EXPERIMENT 8.1

Name: Jvn Ganesh

Roll N.o: 21BDS0085

```
Console Terminal × Background Jobs ×
R 4.4.1 · ~/ ≈
> library(psych)
> library(ggplot2)
Learn more about the underlying theory at https://ggplot2-book.org/
Attaching package: 'ggplot2'
The following objects are masked from 'package:psych':
     %+%, alpha
> # Load the mtcars dataset
> data(mtcars)
> head(mtcars)
> head(mtcars)

mpg cyl disp hp drat wt qsec vs am

Mazda RX4 21.0 6 160 110 3.90 2.620 16.46 0 1

Mazda RX4 wag 21.0 6 160 110 3.90 2.875 17.02 0 1

Datsun 710 22.8 4 108 93 3.85 2.320 18.61 1 1

Hornet 4 Drive 21.4 6 258 110 3.08 3.215 19.44 1 0

Hornet sportabout 18.7 8 360 175 3.15 3.440 17.02 0 0

Valiant 18.1 6 225 105 2.76 3.460 20.22 1 0
                                                              wt qsec vs am gear carb
                                                                                1
> # Measures of Central Tendency
> print("JVNGANESH 21BDS0085")
[1] "JVNGANESH 21BDS0085"
> # Mean
> mean_values <- sapply(mtcars, mean)
> print("Mean Values:")
[1] "Mean Values:"
> print(mean_values)
 mpg cyl disp hp drat wt qsec vs am gear carb
20.090625 6.187500 230.721875 146.687500 3.596563 3.217250 17.848750 0.437500 0.406250 3.687500 2.812500
> # Median
> median_values <- sapply(mtcars, median)
> print("Median Values:")
2.000
> # Quantiles (25%, 50%, 75%)
| γ | quantiles | (25%, 50%, 75%): |
| > quantiles | - apply(mtcars, 2, quantile) |
| > print("quantiles (25%, 50%, 75%): |
| 1 | "quantiles (25%, 50%, 75%): |
| > print(quantiles) |
3
                                                                                              1
```

```
Console Terminal × Background Jobs ×
R 4.4.1 · ~/ ≈
> # Deciles (ntile)
> deciles <- apply(mtcars, 2, function(x) quantile(x, probs = seq(0, 1, 0.1)))
> print("Deciles:")
[1] "Deciles:"
> print(deciles)
          mpg cyl disp hp drat wt qsec vs am

0.40 4 71.10 52.0 2.760 1.5130 14.500 0 0.0

4.34 4 80.61 66.0 3.007 1.9555 15.534 0 0.0

5.20 4 120.14 93.4 3.072 2.3490 16.734 0 0.0

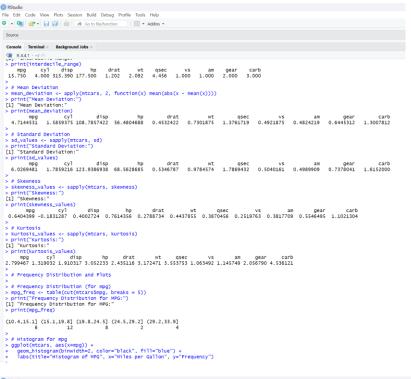
5.98 4 142.06 106.2 3.150 2.7730 17.020 0 0.0
                                                                                 am gear carb
                                                                                            3 1.0
3 1.0
0%
      10.40
10% 14.34
20%
      15.20
30%
       15.98
                                                                                            3 2.0
                    6 160.00 110.0 3.354 3.1580 17.340
                    6 196.30 123.0 3.695 3.3250 17.710
8 275.80 165.0 3.818 3.4400 18.180
50%
      19.20
                                                                              0 0.0
60%
       21.00
                                                                              1 0.6
                   8 303.10 178.5 3.914 3.5550 18.607
8 350.80 200.0 4.048 3.7700 19.332
8 396.00 243.5 4.209 4.0475 19.990
70%
       21.47
                                                                              1 1.0
80% 24.08
                                                                                            4 4.0
                                                                             1 1.0
        30.09
100% 33.90
                   8 472.00 335.0 4.930 5.4240 22.900 1 1.0
                                                                                            5 8.0
> # Percentiles (90th percentile)
> percentiles <- apply(mtcars, 2, function(x) quantile(x, probs = 0.9))
> print("90th Percentiles:")
[1] "90th Percentiles:"
> print(percentiles)
 mpg cyl disp hp drat wt qsec vs am gear carb
30.0900 8.0000 396.0000 243.5000 4.2090 4.0475 19.9900 1.0000 1.0000 5.0000 4.0000
> # Measures of Dispersion
> # Range
> range_values <- sapply(mtcars, function(x) diff(range(x)))</pre>
> print("Range:")
[1] "Range:"
> print(range_values)
 mpg cyl disp hp drat wt qsec vs am gear 23.500 4.000 400.900 283.000 2.170 3.911 8.400 1.000 1.000 2.000
                                                                                                                                    carb
> # Interquartile Range (IQR)
> # Interquartire Range (1907)
> iqr_values <- sapply(mtcars, IQR)
> print("Interquartile Range (IQR):")

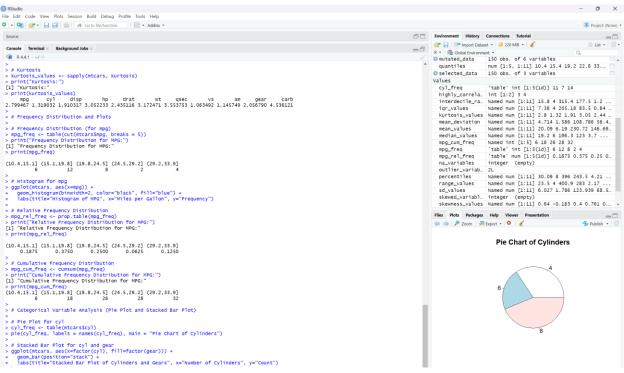
[1] "Interquartile Pange (TQR):"
[1] "Interquartile Range (IQR):
> print(iqr_values)
mpg cyl

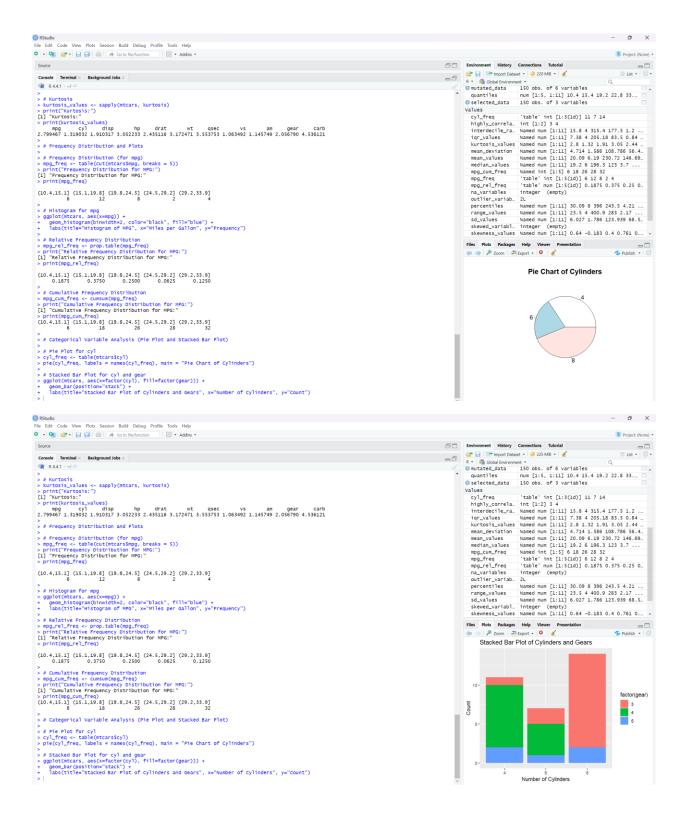
        mpg
        cyl
        disp
        hp
        drat
        wt
        qsec
        vs
        am
        gear
        carb

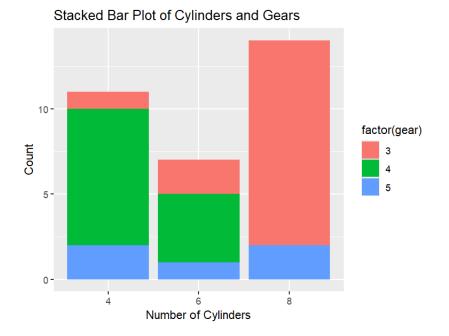
        7.37500
        4.00000
        205.17500
        83.50000
        0.84000
        1.02875
        2.00750
        1.00000
        1.00000
        1.00000
        2.00000

                                        disp
> # Interdecile Range
> interdecile_range <- apply(mtcars, 2, function(x) quantile(x, probs = 0.9) - quantile(x, probs = 0.1))
> print("Interdecile Range:")
mpg cyl disp hp drat wt qsec vs am yea carb
15.750 4.000 315.390 177.500 1.202 2.092 4.456 1.000 1.000 2.000 3.000
                                              hp
                                                       drat
                                                                        wt
```









Code:

Install and load required packages

install.packages("psych")

install.packages("moments")

install.packages("ggplot2")

library(psych)

library(moments)

library(ggplot2)

Load the mtcars dataset

data(mtcars)

head(mtcars)

Measures of Central Tendency

```
print("JVNGANESH 21BDS0085")
# Mean
mean_values <- sapply(mtcars, mean)</pre>
print("Mean Values:")
print(mean_values)
# Median
median_values <- sapply(mtcars, median)</pre>
print("Median Values:")
print(median_values)
# Quantiles (25%, 50%, 75%)
quantiles <- apply(mtcars, 2, quantile)
print("Quantiles (25%, 50%, 75%):")
print(quantiles)
# Deciles (ntile)
deciles <- apply(mtcars, 2, function(x) quantile(x, probs = seq(0, 1, 0.1)))
print("Deciles:")
print(deciles)
# Percentiles (90th percentile)
percentiles <- apply(mtcars, 2, function(x) quantile(x, probs = 0.9))
print("90th Percentiles:")
print(percentiles)
```

```
# Measures of Dispersion
# Range
range_values <- sapply(mtcars, function(x) diff(range(x)))</pre>
print("Range:")
print(range_values)
# Interquartile Range (IQR)
iqr_values <- sapply(mtcars, IQR)</pre>
print("Interquartile Range (IQR):")
print(iqr_values)
# Interdecile Range
interdecile_range <- apply(mtcars, 2, function(x) quantile(x, probs = 0.9) - quantile(x, probs
= 0.1))
print("Interdecile Range:")
print(interdecile_range)
# Mean Deviation
mean_deviation <- apply(mtcars, 2, function(x) mean(abs(x - mean(x))))</pre>
print("Mean Deviation:")
print(mean_deviation)
# Standard Deviation
sd_values <- sapply(mtcars, sd)</pre>
print("Standard Deviation:")
```

```
print(sd_values)
# Skewness
skewness_values <- sapply(mtcars, skewness)</pre>
print("Skewness:")
print(skewness_values)
# Kurtosis
kurtosis_values <- sapply(mtcars, kurtosis)</pre>
print("Kurtosis:")
print(kurtosis_values)
# Frequency Distribution and Plots
# Frequency Distribution (for mpg)
mpg_freq <- table(cut(mtcars$mpg, breaks = 5))</pre>
print("Frequency Distribution for MPG:")
print(mpg_freq)
# Histogram for mpg
ggplot(mtcars, aes(x=mpg)) +
geom_histogram(binwidth=2, color="black", fill="blue") +
labs(title="Histogram of MPG", x="Miles per Gallon", y="Frequency")
# Relative Frequency Distribution
mpg_rel_freq <- prop.table(mpg_freq)</pre>
```

```
print("Relative Frequency Distribution for MPG:")
print(mpg_rel_freq)
# Cumulative Frequency Distribution
mpg_cum_freq <- cumsum(mpg_freq)</pre>
print("Cumulative Frequency Distribution for MPG:")
print(mpg_cum_freq)
# Categorical Variable Analysis (Pie Plot and Stacked Bar Plot)
# Pie Plot for cyl
cyl_freq <- table(mtcars$cyl)</pre>
pie(cyl_freq, labels = names(cyl_freq), main = "Pie Chart of Cylinders")
# Stacked Bar Plot for cyl and gear
ggplot(mtcars, aes(x=factor(cyl), fill=factor(gear))) +
geom_bar(position="stack") +
labs(title="Stacked Bar Plot of Cylinders and Gears", x="Number of Cylinders", y="Count")
OUTPUT:
R version 4.4.1 (2024-06-14 ucrt) -- "Race for Your Life"
Copyright (C) 2024 The R Foundation for Statistical Computing
Platform: x86_64-w64-mingw32/x64
R is free software and comes with ABSOLUTELY NO WARRANTY.
```

You are welcome to redistribute it under certain conditions.

Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.

Type 'contributors()' for more information and

'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or

'help.start()' for an HTML browser interface to help.

Type 'q()' to quit R.

[Workspace loaded from ~/.RData]

- > # Install and load required packages
- > install.packages("psych")

WARNING: Rtools is required to build R packages but is not currently installed. Please download and inst

https://cran.rstudio.com/bin/windows/Rtools/

Installing package into 'C:/Users/lenovo/AppData/Local/R/win-library/4.4'

(as 'lib' is unspecified)

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.4/psych_2.4.6.26.zip'

Content type 'application/zip' length 3730235 bytes (3.6 MB)

downloaded 3.6 MB

package 'psych' successfully unpacked and MD5 sums checked

The downloaded binary packages are in

C:\Users\lenovo\AppData\Local\Temp\RtmpopuGli\downloaded_packages

> install.packages("moments")

WARNING: Rtools is required to build R packages but is not currently installed. Please download and inst

https://cran.rstudio.com/bin/windows/Rtools/

Installing package into 'C:/Users/lenovo/AppData/Local/R/win-library/4.4'

(as 'lib' is unspecified)

trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.4/moments_0.14.1.zip'

Content type 'application/zip' length 57241 bytes (55 KB)

downloaded 55 KB

package 'moments' successfully unpacked and MD5 sums checked

The downloaded binary packages are in

C:\Users\lenovo\AppData\Local\Temp\RtmpopuGli\downloaded_packages

> install.packages("ggplot2")

WARNING: Rtools is required to build R packages but is not currently installed. Please download and inst

https://cran.rstudio.com/bin/windows/Rtools/

Installing package into 'C:/Users/lenovo/AppData/Local/R/win-library/4.4'

(as 'lib' is unspecified)

```
trying URL 'https://cran.rstudio.com/bin/windows/contrib/4.4/ggplot2_3.5.1.zip'
Content type 'application/zip' length 5011572 bytes (4.8 MB)
downloaded 1.6 MB
Warning in install.packages:
 downloaded length 1654221 != reported length 5011572
Warning in install.packages:
 URL 'https://cran.rstudio.com/bin/windows/contrib/4.4/ggplot2_3.5.1.zip': Timeout of 60 seconds was r
Error in download.file(url, destfile, method, mode = "wb", ...):
 download from 'https://cran.rstudio.com/bin/windows/contrib/4.4/ggplot2_3.5.1.zip' failed
Warning in install.packages:
 download of package 'ggplot2' failed
> library(psych)
> library(moments)
> library(ggplot2)
Learn more about the underlying theory at https://ggplot2-book.org/
Attaching package: 'ggplot2'
The following objects are masked from 'package:psych':
 %+%, alpha
```

```
> # Load the mtcars dataset
> data(mtcars)
> head(mtcars)
        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
> # Measures of Central Tendency
> print("JVNGANESH 21BDS0085")
[1] "JVNGANESH 21BDS0085"
># Mean
> mean_values <- sapply(mtcars, mean)
> print("Mean Values:")
[1] "Mean Values:"
> print(mean_values)
   mpg
          cyl
                disp
                             drat
                                         qsec
                                                        am
                                                              gear
                                                                    carb
20.090625 6.187500 230.721875 146.687500 3.596563 3.217250 17.848750 0.437500 0.4062$0 3
> # Median
> median_values <- sapply(mtcars, median)
> print("Median Values:")
```

```
[1] "Median Values:"
> print(median_values)
 mpg cyl disp hp drat wt qsec vs am gear carb
19.200 6.000 196.300 123.000 3.695 3.325 17.710 0.000 0.000 4.000 2.000
> # Quantiles (25%, 50%, 75%)
> quantiles <- apply(mtcars, 2, quantile)
> print("Quantiles (25%, 50%, 75%):")
[1] "Quantiles (25%, 50%, 75%):"
> print(quantiles)
   mpg cyl disp hp drat wt qsec vs am gear carb
0% 10.400 4 71.100 52.0 2.760 1.51300 14.5000 0 0 3 1
25% 15.425 4 120.825 96.5 3.080 2.58125 16.8925 0 0 3 2
50% 19.200 6 196.300 123.0 3.695 3.32500 17.7100 0 0 4 2
75% 22.800 8 326.000 180.0 3.920 3.61000 18.9000 1 1 4 4
100% 33.900 8 472.000 335.0 4.930 5.42400 22.9000 1 1 5 8
> # Deciles (ntile)
> deciles <- apply(mtcars, 2, function(x) quantile(x, probs = seq(0, 1, 0.1)))
> print("Deciles:")
[1] "Deciles:"
> print(deciles)
   mpg cyl disp hp drat wt qsec vs am gear carb
0% 10.40 4 71.10 52.0 2.760 1.5130 14.500 0 0.0 3 1.0
10% 14.34 4 80.61 66.0 3.007 1.9555 15.534 0 0.0 3 1.0
```

```
20% 15.20 4 120.14 93.4 3.072 2.3490 16.734 0 0.0 3 1.2
30% 15.98 4 142.06 106.2 3.150 2.7730 17.020 0 0.0 3 2.0
40% 17.92 6 160.00 110.0 3.354 3.1580 17.340 0 0.0 3 2.0
50% 19.20 6 196.30 123.0 3.695 3.3250 17.710 0 0.0 4 2.0
60% 21.00 8 275.80 165.0 3.818 3.4400 18.180 1 0.6 4 3.0
70% 21.47 8 303.10 178.5 3.914 3.5550 18.607 1 1.0 4 4.0
80% 24.08 8 350.80 200.0 4.048 3.7700 19.332 1 1.0 4 4.0
90% 30.09 8 396.00 243.5 4.209 4.0475 19.990 1 1.0 5 4.0
100% 33.90 8 472.00 335.0 4.930 5.4240 22.900 1 1.0 5 8.0
>
> # Percentiles (90th percentile)
> percentiles <- apply(mtcars, 2, function(x) quantile(x, probs = 0.9))
> print("90th Percentiles:")
[1] "90th Percentiles:"
> print(percentiles)
  mpg cyl disp
                     hp drat
                              wt gsec vs
                                                  am gear carb
30.0900 8.0000 396.0000 243.5000 4.2090 4.0475 19.9900 1.0000 1.0000 5.0000 4.0000
> # Measures of Dispersion
> # Range
> range_values <- sapply(mtcars, function(x) diff(range(x)))
> print("Range:")
[1] "Range:"
> print(range_values)
```

```
mpg cyl disp hp drat wt qsec vs am gear carb
23.500 4.000 400.900 283.000 2.170 3.911 8.400 1.000 1.000 2.000 7.000
> # Interquartile Range (IQR)
> iqr_values <- sapply(mtcars, IQR)
> print("Interquartile Range (IQR):")
[1] "Interquartile Range (IQR):"
> print(iqr_values)
         cyl disp
                           drat
  mpg
                      hp
                                        qsec
                                                VS
                                                      am
                                                           gear
                                                                  carb
7.37500 4.00000 205.17500 83.50000 0.84000 1.02875 2.00750 1.00000 1.00000 1.00000
                                                                                            2.000
>
> # Interdecile Range
> interdecile_range <- apply(mtcars, 2, function(x) quantile(x, probs = 0.9) - quantile(x, probs = 0.1))
> print("Interdecile Range:")
[1] "Interdecile Range:"
> print(interdecile_range)
 mpg cyl disp hp drat wt qsec vs am gear carb
15.750 4.000 315.390 177.500 1.202 2.092 4.456 1.000 1.000 2.000 3.000
>
> # Mean Deviation
> mean_deviation <- apply(mtcars, 2, function(x) mean(abs(x - mean(x))))
> print("Mean Deviation:")
[1] "Mean Deviation:"
> print(mean_deviation)
   mpg
           cyl
                  disp
                          hp
                                drat
                                                                            carb
                                        wt
                                              qsec
                                                       VS
                                                              am
                                                                    gear
```

```
4.7144531 1.5859375 108.7857422 56.4804688 0.4532422 0.7301875 1.3761719 0.4921875 0.48
> # Standard Deviation
> sd_values <- sapply(mtcars, sd)</pre>
> print("Standard Deviation:")
[1] "Standard Deviation:"
> print(sd_values)
   mpg
           cyl
                 disp
                          hp
                                drat
                                        wt
                                              qsec
                                                             am
                                                                    gear
                                                                           carb
                                                       VS
6.0269481 1.7859216 123.9386938 68.5628685 0.5346787 0.9784574 1.7869432 0.5040161 0.49
> # Skewness
> skewness_values <- sapply(mtcars, skewness)
> print("Skewness:")
[1] "Skewness:"
> print(skewness_values)
                disp
   mpg
          cyl
                        hp
                             drat
                                                                      carb
                                     wt
                                           qsec
                                                   ٧S
                                                         am
                                                               gear
0.6404399 -0.1831287 0.4002724 0.7614356 0.2788734 0.4437855 0.3870456 0.2519763 0.3817709
> # Kurtosis
> kurtosis_values <- sapply(mtcars, kurtosis)
> print("Kurtosis:")
[1] "Kurtosis:"
> print(kurtosis_values)
  mpg cyl disp
                                                 am gear carb
                    hp drat
                                wt qsec
2.799467 1.319032 1.910317 3.052233 2.435116 3.172471 3.553753 1.063492 1.145749 2.056790 4.536
```

```
>
> # Frequency Distribution and Plots
> # Frequency Distribution (for mpg)
> mpg_freq <- table(cut(mtcars$mpg, breaks = 5))
> print("Frequency Distribution for MPG:")
[1] "Frequency Distribution for MPG:"
> print(mpg_freq)
(10.4,15.1] (15.1,19.8] (19.8,24.5] (24.5,29.2] (29.2,33.9]
    6
          12
                  8
                        2
> # Histogram for mpg
> ggplot(mtcars, aes(x=mpg)) +
+ geom_histogram(binwidth=2, color="black", fill="blue") +
+ labs(title="Histogram of MPG", x="Miles per Gallon", y="Frequency")
> # Relative Frequency Distribution
> mpg_rel_freq <- prop.table(mpg_freq)
> print("Relative Frequency Distribution for MPG:")
[1] "Relative Frequency Distribution for MPG:"
> print(mpg_rel_freq)
(10.4,15.1] (15.1,19.8] (19.8,24.5] (24.5,29.2] (29.2,33.9]
  0.1875
          0.3750 0.2500 0.0625 0.1250
```

```
>
> # Cumulative Frequency Distribution
> mpg_cum_freq <- cumsum(mpg_freq)
> print("Cumulative Frequency Distribution for MPG:")
[1] "Cumulative Frequency Distribution for MPG:"
> print(mpg_cum_freq)
(10.4,15.1] (15.1,19.8] (19.8,24.5] (24.5,29.2] (29.2,33.9]
    6
          18
                  26
                         28
                                32
>
> # Categorical Variable Analysis (Pie Plot and Stacked Bar Plot)
> # Pie Plot for cyl
> cyl_freq <- table(mtcars$cyl)</pre>
> pie(cyl_freq, labels = names(cyl_freq), main = "Pie Chart of Cylinders")
>
> # Stacked Bar Plot for cyl and gear
> ggplot(mtcars, aes(x=factor(cyl), fill=factor(gear))) +
+ geom_bar(position="stack") +
+ labs(title="Stacked Bar Plot of Cylinders and Gears", x="Number of Cylinders", y="Count")
```