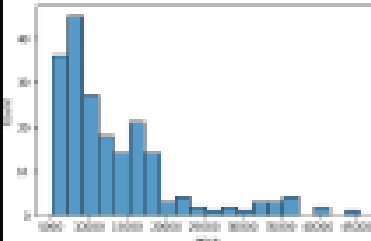
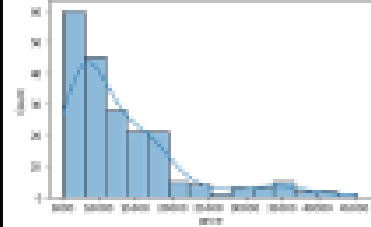
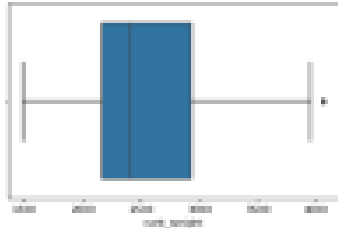


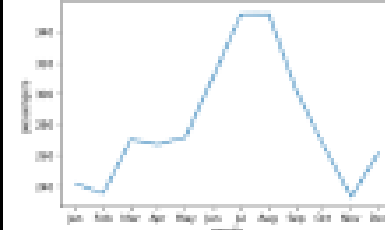
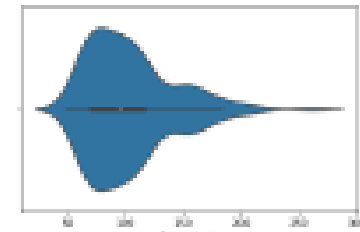
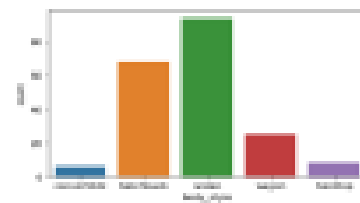
Data Visualization

- Data visualization is the process of translating data and metrics into charts, graphs and other visuals.
- The resulting visual representation of data makes it easier to identify patterns, trends, and outliers hidden in the data, enabling us to gain better insights.
- We can use different charts/plots to visualize different kinds of data
- Each chart/plot helps us gain insights from a different perspective

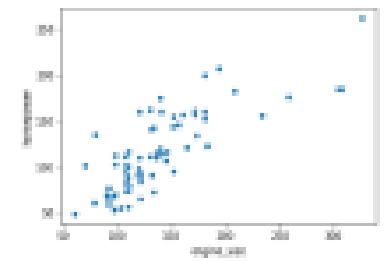
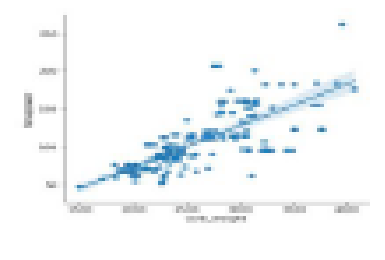
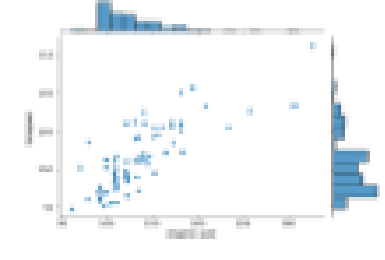
Visualization – One Variable

Plot	Type of Data	Usage	Example
Histogram	Numerical	Helps us understand data distribution by dividing it into bins and showing the number of observations in each bin via bars	
Histogram with density curve	Numerical	Helps us understand data distribution by displaying a distribution curve on top of the histogram bars	
Boxplot	Numerical	Helps us understand data distribution and skewness by displaying the data in the form of a box divided by quartiles	

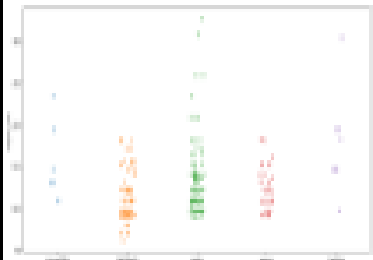
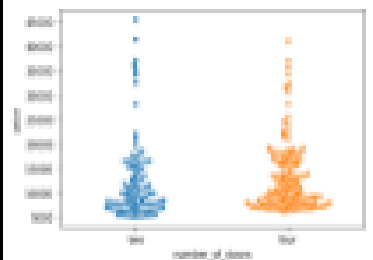
Visualization – One Variable

Plot	Type of Data	Usage	Example
Line Plot	Numerical	Helps us understand the trend or pattern in the data by displaying it as straight lines formed by connecting individual data points	 A line plot showing temperature on the y-axis (ranging from 100 to 140) against months on the x-axis (Jan to Dec). The line starts at approximately 105 in Jan, rises to 125 in Mar, dips slightly in Apr, rises to 135 in Jun, peaks at 140 in Jul and Aug, then falls to 115 in Nov and rises to 120 in Dec.
Violin Plot	Numerical	Helps us understand data distribution by plotting a density curve symmetrically around a boxplot	 A violin plot showing the distribution of measurements on the x-axis (ranging from 50 to 350). The plot is symmetric around a central boxplot, with a wider body on the left (lower measurements) and a long tail extending to the right (higher measurements).
Bar Graph	Categorical	Helps us understand data distribution by showing the counts of observations in each level (or group) using bars	 A bar graph showing counts on the y-axis (ranging from 0 to 40) for different body types on the x-axis: Underweight, Normal, Overweight, Obese, and Morbidly Obese. The counts are approximately: Underweight (5), Normal (35), Overweight (45), Obese (20), and Morbidly Obese (10).

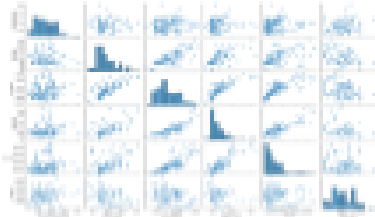
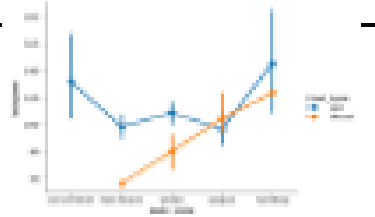
Visualization – Two Variable

Plot	Type of Data	Usage	Example
Scatter Plot	Numerical	Helps us understand potential relationship between two numerical variables	
Implot	Numerical	Helps us understand and measure the relationship between the two variables quantitatively	
Joint Plot	Numerical	Helps us understand the distribution and relationship between two numerical variables on the same plot.	

Visualization – Two Variable

Plot	Type of Data	Usage	Example
Strip Plot	Categorical	Helps us to visualize the distribution of a numerical variable wrt different categories of a categorical variable	
Swarm Plot	Numerical	Helps us to visualize the distribution of a numerical variable wrt different categories of a categorical variable and avoids overlapping of data points	

Visualization – Multiple Variable

Plot	Type of Data	Usage	Example
Pair Plot	Numerical	Helps us understand the relationship between two or more pairs of numerical variables	
Cat Plot	Numerical	Helps us understand relationship between a numerical variable and one or more categorical variables	
Heatmap	Numerical	Helps us understand the correlation between pairs of columns in the data	