INNOMATICS®
RESEARCH LABS

INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Analysis of AMCAT Data

About me

I am **Chirag Gohil** below is my background information:

1.Background:

1. Currently pursuing Data Science Master Certificate Program at Fingertips & Jain University, Ahmedabad .
2. Completed BCA from GLS University , Ahmedabad (CGPA : 8.2)

2.Why Data Science:

1. Passion for leveraging technology to derive insights and solve complex problems.
2. Aiming to apply data science to innovate and contribute to real-world solutions.

3.Work Experience:

1. Data Science Intern at Innomatics Research Labs, specializing in Data Science.

4.LinkedIn & GitHub:

1. Linkedin : https://www.linkedin.com/in/chiraggohil28
2. Github : https://github.com/Gohil28

INNOMATICS RESEARCH LABS

Introduction

Brief Overview:

- The project involves conducting Exploratory Data Analysis (EDA) on a salary dataset to extract valuable insights.

Objective:

- Uncover patterns, outliers, and relationships within the data.

Dataset Overview

- The dataset consists of 3998 entries and 39 columns.
- Columns include information such as ID, Salary, Date of Joining (DOJ), Date of Leaving (DOL), Designation, JobCity, Gender, Date of Birth (DOB), 10th and 12th percentage, and many more.
- **Head of the Dataset:**

	Unnamed: 0	ID	Salary	DOJ	DOL	Designation	JobCity	Gender	DOB	10percentage	...	ComputerScience	MechanicalEngg	ElectricalEngg	TelecomEngg	CivilEngg	conscientiousness
0	train	203097	420000.0	6/1/12 0:00	present	senior quality engineer	Bangalore	f	2/19/90 0:00	84.3	...	-1	-1	-1	-1	-1	0.9737
1	train	579905	500000.0	9/1/13 0:00	present	assistant manager	Indore	m	10/4/89 0:00	85.4	...	-1	-1	-1	-1	-1	-0.7335
2	train	810601	325000.0	6/1/14 0:00	present	systems engineer	Chennai	f	8/3/92 0:00	85.0	...	-1	-1	-1	-1	-1	0.2718
3	train	267447	1100000.0	7/1/11 0:00	present	senior software engineer	Gurgaon	m	12/5/89 0:00	85.6	...	-1	-1	-1	-1	-1	0.0464
4	train	343523	200000.0	3/1/14 0:00	3/1/15 0:00	get	Manesar	m	2/27/91 0:00	78.0	...	-1	-1	-1	-1	-1	-0.8810

5 rows x 39 columns

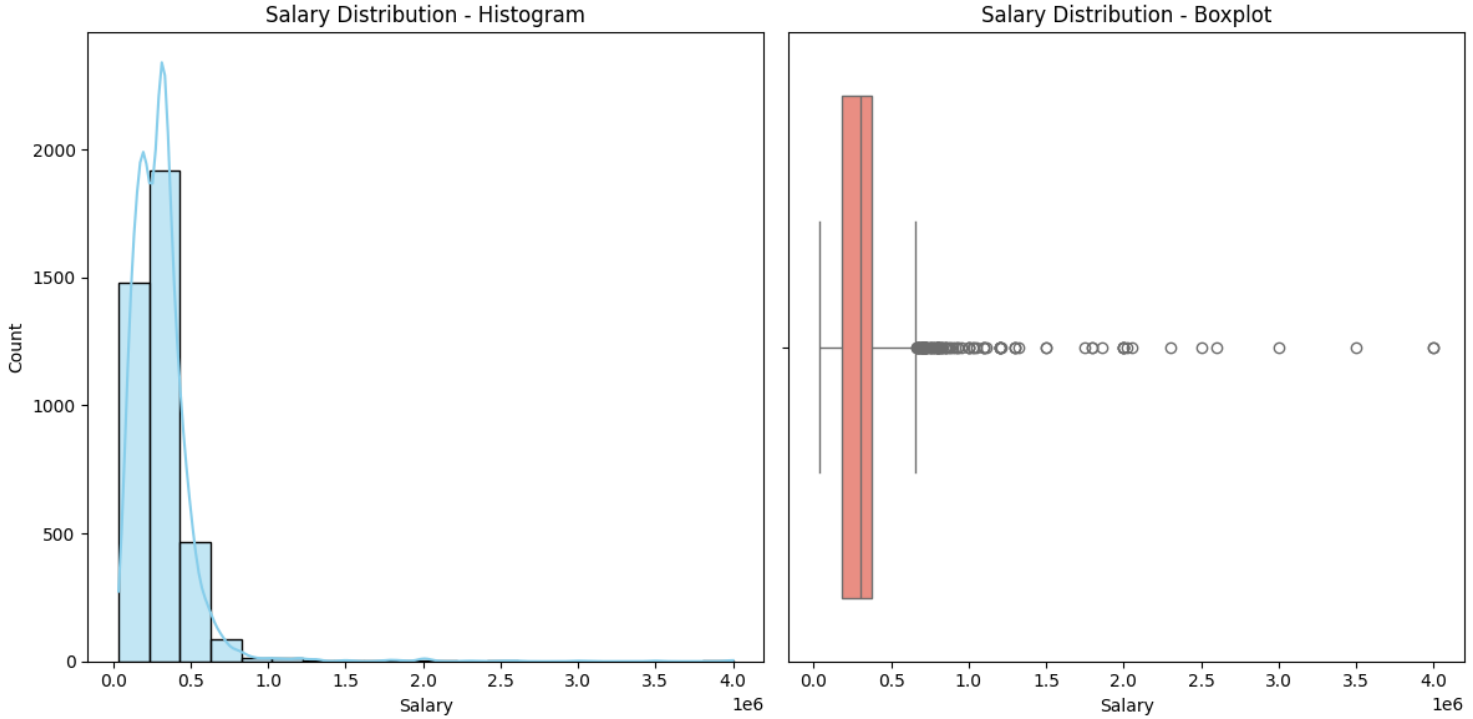
- **Shape:**

  Rows: 3998, Columns: 39

- **Description:**
- The dataset provides a comprehensive overview of candidates' educational backgrounds, professional experiences, and personality traits.
- Various data types include float64, int64, and object.

INNOMATICS RESEARCH LABS

Univariate Analysis - Salary Distribution



The Salary distribution exhibits positive skewness with a value of **6.45**, indicating that the majority of salaries are concentrated on the lower end, while a few instances have exceptionally high salaries.

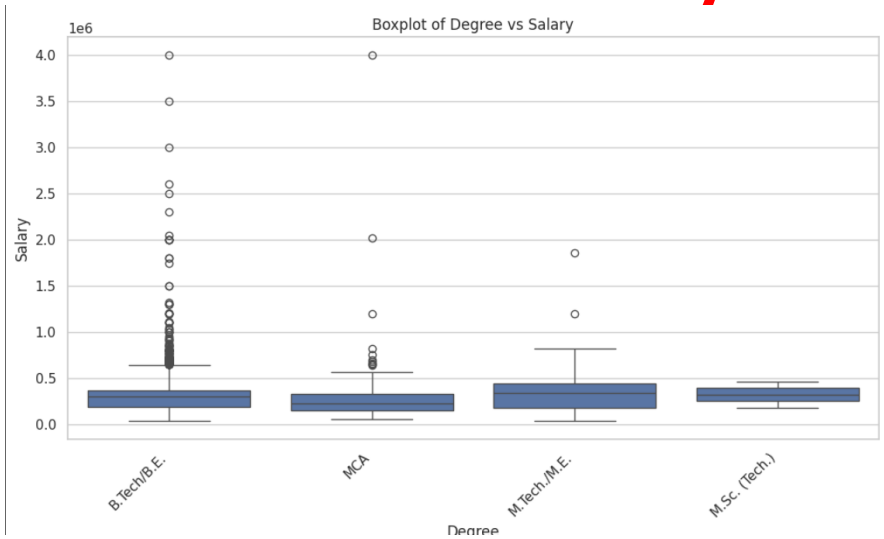
A total of 109 outliers were identified, with salaries significantly exceeding the general trend. Noteworthy examples include salaries of 1,100,000, 800,000, and 1,500,000.

The histogram and boxplot visually represent the skewed nature of the distribution and the presence of outliers. Further investigation into the context of these high salaries is recommended to determine whether they are valid data points or anomalies.

This analysis serves as a foundation for understanding the distribution of salaries in the dataset.

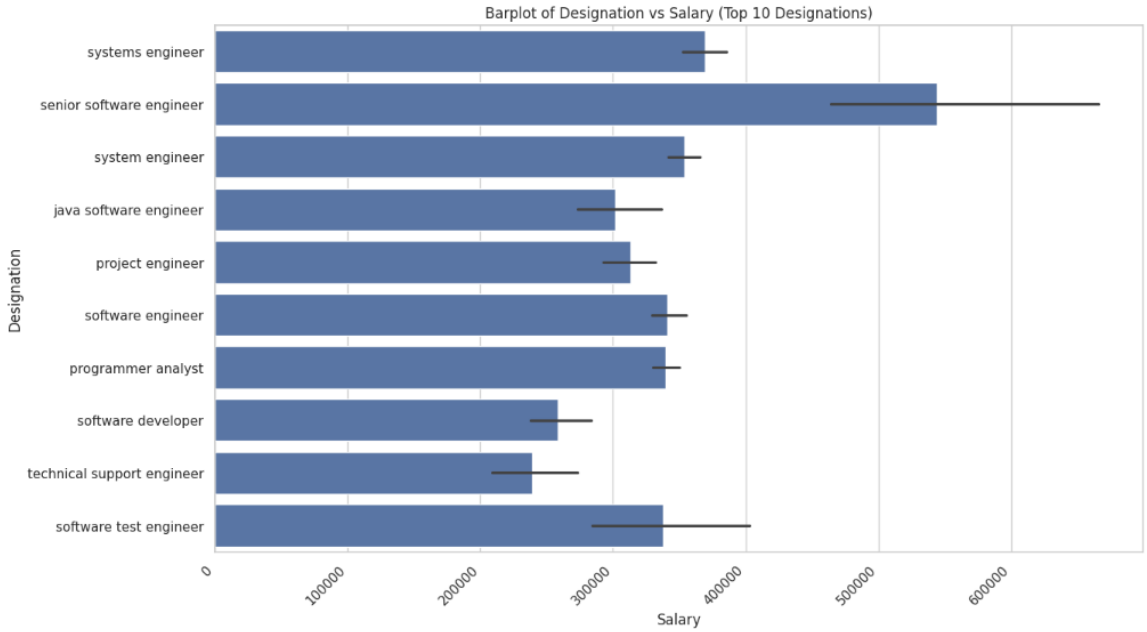
INNOMATICS RESEARCH LABS

Bivariate Analysis



In bar plot of Degree vs Salary , the **M.Tech/M.E**. Degree has slightly High salary compare to others.

There certainly more outliers for B.Tech/B.E. Degree which tells it has chances of getting more higher salary in exceptional cases



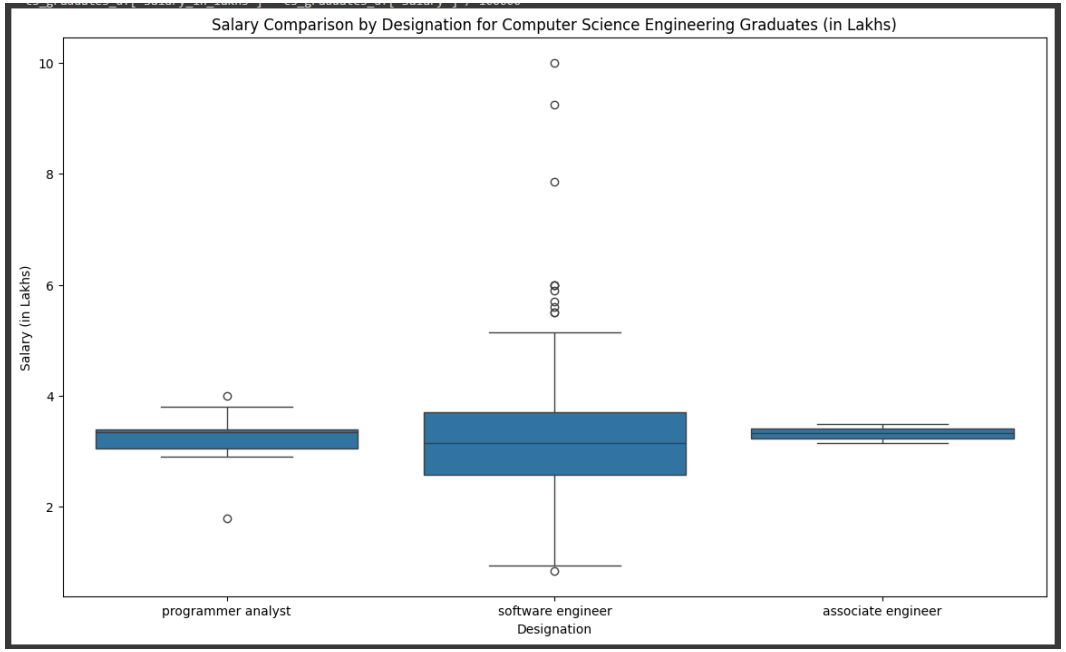
In Bar plot , It can be seen that senior software engineer having usually more salary compared to other designations

Technical support engineer having least salary among top 10 highest paid designations

INNOMATICS RESEARCH LABS

Research Questions

Times of India article dated Jan 18, 2019 states that "*After doing your Computer Science Engineering if you take up jobs as a Programming Analyst, Software Engineer, Hardware Engineer and Associate Engineer you can earn up to 2.5-3 lakhs as a fresh graduate.*"

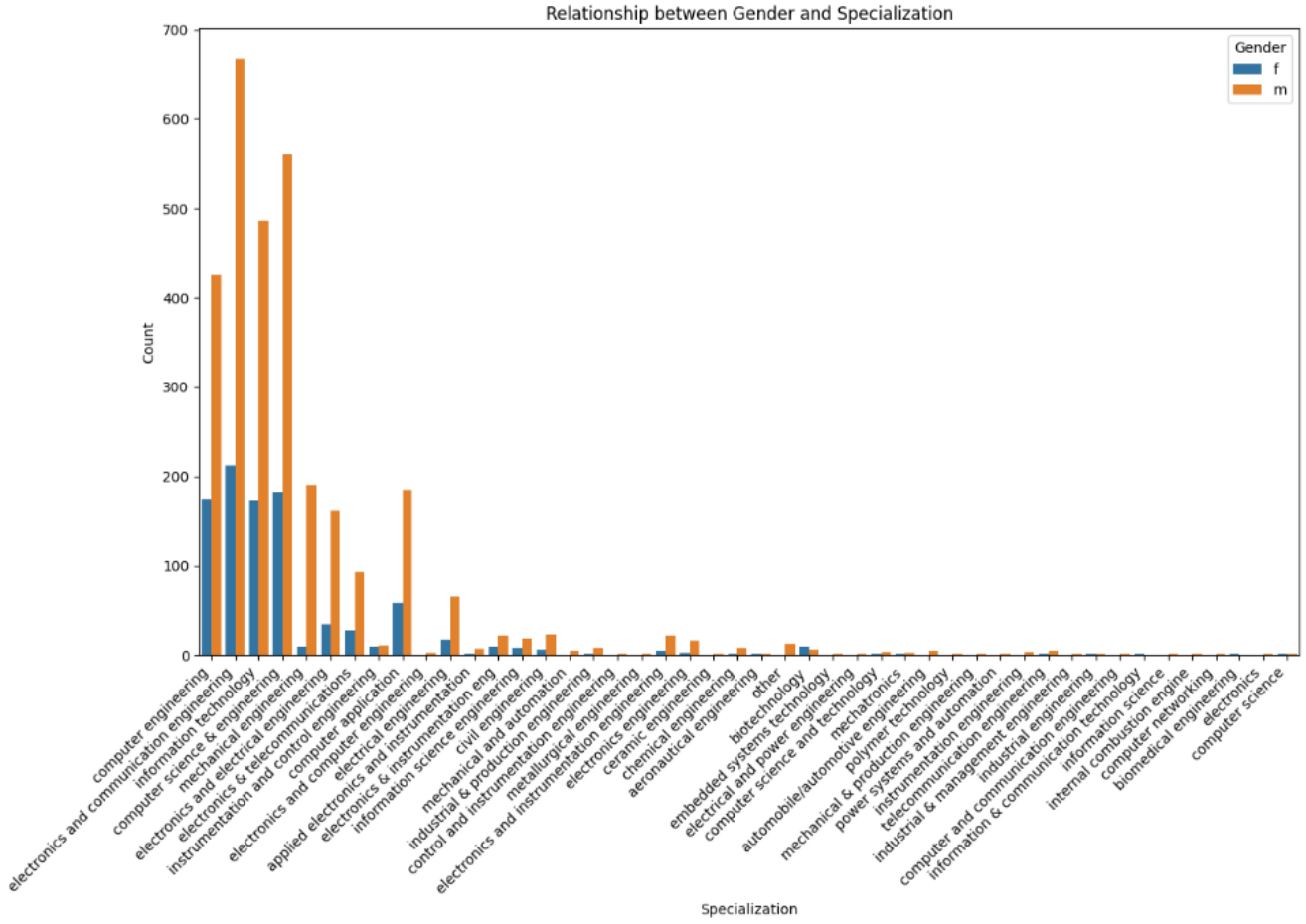


Conclusion:

- Yes, after doing Computer Science Engineering if you take up jobs as a Programming Analyst, Software Engineer and Associate Engineer you can earn up to **2.5-3 lakhs** as a fresh graduate
- However ,there was no data present to support the fact that after Computer Science Engineering if you take up jobs as a Hardware Engineer you can earn up to 2.5-3 lakhs as a fresh graduate

INNOMATICS RESEARCH LABS

Is there a relationship between gender and specialization? (i.e. Does the preference of Specialisation depend on the Gender?)



- By looking at plot we can see males are doing more specialization than females
- Females are preferring less specialization, in some cases it is less than half the number of males

INNOMATICS RESEARCH LABS

Conclusions

- In conclusion, our exploratory data analysis (EDA) sheds light on several key aspects of the dataset related to salary and various features. Here are the main takeaways:

1. **Salary Distribution:**
    1. The distribution of salaries exhibits some degree of skewness, indicating variations in compensation levels among individuals.
    2. Through histogram and boxplot visualizations, we identified outliers that require further investigation.
2. **Correlation Analysis:**
    1. Bivariate analysis revealed correlations between salary and specific features such as 'Quant,' 'Logical,' 'English,' '10percentage,' and '12percentage.'
    2. Certain features, like 'Computer Science,' 'ID,' and 'CollegeTier,' show negative correlations with salary.
3. **Research Questions:**
    1. The Times of India's claim about salary expectations for fresh graduates in the field of Computer Science Engineering was tested against the dataset.
    2. Additionally, we explored the relationship between gender and specialization preferences.

THANK YOU

INNOMATICS RESEARCH LABS