



STATISTIC COURSE SUMMARIZATION PART 3 - TASK 6

Made by
Rawan Hatem

TYPES OF GRAPHS & WHEN TO USE THEM

- Bar graph: Shows the number in particular categories.
- Circle graph: Compares parts of the data to the whole.
- Double bar graph: Good for comparing data.
- Box and whiskers: Show measures of variation.
- Histogram: Good to show frequency of data divided into intervals.
- Line graph: Shows change over time.
- Line plot: Shows frequency on a number line.

Histogram

HISTOGRAM & HOW TO MAKE IT

- A histogram is a type of graph that shows how often different values occur in a set of data.
- To create a histogram, you need to collect data and organize it into a frequency table, which shows the range of values and the number of occurrences in each range.
- The horizontal (x) axis of a histogram represents the range of values, while the vertical (y) axis represents the frequency or "how many" in each range.
- The data from the frequency table is plotted onto the histogram using bars that touch each other, unless there is a range with zero frequency.
- The completed histogram provides a visual representation of the data and can help identify patterns and trends.

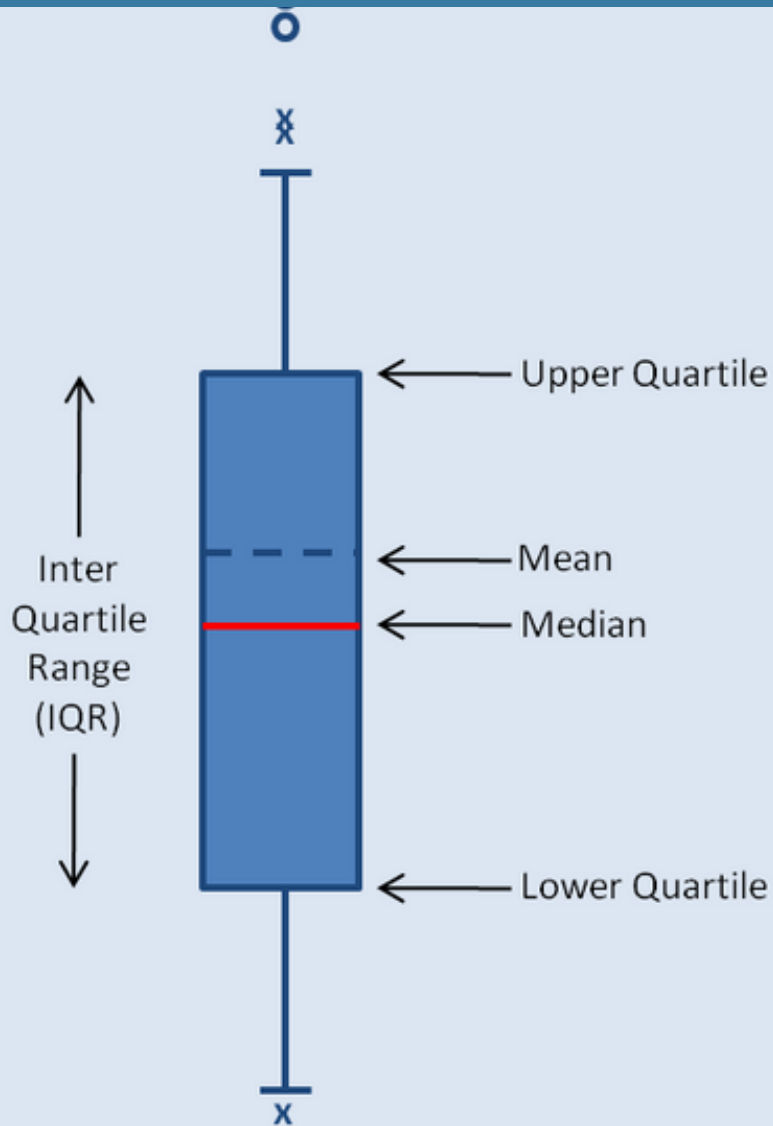
STEM & LEAF PLOT/DIAGRAM

- A **Stem and Leaf** Plot is: a graph that organizes data by using the place values of numbers.
- The **left side** of the plot is the "**stem**," and the **right side** is the "**leaf**."
- **To create a Stem and Leaf Plot:**
 - a. **Order the numbers** from least to greatest.
 - b. **Draw a t-chart** with a vertical line separating the stem and leaf.
 - c. The **stem** represents the **tens place**, and the **leaf** represents the **ones place**.
 - d. Write the **stem values** in **counting order**.
 - e. Write the **leaf values** in order from **least to greatest**.
- A **key** is essential for **understanding the Stem and Leaf Plot** and show the **relationship between the stem and leaf values**.
- **Stem and Leaf** Plots can be **used for larger numbers and decimal numbers**.
- For larger numbers, the **stem** can have **single and double digits**, but the **leaf** can **only** have **single digits**.
- For **decimal numbers**, the **stem** represents the **whole number**, and the **leaf** represents the **decimal part**.

Stem	Leaf
1	2 4 8
2	2 2 3 5
3	
4	5 7 8

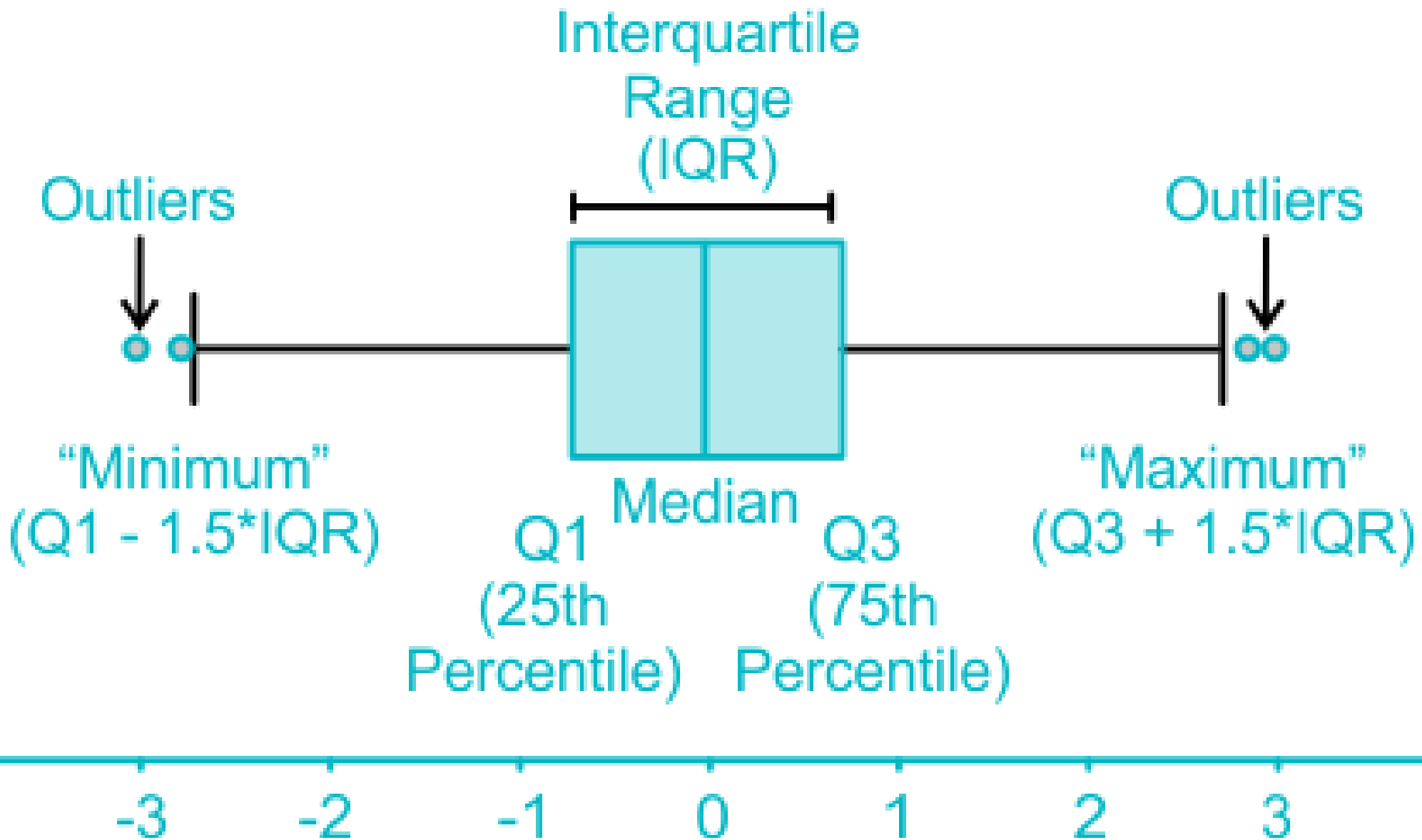
Key: 4 | 7 = 47

BOX & WHISKER PLOT



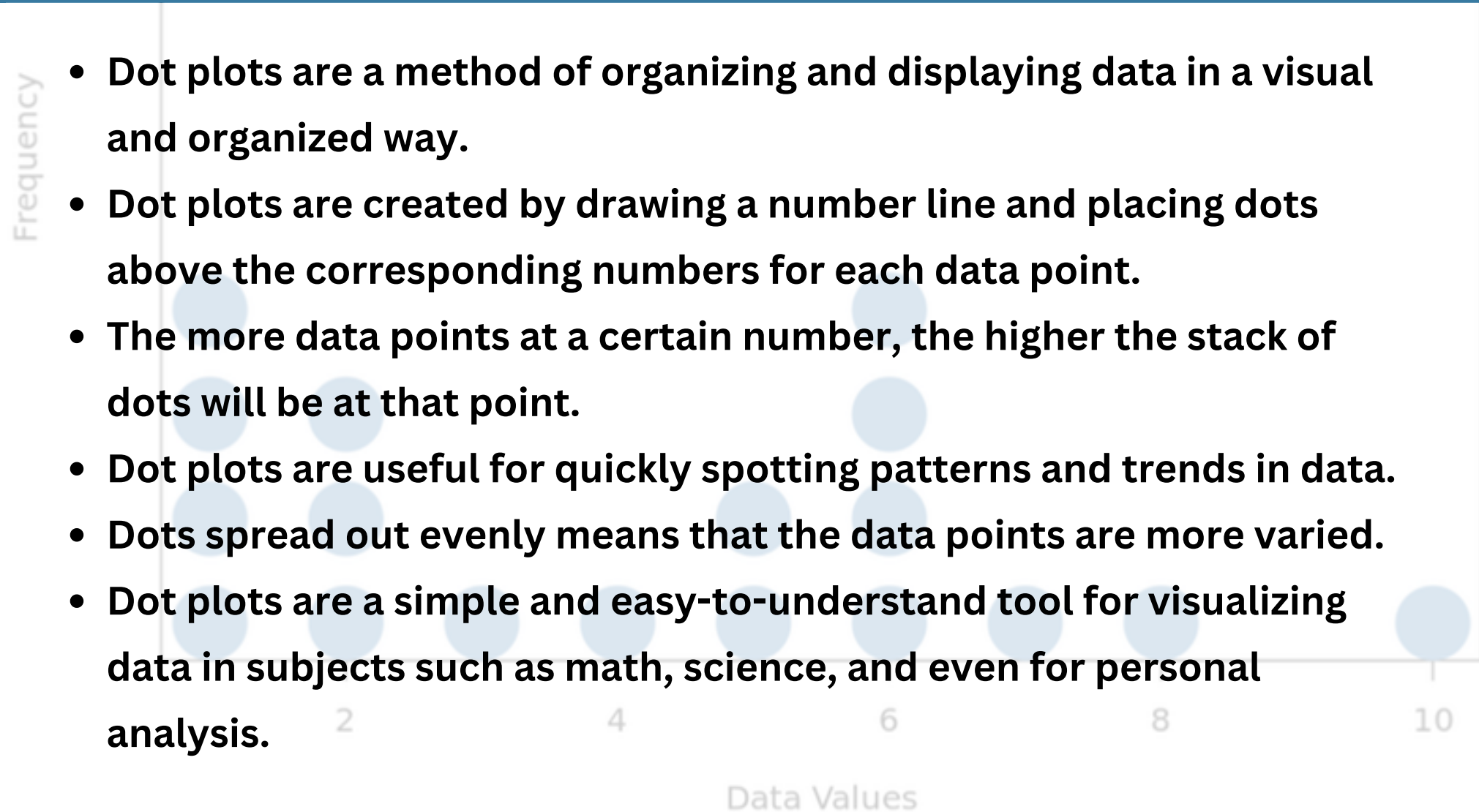
- The five-number summary is used to create a box-and-whisker plot, which visually represents the data set divided into four equal parts.
- The box in the box-and-whisker plot encompasses the three quartiles, with a line drawn in the middle to separate it into two equal parts.
- Whiskers are drawn from quartile 1 to the minimum value and from quartile 3 to the maximum value.
- Each section of the box-and-whisker plot represents 25% of the data, and the whiskers also represent 25% of the data.
- interquartile range the difference between quartile 3 and quartile 1, representing the middle 50% of the data without variation greater than the IQR value.

IQR REPRESENTATION IN BOX PLOT



DOT PLOT

- Dot plots are a method of organizing and displaying data in a visual and organized way.
- Dot plots are created by drawing a number line and placing dots above the corresponding numbers for each data point.
- The more data points at a certain number, the higher the stack of dots will be at that point.
- Dot plots are useful for quickly spotting patterns and trends in data.
- Dots spread out evenly means that the data points are more varied.
- Dot plots are a simple and easy-to-understand tool for visualizing data in subjects such as math, science, and even for personal analysis.



PIE CHART

- A pie chart is a type of graph divided into slices.
 - Each slice in a pie chart represents a percentage of the whole.
 - how to calculate the percentage for each slice in a pie chart:
 - Add up all the values in the data set to find the total.
 - Divide each value by the total to find the decimal equivalent.
 - Multiply the decimal equivalent by 100 to find the percentage.
-



BAR CHART

- A bar chart, also known as a bar graph, compares different types of information using bars or columns to reflect the data.
- The horizontal axis (x-axis) of a bar chart typically represents a type of variable that is qualitative, meaning it is categorical and can only be expressed in words.
- The vertical axis (y-axis) of a bar chart represents values that determine the height of each bar, with the numbers on the y-axis reflecting frequency or the number of times something occurs.
- When creating a bar chart, it's important to ensure that the width of the bars is the same, as well as the space between each one.
- When choosing the increments on the y-axis, it's important to consider the highest frequency to ensure the chart is easy to read.

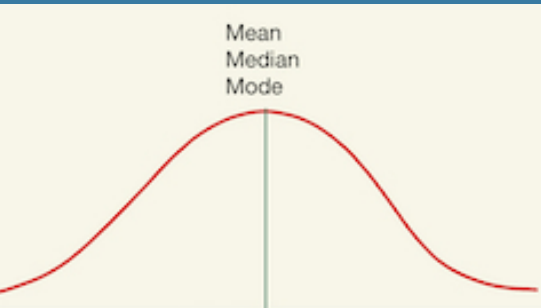


SYMMETRY & SKEWNESS



- Symmetry and skewness refer to the shape of a distribution.
 - A distribution is symmetrical if it can be divided into two equal sizes of the same shape.
 - A skewed distribution has asymmetry and can be skewed to the left or right.
 - Skewness is determined by the direction in which the data points cluster.
 - A distribution is skewed to the left if it has a long tail trailing towards the left.
 - A distribution is skewed to the right if it has a long tail trailing towards the right.
-

SYMMETRY & SKEWNESS.CONT



Zero Skewness



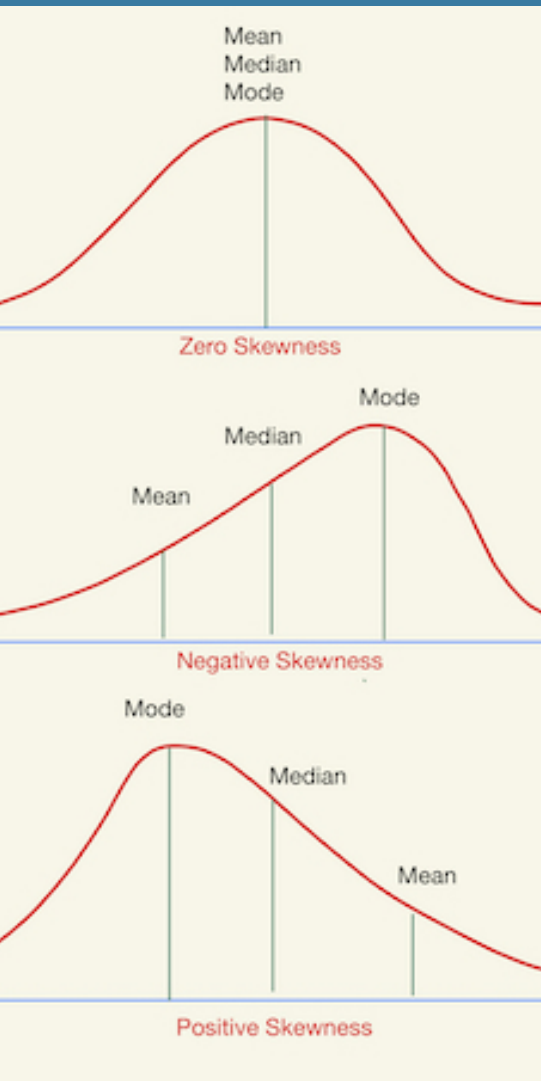
Negative Skewness



Positive Skewness

- In a symmetrical distribution, the plane of symmetry is at the median.
 - If a distribution is skewed to the left, the mean is less than the median.
 - If a distribution is skewed to the right, the mean is greater than the median.
-

DETERMINE SYMMETRY & SKEWNESS FROM STEM-LEAF PLOT & BOX PLOT



- To determine skewness in a stem plot, flip it onto its side and ensure the stems are positioned like a regular number line.
 - In box plots, the presence of outliers may affect skewness interpretation.
 - If boxes in a box plot are unequal, the larger box determines the skew.
 - If boxes are equal in size, look at the whiskers to determine skewness.
 - In a symmetrical distribution, the plane of symmetry is at the median.
-



HEATMAPS

- A heatmap is a **graphical representation of data** with values in a **matrix are represented as colors**.
 - heatmaps use a color scale **to visualize the intensity of values**, with **warmer colors** indicating **higher values** and **cooler colors** indicating **lower values**.
 - **Heatmaps** are commonly **used to visualize numerical data** across **two dimensions**, such as time and categories, or spatial data on a grid.
 - They are useful for identifying patterns, trends, and correlations in data, especially in large datasets.
 - Heatmaps are widely used in various fields including data analysis, biology, finance, and geospatial analysis.
-

VIOLIN PLOT

- A violin plot is a method of plotting numeric data, combining a box plot and a kernel density plot.
- The outer layer of the violin plot, the kernel density estimation, displays the density of the data at different values.
- The width of the plot at different values indicates the density of the data; a wider section suggests higher density, while a narrower section indicates lower density.
- A central box plot, often included inside the violin, provides additional details about the data's distribution.
- The central line in the box plot represents the median of the data, and the edges of the box represent the interquartile range, giving insights into the data spread.
- Whiskers in the violin plot indicate variability outside the upper and lower quartiles and can help identify outliers.
- Violin plots are useful for comparing distributions across different categories or groups.

Young Adults

Middle-aged

Seniors