

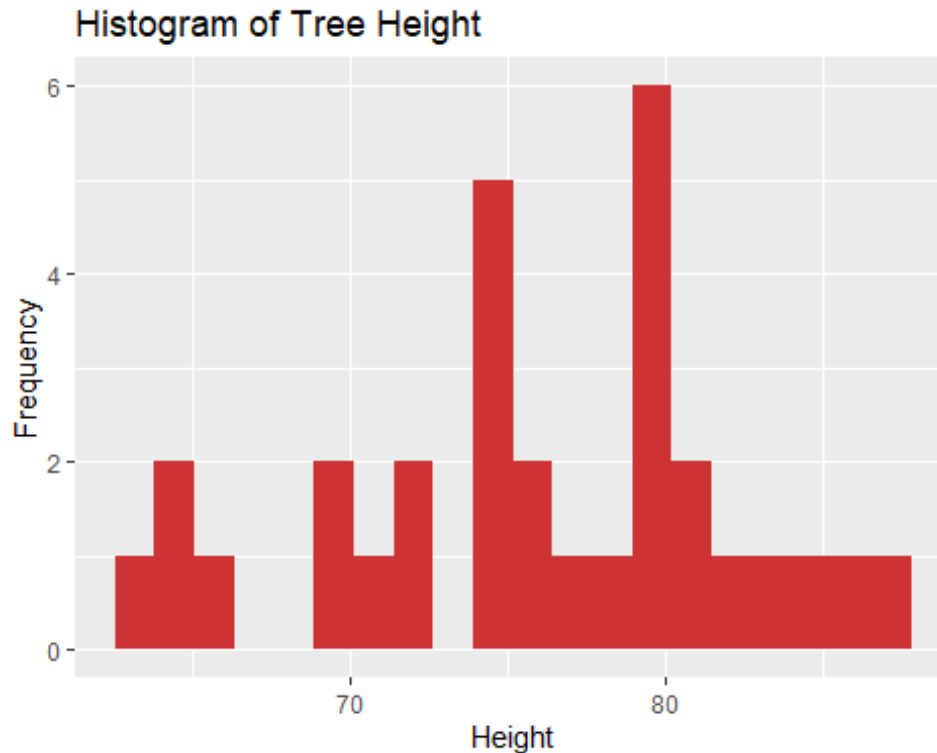
#R Module 2: Working with R Scripts and Projects

This R Module is intended to introduce summary statistics, measures of central tendency, and measures of dispersion. In terms of R concepts, it will teach students about how to work with files in the working directory, create and use scripts (rather than type all the code in the console!) and organize their R Modules into an R Project (.rproj).

```
# Load in the data
library(datasets)
tree_data <- trees
```

1. Create your own histogram of the tree_data dataset with ggplot2, showing:
 - Distribution of the Height variable,
 - 20 bins,
 - A color of your choosing,
 - Labels (title and x-axis) that reflect the change in variable

```
library(ggplot2)
ggplot(
  data = tree_data,
  aes(x = Height)
) +
  geom_histogram(bins = 20, fill = "brown3") +
  labs(
    x = "Height",
    y = "Frequency",
    title = "Histogram of Tree Height"
  )
```



Question 2:

Visualize and Explore the `mtcars` dataset: 1. Create a new script called `mtcars.R` 2. Assign `mtcars` to an object (e.g., “cars”) using the `<-` operator 3. Choose 3 variables from `mtcars` and for each: - Record the Mean, Median, and Mode - Create a histogram in `ggplot2`, with appropriate labels, title, etc.

```
# The code for this one should include their mtcars.R script as well

cars <- mtcars

# They should have three of the following, for each of the variables they choose:
mean(mtcars$mpg)

## [1] 20.09062

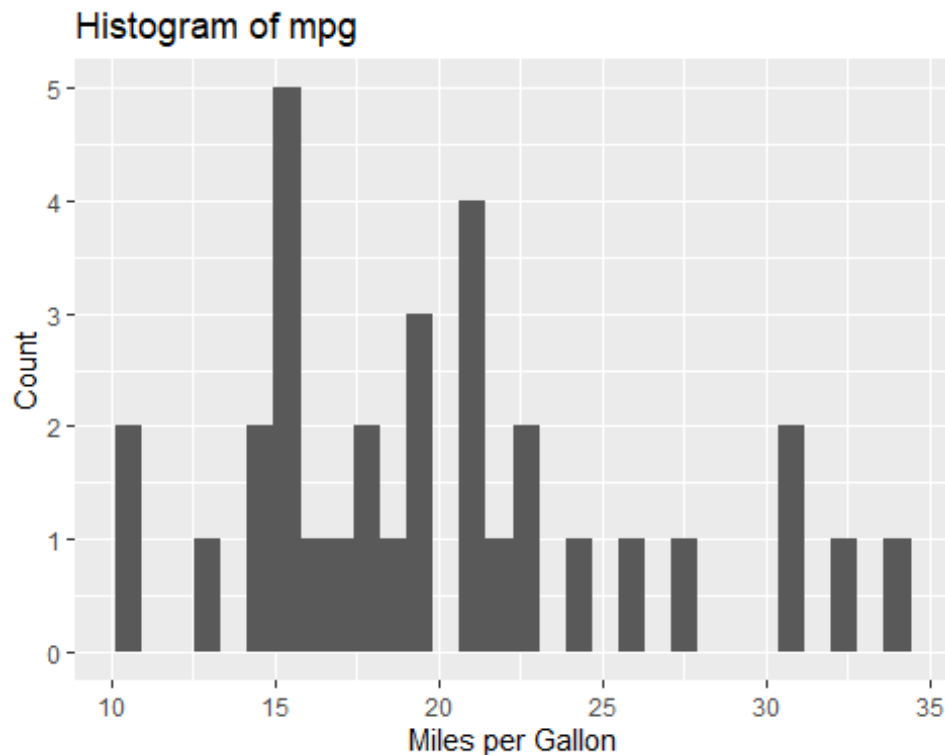
median(mtcars$mpg)

## [1] 19.2

# Any way they want to find the mode is fine. Oddly, base R doesn't have a function to find the mode.
find_mode <- function(x) {
  unique_data <- unique(x)
  unique_data[which.max(tabulate(match(x, unique_data)))]
}
```

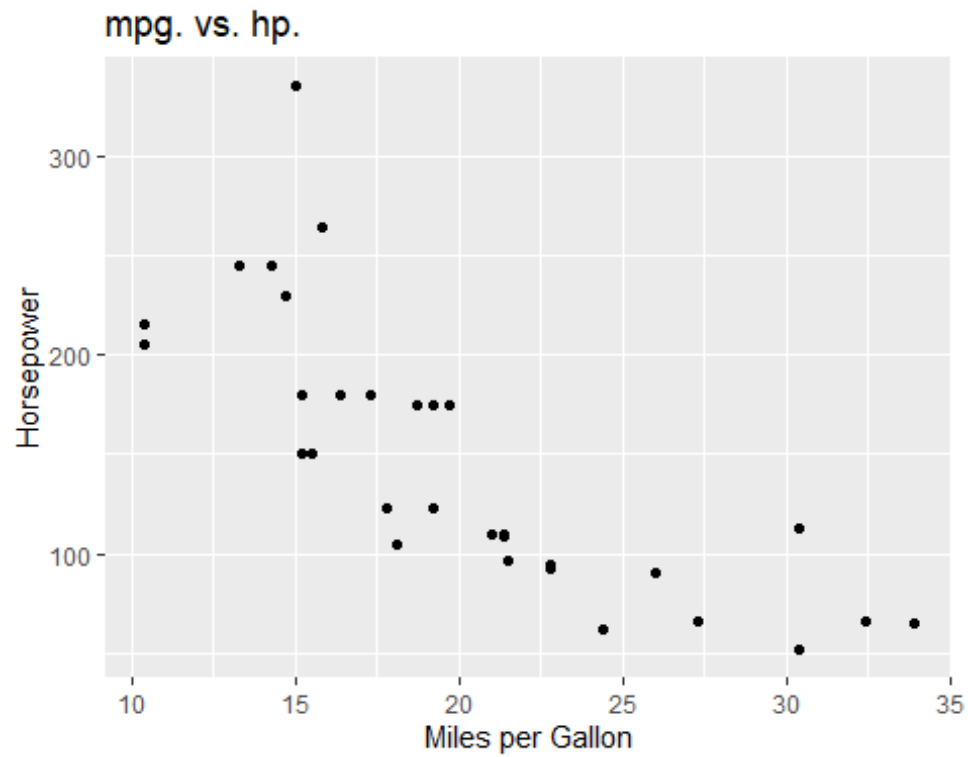
```
find_mode(mtcars$mpg)
## [1] 21

ggplot(mtcars, aes(x = mpg)) +
  geom_histogram() +
  labs(
    x = "Miles per Gallon",
    y = "Count",
    title = "Histogram of mpg"
  )
)
```



- d. Choose two variables from `mtcars` and create a scatter plot in `ggplot2` showing their relation:
- Include title, labels, etc.
 - *Hint: Use the `geom_point()` function*
 - Use Google, Stack Exchange, and the built-in Help for examples on how to use these plots!

```
ggplot(data = mtcars, aes(x = mpg, y = hp)) +
  geom_point() +
  labs(x = "Miles per Gallon",
    y = "Horsepower",
    title = "mpg. vs. hp.")
```



Include a screenshot of your `mtcars` R script Submit all questions to Google Classroom; include Name, Question Number, and all figures