Display Type	Best Used to...
Bar Graph	Show the number in categories
Circle Graph	Compare parts of the data to the whole
Double bar Graph	Compare two or more sets of data
Box Whiskers Plot	Show measures of variation
Histogram	Show frequency of data divided into intervals
Line Graph	Show change over time
Line Plot	Show frequency data on a number line

1) **<u>Histograms:</u>**

In a histogram data is grouped into continuous number ranges and each range corresponds to a vertical bar.

- Horizontal axis displays the number range.
- Vertical axis (frequency) represents the amount of data present in each range.

It allows us to assess where the values are concentrated, what the extremes are, and whether there are any gaps or anomalous values. A Histogram is similar to a vertical bar graph, however, the distinction is that the Histogram has no space between the bars, but a bar graph has.

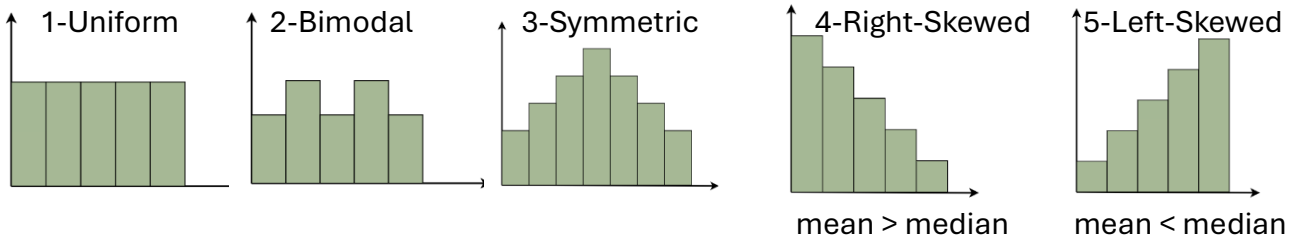
A histogram is a graph that shows the distribution of data. It resembles a sequence of interconnected bars. Each bar represents a range of values, and its height indicates how many data points are inside that range.

Parts of a Histogram

A histogram is a graph that represents the distribution of data. Here are the essential components, presented in simple terms:

- Title: This is similar to the name of the histogram. It explains what the histogram is about and what data it displays.
- X-axis: X-axis is a horizontal line at the bottom of the histogram. It displays the categories or groups that the data is sorted into. For example, if you're measuring people's heights, the X-axis may indicate several height ranges such as "5-6 feet" or "6-7 feet".
- Y-axis: The Y-axis is a vertical line on the side of the histogram. It displays the number of times something occurs in each category or group shown on the X-axis. So, if you're measuring heights, the Y-axis may display how many individuals are in each height range.
- Bars: Bars are the vertical rectangles you see on the chart. Each bar on the X-axis represents a category or group, and its height indicates how many times something occurs inside that category and the width indicates the range covered by each category on the X-axis. So, higher bars indicate more occurrences, whereas shorter bars indicate fewer occurrences.

Types of Histogram:



Difference between Bar Graph And Histogram

Feature	Bar Graph	Histogram
Purpose	Used to show comparisons among discrete categories.	Used to show the distribution of continuous data over intervals.
Data Type	Categorical or discrete.	Continuous, but binned into discrete intervals.
Orientation	Bars can be oriented horizontally or vertically.	Bars are typically vertical.
Spacing Between Bars	Spaces between bars to indicate that categories are distinct.	No space between bars (except for gaps indicating no data for a bin) to signify continuous data range.
Order of Bars	Can be arranged in any order, often sorted by frequency.	Arranged in ascending order of the variable.
X-axis	Represents different categories.	Represents the intervals or "bins" of the continuous data.
Y-axis	Represents the value (count, percentage, etc.) for each category.	Represents the frequency or count of data points within each bin.
Use Cases	Comparing population sizes in different cities, showing sales by product category.	Showing the distribution of exam scores, ages of participants in a study.

2)Pie chart:

The “pie chart” is also known as a “circle chart”, dividing the circular statistical graphic into sectors or sections to illustrate the numerical problems. Each sector denotes a proportionate part of the whole. To find out the composition of something, Pie-chart works the best at that time.

3)Bar charts:

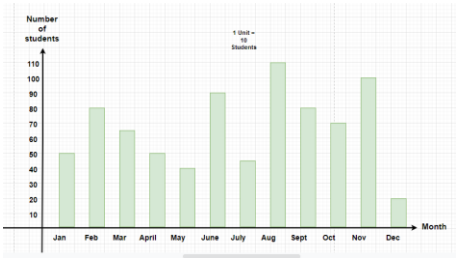
visually represents categorical data using rectangular bars. The height or length of each bar corresponds to the value it represents. These bars can be arranged vertically or horizontally. When plotted vertically, the bar chart is often referred to as a column chart.

Key elements of a bar graph include:

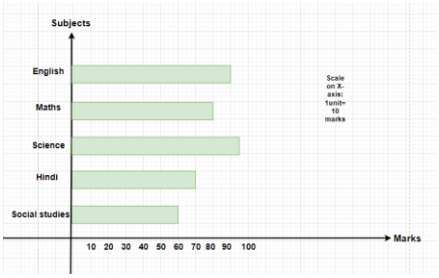
- **Axes**: Most bar graphs have two axes.
    - The x-axis typically lists the categories being compared, and the y-axis represents the measured values.
- **Bars**: Each bar’s length or height varies according to the data it represents.
- **Labels**: Categories and values are labelled to make the data easy to understand.
- **Title**: The title of the bar graph provides an overview of what the data represents.
- **Legend:** Explains the meaning of different colours or patterns if multiple data sets are presented.
- **Scale:** The units or intervals used on the axes to measure and represent the data accurately.

Types:

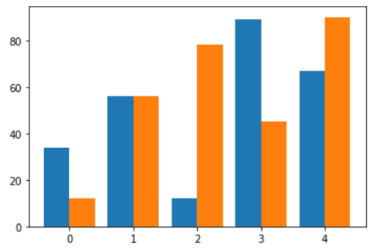
1-Vertical

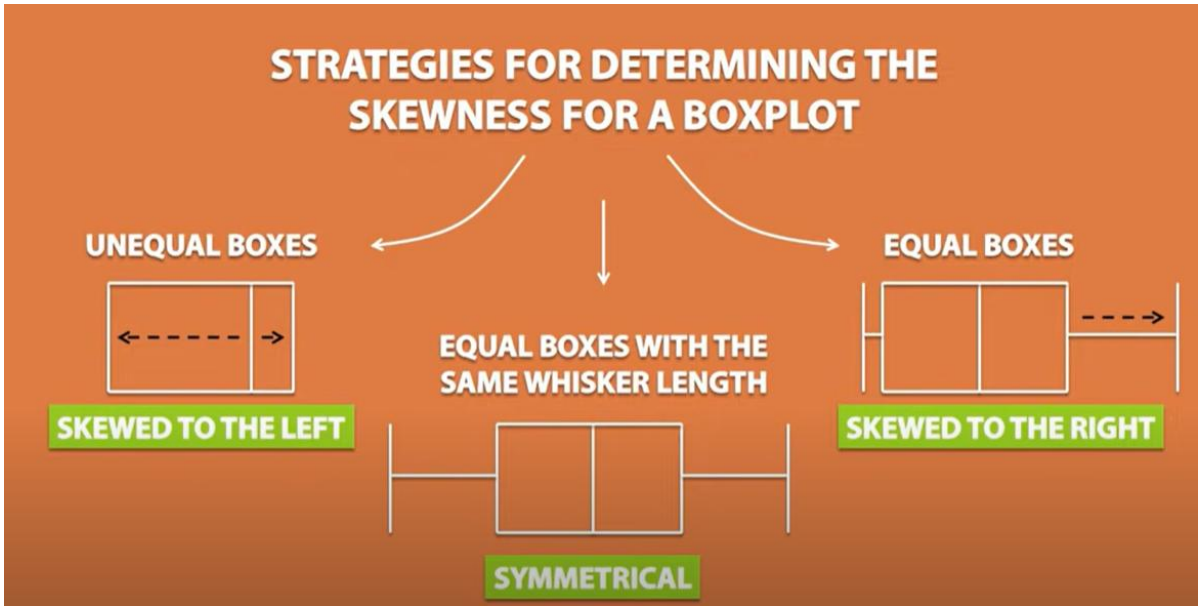


2-Horizontal



3-Grouped





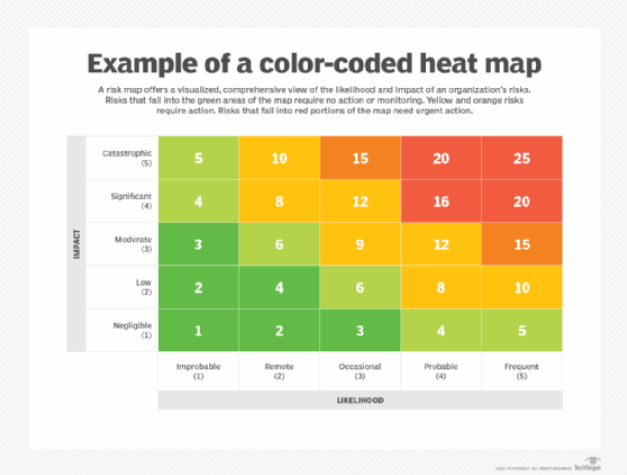
3)heatmap:

A heat map is a two-dimensional representation of data in which various values are represented by colors. A simple heat map provides an immediate visual summary of information across two axes, allowing users to quickly grasp the most important or relevant data points. More elaborate heat maps allow the viewer to understand complex data sets.

A heat map is a way to represent data points in a data set in a visual manner. All heat maps share one thing in common -- they use different colors or different shades of the same color to represent different values and to communicate the relationships that may exist between the variables plotted on the x-axis and y-axis. Usually, a darker color or shade represents a higher or greater quantity of the value being represented in the heat map.

For instance, a heat map showing the rain distribution (range of values) of a city grouped by month may use varying shades of red, yellow and blue. The months may be mapped on the y axis and the rain ranges on the x axis. The lightest color (i.e., blue) would represent the lower rainfall. In contrast, yellow and red would represent increasing rainfall values, with red indicating the highest values.

The changing colors across the two axes will reveal patterns in one or both variables and whether there are relationships between them. Thus, there may be blues on the heat map for some months, indicating that rainfall was in the lower range(s) in these months. A few yellows and reds for some other months would show higher rainfall for those months. From this simple heat map, any user can see how much rain the city received in each month and whether it was higher or lower than the rainfall in other months.

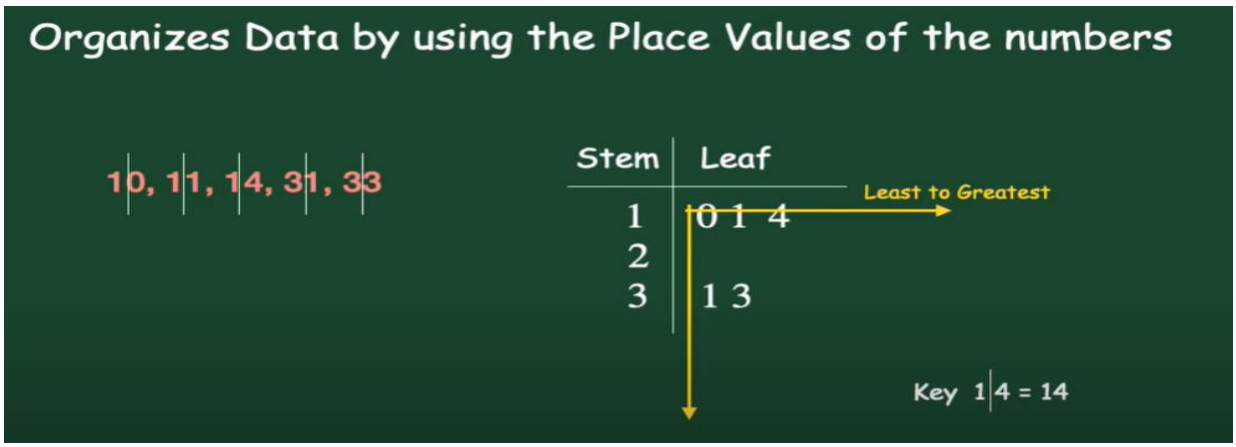


Benefits of heat maps

The values represented on a heat map would normally be much harder for a human to understand if presented numerically in a spreadsheet, especially if the data set is large or complex. Another way of saying this is that a heat map enables viewers to visualize data and understand it easily.

Since heat maps are fairly self-explanatory, complex or detailed data analysis is generally not required. Therefore, any user can gather insights from it and even see key trends from the data set. Also, heat maps can be combined with other types of data visualization to display important data insights to a viewer at a glance.

4) Stem and Leaf Plot :



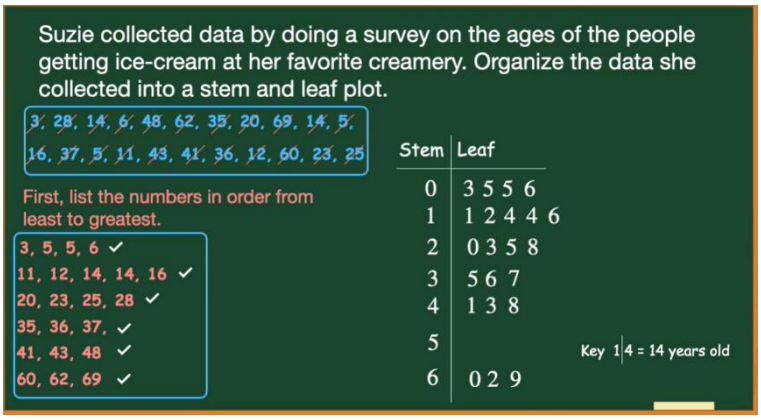
*They can help you identify the central tendency, variability, skewness of your distribution, and outliers. Stem and leaf plots are also known as stemplots.

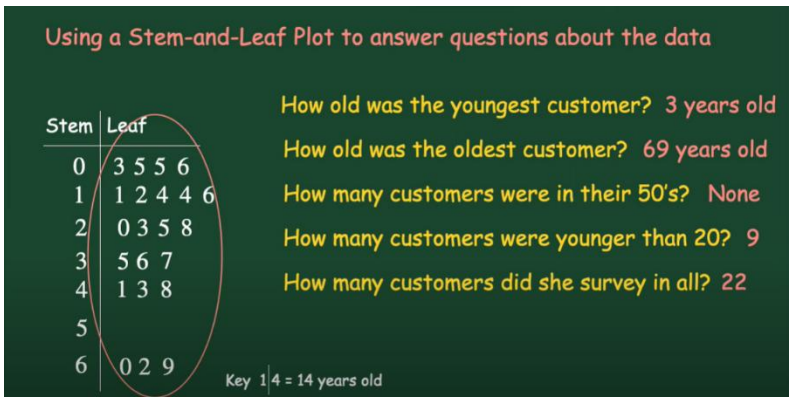
*Stem and leaf plots have one advantage over histograms because they display the original data, while histograms only summarize them.

-To make a stem and leaf plot, do the following:

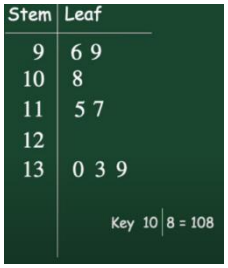
1. Sort your data in ascending order and round the values.
2. Divide your raw data into stem and leaf values.
3. Write down your stem values to set up the groups.
4. Add the leaf values in numerical order to create the depths for each stem value group.

EX:

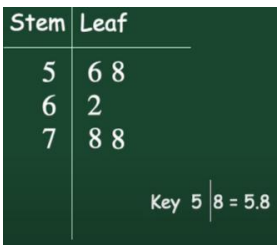




*You can do it for :1) large numbers : || 2) decimals :



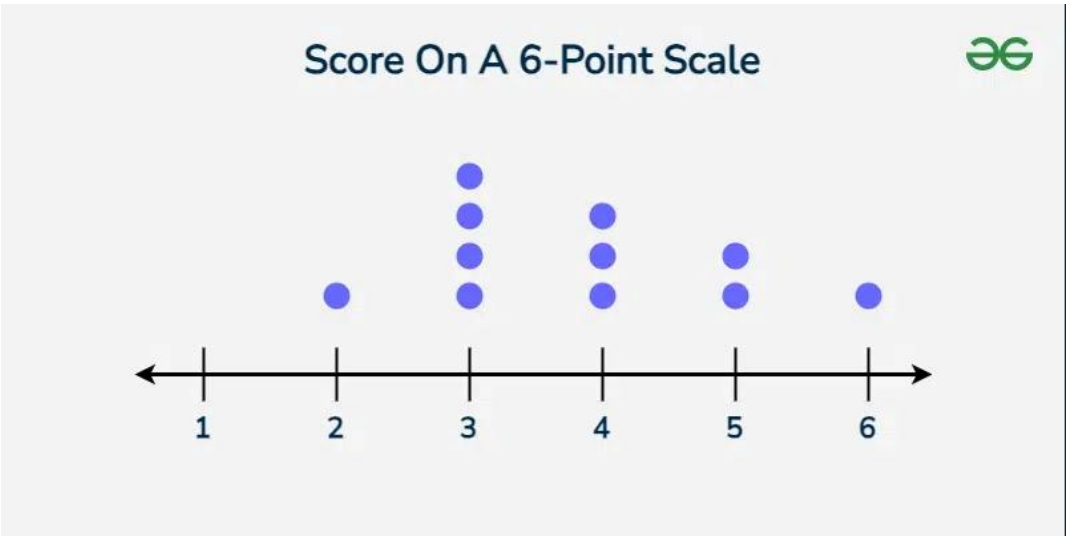
You must put the key.



5) Dot plot : is also known as a **dot chart**, and it is considered to be one of the easiest methods of representing data, with each dot indicating one value. It is very useful where the data set is not too large or small. A dot plot is a simple form of frequency display in that numbers are represented by evenly spaced dots positioned on a statement of a number line. This method makes it easier to visualize the distribution of the-data, clusters, gaps, and outlier- phenomena in the data. Dot plots are present in statistics, education, and other branches to show detailed information about the distribution of data, which makes them very useful for data visualization and analysis.

-How to Make a Dot Plot?

1. Gather Your Data: Summarize the numbers you want to present into a format suitable for representation. Check to see if it is accurate and in the proper order. Dot plots are used where quantitative numbers are involved.
2. Organize the Data: Order the data in a specified ascending sequence. It is important to plot and identify patterns or outliers.
3. Choose a Scale: Choose the scale for the plot; it must be appropriate to the range of the data used. The scale should be equally subdivided and extend from the lowest value in the set to the highest value of the given data.
4. Draw the Axes: Create a horizontal line (X axis) on graphing paper or graphing software. Make sure to use the axis labels and the graph itself as the data values.
5. Analyze the data points plotted: In other words, for every data point, plot a dot above the value associated with that on the x-axis. Repeat the same dots vertically wherever a data point is common to two or more groups.
6. Label the plot: Title your dot plot to share what is being represented in the dot plot. Axis x should show a variable that is measurable in the experiment.



The difference between Dot Plot and Line Plot:

Aspect	Dot Plot	Line Plot
Representation	Uses dots to represent individual data points	Connects data points with lines
Data Type	Typically used for discrete data	Suitable for both discrete and continuous data
Data Visualization	Shows frequency or distribution of data	Emphasizes trends and patterns over time
X-axis	May or may not have a continuous scale	Usually represents time, continuous variable
Y-axis	Represents frequency or count	Represents the value of the variable
Clarity	Suitable for small to moderate datasets	Effective for large datasets or time series
Outliers	Outliers are visible as individual dots	Outliers may not be as evident with connected lines
Interpretation	Focuses on data distribution and patterns	Emphasizes trends, changes, or relationships

6)Box and Whisker Plot :

is defined as a visual representation of the five-point summary. The Box and Whisker Plot is also called as **Box Plot**. It consists of a rectangular "box" and two "whiskers." Box and Whisker Plot contains the following parts:

- **Box:** The box in the plot spans from the first quartile (Q1) to the third quartile (Q3). This box contains the middle 50% of the data and represents the interquartile range (IQR). The width of the box provides insights into the data's spread.
- **Whiskers**: The whiskers extend from the minimum value to Q1 and from Q3 to the maximum value. They signify the range of the data, excluding potential outliers. The whiskers can vary in length, indicating the data's skewness or symmetry.
- **Median Line**: A line within the box represents the median (Q2). It divides the data into two halves, revealing the central tendency.
- **Outliers**: Individual data points lying beyond the whiskers are considered outliers and are often plotted as individual points.

-What is a Five-Point Summary?

- The five-point summary rundown comprises five key measurements: the base worth, the principal quartile (Q1), the middle (Q2), the third quartile (Q3), and the greatest worth. These measurements partition a dataset into four similarly estimated parts, uncovering important data about the dataset's focal inclination, spread, and skewness

-Uses of Box and Whisker Plot

1. Imagining Information Dispersion: Box plots are brilliant instruments for acquiring a visual comprehension of the circulation of a dataset. They give a speedy outline of the central tendency, spread, and state of the information dissemination, assisting with distinguishing whether the information is symmetric, slanted, or contains exceptions.

2. Contrasting Distributions: Box plots are valuable for looking at the circulations of different datasets one next to the other. This is especially important when you need to think about the qualities of various gatherings, populaces, or classes. **For instance,** Contrasting the grades of understudies from various schools or locales, examining the exhibition of different items or medicines in a review, etc.

3. Estimating Skewness: By looking at the box and whiskers' general lengths and positions, an individual can evaluate the skewness of the information. A more drawn-out tail on one side of the box recommends skewness that way.

4. Information Investigation: Box plots can act as starting tools for information investigation. They give a compact rundown of a dataset's key qualities, assisting with settling on the proper information investigation techniques or changes.

5. Statistical Analysis: Box plots are much of the time utilised close by measurable tests and investigations. They can assist with picturing the circulation of information before directing speculation testing or looking at the method for various gatherings.

6. Quality Control: In assembling and quality control processes, box plots are utilised to screen varieties in item determinations and distinguish imperfections or deviations from quality guidelines. They help recognise when an interaction is working inside satisfactory cutoff points or when it needs changes.

7. Navigation: Box plots furnish chiefs with an unmistakable and instinctive method for surveying information qualities. They are utilised in business, money, and medical care to go with informed choices given information synopses.

8. Risk Appraisal: In fields like finance and insurance, box plots can be utilised to envision the gamble related to various speculations or protection contracts. They assist partners with figuring out the possible fluctuation in returns or misfortunes.

9. General Wellbeing and Epidemiology: Box plots are utilised to imagine and think about well-being-related information. **For example,** the circulation of illness rates among various districts or segment gatherings.

10. Ecological Sciences: Box plots can be applied to examine natural information. **For example,** air quality estimations or water contamination levels, and survey varieties across time or areas.

--When to Use Box and Whisker Plot

1. Comparing Scores: When there is a need to think about the performance of students from various classes or schools, a box plot can assist with surveying the dispersion of test scores in each gathering and recognise whether one gathering beats the others.

2. Analysing Worker Compensations: While examining the pay rates of representatives in an organisation, one can utilises box plots to look at the compensation circulations among various divisions or occupation jobs, assisting with recognising differences or exceptions.

3. Evaluating Product Quality: In assembling, if one needs to screen the nature of an item, one can make box plots of estimations taken at different creation runs. This recognises varieties and whether the item satisfies quality guidelines.

4. Distinguishing Anomalies in Financial Data: While examining monetary information, like stock returns, one can utilise box plots to identify exception exchanging days or uncommon cost developments, which might show huge occasions or blunders in information.

5. Comparing Patient Recuperation Times: In medical care, one could utilise box plots to think about the recuperation seasons of patients who have various therapies or medical procedures. This can assist with figuring out which treatment approach is more compelling.

6. Assessing Marketing Campaigns: Marketers can utilise box plots to evaluate the effect of various publicising efforts by contrasting measurements like navigate rates or change rates across crusade varieties.

7. Observing Air Quality: Ecological researchers and offices use box plots to envision air quality information, contrasting pollutant concentrations across various monitoring stations or locales.

8. Assessing Investment Portfolios: Financial experts can utilise box plots to think about the circulations of profits for various venture portfolios, assisting investors and backers with understanding gamble and return compromises.

9. Comparing Housing Prices: Real estate marketers can utilise box plots to think about the costs of houses in various areas or urban communities, giving experiences in real estate market varieties.

10. Breaking down Crime Percentages: Law enforcement agencies can utilise box plots to look at crime percentages in various regions or after some time, distribute assets and focus on mediations.

--How to Make Box and Whisker Plot

1. Gather Information: Accumulate the dataset of which the envision is needed.

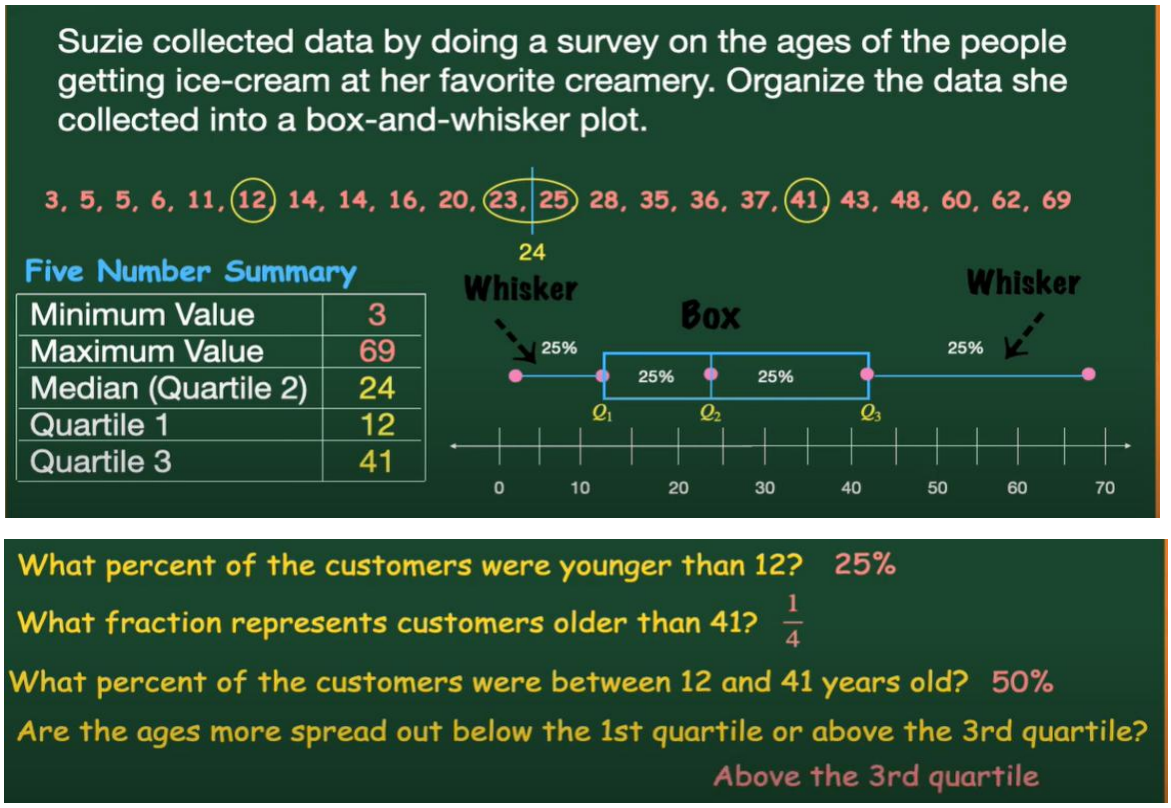
2. Work out Quartiles: Track down the main quartile (Q1), third quartile (Q3), and median (Q2) from the given information.

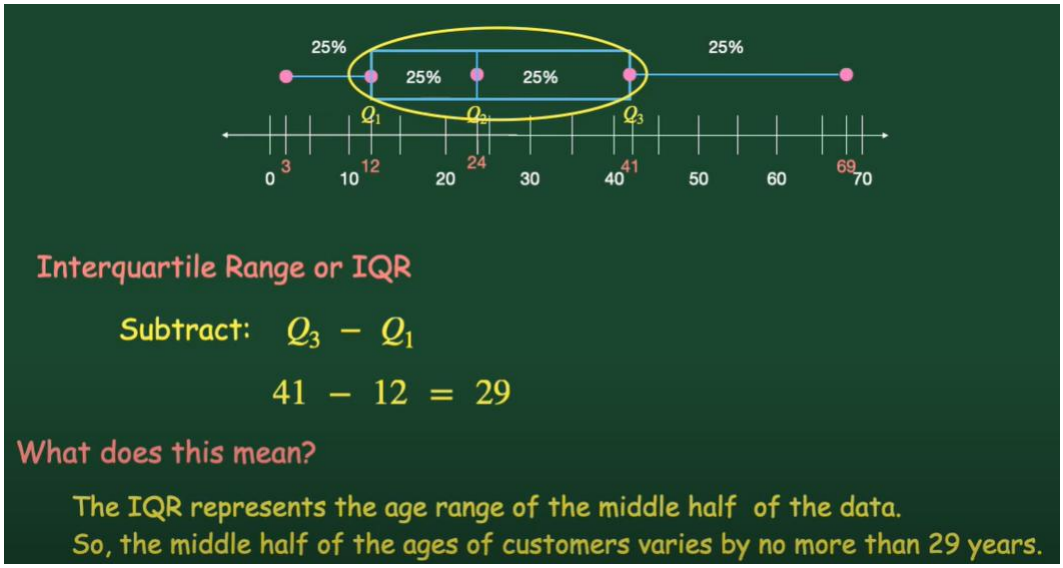
3. Decide Whiskers: Ascertain the base and most extreme qualities, barring anomalies.

4. Plot the Box and Whiskers: Draw a case from Q1 to Q3, a line inside the crate at Q2, and hairs from the base to Q1 and from Q3 to the greatest.

5. Recognise Outliers: Plot any pieces of information outside the stubbles as individual focuses.

EX:

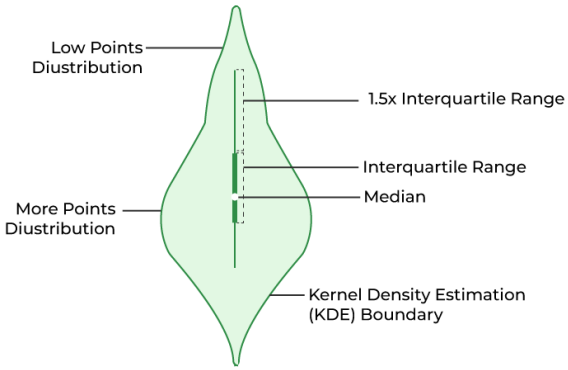




7) Violin: is a method to visualize the distribution of numerical data of different variables. It is quite similar to Box Plot but with a rotated plot on each side, giving more information about the density estimate on the y-axis. The density is mirrored and flipped over, and the resulting shape is filled in, creating an image resembling a violin. The advantage of a violin plot is that it can show nuances in the distribution that aren't perceptible in a boxplot. On the other hand, the boxplot more clearly shows the outliers in the data. Violin Plots hold more information than box plots, they are less popular. Because of their unpopularity, their meaning can be harder to grasp for many readers not familiar with the violin plot representation.

***How to read:**

The violin plot uses a kernel density estimation technique for deciding the boundary of the plot. A Kernel density estimation (KDE) is a statistical technique that is used to estimate the probability density function (PDF) of a random variable based on a set of observed data points. It provides a smooth and continuous estimate of the underlying distribution from which the data is assumed to be generated.



- **A white Centered Dot at the middle of the graph** – The white dot point at the middle is the median of the distribution.
- **A thin gray bar inside the plot –** The bar in the plot represents the Quartile range of the distribution.
- **A long thin line coming outside from the bar –** The thin line represents the rest of the distribution which is calculated by the formulae Q1-1.5 IQR for the lower range and Q3+1.5 IQR for the upper range. The point lying beyond this line are considered as outliers.
- **A line boundary separating the plot-** A KDE plot is used for defining the boundary of the violin plot it represents the distribution of data points.