Project Report on AMCAT Data Analysis

1. Introduction

This report presents an exploratory analysis of the AMCAT dataset, which provides insights into the employability and salary trends of fresh graduates in various job designations across different cities in India. The analysis aims to uncover relationships between salary and various features, including date of joining (DOJ), job designation, educational background, and personality traits.

2. Objectives

- To analyse the salary distribution across different job designations and locations.
- To explore the relationship between salary and various factors such as education percentage, gender, and personality traits.
- To identify trends in employment for fresh graduates based on the dataset.

3. Data Description

The dataset contains the following columns:

- **ID:** Unique identifier for each record.
- **Salary:** The annual salary of the employee.
- **DOJ (Date of Joining):** The date when the employee joined the organization.
- **DOL (Date of Leaving):** The date when the employee left the organization (if applicable).
- **Designation:** The job title held by the employee.
- **JobCity:** The city where the job is located.
- **Gender:** The gender of the employee (male or female).
- **DOB (Date of Birth):** The birth date of the employee.
- **10percentage:** The percentage scored in the 10th-grade examinations.
- **Engineering Specializations:** Various columns indicating the scores or participation in engineering disciplines (e.g., Computer Science, Mechanical Engineering).
- **Personality Traits:** Columns indicating personality traits such as conscientiousness, agreeableness, extraversion, neuroticism, and openness to experience.

4. Methodology

The analysis was conducted using the following steps:

1. **Data Preprocessing:** Cleaned the data to handle missing values and convert date formats.

2. **Univariate Analysis:** Examined individual columns to understand distributions and identify outliers.
3. **Bivariate Analysis:** Explored relationships between salary and other variables, including job designation, educational percentage, and personality traits.

5. Findings

5.1 Salary Distribution

- The salary data ranges from a minimum of ₹200,000 to a maximum of ₹1,100,000, indicating a wide disparity in earnings based on designation and experience.
- Most salaries are concentrated between ₹300,000 to ₹600,000, reflecting typical starting salaries for fresh graduates.

5.2 Designation and Salary Correlation

- Certain designations, such as "senior software engineer," exhibit higher salary averages compared to roles like "assistant manager."
- Locations like Bangalore and Gurgaon show higher salary offerings, likely due to the presence of tech hubs.

5.3 Impact of Educational Background

- There is a positive correlation between the 10th-grade percentage and salary, suggesting that better academic performance may lead to higher salary offers.
- An analysis of engineering specialization shows that graduates from Computer Science tend to command higher salaries compared to other engineering disciplines.

5.4 Gender Analysis

- The dataset shows a mix of salaries for both genders, but a further investigation is necessary to determine if gender influences salary levels significantly.
- Notably, female employees exhibit fewer outliers in salary compared to their male counterparts, which may indicate a need for deeper analysis regarding gender pay gaps.

6. Conclusion

The AMCAT dataset reveals insightful trends regarding the employment landscape for fresh graduates in India. Salary distribution highlights a significant variance based on designation, educational background, and job location. Future work can focus on addressing gender disparities and exploring the long-term career trajectories of these graduates.

INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Analyzing and Prediction of AMCAT Scores

About me

- I’m Thalla Samiksha, I have recently completed my Graduation from Vignana Jyothi Institute of Arts And Sciences

Why do you want to do Data Science?

I want to learn data science because it combines my passion for problem-solving with the ability to uncover patterns and insights from data. In today's world, data science skills are highly sought after across various industries, opening up numerous exciting career opportunities. By mastering data science, I can contribute to better decision-making and drive innovations that have a positive impact on businesses and society. Additionally, the field of data science is continuously evolving with new technologies and methods, offering endless opportunities for learning and growth, which keeps me motivated and engaged.

LinkedIn URL: https://www.linkedin.com/in/samikshathalla/

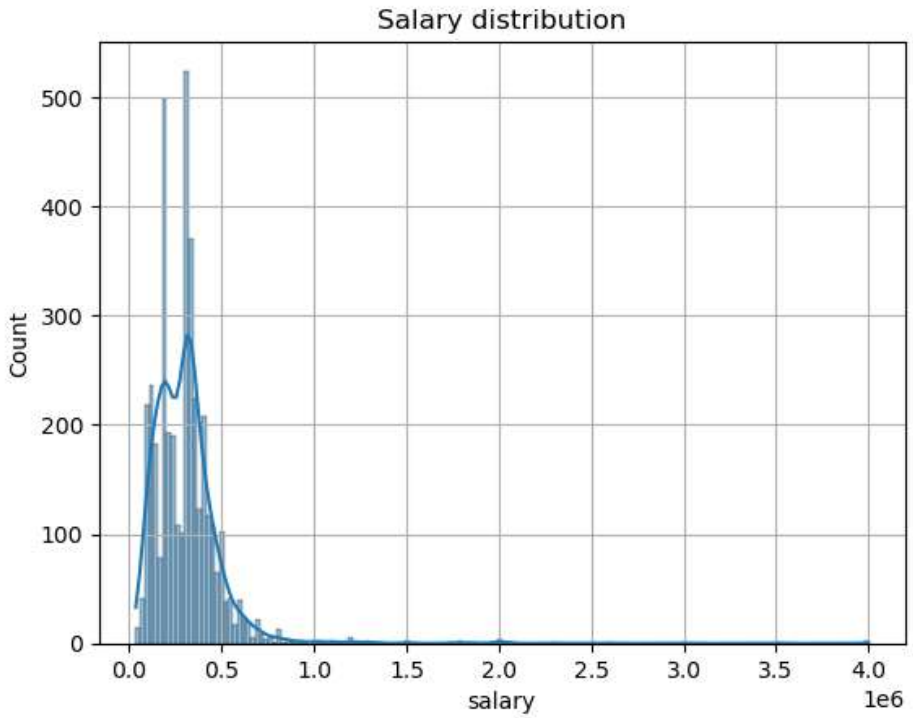
Objective of the project

- Perform comprehensive univariate and bivariate analyses to understand data distributions, detect outliers, and find relationships among variables.

Data Set Overview

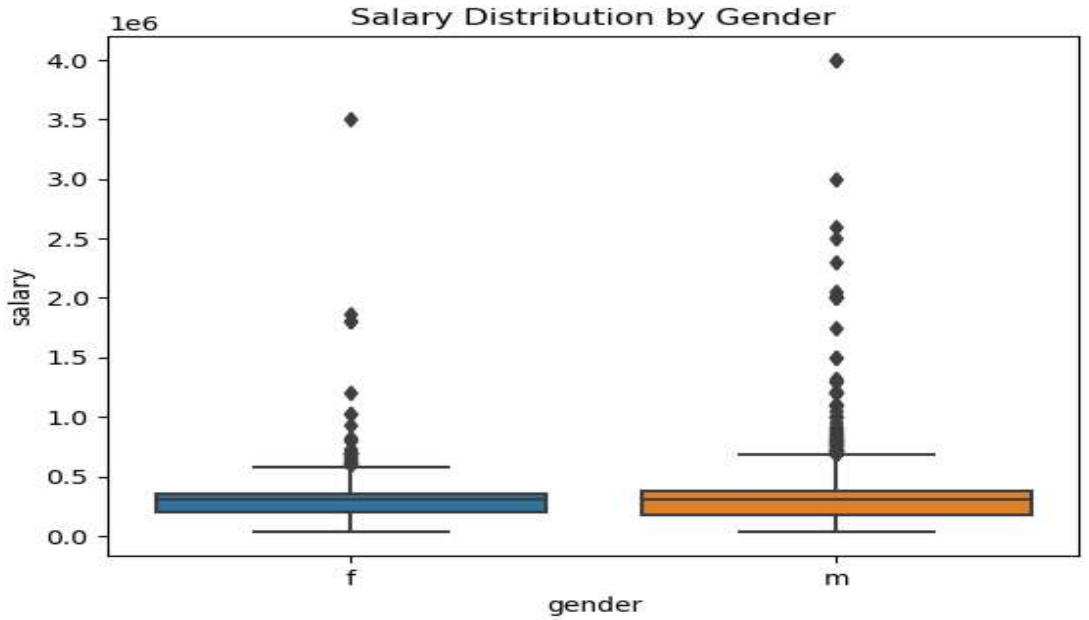
- The dataset contains around 40 independent variables and 4000 data points. The independent variables are both continuous and categorical in nature. The dataset contains a unique identifier for each candidate.

Univariate Analysis



- The salary distribution is highly right-skewed, with the majority of salaries concentrated below 500,000.
- A significant peak is observed around 400,000, indicating a large number of employees earning within this range.
- The long tail extending to 4,000,000 suggests the presence of high-earning outliers.

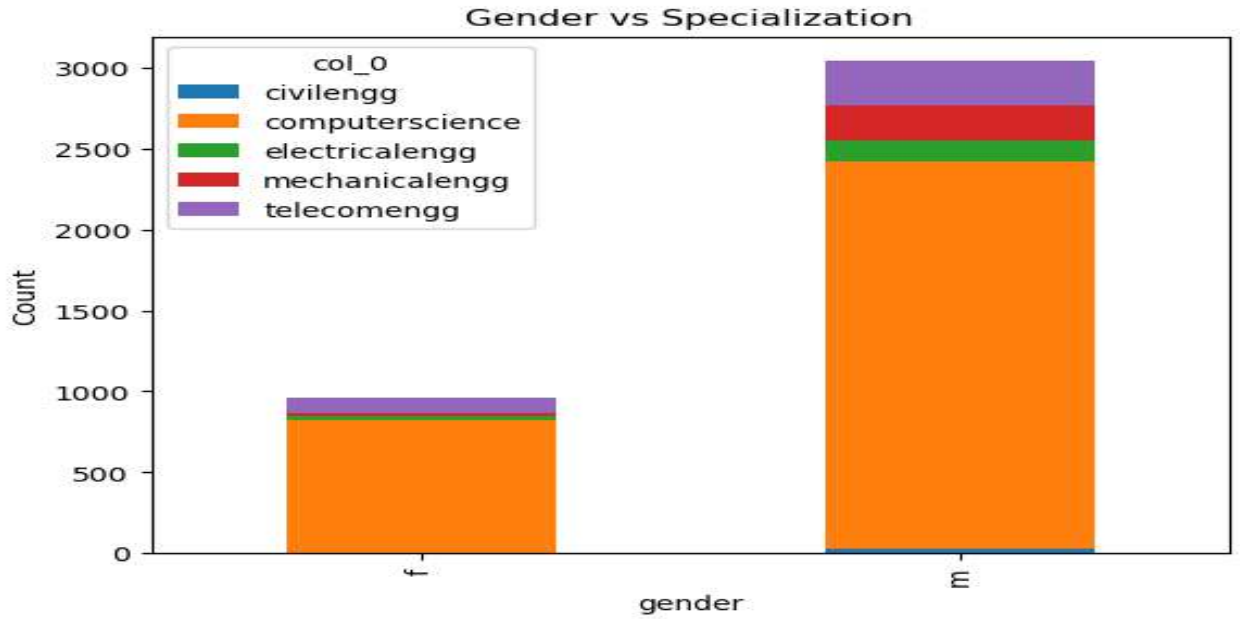
Bivariate Analysis



When looking at the data, we see that **males** have more salaries that are much higher than the typical range, which are called outliers. This means that some males are earning a lot more compared to the average salary. On the other hand, **females** have fewer extreme salaries, and their earnings tend to stay closer to the average.

INNOMATICS RESEARCH LABS

Is there any Relationship between Gender and Specialization



Chi-Square test statistics: 45.25649, P-Value: 3.516526, There is a Relationship between Gender And Specialization

Conclusion

- Salary Distribution: The salary distribution shows a right skew, with outliers at the higher end, indicating a few individuals earn significantly more than others.
- Gender Disparity: Males tend to occupy more positions across cities, and there might be a salary gap favoring males.
- Salary Claim Validation: Based on the dataset, the average salary for Computer Science engineers does align with or slightly exceed the claim made in the Times of India article.
- Job City Insights: Certain cities, like Bangalore, have a higher concentration of employees, potentially offering better salary prospects.

THANK YOU

INNOMATICS RESEARCH LABS