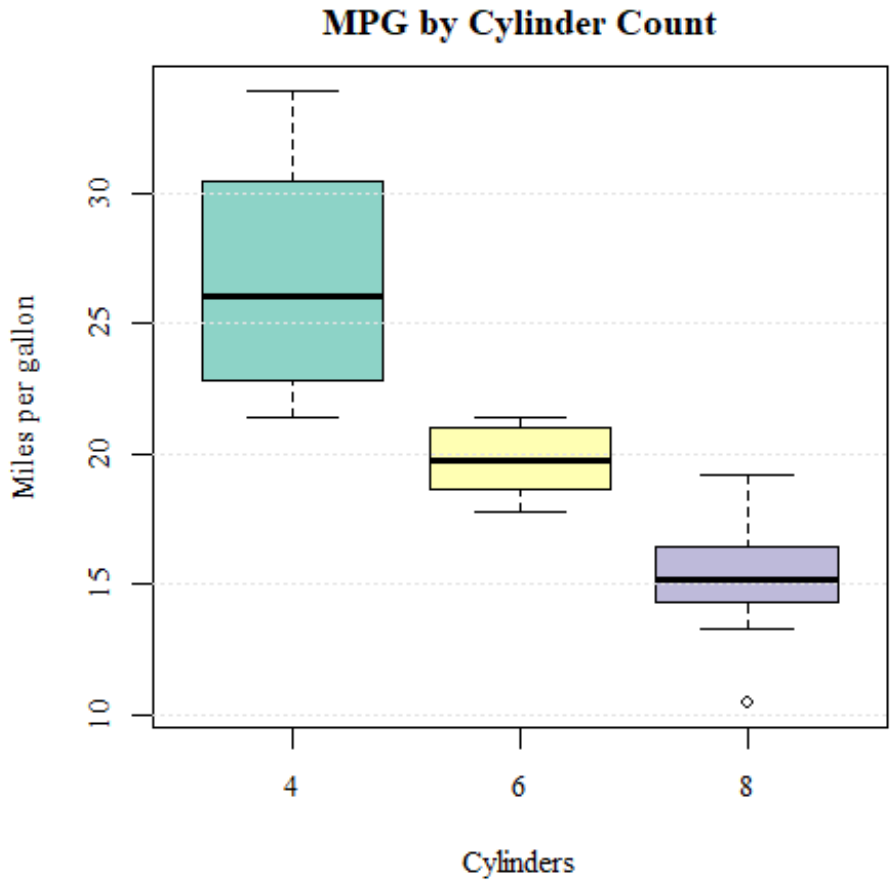
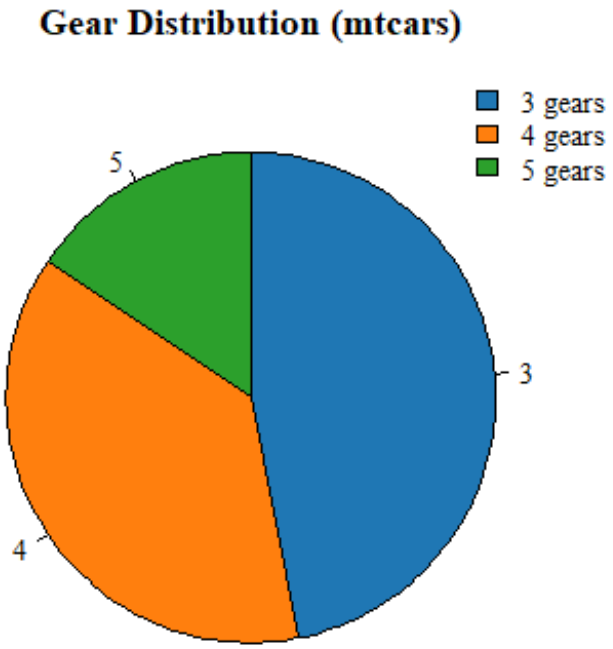
EPPS 6356 — Assignment 5: Base R vs ggplot2

This report compiles Task 1 (Base R) and Task 2 (ggplot2) charts for the mtcars dataset. Each plot was shown in RStudio and exported in five formats (PDF, JPEG, SVG, TIFF, BMP).

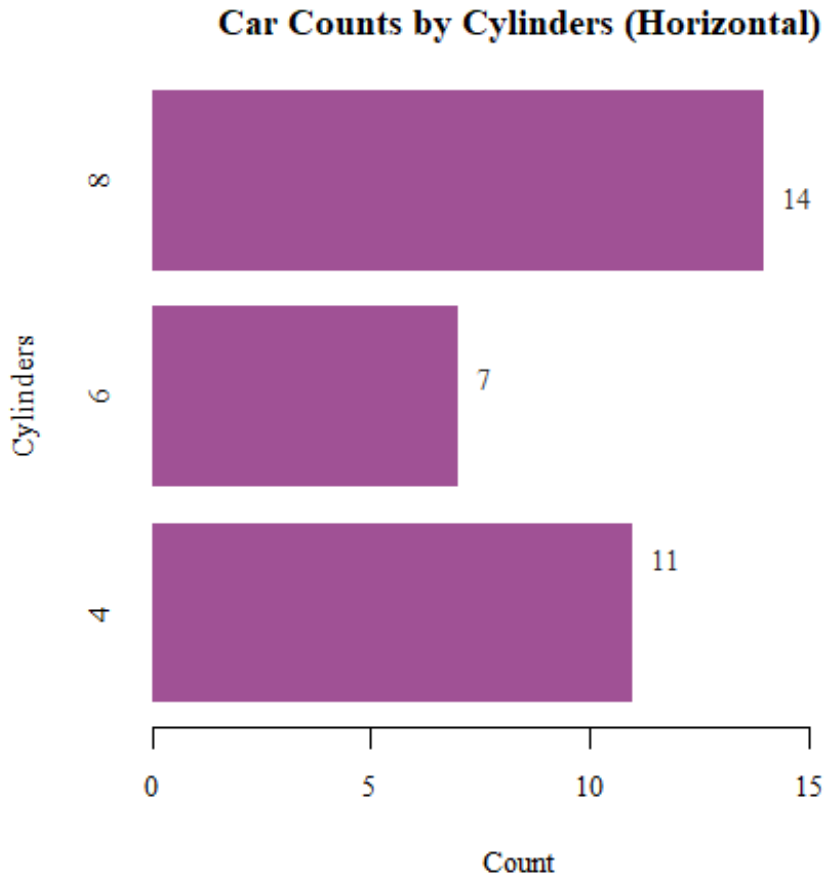
Task 1 — Base R Charts



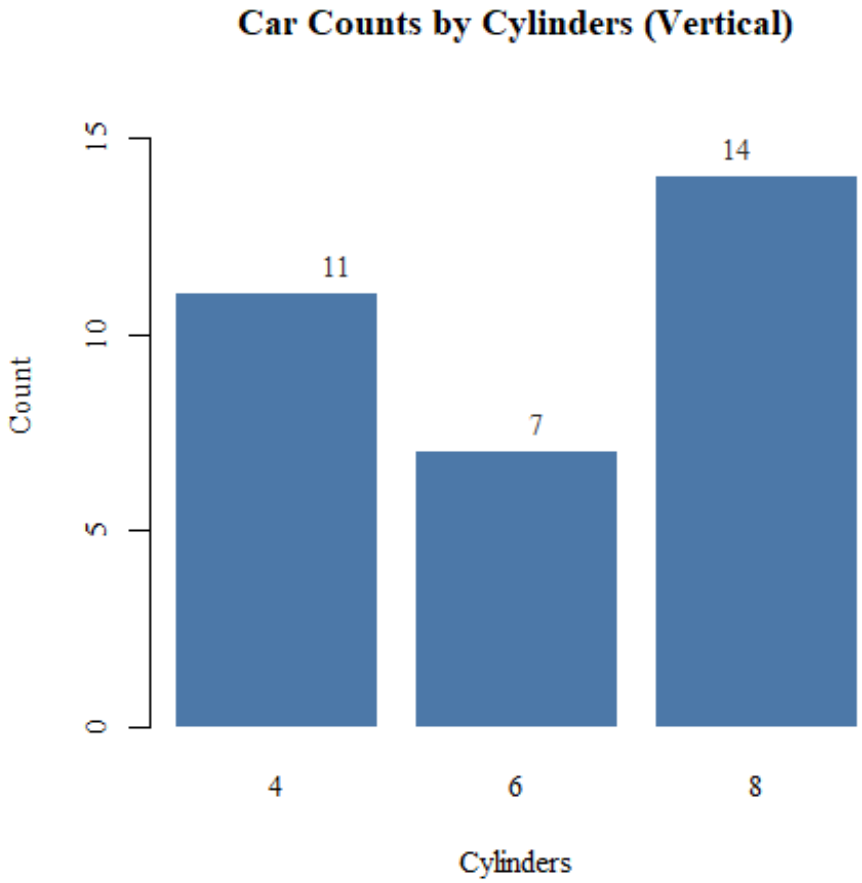
Histogram — Distribution of MPG (Base R)



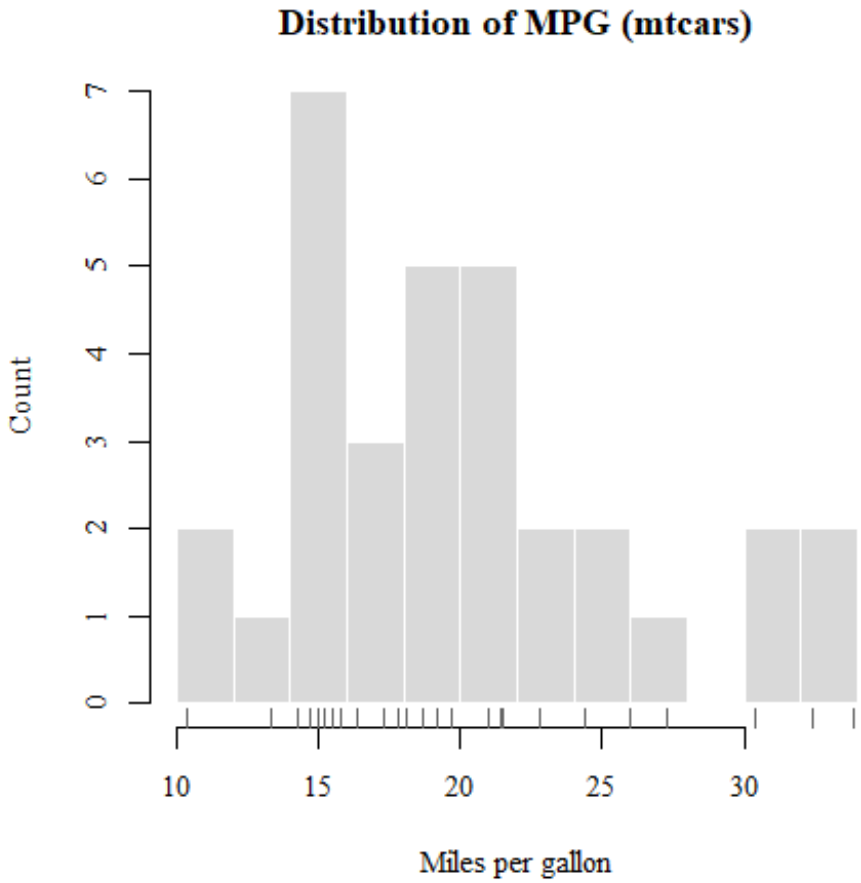
Bar Chart (Vertical) — Car Counts by Cylinders (Base R)



Bar Chart (Horizontal) — Car Counts by Cylinders (Base R)

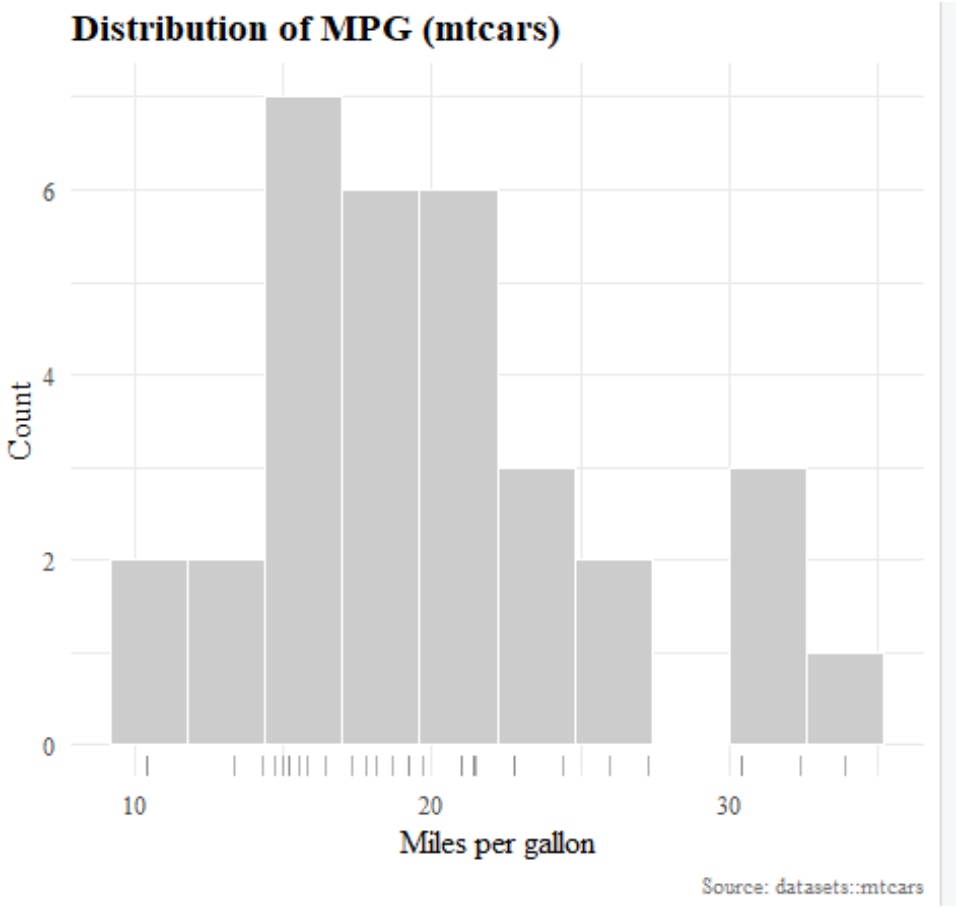


Pie Chart — Gear Distribution (Base R)

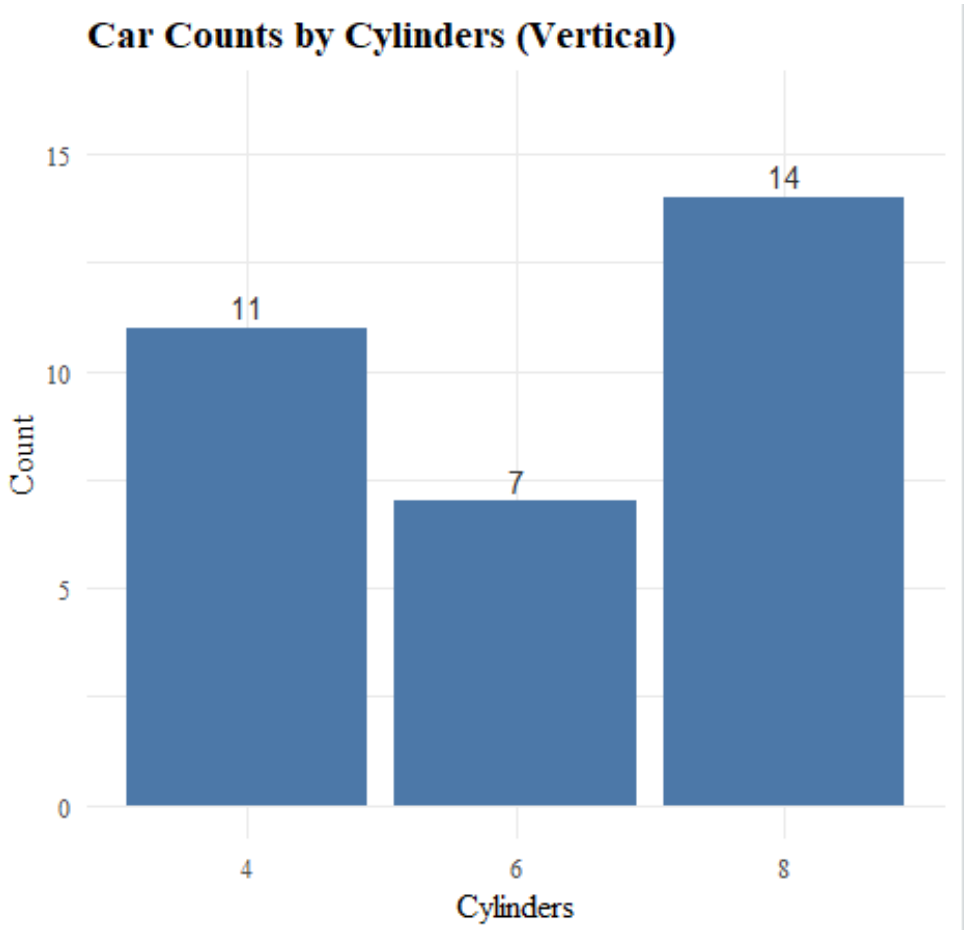


Box Plot — MPG by Cylinder Count (Base R)

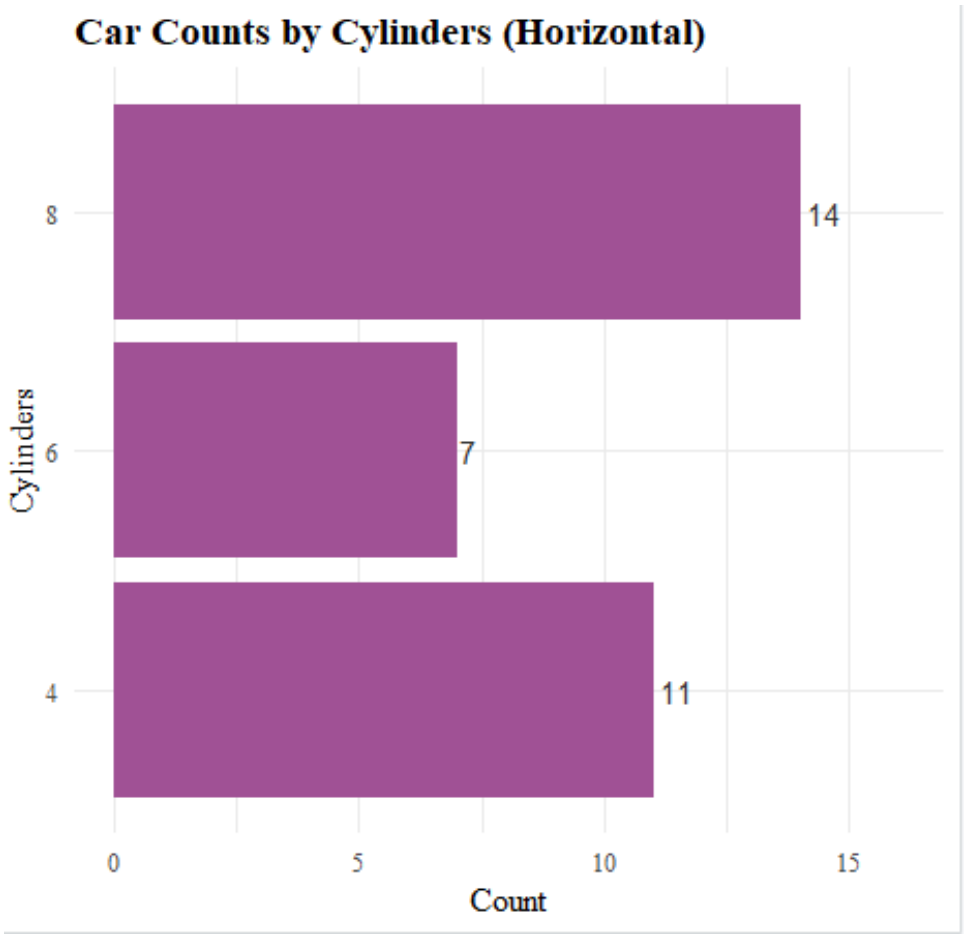
Task 2 — ggplot2 Charts



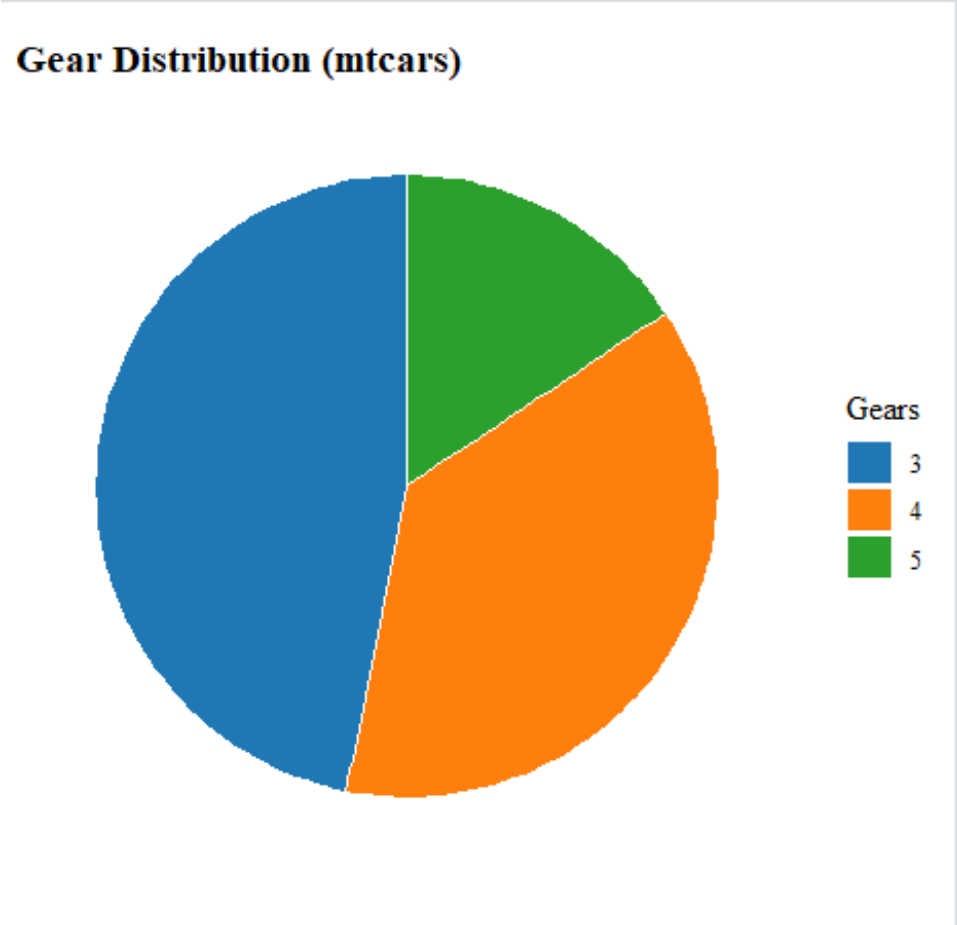
Histogram — Distribution of MPG (ggplot2)



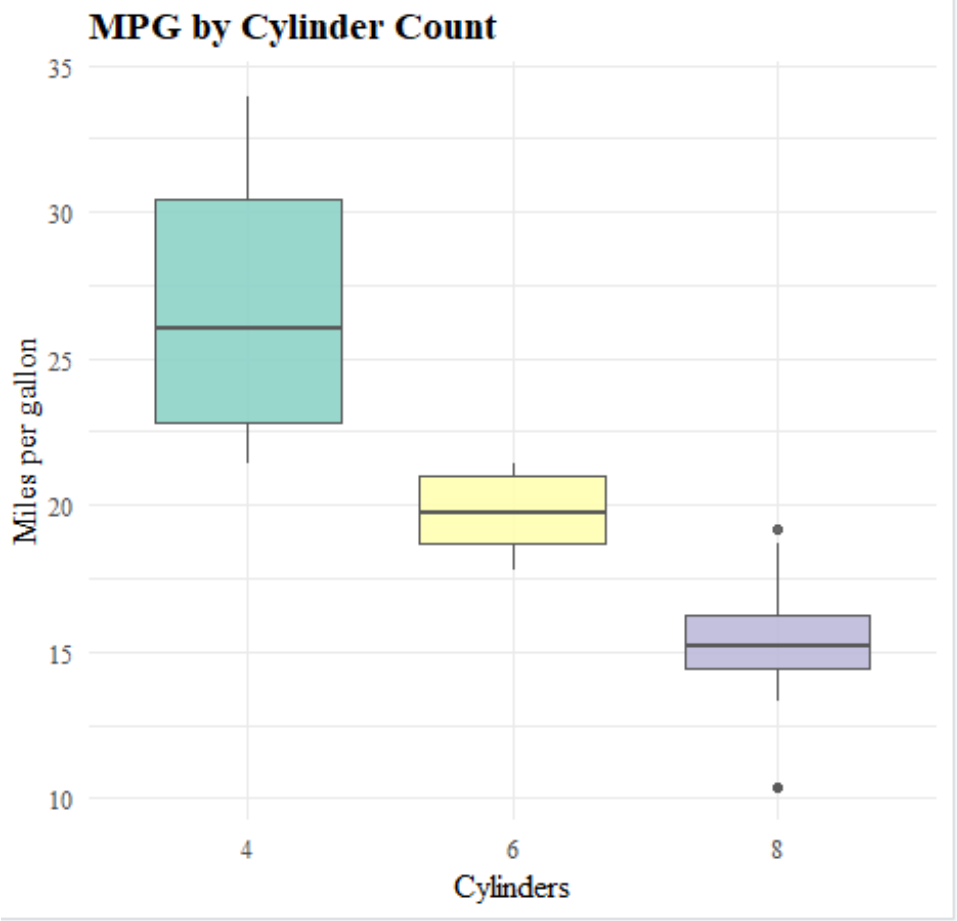
Bar Chart (Vertical) — Car Counts by Cylinders (ggplot2)



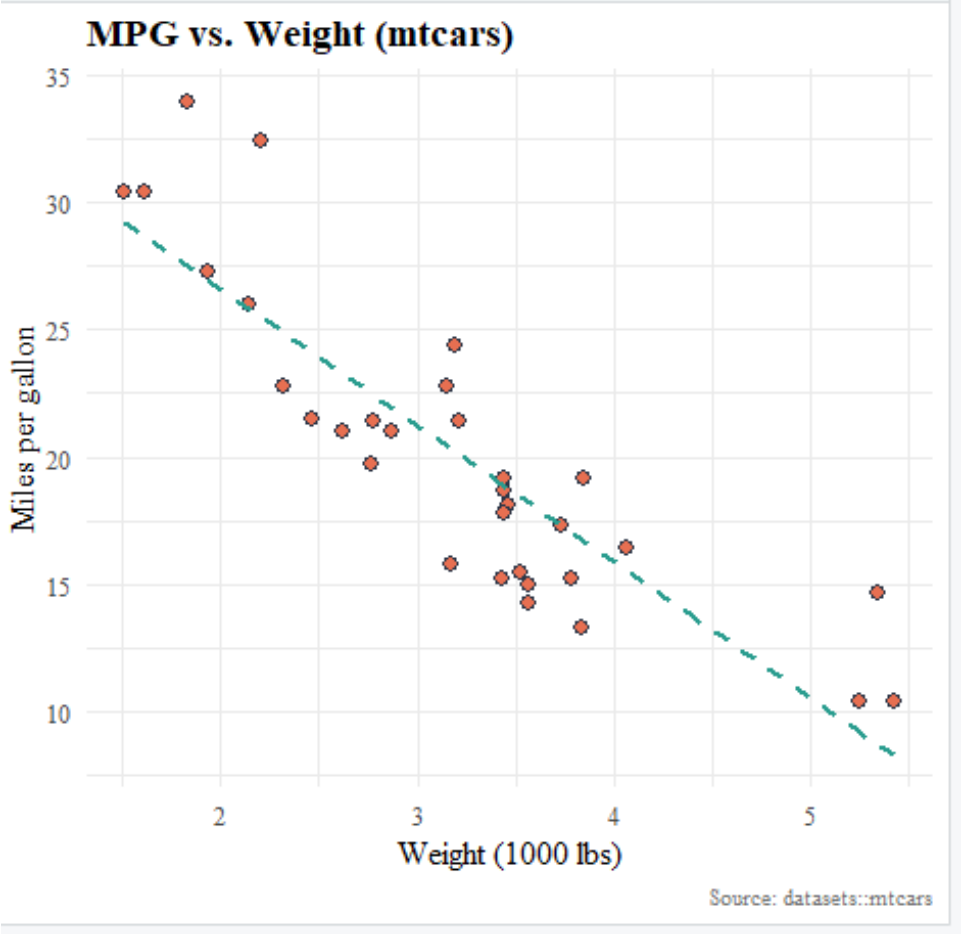
Bar Chart (Horizontal) — Car Counts by Cylinders (ggplot2)



Pie Chart — Gear Distribution (ggplot2)



Box Plot — MPG by Cylinder Count (ggplot2)



Scatter — MPG vs. Weight (ggplot2)

Format Comparison

• PDF & SVG (vector): Crisp at any zoom/print size; usually smallest for simple shapes.
• JPEG (raster, lossy): Smaller files; may show artifacts around lines/text; better for photos.
• TIFF (raster, lossless LZW): Larger but publication-friendly; preserves edges cleanly.
• BMP (raster, uncompressed): Very large; rarely used unless explicitly required.

File Format Differences

In this assignment, each plot was exported in five different file formats — **PDF, SVG, JPEG, TIFF, and BMP** — to illustrate how the same visualization behaves across various output types.

1. Vector Formats (PDF and SVG):

Both PDF and SVG are *vector-based*, meaning they store the plot as mathematical shapes rather than pixels. This makes them **infinitely scalable** — lines, text, and symbols remain sharp at any zoom level or print size.

- **PDF** is ideal for professional documents, academic papers, and printing.
- **SVG** is widely used for web graphics and interactive dashboards because of its browser compatibility and small file size.
  These two formats are the best choices for maintaining high-quality visuals with minimal distortion.

2. Raster Formats (JPEG, TIFF, BMP):
These formats store plots as fixed-resolution images composed of pixels, which can lose clarity when zoomed in or printed at larger sizes.

- **JPEG** uses *lossy compression*, reducing file size but potentially blurring fine lines or text — suitable for online presentations but not for detailed analysis.
- **TIFF** uses *lossless compression* (LZW in this case), preserving quality while producing larger files. It is the preferred format for academic journal submissions and professional printing.
- **BMP** is an *uncompressed* raster format, creating very large files. It preserves all details but offers no compression advantage, making it rarely used in modern workflows.