

STEPHEN O OMONDI | Data Science@ Syracuse | SUID: 946934043

Thursday, January 4, 2019

Week 2: Home Work 2

1. **Step 1: What is the hp (hp stands for "horse power")**

    1) What is the highest hp? ***335***

    2) Which car has the highest hp? ***Maserati Bora***

2. **Step 2: Explore mpg (mpg stands for "miles per gallon")**

    1) What is the highest mpg? ***33.9***

    2) Which car has the highest mpg? ***Toyota Corolla***

    3) Create a sorted data frame, based on mpg.

3. **Step 3: Which car has the best combination of mpg and hp?**

    1) What logic did you use?

    I created a new column that is the average of mpg and hp for each vehicle. I bound this new column to the parent data frame then sorted the new data frame using the new average column in a descending order.

    2) Which Car?

    Maserati Bora emerged at the top of the list with 335hp and 15mpg for an average of 175.00 horse power miles per gallon.

```
# copying mtcars into a new variable: myCars
myCars <- mtcars

# Highest horse power - hp
maxhp <- max(myCars$hp)

# print maxhp
maxhp

# Car with highest hp
myCars[myCars$hp == maxhp,]



# print maxhpCar
maxhpCar

# Highest mpg
maxmpg <- max(myCars$mpg)

# Print highest mpg
maxmpg

# Car with highest mpg
maxmpgCar <- myCars[myCars$mpg == maxmpg,]

# Print car with highest mpg
maxmpgCar

# sorted dataframe based on mpg
mpgSort <- data.frame(myCars, stringsAsFactors = FALSE)
mpgSort[order(-mpgSort$mpg),] # decreasing order defined
```

```
# Best combination of mpg and hp
#calculate the mean of mpg and hp
mpghpAvg <- function (a,b) {
  (a + b)/2
}

# add a mean column to the dataframe as newCol
AVGmpghp <- mpghpAvg(myCars$mpg, myCars$hp)

# bind the new column to the myCars data frame
bindMeanCol <- cbind(myCars, AVGmpghp)

# sort the new dataframe using the new average column: bestMpgHp
bestMpgHp <- bindMeanCol[order(-bindMeanCol$AVGmpghp),]

# print the newly sorted data to ordered by best mpg hp average
bestMpgHp
```