

**AMCAT EDA Project**

**Aspiring Mind**

**Employment**

**Outcome 2015 (AMEO)**

**About me**

* Data Science student at Innomatics research labs , showcasing my resilience and adaptability in my academic and professional journey, certified in Python programming language and data analysis. Currently serving as a Data Science Intern at Innomatics Research Labs. I chose Data Science for its power in decisionmaking, Analyzing data, deriving insights, and solving problems.
* **LinkedIn -** <https://www.linkedin.com/in/lakshmaiah-rusum-551951289/>
* **GitHub -**https://github.com/Rusum-Lakshmaiah

# Business Problem

The business problem revolves around in this project involves understanding the employment outcomes of engineering graduates, focusing on factors such as salary, job titles, job locations, and various skills. This analysis aims to provide insights into the factors influencing salary, job preferences, and the overall employment landscape for engineering graduates.

**Objective:**

• The main objective of the project is to conduct an Exploratory Data

Analysis (EDA) on the Aspiring Mind Employment Outcome 2015 (AMEO) dataset. The analysis aims to uncover patterns, relationships, and insights regarding the employment outcomes of engineering graduates.

# Summary of the Data

* The dataset, from Aspiring Minds, reveals insights into the employment outcomes of engineering graduates. With around 4000 records and 40 variables, it includes a unique identifier (UID), annual salary (CTC), dates of joining (DOJ) and leaving (DOL), demographic data, academic performance metrics, college details, skill scores (English, Logical, Quantitative, Domain, Computer Programming), and personality test scores (Conscientiousness, Agreeableness, Extraversion, Neuroticism, Openness to Experience). This comprehensive dataset sets the stage for an in-depth exploratory data analysis, providing valuable information on various facets of graduates' professional journeys.

**Exploratory Data Analysis:**

***a. Data Cleaning Steps:***

**1. Column Removal:**

* Dropped columns 'Data Source' and 'ID' as they were not providing relevant information for analysis.

**2. Checking Null Values:**

• There is no null values present in the given dataset

**Data Manipulation:**

1. **Date Conversion:**

1. Converted date columns ('DOJ', 'DOL', 'DOB') to datetime format.

1. **Creating 'EmploymentStatus' Column:**

1. Added a new column 'EmploymentStatus' based on the presence of 'DOL' values, indicating whether the employee is currently employed or has left.

1. **Calculating 'YearsOfExperience':**
   * 1. Calculated the 'Years Of Experience' based on the 'DOJ' column, representing the number of years an employee has been working.
     2. And creating a new column to find the 'Years Of Experience
2. **Handling 'Present' Values in 'DOL':**
   * Replaced 'present' values in the 'DOL' column with '9999-12-31' to represent ongoing or indefinite employment.

**Univariate Analysis:**

The analysis includes histograms and Displots, boxplots, and QuantileQuantile (QQ) plots, offering insights into the distribution and normality of numerical features. A statistical summary provides key metrics. The frequency distribution of categorical variables, such as DOL, Designation, and JobCity, is explored. Interactive count plots highlight Gender and Degree distribution. Additional visualizations cover top designations, JobCity counts, and box plots for College GPA and Salary. The highest salary and "Senior Software Engineer" count are identified, and the top 10 designations are also identified, presenting a comprehensive overview of the dataset succinctly.

**Bivariate Analysis steps:-**

* + Explored the relationship between two variables:-
  + Salary vs Gender:- In my analysis of Salary vs Gender, I found that males tend to earn more than females, despite the mean ages being equal.This highlights a potential gender wage gap in the industry,an important issue that merits further investigation and action
  + Salary vs JobRole:-In my analysis of salary and job role, I found that managers, system engineers, software engineers, and test engineers command the highest salaries. This suggests that these roles are highly valued in the industry, reflecting the complexity of the work and the skills required. It's an important insight for those considering these career paths.
  + Salary vs Degree:- In my analysis of Salary and Degree, I found an interesting trend. Despite the common perception that higher education leads to higher earnings, my data suggests that individuals with a Bachelor's degree have more chances to earn better than those with a Master's degree. This could be due to various factors such as industry demand, practical experience, or the specific roles they are in. It's a valuable insight that challenges traditional beliefs about education and earnings.
  + Salary vs SpecializationIn:- my analysis of salary and specialization, I found that individuals with a Bachelor's degree tend to earn better than those with a Master's degree. Interestingly, students who have specialized in Industrial Engineering earn more compared to those who are Computer Science Engineers. This could be due to the industry demand for specific skills or the nature of the roles they are in. It's a valuable insight for students when considering their specialization and potential earnings.

.

**Key Bussiness Question**

* + An attempt to verify a Times of India claim regarding Computer Science Engineering jobs yielded an empty dataset, The dataset didn't contain information for specific roles like Programming Analyst, Software Engineer, Hardware Engineer, or Associate Engineer within the Computer Science domain.
  + Meanwhile, an analysis of the relationship between gender specialization using a chi-squared test revealed a significant association (Chi-squared =2.77, p-value = 0.24). The null hypothesis was rejected, The p-value of the statistic is greater than the significance level, indicating that the variables are independent. Hence, there's no association between gender and specialization.

**Agenda (This should be the PPT flow)**

* + **Conclusion (Key finding overall)**
  + **Q&A Slide**
  + **Your Experience/Challenges working on Web Scraping – Data Analysis Project.**

**Key Insights:**

* + **Most of the Students working in Bangalore.**
  + **senior software engineer gets more Salary.**
  + **Tier2 students are getting more job opportunities.**
  + **When Experience Increases Salary Increases.**
  + **Electronics And Communication Engineering background Students gets more Salary.**
  + **Most of the students are belonging to B.Tech and B.E background**

**Conclusion:**

"This Exploratory Data Analysis (EDA) project has been a valuable learning experience for me.Throughout the project, I have honed my skills in data analysis, gaining proficiency in applying statistical concepts to uncover meaningful insights. The hands-on nature of the analysis has enhanced my ability to navigate and make sense of large datasets, allowing me to extract valuable information efficiently.

Overall, this project has significantly contributed to my analytical skill set, providing a solid foundation for interpreting data and drawing meaningful conclusions."

# THANK YOU