EXPERIMENT 8.1

Taniya Ahmed

21BDS0059

1. **Loading Dataset and Packages.**

CODE:

```
install.packages("ggplot2")
install.packages("moments")
install.packages("psych")
library(ggplot2)
library(moments)
library(psych)

df = mtcars
head(df)
print("Taniya Ahmed 21BDS0059")
```

OUTPUT:

```
> head(df)
                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
> print("Taniya Ahmed 21BDS0059")
[1] "Taniya Ahmed 21BDS0059"
>
```

2. Measures of Central Tendency

CODE:

```
arith_mean = mean(df$disp)

geom_mean = exp(mean(log(df$disp)))

harm_mean = length(df$disp) / sum(1/df$disp)


median_val = median(df$disp)

quantile_val = quantile(df$disp)

decile_val = quantile(df$disp, probs = seq(0.1, 1, by = 0.1))

percentile_val = quantile(df$disp, probs = seq(0.01, 1, by = 0.01))


cat("Means are: \nArithmetic Mean:", arith_mean, "\nGeometric Mean:", geom_mean,
"\nHarmonic Mean:", harm_mean)

print("Taniya Ahmed 21BDS0059")


cat("Median and Quartiles are:\nMedian value:", median_val, "\n\nQuantile value:",
quantile_val, "\n\nDecile value:", decile_val, "\n\nPercentile value", percentile_val)

print("Taniya Ahmed 21BDS0059")
```

OUTPUT:

```
> arith_mean = mean(df$disp)
> geom_mean = exp(mean(log(df$disp)))
> harm_mean = length(df$disp) / sum(1/df$disp)
>
> median_val = median(df$disp)
> quantile_val = quantile(df$disp)
> decile_val = quantile(df$disp, probs = seq(0.1, 1, by = 0.1))
> percentile_val = quantile(df$disp, probs = seq(0.01, 1, by = 0.01))
>
> cat("Means are: \nArithmetic Mean:", arith_mean, "\nGeometric Mean:", geom_mean, "\nHarmonic Mean:", harm_mean)
Means are:
Arithmetic Mean: 230.7219
Geometric Mean: 197.3218
Harmonic Mean: 166.7994> print("Taniya Ahmed 21BDS0059")
[1] "Taniya Ahmed 21BDS0059"
>
> cat("Median and Quartiles are:\nMedian value:", median_val, "\n\nQuantile value:", quantile_val, "\n\nDecile value:", d
ecile_val, "\n\nPercentile value", percentile_val)
Median and Quartiles are:
Median value: 196.3

Quantile value: 71.1 120.825 196.3 326 472

Decile value: 80.61 120.14 142.06 160 196.3 275.8 303.1 350.8 396 472

Percentile value 72.526 73.952 75.378 76.42 77.35 78.28 78.751 78.844 78.937 80.61 85.601 90.592 95.487 99.486 103.485 10
7.484 111.267 115.018 118.769 120.14 120.202 120.264 120.391 120.608 120.825 122.188 128.326 134.464 140.602 142.06 143.3
62 144.664 145.391 145.918 146.445 148.828 152.951 157.074 160 160 160 160.152 162.508 164.864 167.22 167.6 167.6 167.6 1
78.506 196.3 214.094 228.96 239.19 249.42 258.89 264.408 269.926 275.444 275.8 275.8 275.8 275.8 275.8 275.8 279.58 287.3
92 295.204 301.24 302.17 303.1 304.14 308.48 312.82 317.16 326 335.92 345.84 350.18 350.49 350.8 351.99 354.78 357.57 360
360 360 360 371.2 383.6 396 408.4 420.8 433.2 442.8 449 455.2 460.84 464.56 468.28 472> print("Taniya Ahmed 21BDS0059")
[1] "Taniya Ahmed 21BDS0059"
```

3. **Measure of dispersion and Probability distribution description**

CODE:

```
range_val = range(df$hp)
range_diff = diff(range_val)

iqr_val = IQR(df$hp)
interdecile_range = quantile(df$drat, 0.9) - quantile(df$drat, 0.1)

mean_deviation = mean(abs(df$hp - mean(df$hp)))
sd_val = sd(data)

skew_val = skewness(df$drat)
kurt_val = kurtosis(df$drat)

cat("Range value: ", range_val)
cat("Range diff: ", range_diff)
cat("IQR value: ", iqr_val)
cat("Interdecile range: ", interdecile_range)
cat("Mean deviation: ", mean_deviation)
cat("Standard Deviation value: ", sd_val)
cat("Skewness value: ", skew_val)
cat("Kurtosis value: ", kurt_val)
print("Taniya Ahmed 21BDS0059")
```

OUTPUT:

```
> range_val = range(df$hp)
> range_diff = diff(range_val)
>
> iqr_val = IQR(df$hp)
> interdecile_range = quantile(df$drat, 0.9) - quantile(df$drat, 0.1)
>
> mean_deviation = mean(abs(df$hp - mean(df$hp)))
> sd_val = sd(data)
>
> skew_val = skewness(df$drat)
> kurt_val = kurtosis(df$drat)
>
> cat("Range value: ", range_val, "\n")
Range value:  52 335
> cat("Range diff: ", range_diff, "\n")
Range diff:  283
> cat("IQR value: ", iqr_val, "\n")
IQR value:  83.5
> cat("Interdecile range: ", interdecile_range, "\n")
Interdecile range:  1.202
> cat("Mean deviation: ", mean_deviation, "\n")
Mean deviation:  56.48047
> cat("Standard Deviation value: ", sd_val, "\n")
Standard Deviation value:  30.15958
> cat("Skewness value: ", skew_val, "\n")
Skewness value:  0.2788734
> cat("Kurtosis value: ", kurt_val, "\n")
Kurtosis value:  2.435116
> print("Taniya Ahmed 21BDS0059")
[1] "Taniya Ahmed 21BDS0059"
>
```

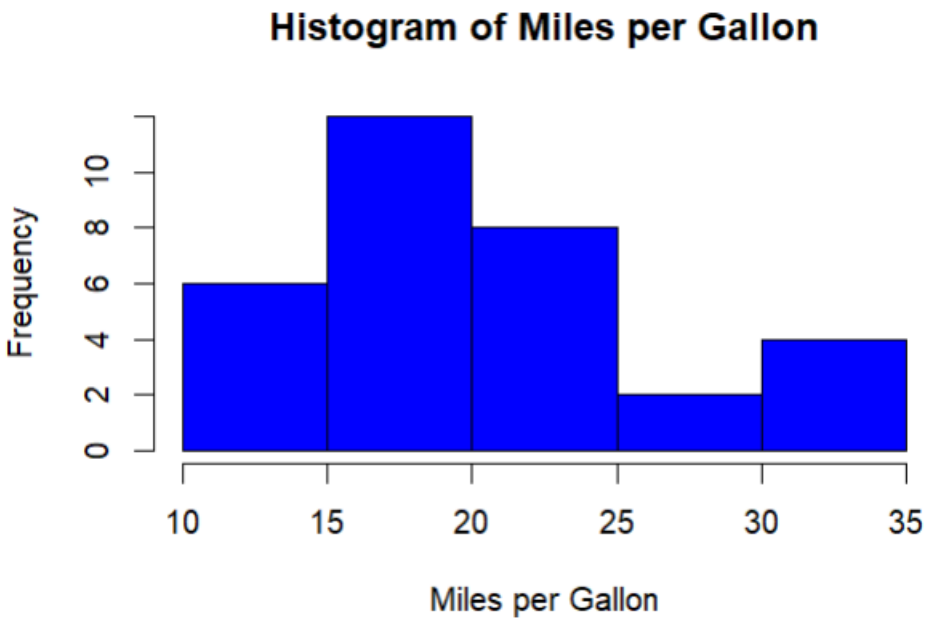
4. **Frequency distribution with plots.**

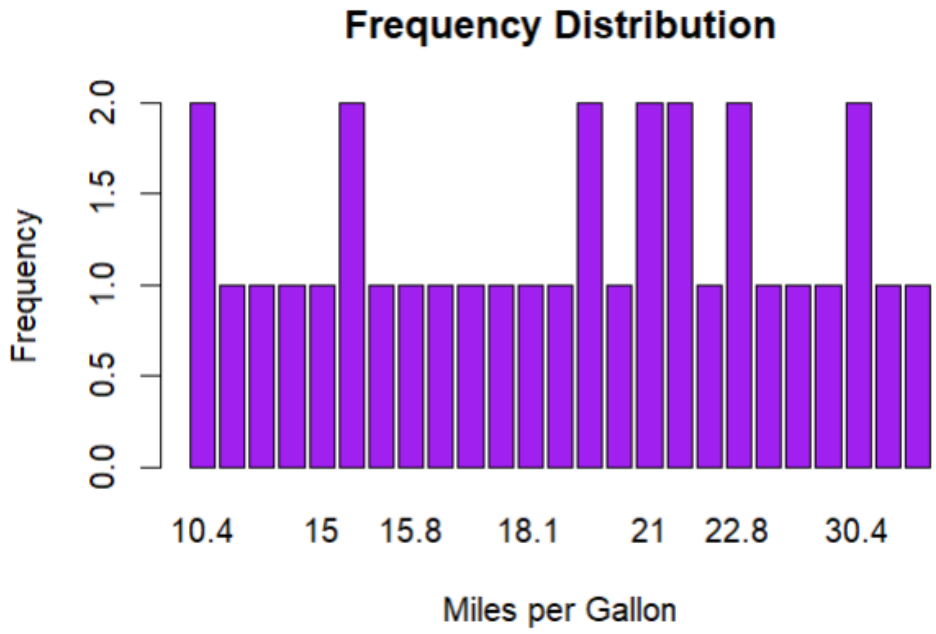
CODE:

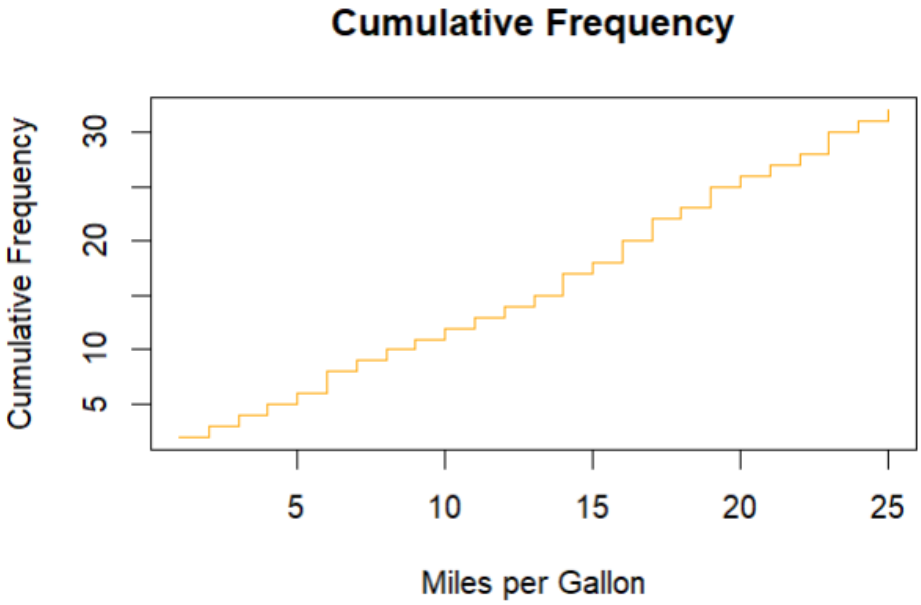
```
table_val = table(df$mpg)

relative_freq = prop.table(table_val)

cumulative_freq = cumsum(table_val)

print("Taniya Ahmed 21BDS0059")


hist(df$mpg, main = "Histogram of Miles per Gallon", xlab = "Miles per Gallon", ylab = "Frequency", col = "blue")

barplot(table_val, main = "Frequency Distribution", xlab = "Miles per Gallon", ylab = "Frequency", col = "purple")

plot(cumulative_freq, type = "s", main = "Cumulative Frequency", xlab = "Miles per Gallon", ylab = "Cumulative Frequency", col = "orange")

print("Taniya Ahmed 21BDS0059")
```

OUTPUT:

```
> table_val = table(df$mpg)
> relative_freq = prop.table(table_val)
> cumulative_freq = cumsum(table_val)
> print("Taniya Ahmed 21BDS0059")
[1] "Taniya Ahmed 21BDS0059"
```




Frequency Distribution
Frequency
0.0
0.5
1.0
1.5
2.0
10.4
15
15.8
18.1
21
22.8
30.4
Miles per Gallon


Cumulative Frequency
Cumulative Frequency
5
10
20
30
5
10
15
20
25
Miles per Gallon