INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

AMCAT – Exploratory Data Analysis

About me :

Hi! I'm **Gopisetty Sai Charan** (IN9240393) from Sri Venkateswara University College of Engineering, Tirupati currently pursuing 4th Year of Bachelors in Electronics and Communication Domain.

I want to learn data science because I love solving problems and working with data to find patterns and insights. By learning data science, I aim to develop the skills needed to contribute to innovative projects and drive meaningful outcomes in the industry.

I have previously worked at Zettabyte Plus as a Front-end Web Developer and at YBI Foundation as a Data Analyst Intern.

My Github Profile : https://www.github.com/Sai34Charan/
My Linkedin Profile : https://www.linkedin.com/in/gopisetty-sai-charan-gsc12104016/

INNOMATICS RESEARCH LABS

Business Problem :

- The main aim of this project is with respect to current job market over higher competition, how likely do freshers as well as experienced candidates can get the job in concerned domain of interest. In which how the salary will deviate with respect to the Job location, Specialization, Skills, Exam Scores, College Tiers, Experience and more parameters are involved.

Objective of the project :

- To Perform Univariate and Bivariate Analysis to understand data distribution in order to identify patterns and relationship with the data, offering valuable insights for employment outcomes of engineering graduates.

Exploratory Data Analysis:

This process involves concerned points in order extract insights through Visual and Non-Visual approaches.

1. **Data Cleaning Steps** :

   Removing duplicates, handling missing values, correcting errors.

2. **Data Manipulation Steps** :

   Sorting data, creating new variables, merging datasets.

3. **Univariate Analysis Steps** :

   Non-Visual: Calculating mean, median, mode, and standard deviation.

   Visual: Plotting histograms and box plots to visualize data distribution.

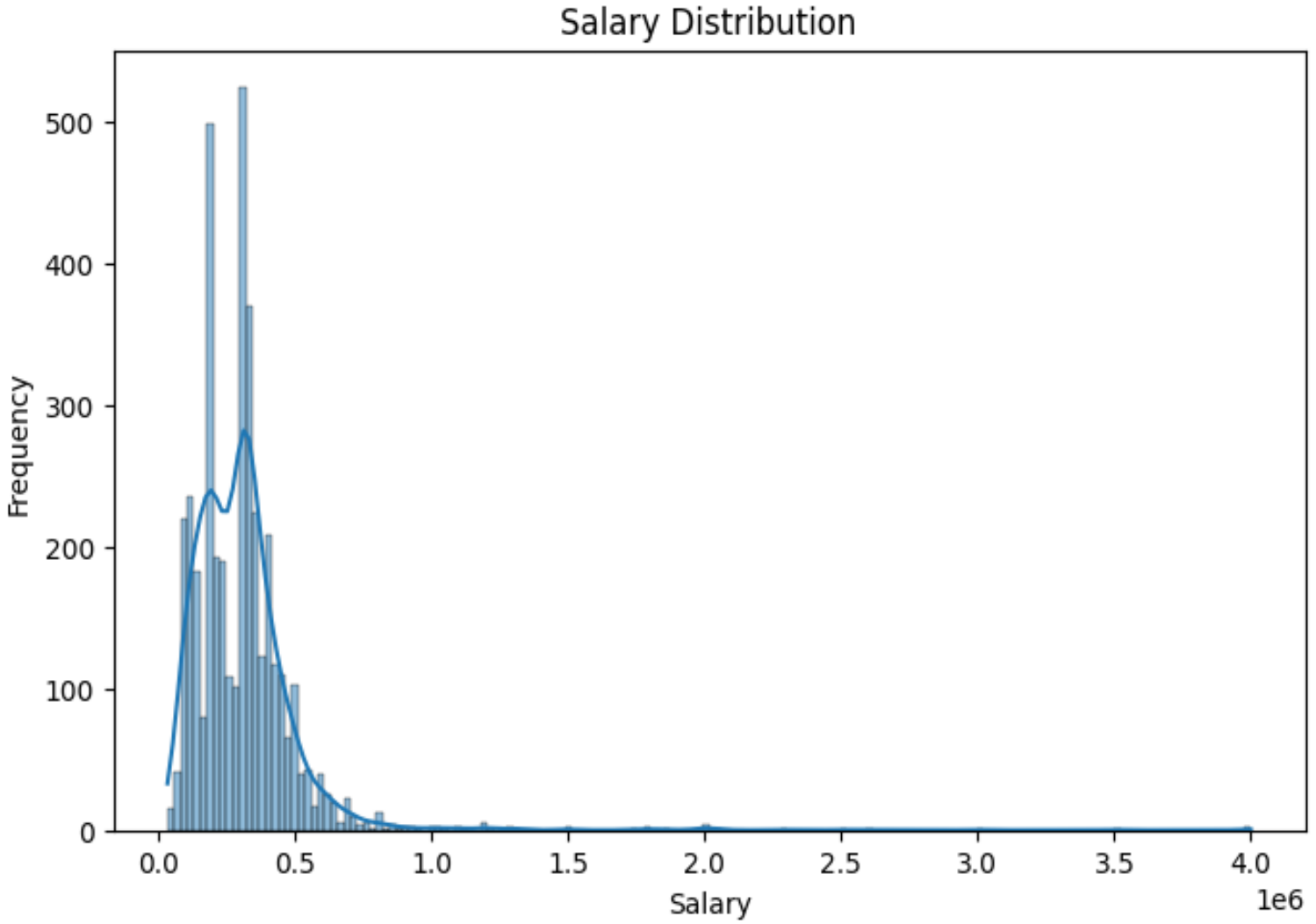
4. **Bivariate Analysis Steps** :

   Non-Visual: Calculating correlation coefficients, performing regression analysis.

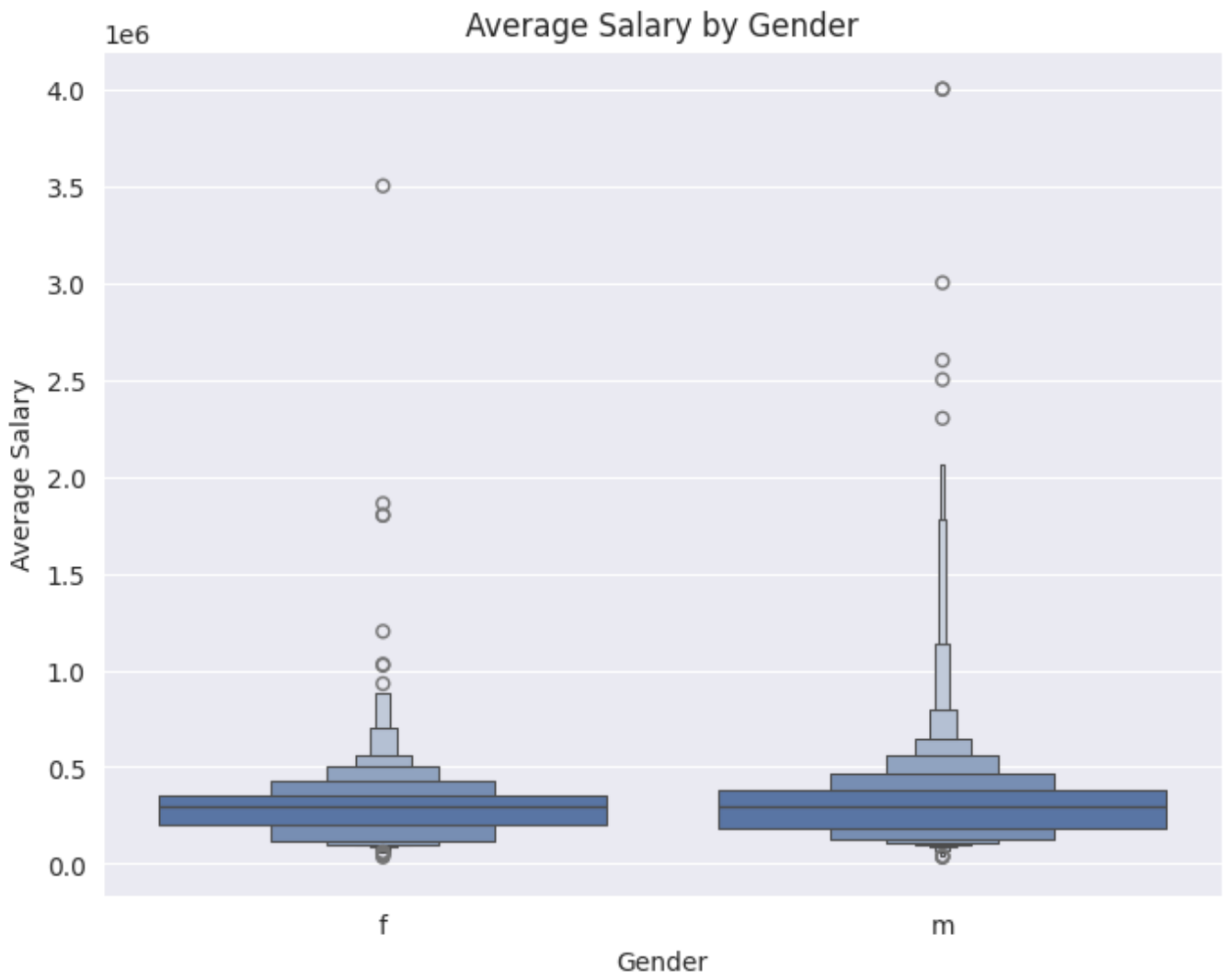
   Visual: Creating scatter plots and heatmaps to explore relationships between variables

Univariate Analysis :

- The targeted salary attribute having outliers that are skewed right . It indicates that most of the candidates having higher salaries than the average.
- A significant peak at 4,000,000 suggest that larger number of employees work at this range.



INNOMATICS RESEARCH LABS



Bivariate Analysis :

- From the Box plot it is seen that males are tend to earn a higher distributed salaries compared to the females though the average salary tend to approximately same.
- A similar bar plot is drawn between the other columns in this bivariate analysis.

INNOMATICS RESEARCH LABS

Research Questions :

Is there any Relationship between Gender and Specialization?

Ans.

From, Chi- Square test statistics : 45.25649, P- value : 3.51, There is a relationship between Gender and Specialization.

How do quantitative ability scores correlate with computer programming scores?

Ans.

A strong positive correlation value of 0.64 can be seen that there is linear relationship between programming scores and quantitative scores .

Conclusion :

Through this entire Exploratory Data Analysis, the drawn insights deals with the salary field which gives the hypothetical ideas about the how the particular parameters contributes the job opportunities that may be the specialization, exam scores, skills and more in the path. Few key points are...

1. **Salary Distribution**:
    - The salary attribute shows right-skewed outliers, indicating that a few candidates earn significantly more than the average salary.

2. **Engineering Scores**:
    - Fields like Computer Science, Mechanical Engineering, Civil Engineering, Electrical Engineering, and Telecom Engineering have right-skewed outliers, suggesting higher scores pulling the mean to the right.

3. **Personality Scores**:
   - Standardized scores for conscientiousness, agreeableness, and extraversion are left-skewed, indicating most students scored lower in these traits.

4. **Gender Distribution**:
   - Approximately 76% of AMCAT exam takers are male, while 24% are female.

5. **Specializations**:
   - Electronics and Communication Engineering has the highest number of candidates (800), while Electronics and Instrumentation Engineering has the least (32).

6. **State-wise Participation**:
   - Uttar Pradesh has the highest number of candidates (915) taking the AMCAT exam, compared to other states.

THANK YOU

INNOMATICS RESEARCH LABS