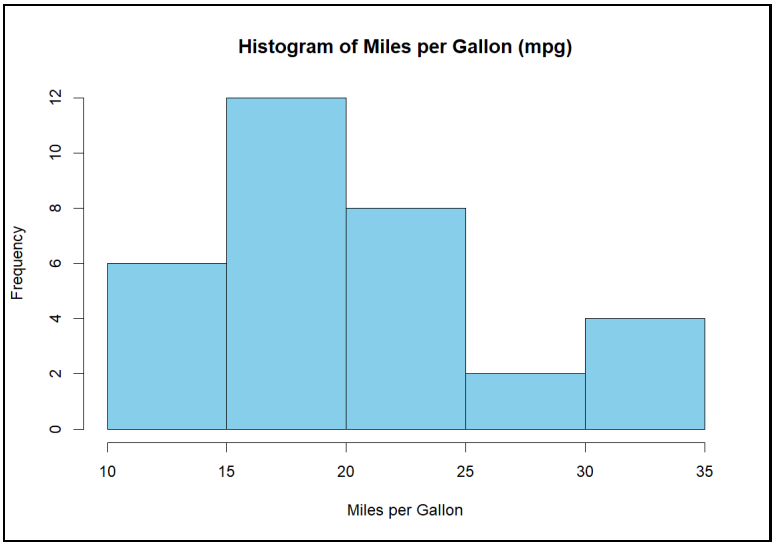
Assignment 5

1. Using sample datasets or own data, create the following charts using only R graphics functions (i.e. without using any other packages). Be sure you customize the chart with your own style/theme (e.g. font, color, pch, etc.)

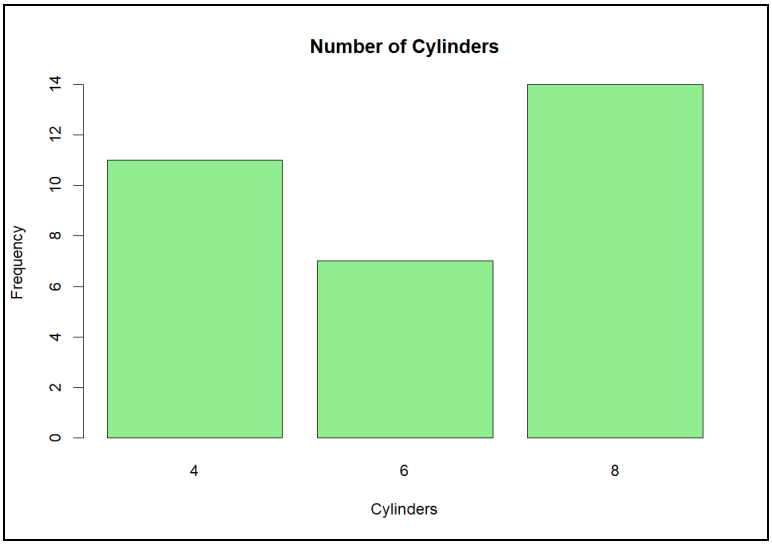
Histogram

```
> hist(mtcars$mpg,
+      col = "skyblue",
+      border = "black",
+      main = "Histogram of Miles per Gallon (mpg)",
+      xlab = "Miles per Gallon")
```



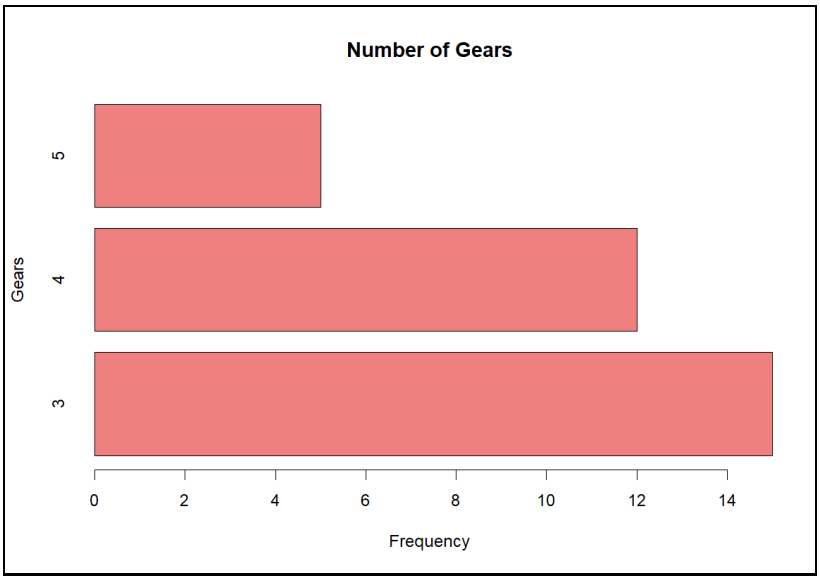
Barchart Vertical

```
> barplot(table(mtcars$cyl),
+         col = "lightgreen",
+         main = "Number of Cylinders",
+         xlab = "Cylinders",
+         ylab = "Frequency")
```



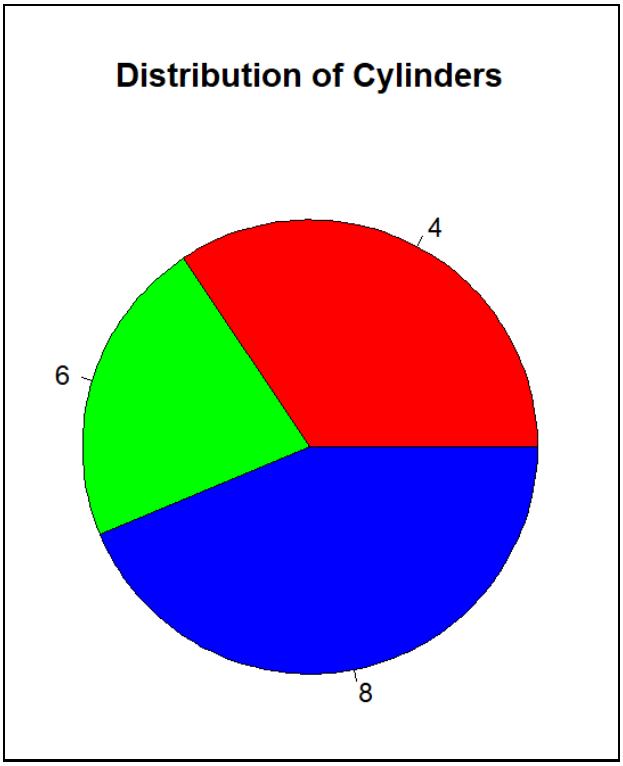
Barchart Horizonal

```
> barplot(table(mtcars$gear),
+         horiz = TRUE,
+         col = "lightcoral",
+         main = "Number of Gears",
+         xlab = "Frequency",
+         ylab = "Gears")
```



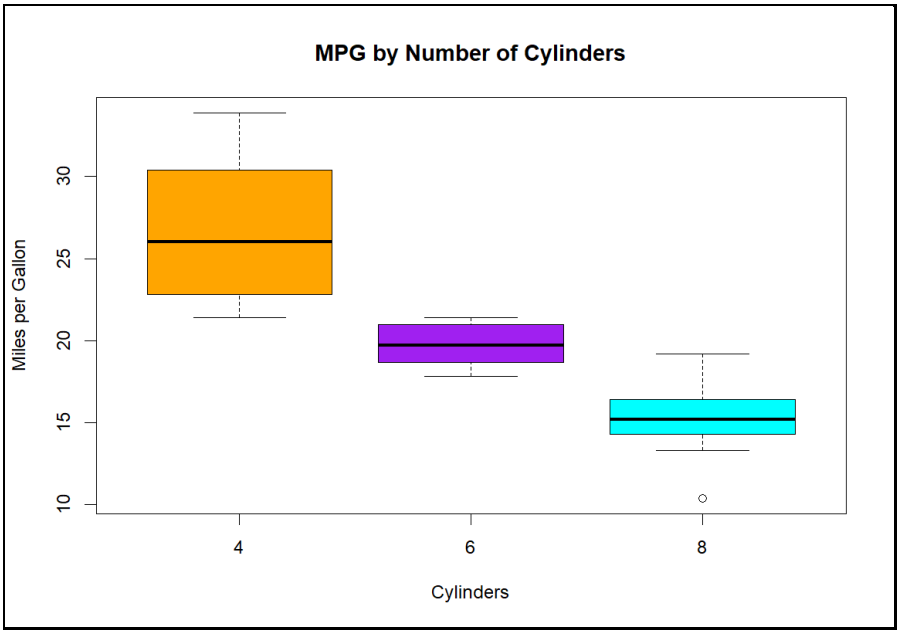
Piechart

```
> pie(table(mtcars$cyl),
+     col = c("red", "green", "blue"),
+     main = "Distribution of Cylinders")
```



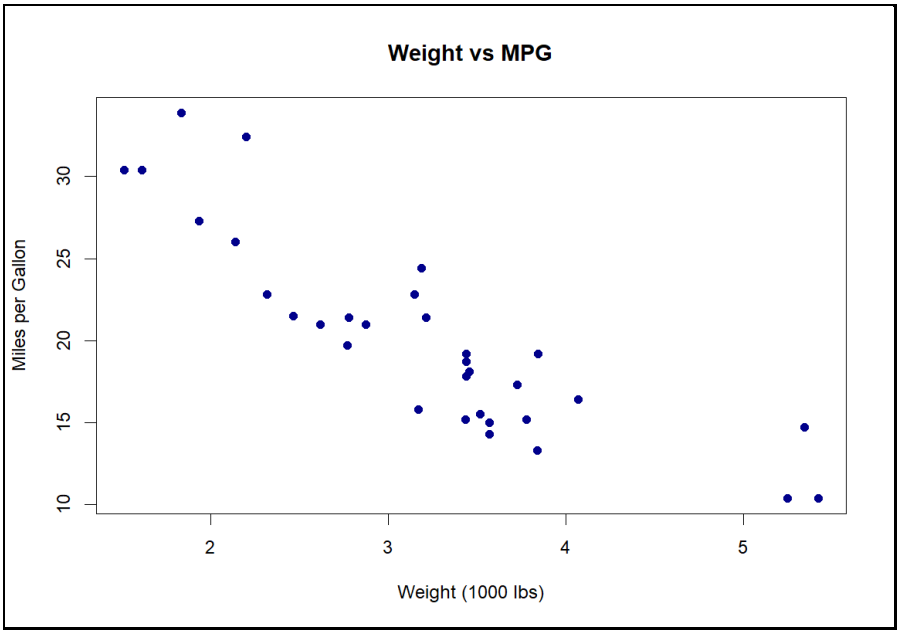
Boxplot

```
> boxplot(mpg ~ cyl, data = mtcars,
+         col = c("orange", "purple", "cyan"),
+         main = "MPG by Number of Cylinders",
+         xlab = "Cylinders",
+         ylab = "Miles per Gallon")
```



Scatterplot

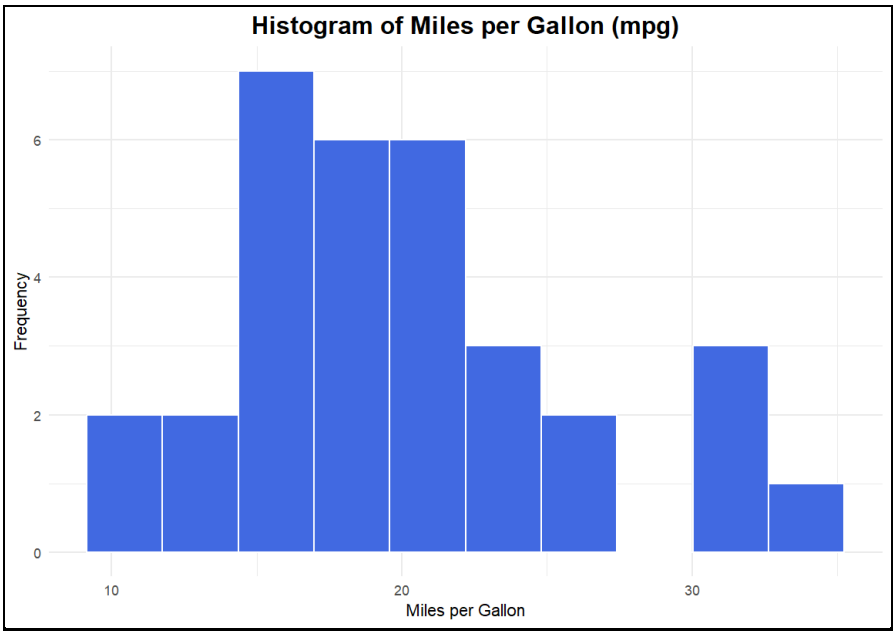
```
> plot(mtcars$wt, mtcars$mpg,
+      pch = 19,
+      col = "darkblue",
+      main = "Weight vs MPG",
+      xlab = "Weight (1000 lbs)",
+      ylab = "Miles per Gallon")
```



2. Repeat 1 using ggplot2, with your own style.

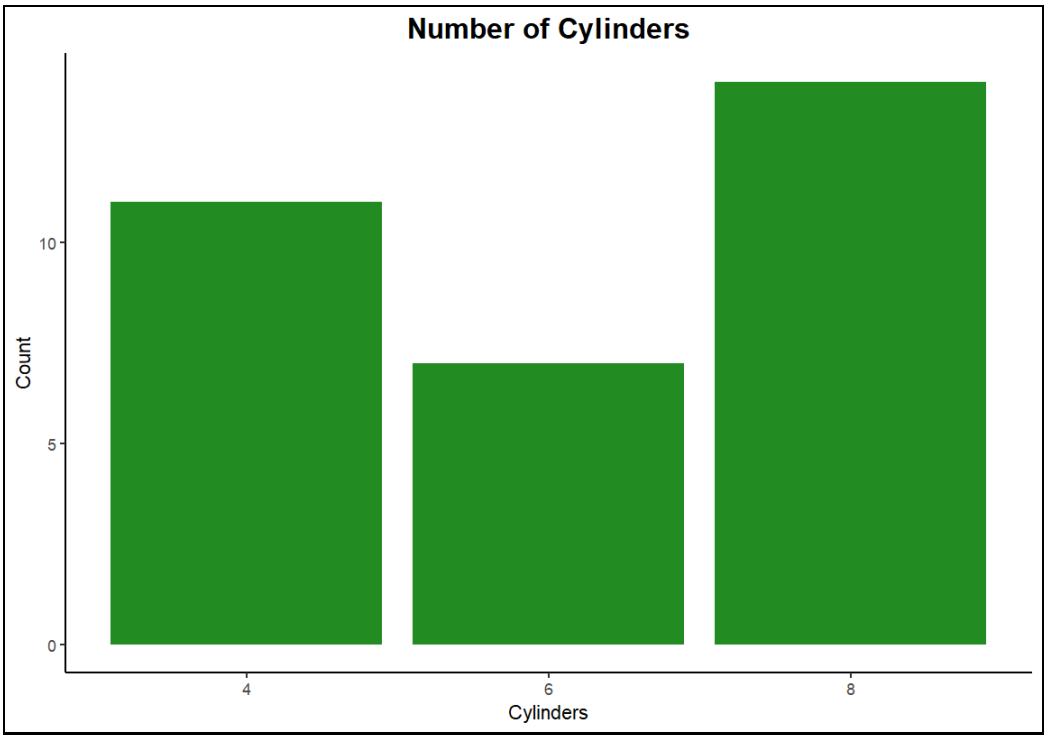
Histogram

```
> ggplot(mtcars, aes(x = mpg)) +
+     geom_histogram(fill = "royalblue", color = "white", bins = 10) +
+     theme_minimal() +
+     labs(title = "Histogram of Miles per Gallon (mpg)",
+          x = "Miles per Gallon", y = "Frequency") +
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
```



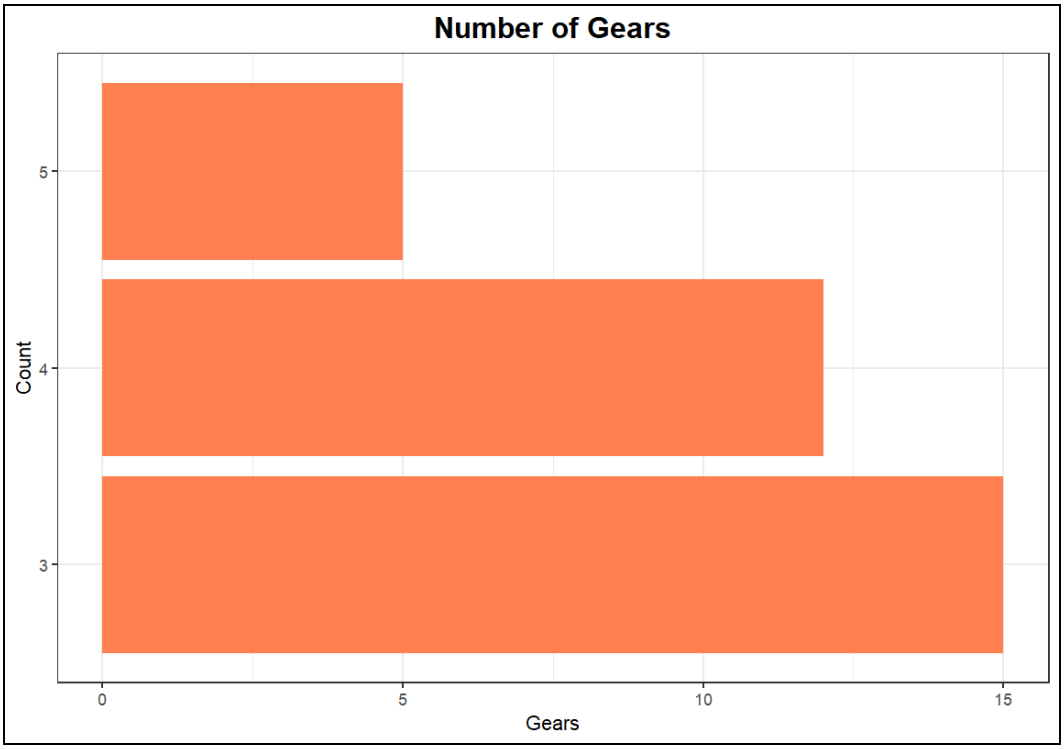
Barchart vertical

```
> ggplot(mtcars, aes(x = factor(cyl))) +
+     geom_bar(fill = "forestgreen") +
+     theme_classic() +
+     labs(title = "Number of Cylinders", x = "Cylinders", y = "Count") +
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
```



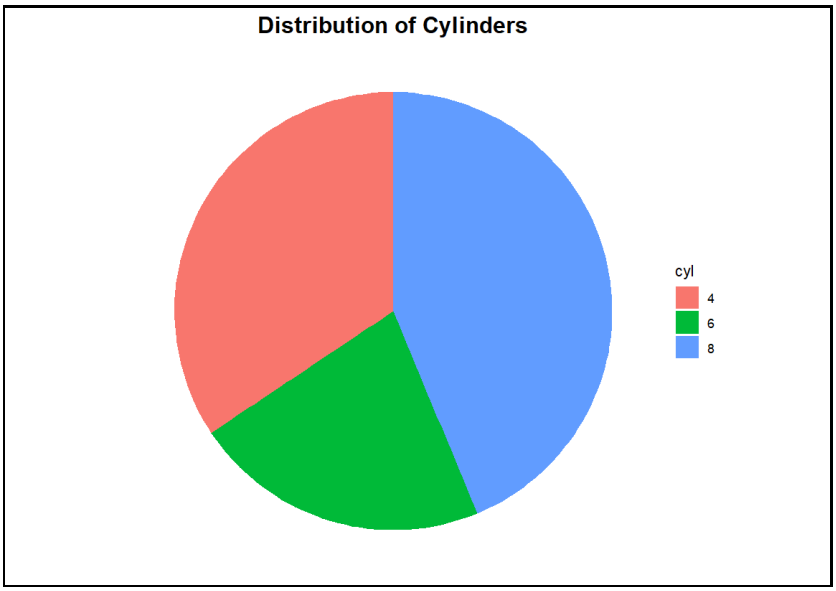
Barchart horizontal

```
> ggplot(mtcars, aes(x = factor(gear))) +
+     geom_bar(fill = "coral") +
+     coord_flip() +
+     theme_bw() +
+     labs(title = "Number of Gears", x = "Count", y = "Gears") +
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
```



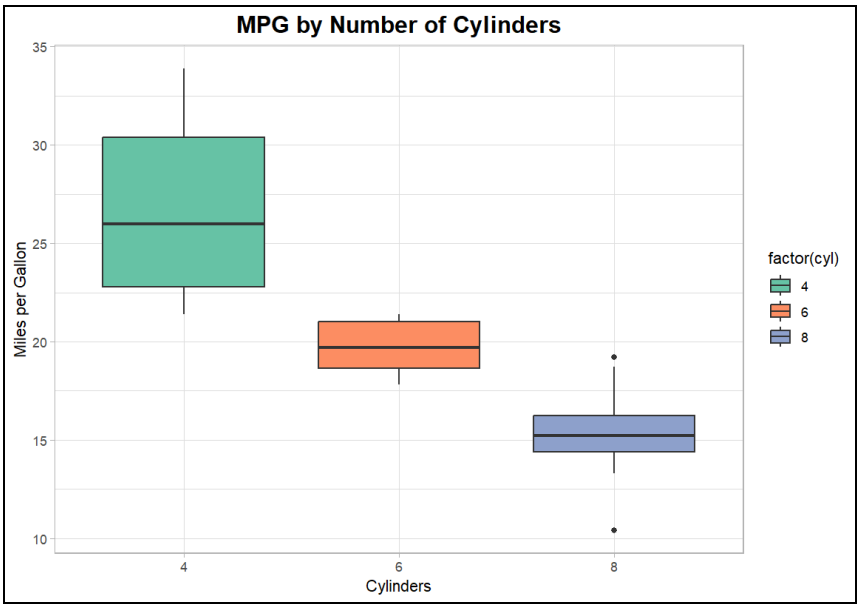
Pie chart

```
> mtcars$cyl <- factor(mtcars$cyl)  # Convert to factor for pie chart
> ggplot(mtcars, aes(x = "", fill = cyl)) +
+     geom_bar(width = 1) +
+     coord_polar(theta = "y") +
+     theme_void() +
+     labs(title = "Distribution of Cylinders") +
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
```



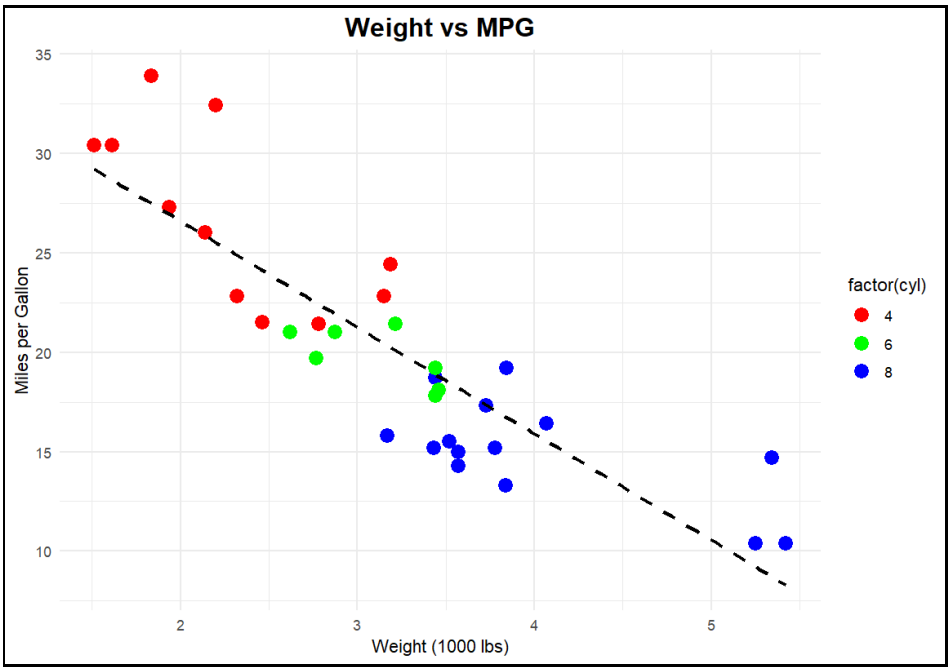
Box plot

```
> ggplot(mtcars, aes(x = factor(cyl), y = mpg, fill = factor(cyl))) +
+     geom_boxplot() +
+     theme_light() +
+     scale_fill_brewer(palette = "Set2") +
+     labs(title = "MPG by Number of Cylinders", x = "Cylinders", y = "Miles per Gallon")
+
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
```



Scatter plot

```
> ggplot(mtcars, aes(x = wt, y = mpg)) +
+     geom_point(aes(color = factor(cyl)), size = 4) +
+     geom_smooth(method = "lm", se = FALSE, linetype = "dashed", color = "black") +
+     theme_minimal() +
+     labs(title = "Weight vs MPG", x = "Weight (1000 lbs)", y = "Miles per Gallon") +
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold")) +
+     scale_color_manual(values = c("red", "green", "blue"))
```



3. Export the charts using different formats such as:

Export Histogram as .pdf

```
> pdf("histogram_mpg.pdf")
> ggplot(mtcars, aes(x = mpg)) +
+     geom_histogram(fill = "royalblue", color = "white", bins = 10) +
+     theme_minimal() +
+     labs(title = "Histogram of Miles per Gallon (mpg)",
+          x = "Miles per Gallon", y = "Frequency") +
+     theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
> dev.off()
```

Export vertical barchart as .jpg

```
> jpeg("C:/Users/jsnor/OneDrive/Desktop/SCHOOL/UT-Dallas/Courses/EPPS6356-
DataVisualization/Assignments/Assignment 5/Exports/vertical_bar_cyl.jpg")
> print(
+     ggplot(mtcars, aes(x = factor(cyl))) +
+         geom_bar(fill = "forestgreen") +
+         theme_classic() +
+         labs(title = "Number of Cylinders", x = "Cylinders", y = "Count") +
+         theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
+ )
> dev.off()
```

Export horizontal barchart as .svg

```
> svg("C:/Users/jsnor/OneDrive/Desktop/SCHOOL/UT-Dallas/Courses/EPPS6356-
DataVisualization/Assignments/Assignment 5/Exports/horizontal_bar_gear.svg")
> print(
+     ggplot(mtcars, aes(x = factor(gear))) +
+         geom_bar(fill = "coral") +
+         coord_flip() +
+         theme_bw() +
+         labs(title = "Number of Gears", x = "Count", y = "Gears") +
+         theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
+ )
> dev.off()
```

Export pie chart as .tiff

```
> tiff("C:/Users/jsnor/OneDrive/Desktop/SCHOOL/UT-Dallas/Courses/EPPS6356-
DataVisualization/Assignments/Assignment 5/Exports/pie_chart_cyl.tiff")
> print(
+     ggplot(mtcars, aes(x = "", fill = cyl)) +
+         geom_bar(width = 1) +
+         coord_polar(theta = "y") +
+         theme_void() +
+         labs(title = "Distribution of Cylinders") +
+         theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
+ )
> dev.off()
```

Export box plot as .bmp

```
> bmp("C:/Users/jsnor/OneDrive/Desktop/SCHOOL/UT-Dallas/Courses/EPPS6356-
DataVisualization/Assignments/Assignment 5/Exports/boxplot_mpg_cyl.bmp")
> print(
+     ggplot(mtcars, aes(x = factor(cyl), y = mpg, fill = factor(cyl))) +
+         geom_boxplot() +
+         theme_light() +
+         scale_fill_brewer(palette = "Set2") +
+         labs(title = "MPG by Number of Cylinders", x = "Cylinders", y = "Miles per
Gallon") +
+         theme(plot.title = element_text(hjust = 0.5, size = 16, face = "bold"))
+ )
> dev.off()
```

Export scatter plot as .png

```
> png("C:/Users/jsnor/OneDrive/Desktop/SCHOOL/UT-Dallas/Courses/EPPS6356-
DataVisualization/Assignments/Assignment
5/Exports/scatterplot_wt_mpg.png")
> print(
+     ggplot(mtcars, aes(x = wt, y = mpg)) +
+         geom_point(aes(color = factor(cyl)), size = 4) +
+         geom_smooth(method = "lm", se = FALSE, linetype = "dashed",
color = "black") +
+         theme_minimal() +
+         labs(title = "Weight vs MPG", x = "Weight (1000 lbs)", y =
"Miles per Gallon") +
+         theme(plot.title = element_text(hjust = 0.5, size = 16, face =
"bold")) +
+         scale_color_manual(values = c("red", "green", "blue"))
+ )
`geom_smooth()` using formula = 'y ~ x'
> dev.off()
```

Note the differences in these file format.

In this assignment, I created a variety of visualizations using both base R and ggplot2, making sure to customize each chart to reflect my personal style. I then exported the charts in multiple formats—PDF, JPG, SVG, TIFF, BMP, and PNG—to explore the differences between them and better understand their strengths and limitations. Each file type serves a specific purpose, and my experience working with them highlighted when and where they are most useful.

PDF stood out as the ideal format for academic work and print publications. Since it's vector-based, the quality remains excellent no matter how much the image is scaled, ensuring that even the smallest details are preserved. This makes PDFs particularly useful for professional documents or papers where precision matters. On the other hand, JPG proved to be better suited for web content. The compression helps reduce file size, but I noticed a slight degradation in quality and some subtle changes in color. It works well in situations where size efficiency is more important than perfect image fidelity.

I found SVG to be an effective option for scalable web graphics. Like PDF, it's a vector format, but it's more interactive and designed for web use. It keeps the file size small without compromising on quality, which makes it perfect for charts or diagrams that need to be scalable. However, it isn't always as widely supported in all applications, which can be a drawback in some situations.

TIFF offered an entirely different value—maintaining every detail of the image through either uncompressed or lossless compression. It's particularly useful for publication-quality images where preserving every pixel matters, though the tradeoff is a much larger file size. I could see myself using TIFF for journals or reports that require precise visuals, but it's less practical for general sharing due to the size.

BMP, while it offers high-quality, uncompressed visuals, felt outdated. The large file sizes and limited use beyond Windows-based applications make it an impractical option in most cases today. In contrast, PNG offered a much better balance between quality and size. As a lossless format, PNG preserves detail without sacrificing efficiency and supports transparency, making it versatile for web use. I found it particularly useful for graphics with fewer colors or for sharing online.

Each of these formats has its role. For academic and print work, PDF or TIFF provides the quality needed for professional presentation. For web or digital use, PNG and SVG are the most practical, with PNG being especially reliable when transparency is required. JPG works well when minimizing file size is a priority, though I'd avoid it if maintaining full color fidelity is essential. BMP, given its limitations, is only useful in very specific scenarios.

In the end, this exercise showed me how critical it is to choose the right format based on the context. Whether the goal is to preserve image quality, minimize file size, or ensure scalability, understanding the purpose behind each format ensures that visualizations serve their intended purpose effectively.