SECOND YEAR B. TECH THIRD SEMESTER

ELECTRONICS AND COMMUNICATION ENGINEERING DEPARTMENT

MINI PROJECT REPORT(ITW-1)

ON

Data Analysis over data set of Job and Employees

By

- **Shubhendu Kumar (BT21ECE066)**
- **Harsh Bajaj (BT21ECE031)**

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY,

NAGPUR

(An Institution of National Importance by Act of Parliament)

BACKGROUND OF PROJECT:

After seeing the excel sheet of some data having a huge number of entries is difficult for us human beings to analyze and make good output out of that data. So we decided to work on this

METHODOLOGY:

We used four python libraries to get our job done.

- Numpy
- Pandas
- Seaborn
- Matplotlib

We started our project by exploring the data set. In this, we used Pandas data frame.

Now coming to the next we made data understandable by calculating median salary for different job roles, then we focused on How remote is their work? by plotting histogram with the help of Matlplotlib and Seaborn. After this we found out the highest and lowest salary.

Again, we came up with distribution of salary data visualization with graphs.

RESULTS:

We successfully implemented python libraries

- Numpy (we used it for array (salaries))
- Pandas (we used it for loading tabular dataframe)
- Seaborn & Matplotlib (we used them for making graphs more interactive through bar graph and histogram)

By this project we got more deeper understanding about these Python libraries.