

The logo of Galgotias University is a stylized 'G' composed of three overlapping, curved bands. The top band is yellow, the middle band is light blue, and the bottom band is light red. The 'G' is centered in the upper half of the slide.

## Statistical Foundations of Marketing

GALGOTIAS  
UNIVERSITY

## Topics covered

- Statistics in Marketing
- Descriptive Statistics
- Measures of central tendency
- Measures of dispersion

GALGOTIAS  
UNIVERSITY

## Statistics in Marketing

- Statistics are important for making decisions, new discoveries, investments, and predictions.
- Statistics are used in real-world scenarios from the worlds of business, sports, education, entertainment, and more.
- Statistics are applied in marketing to identify market trends, and to measure and evaluate the potential and success of marketing programs. The secret to successful marketing is to identify the target market accurately and to use effective marketing communications channels and tactics to reach it.
- So, what is **statistics**?

## About Statistics

- **Statistics** is **the discipline that concerns collecting and analysing numerical data** in large quantities\_\_\_\_\_
- 
- In applying statistics to a scientific, industrial, or social problem, it is conventional to begin with a **population** to be studied.
- Populations can be diverse groups of people or objects such as "all car buyers living in a country".

## Descriptive Statistics

- **Descriptive statistics** is the term given to the analysis of data that helps describe, show or summarize data in a meaningful way such that, for example, patterns might emerge from the data.
- **Descriptive statistics** are broken down into:
  - Measures of central tendency and
  - Measures of variability (spread).

GALGOTIAS  
UNIVERSITY

## Measures of central tendency

- **Central Tendency.** The central tendency of a distribution is an estimate of the "center" of a distribution of values.
- 
- There are three major types of estimates/measures of central tendency:
  - **Mean** --- -- Average value.
  - **Median** --- Middle value in the list arranged in an order.
  - **Mode** ----- Value that occurs most often in the list. If there is tie--  
- two modes.

## Measures of variability (dispersion)

- **Dispersion.** Dispersion refers to the spread of the values around the central tendency.
- **Measures of variability (dispersion)** , or the measures of spread, aid in analyzing how spread-out the distribution is for a set of data.
- The common measures of dispersion - the range, variance and the standard deviation.

GALGOTTIAS  
UNIVERSITY

## Standard Deviation (SD)

- The standard deviation **measures the dispersion of a dataset relative to its mean and is calculated as the square root of the variance.**
- The standard deviation is expressed in the **same units as the mean** is, whereas the variance is expressed in squared units.

GALGOTIAS  
UNIVERSITY



## SD Formula for values from Population and Sample

The "**Population** Standard Deviation":

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$$

The "**Sample** Standard Deviation":

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}$$

Looks complicated, but the important change is to divide by **N-1** (instead of **N**) when calculating a Sample Variance.

## Data Analysis - Statistical techniques

1. **Univariate techniques**– Is the simplest form of data analysis where the data being analyzed contains **only one variable**.
2. **Bivariate techniques**– Involves **two different variables**. The analysis of this type of data deals with causes and the analysis is done to find out the relationship among the two variables.
3. **Multivariate techniques**– Involves **three or more variables**, it is categorized under multivariate.

## The Distribution

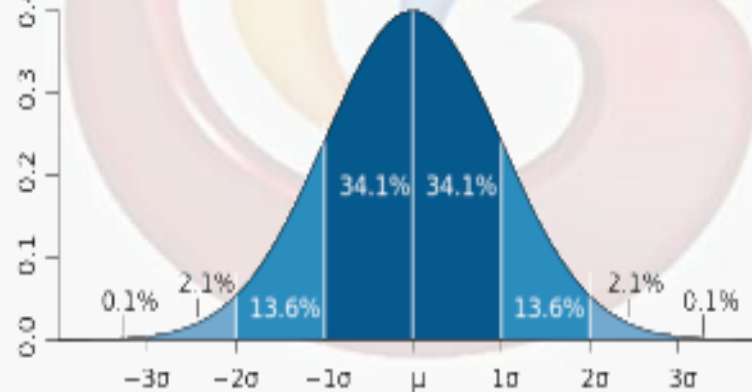
- Distribution is the way in which something is shared out among a group or spread over an area.
- The **distribution** is a summary of the frequency of individual values or ranges of values for a variable.
- The simplest distribution would list every value of a variable and the number of persons who had each value

## Distribution

- A **distribution in statistics** is a function that shows the possible values for a variable and how often they occur.
- There are many different types of probability distributions in statistics including: Basic probability distributions which can be shown on a **probability distribution** table. **Binomial** distributions, which have “Successes” and “Failures.” **Normal** distributions, sometimes called a **Bell Curve**.

## Normal distribution curve

- Normal distribution curves are sometimes designed with a histogram inside the curve.



GALGOTIAS  
UNIVERSITY

## General Linear Model

- The term “**linear**” refers to the fact that we are fitting a line. The term **model** refers to the equation that summarizes the line that we fit.
- The **general linear model (GLM)** is a statistical model that provides a general framework to explain or predict a **continuous dependent variable** by a set of **independent variables that can be categorical or continuous**.
- **$Y = a + bx$**  --- equation.

## References:

- Rackley, Jerry (2015). Marketing Analytics Roadmap: Methods, Metrics, and Tools. Apress, 2015
- Winston, Wayne L (2014). Marketing Analytics: Data-Driven Techniques with Microsoft Excel. Hohn Wiley & Sons, 1st Ed.
- Grigsby, Michael (2015). Marketing Analytics: A practical guide to real marketing science. Kogan page.