**Week 6 Lecture Notes**

**Types of Analysis**:

**Univariate Analysis**: Basic form of statistical data analysis Technique. Defined as analysis carried out on only one (“uni”) variable (“variate”) to describe the variable.

**Univariate Analysis is performed several ways:**

Frequency Distribution Tables

Histograms

Pie Charts

Bar Charts

**Bivariate Analysis**: One of the statistical analysis where two variables are observed One variable here is dependant while the other is independent.

**Multivariate Analysis**: any type of statistical analysis that is used to analyse more complex sets of data.

**Ways to conduct Multivariate analysis**:

Cluster Analysis

Variance Analysis

Multiple Regression Analysis

Factor Analysis

**Causal Analysis**: is where we try to establish cause and effect between variables. Find the cause of or/and effect in studies.

**Prescriptive Analysis**: You know what will happen but what sort of actions will you take in response.

**Inferential Analysis**: The use of a small sample of data to understand a larger population.

**Time Series Analysis**:

Ordered Sequence of values of a variable at equally spaced time intervals: **Financial Analysis**, **Weather Analysis**, **Network Data Analysis**.

Univariate: Perform Time Series Analysis with one variable

Chart, line chart, histogram

Description automatically generated

Multivariate: Perform Time Series Analysis with multiple variables

Graphical user interface

Description automatically generated with medium confidence

**Correlation Analysis**

Statistical method used to evaluate strength of relationship between two quantitative variables (high means strong relationship, Low means Hardly related)

Co-efficient of correlation is a numerical index telling us what extent the two are related and what extent the variations in one variable changes with the variation. This is symbolized either by r or p (Greek p) (Rho)

Positive, Negative, or Zero Correlation.

Pearson’s Correlation Coefficients (r): to measure strength of association between two variables in a correlation

**Regression:** A Method of modelling a target based on independent predictors.

**Linear Regression**: Modelling a linear relationship between two variables, Independent and Dependent. (Uses a straight line)

A picture containing indoor, light, dark

Description automatically generated

Text

Description automatically generated

**Multiple (Linear) Regression**: Its purpose is to predict the likely outcome based on several variables, plotting the relationship between these multiple independent variables and single dependent variables. **Uses several variables in order to predict the response variable.**

**Logistic regression**: Used when working with Binary Data (yes or no). It is a classification algorithm, used to predict a binary outcome based on independent variables.

**R-squared** is a goodness of fit measure for linear regression.