

ST 3011- Statistical Programming

Project Proposal - Group 06

Background of the study

There are many factors influencing placement. Some of them are Work Experience, Entrance Examination Score, Personal Interview Score, Grade Point Average, Internship, and Rating on Extra-curricular Activities. Nowadays Academic and Employability factors play a major role in placement. When considering this dataset, this consists of placement data of students in Jain University of Bangalore. It includes secondary and higher secondary school percentage and specialization. It also includes degree specialization, type and work experience and salary offers to the placed students. Finding a placement is not that much easy, it is necessary to find a placement with a better salary. For that, it is needed to find out the Academic and Employability factors which influencing placement.

Objectives

- To check whether there is the difference in salary between the Genders.
- To identify the percentage factors influenced a candidate in getting placed with high salary.
- To identify which degree specialization is much demanded by corporate.

Methodology

- Data Cleaning - Once the data is processed and organized, the data may be incomplete, contain duplicates, or contain errors. The need for data cleaning will arise from problems in the way that the data are entered and stored. Data cleaning is the process of preventing and correcting these errors. Change column names, Replacing NAs in the salary column and checking duplications can do for this dataset.
- Exploratory data analysis (Descriptive Statistics)

Descriptive statistics provide simple summaries about the sample and about the observations that have been made. Such summaries may be either qualitative or visual. Under descriptive analysis, there are two types of analyses as univariate and bivariate analyses.

❖ Univariate analysis - involves describing the distribution of a single variable. E.g.: Pie charts, Bar charts

❖ Bivariate analysis - When a sample consists of more than one variable, descriptive statistics are used to describe the relationship between pairs of variables. E.g.: Scatter plots, Box plots

- Data visualization in a Dashboard - A data dashboard is an information management tool that visually tracks, analyzes and displays key performance indicators (KPI), metrics and key data points to monitor the health of a business, department or specific process. For clearly shows the achieved results from the analysis to build a dashboard “Python” is used here.