**Scatter Plot (geom\_point())**

**When to Use?**

* When analyzing the **relationship between two numerical variables**.
* Best for checking **correlations, clusters, and outliers**.

**Example: Relationship between Car Weight (wt) and Miles per Gallon (mpg)**

ggplot(mtcars, aes(x = wt, y = mpg)) +

geom\_point(color = "blue", size = 3, alpha = 0.7) +

labs(title = "MPG vs Weight of Cars",

x = "Weight (1000 lbs)",

y = "Miles per Gallon") +

theme\_minimal()

**Insight**: Heavier cars tend to have lower mileage.

**Line Plot (geom\_line())**

**When to Use?**

* When showing **trends over time**.
* Commonly used in **time-series data**.

**Example: Unemployment Rate Over Time (economics dataset)**

ggplot(economics, aes(x = date, y = unemploy)) +

geom\_line(color = "red", size = 1) +

labs(title = "Unemployment Trend Over Time",

x = "Year",

y = "Unemployed (in thousands)") +

theme\_light()

**Insight**: **Helps visualize periods of high unemployment.**

**Bar Chart (geom\_bar())**

**When to Use?**

* When **comparing categories**.
* Good for **count-based data** or **categorical distributions**.

**Example: Count of Diamond Cuts (diamonds dataset)**

ggplot(diamonds, aes(x = cut, fill = cut)) +

geom\_bar() +

labs(title = "Distribution of Diamond Cuts",

x = "Cut Type",

y = "Count") +

theme\_classic()

**Insight**: The majority of diamonds have an "Ideal" cut.

**Histogram (geom\_histogram())**

**When to Use?**

* When analyzing the **distribution of a single numerical variable**.
* Helps detect **skewness, outliers, and spread**.

**Example: Distribution of mpg in mtcars**

ggplot(mtcars, aes(x = mpg)) +

geom\_histogram(binwidth = 2, fill = "skyblue", color = "black") +

labs(title = "MPG Distribution of Cars",

x = "Miles per Gallon",

y = "Count") +

theme\_bw()

**Insight**: Most cars have an MPG between 15-25.

**Box Plot (geom\_boxplot())**

**When to Use?**

* When checking **distribution, median, quartiles, and outliers**.
* Useful for **comparing numerical data across categories**.

**Example: Comparing mpg Across Cylinder Types (mtcars dataset)**

ggplot(mtcars, aes(x = as.factor(cyl), y = mpg, fill = as.factor(cyl))) +

geom\_boxplot() +

labs(title = "MPG Distribution by Cylinder Count",

x = "Number of Cylinders",

y = "Miles per Gallon") +

theme\_minimal()

**Insight**: Cars with more cylinders tend to have lower MPG.

**Density Plot (geom\_density())**

**When to Use?**

* When checking the **distribution of numerical variables**.
* Alternative to **histograms but smoother**.

**Example: Density of MPG**

ggplot(mtcars, aes(x = mpg, fill = "blue")) +

geom\_density(alpha = 0.5) +

labs(title = "Density Plot of MPG",

x = "Miles per Gallon",

y = "Density") +

theme\_minimal()

**Insight**: Helps in identifying the peak values and spread.

**Violin Plot (geom\_violin())**

**When to Use?**

* When **combining boxplot and density plots**.
* Helps visualize **distribution within categories**.

**Example: Violin Plot of mpg Across Cylinders**

ggplot(mtcars, aes(x = as.factor(cyl), y = mpg, fill = as.factor(cyl))) +

geom\_violin() +

labs(title = "MPG Distribution Across Cylinder Count",

x = "Number of Cylinders",

y = "Miles per Gallon") +

theme\_light()

**Insight**: Shows detailed distribution within each category.

# ****Pie Chart (****geom\_bar() ****with**** coord\_polar()****)****

### ****When to Use?****

* When showing **proportions of categories**.
* Alternative: Use bar charts instead for better readability.

### ****Example: Proportion of Diamond Cut Types****

ggplot(diamonds, aes(x = "", fill = cut)) +

geom\_bar(width = 1) +

coord\_polar(theta = "y") +

labs(title = "Proportion of Diamond Cuts") +

theme\_void()

**Insight**: Helps understand categorical distributions.

# ****Facet Grid (****facet\_wrap()****)****

### ****When to Use?****

* When comparing **subgroups in the dataset**.
* Useful for **breaking down complex plots into smaller charts**.

### ****Example: Scatter Plot of**** mpg ****vs**** wt ****for Each Cylinder Count****

### ggplot(mtcars, aes(x = wt, y = mpg)) +

### geom\_point() +

### facet\_wrap(~cyl) +

### labs(title = "MPG vs Weight for Different Cylinders") +

### theme\_light()

### **Insight**: Shows how the relationship between weight and MPG differs across cylinder counts.

# ****Heatmap (****geom\_tile()****)****

### ****When to Use?****

* When visualizing **relationships between two categorical variables**.
* Commonly used for **correlation matrices**.

### ****Example: Heatmap of Car Cylinder vs Gear Count****

### ggplot(mtcars, aes(x = as.factor(gear), y = as.factor(cyl), fill = mpg)) +

### geom\_tile() +

### labs(title = "Heatmap of Cylinders vs Gears",

### x = "Number of Gears",

### y = "Number of Cylinders") +

### theme\_minimal()

### **Insight**: Shows how MPG varies across different gear-cylinder combinations.

|  |  |
| --- | --- |
| **Plot Type** | **When to Use?** |
| **Scatter Plot** | Relationship between two numeric variables |
| **Line Plot** | Trends over time |
| **Bar Chart** | Comparing categories |
| **Histogram** | Distribution of a single variable |
| **Box Plot** | Comparing distributions & outliers |
| **Density Plot** | Smooth distribution visualization |
| **Violin Plot** | Detailed comparison of distributions |
| **Pie Chart** | Showing proportions (use sparingly) |
| **Facet Grid** | Multiple subplots for categories |
| **Heatmap** | Visualizing relationships between two categorical variables |