

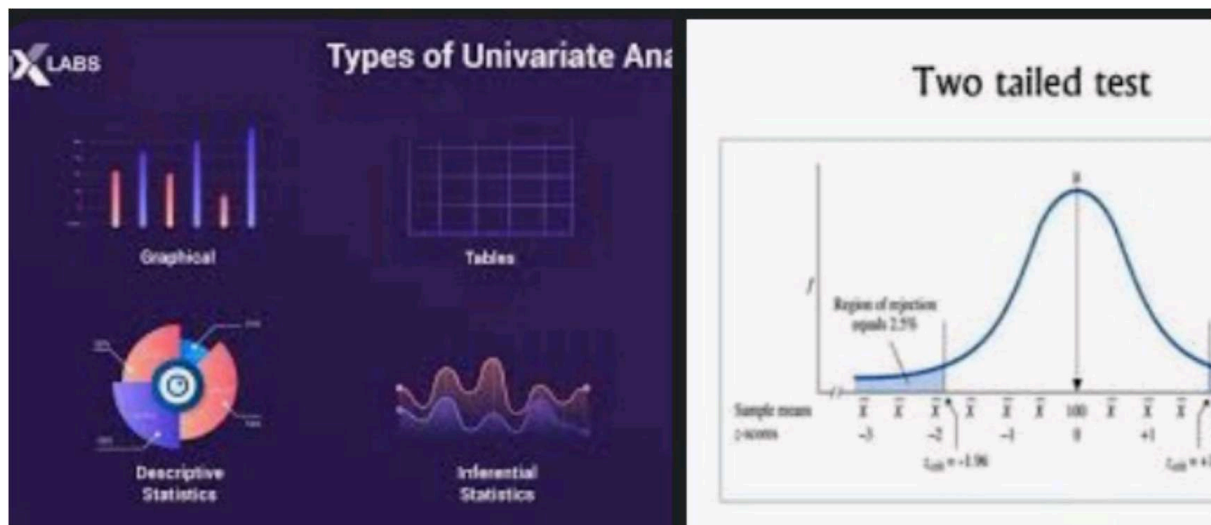
Date	01 May 2023
Team ID	NM2023TMID00732
Project Name	Data set
Maximum Marks	

Data set:

A data set is a collection of data. In the case of tabular data, a data set corresponds to one or more database tables, where every column of a table represents a particular variable, and each row corresponds

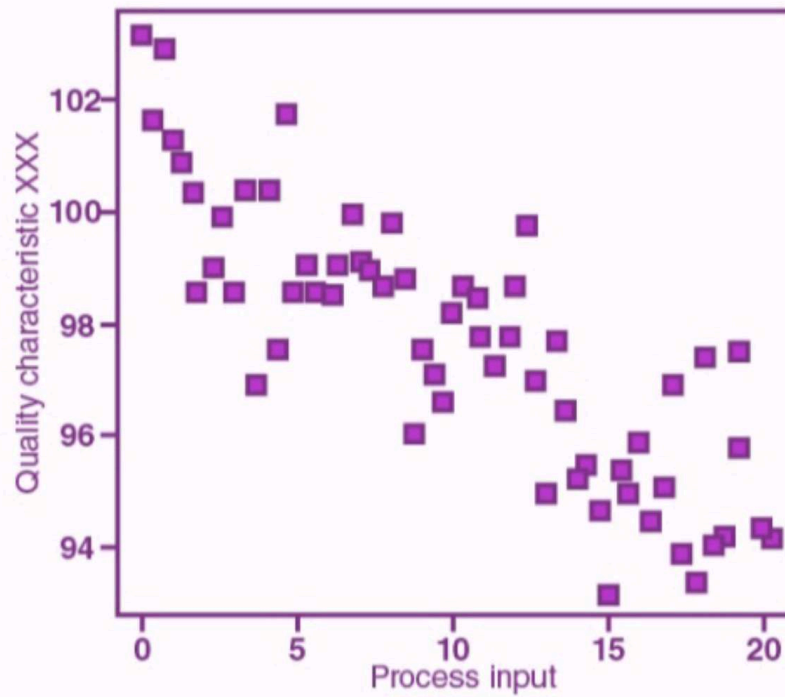
Univariate Analysis:

Univariate analysis explores each variable in a data set, separately. It looks at the range of values, as well as the central tendency of the values.



bivariate Analysis :

bivariate analysis is one of the simplest forms of quantitative analysis. It involves the analysis of two variables, for the purpose of dete

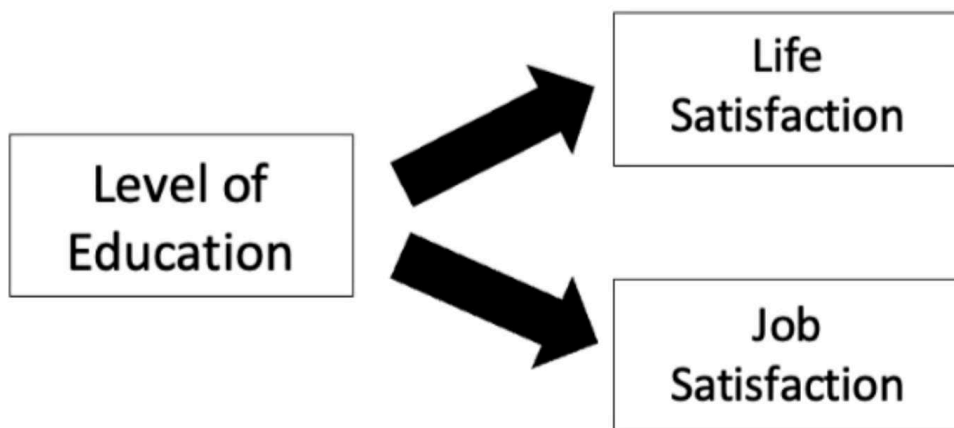


Scatterplot for quality characteristic XXX

Multivariate Analysis:

Multivariate analysis is defined as: The statistical study of data where multiple measurements are made on each experimental unit and where the relationships among multivariate measurements and their structure are important.

Multivariate

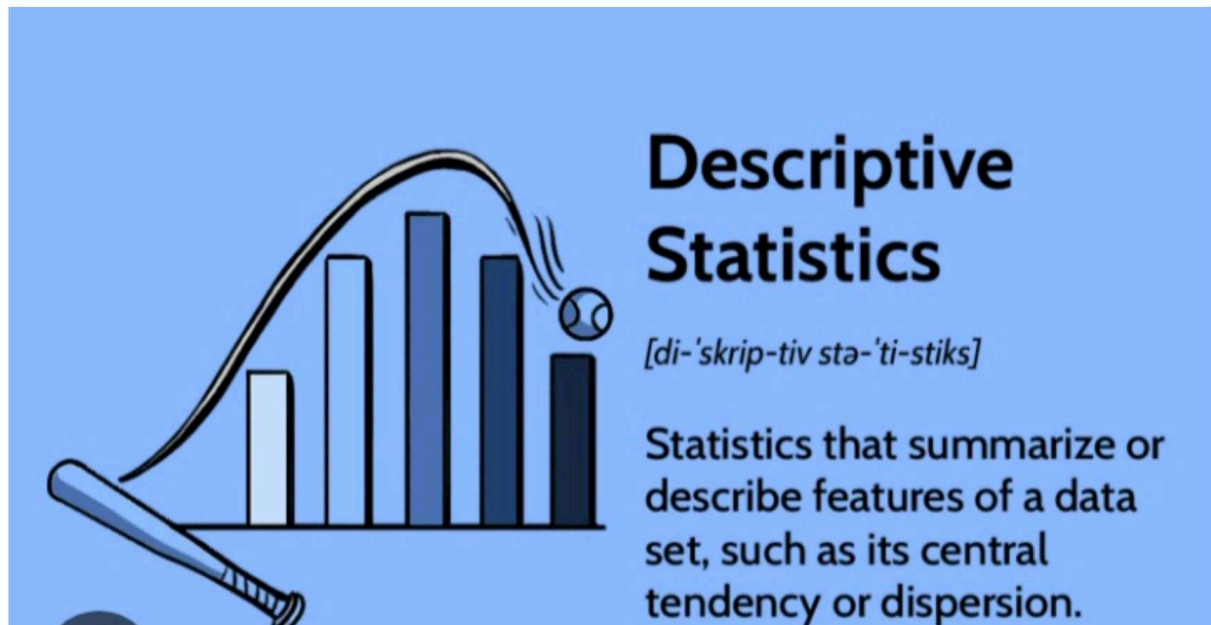


Univariate



Perform Descriptive Statistics of the data:

Descriptive statistics consists of three basic categories of measures: measures of central tendency, measures of variability (or spread), and frequency distribution.



Handle the missing Value:

	id	item	quantity	price	bought	forenoon	afternoon
0	1	milk	2.0	67.0	672.0	456.0	NaN
1	2	sugar	1.0	NaN	453.0	234.0	NaN
2	3	chips	NaN	45.0	456.0	322.0	NaN
3	4	coffee	2.0	45.0	672.0	564.0	NaN
4	5	meat	4.0	56.0	786.0	221.0	NaN
5	6	chocos	3.0	NaN	345.0	NaN	213.0
6	7	juice	1.0	78.0	765.0	NaN	344.0
7	8	jam	NaN	65.0	665.0	NaN	333.0
8	9	bread	3.0	NaN	NaN	NaN	567.0
9	10	butter	4.0	NaN	NaN	NaN	322.0

	A	B	C	D	E	F
1	Height	Weight	Country	Place	Number of days	Some column
2	12	35	India	Bengaluru	1	
3		36	US	New York	2	
4	13	32	UK	London		
5	15		France	Paris	4	
6	16	39	US	California	5	12
7				Mumbai		
8					6	
9						