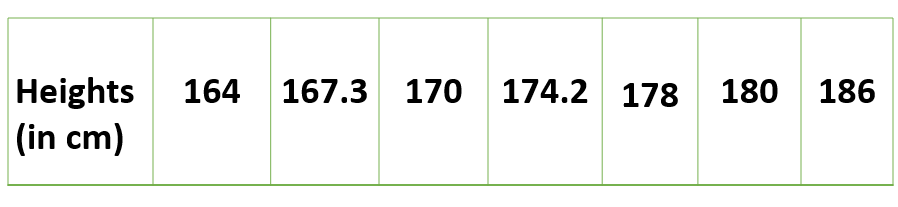
**Univariate and Bivariate Analysis**

Data analysis is an essential part of any business or research project. One important aspect of data analysis is understanding the concepts of univariate and bivariate analysis.

**Univariate Analysis**

Univariate analysis is the study of a single variable. It helps to understand the distribution of the data and how it varies. It is essential for summarizing data and making it more understandable. the main purpose of the analysis is to describe the data and find patterns that exist within it. Some examples of univariate analysis include measures of central tendency (mean, median, mode) and measures of dispersion (range, variance, standard deviation). Univariate analysis is useful when there is only one variable to consider. The example of a univariate can be height.



**Measures of Central Tendency**

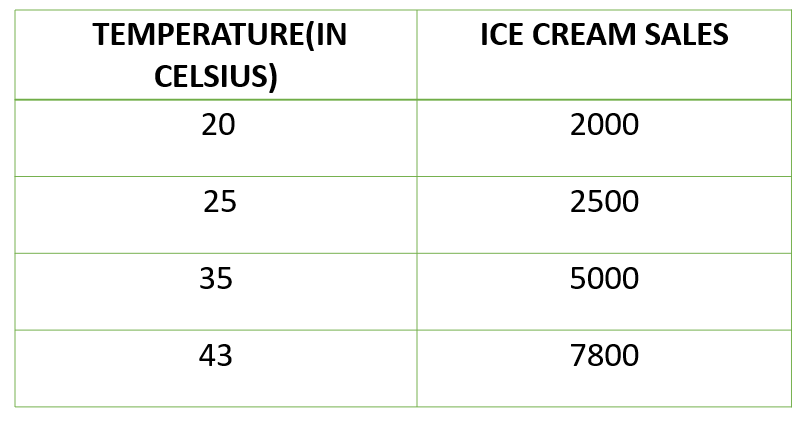
Measures of central tendency are used to summarize the data by finding the "center" or "typical" value of a dataset. The three most common measures of central tendency are the mean, median, and mode. The mean is the sum of all the values in a dataset divided by the number of values. The median is the middle value in a dataset, with half the values above it and half the values below it. The mode is the most frequent value in a dataset.

**Measures of Dispersion**

Measures of dispersion are used to describe how much the data varies from the "center" or "typical" value. The three most common measures of dispersion are the range, variance, and standard deviation. The range is the difference between the largest and smallest values in a dataset. The variance is the average of the squared differences between each data point and the mean. The standard deviation is the square root of the variance.

**Bivariate Analysis**

Bivariate analysis is the study of the relationship between two variables. It helps to understand how they are related and how changes in one variable affect the other. These variables are often plotted on X and Y axis on the graph for better understanding of data and one of these variables is independent while the other is dependent. It is useful for predicting future outcomes based on past data. Some examples of bivariate analysis include correlation analysis and regression analysis. The example of bivariate can be temperature and ice sales in summer vacation.



**Correlation Analysis**

Correlation analysis is used to determine the strength and direction of the relationship between two variables. It measures how much one variable changes when the other variable changes. The correlation coefficient ranges from -1 to 1, where -1 indicates a perfect negative correlation, 0 indicates no correlation, and 1 indicates a perfect positive correlation.

**Regression Analysis**

Regression analysis is used to predict the value of one variable based on the value of another variable. It is useful for understanding the relationship between two variables and predicting future outcomes. Simple linear regression uses a straight line to model the relationship between two variables, while multiple regression uses more than one predictor variable.

**Conclusion**

Understanding the concepts of univariate and bivariate analysis is essential for anyone working with data. Univariate analysis helps to summarize the data and make it more understandable, while bivariate analysis helps to understand the relationship between two variables. By using these techniques, we can gain valuable insights into the data and make more informed decisions.

**Extra note**

### **Multivariate data**

When the data involves **three or more variables**, it is categorized under multivariate. Example of this type of data is suppose an advertiser wants to compare the popularity of four advertisements on a website, then their click rates could be measured for both men and women and relationships between variables can then be examined. It is similar to bivariate but contains more than one dependent variable. The ways to perform analysis on this data depends on the goals to be achieved. Some of the techniques are regression analysis, path analysis, factor analysis and multivariate analysis of variance (MANOVA).

There are a lots of different tools, techniques and methods that can be used to conduct your analysis. You could use software libraries, visualization tools and statistic testing methods. However, this blog we will be compare Univariate, Bivariate and Multivariate analysis.

|  |  |  |
| --- | --- | --- |
| Univariate | Bivariate | Multivariate |
| It only summarize single variable at a time. | It only summarize two variables | It only summarize more than 2 variables. |
| It does not deal with causes and relationships. | It does deal with causes and relationships and analysis is done. | It does not deal with causes and relationships and analysis is done. |
| It does not contain any dependent variable. | It does contain only one dependent variable. | It is similar to bivariate but it contains more than 2 variables. |
| The main purpose is to describe. | The main purpose is to explain. | The main purpose is to study the relationship among them. |
| The example of a univariate can be height. | The example of bivariate can be temperature and ice sales in summer vacation. | Example, Suppose an advertiser wants to compare the popularity of four advertisements on a website.  Then their click rates could be measured for both men and women and relationships between variable can be examined |