



INNOVATION. AUTOMATION. ANALYTICS

PROJECT ON

Exploratory Data Analysis
(AMCAT Dataset)

About me

- I am R.Radhika , I completed my Masters in Data science in the year 2023.
- To acquire more knowledge in coding,projects I joined in Innomatics
- **Linked in and Github urls:**

❑ <https://www.linkedin.com/in/ramachandrapuram-radhika/>

❑ <https://github.com/ramachandrapuram-radhika/>

OBJECTIVE OF THE PROBLEM

- This exploratory data analysis of “AMCAT DATASET” focuses on understanding various factors that might influence the level of salaries indicated in the dataset. We consider education and experience, gender, specialization, and job roles and observe how they are related in order to understand a factor that influences higher or lower levels of salaries. The critical steps which indicate the analysis involved creating a mental image of the data, establishing trends and patterns, testing many hypotheses post observations to finally build insightful results which could be used as guidelines for any decision making process that could further calibrate salary prediction models.

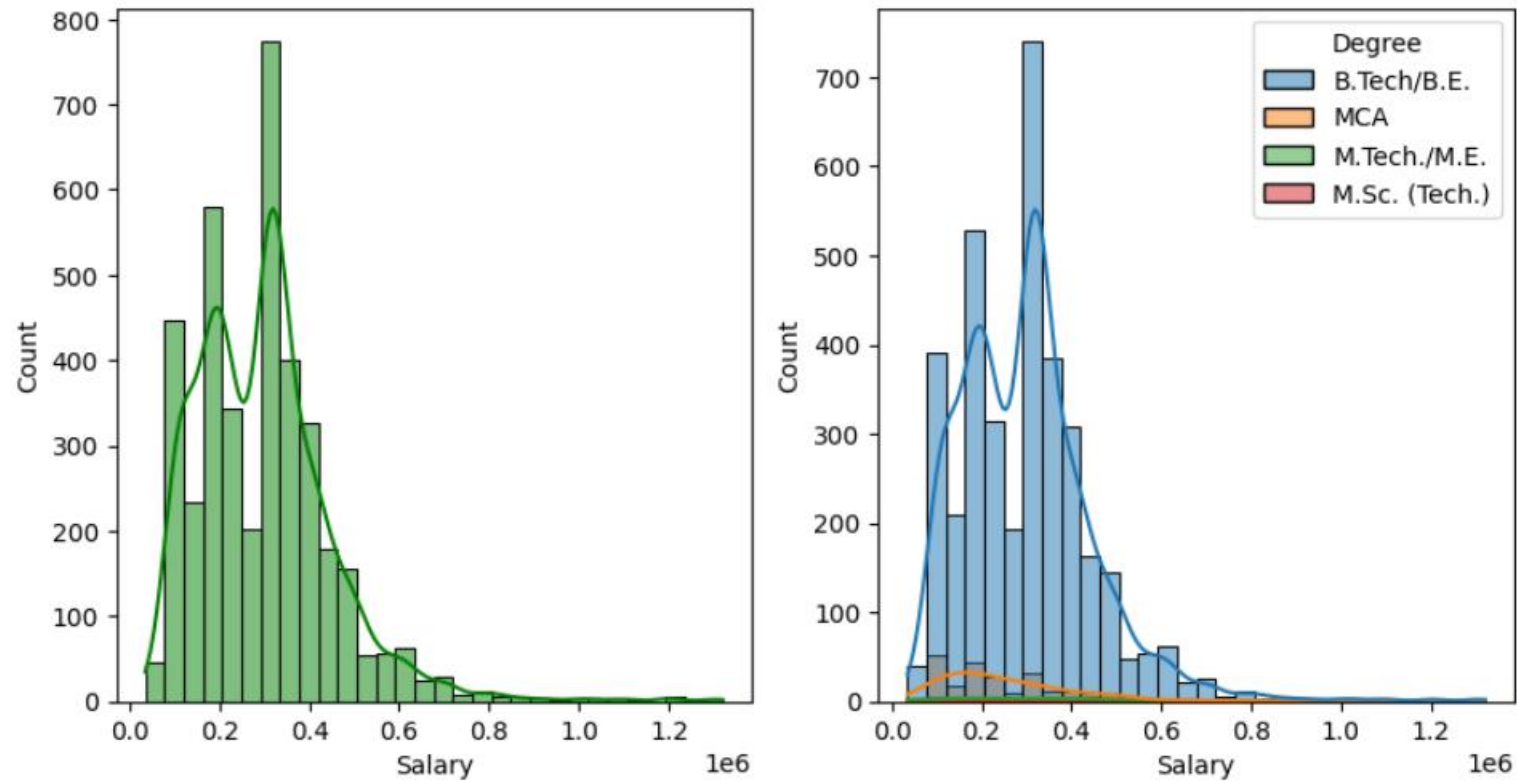
SUMMARY OF THE DATASET

- There are 38 columns in dataset that are used to find the individual impacts on salary.
- Out of 38 columns, there are 29 numerical columns and 9 categorical columns.
- With 3998 Datapoints that make our analysis to the optimal insights with all the necessary information.

DATA CHECKS TO PERFORM

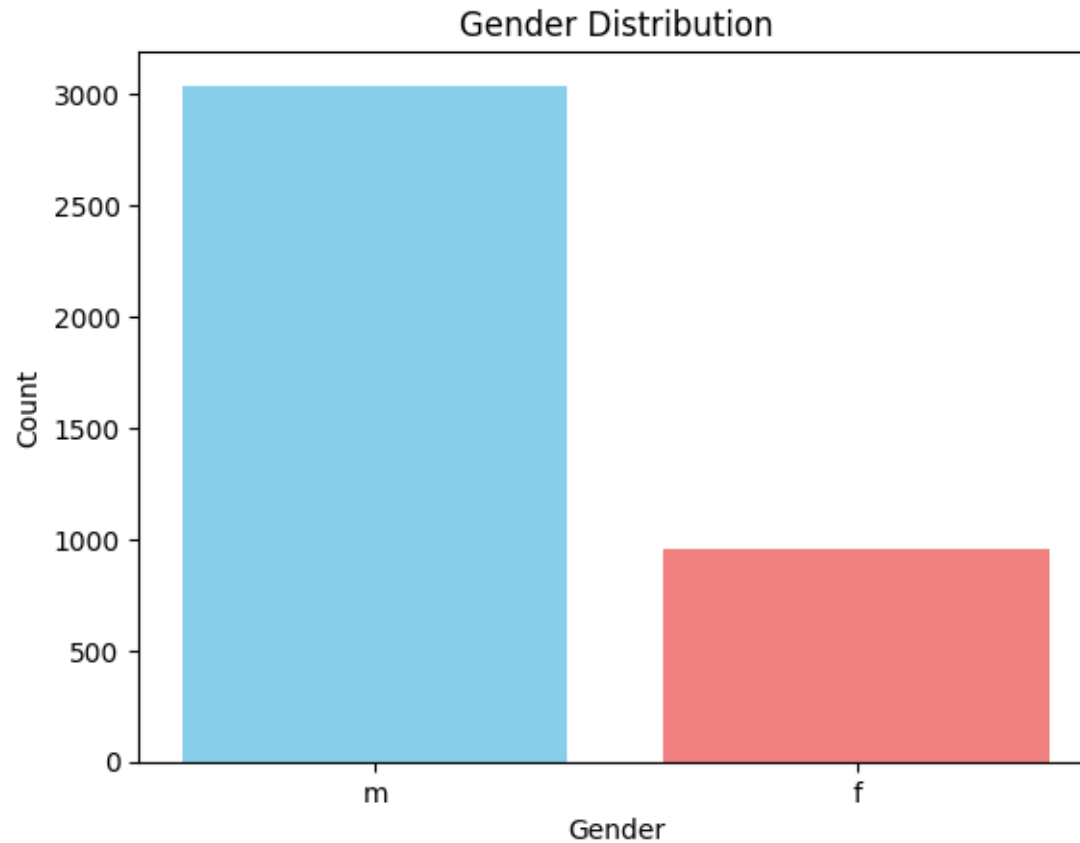
- Check missing values, duplicated values and various different columns.
- Check the datatypes and also look at the unique number of columns.
- Check statistics of data set
- Check various categories present in the different categorical column.
- Drop unnecessary columns

Salary vs Degree



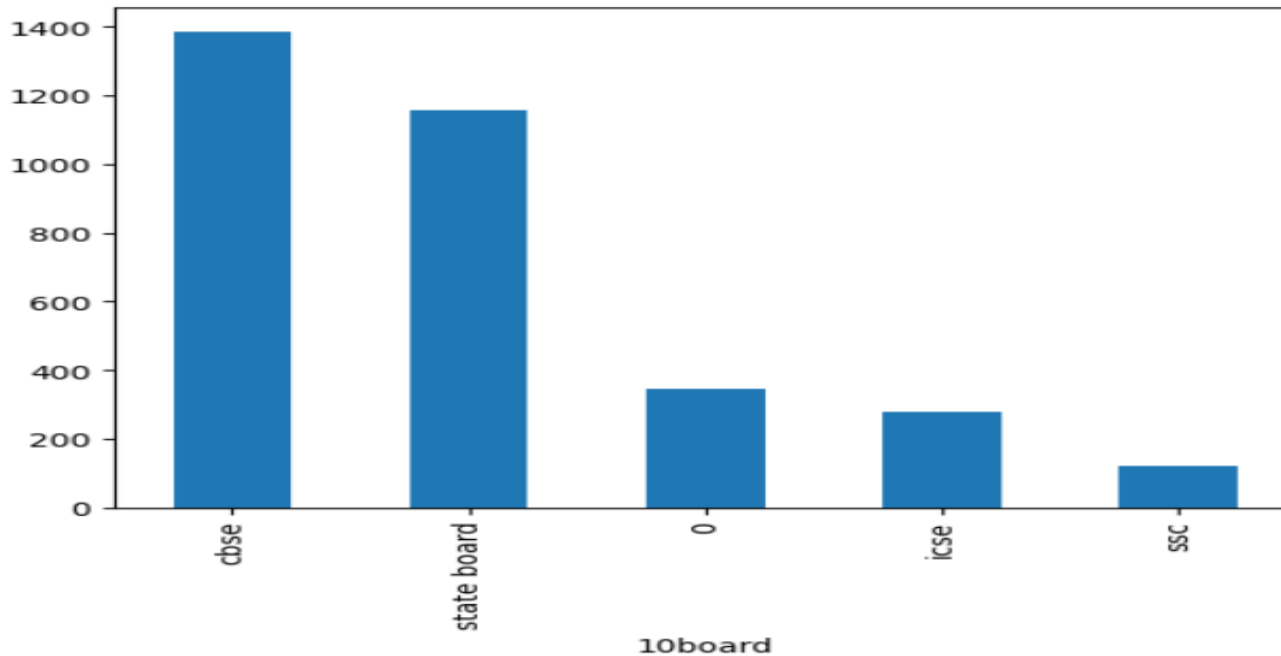
- Average salary is highest for BE/B.tech graduates as compared to any other degree graduates.

Gender Distribution



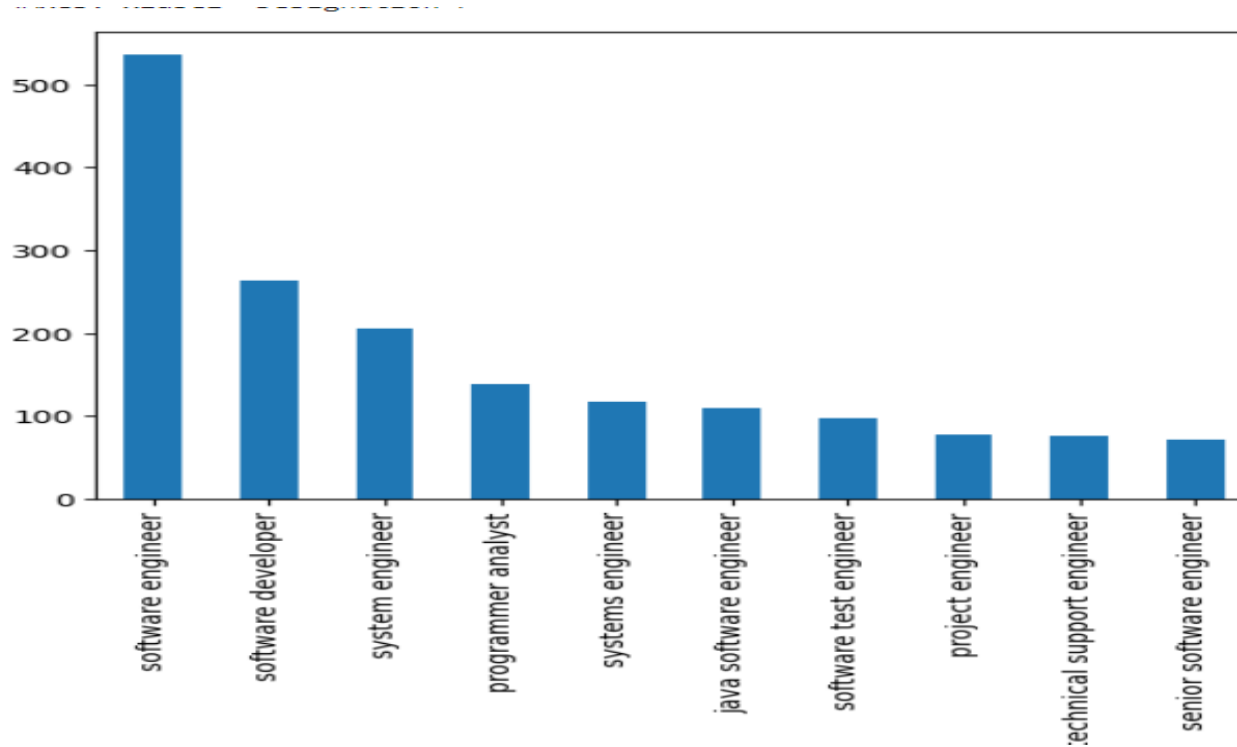
- There are more number of males compared to females.

Distribution of 10board



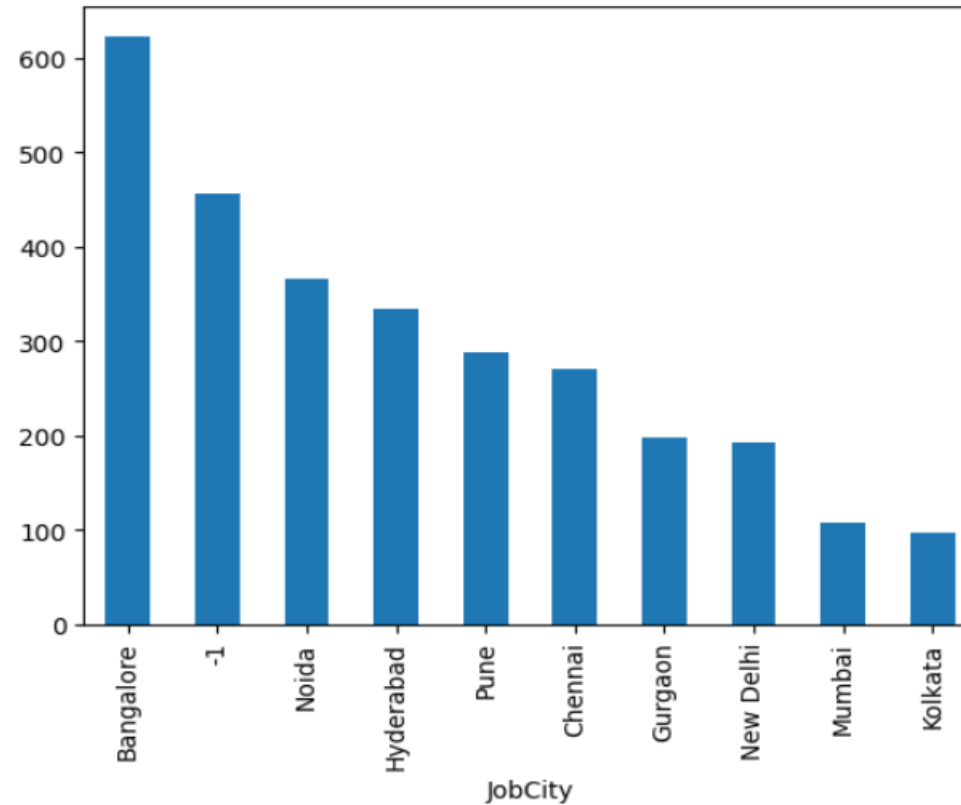
- Most of the people are from cbse
- compare to cbse ,less people are from ssc

Designation Distribution



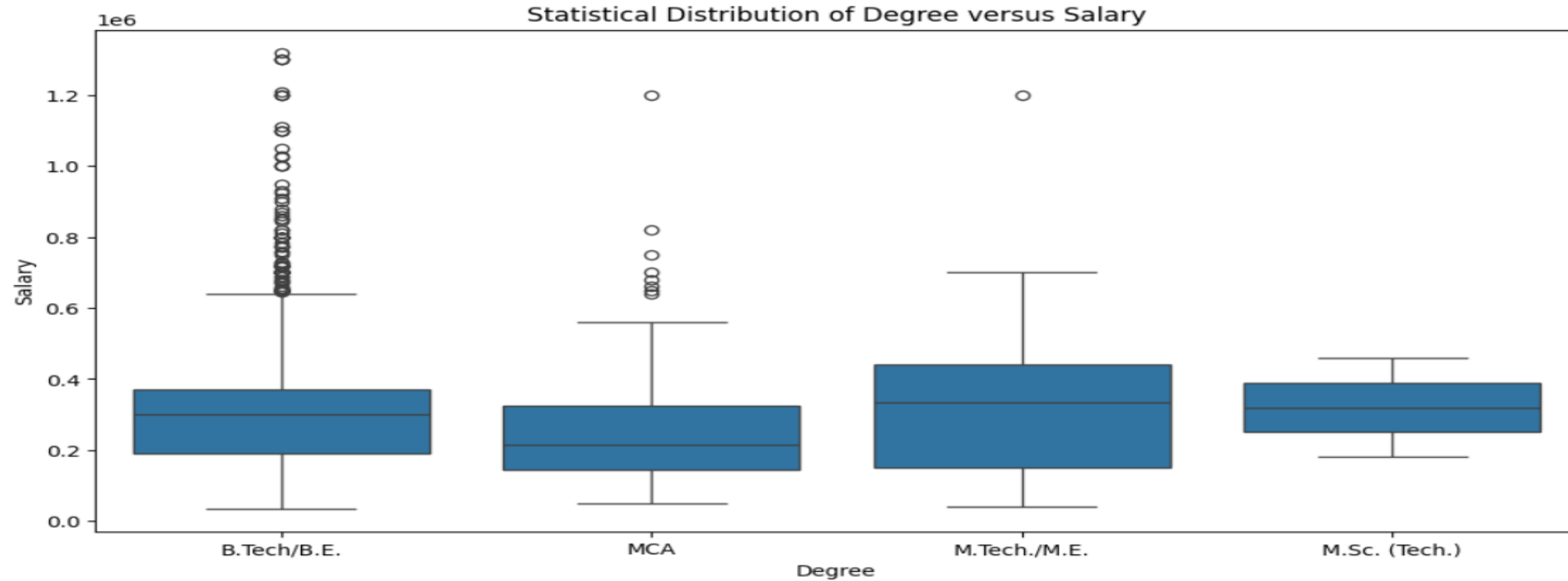
- Most of people are software Developer
- less people are senior software engineer

Jobcity Distribution



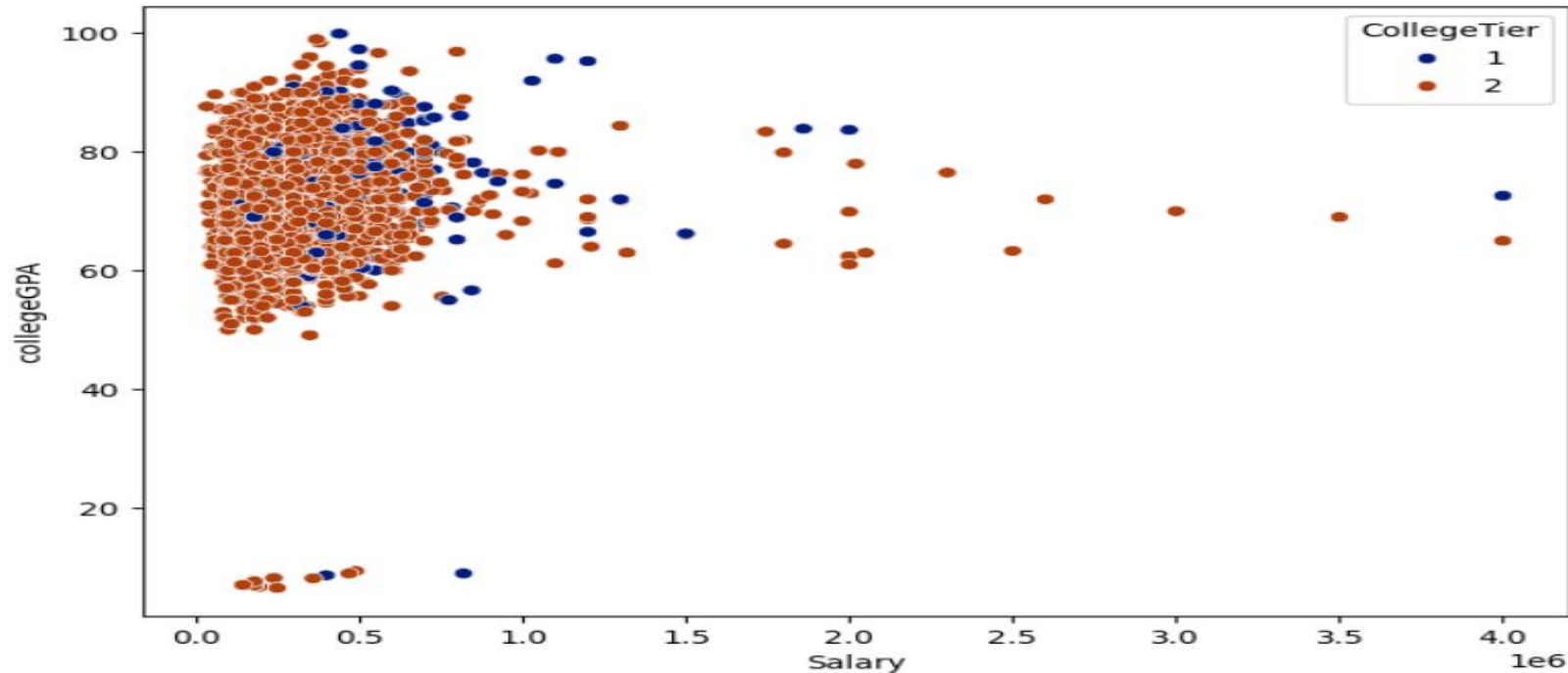
These are the 5 most cities which are popular Bangalore,Noida,Hyderabad,Pune,Chennai

Salary vs Degree



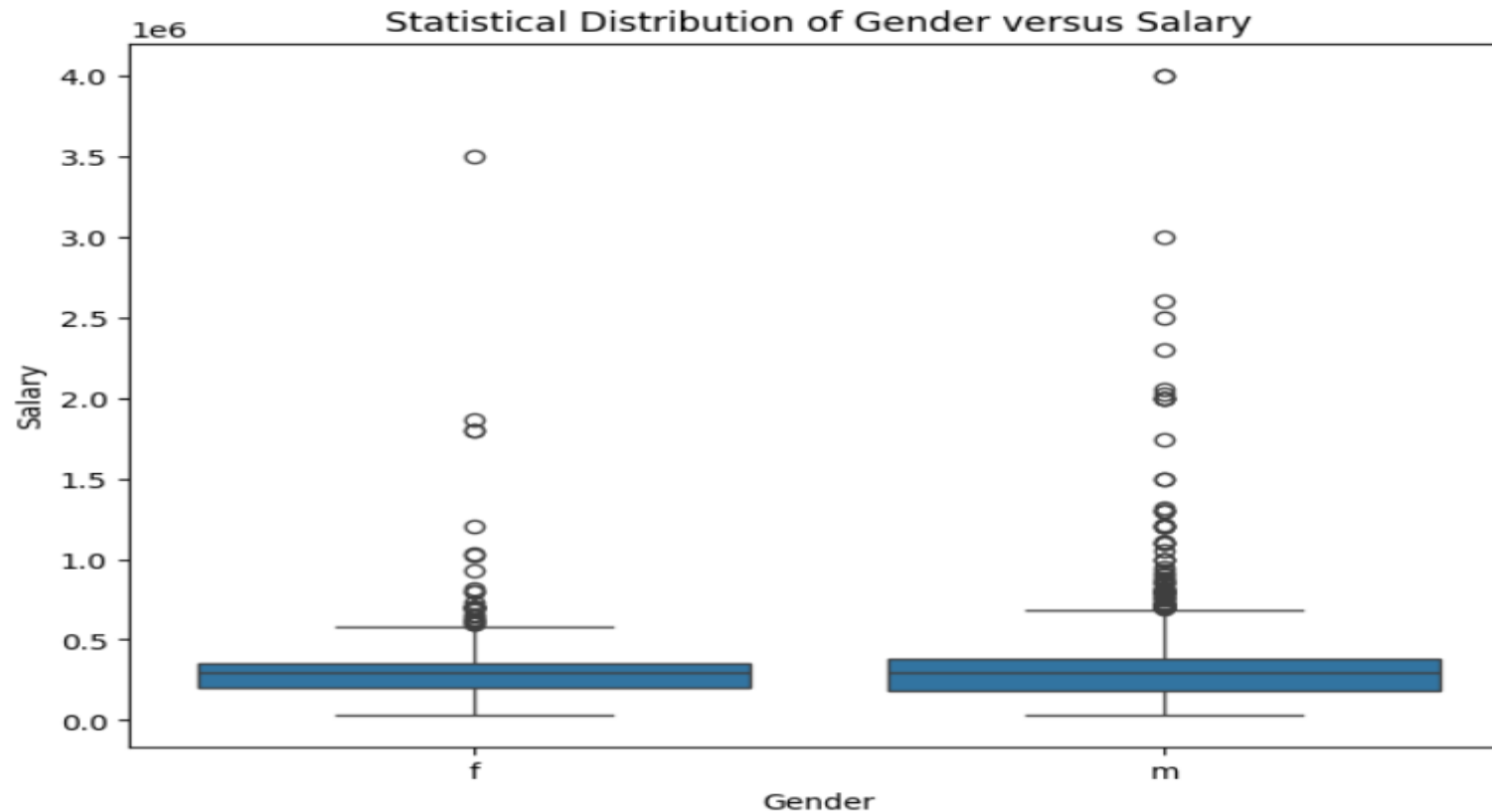
More outliers in btech compare to other degrees

Salary vs collegeGPA



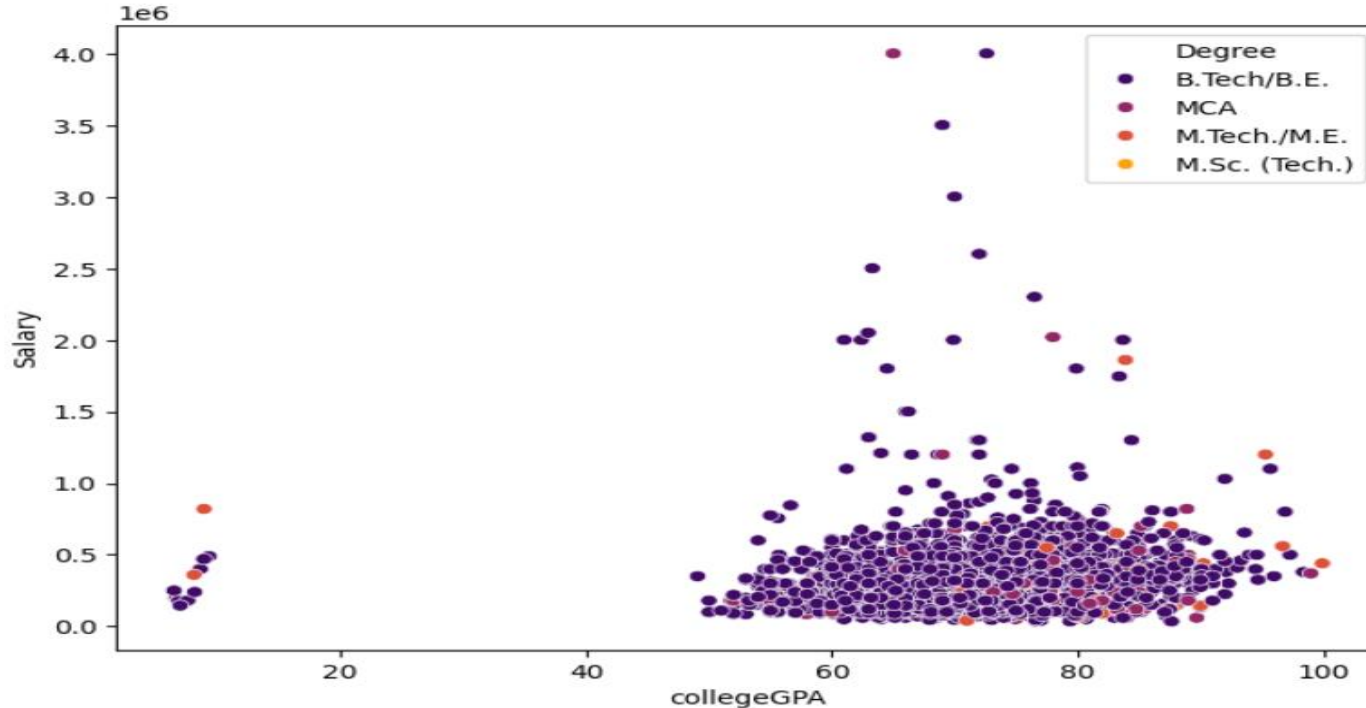
- There are more outliers so we need to drop some extreme values of salary and collegeGPA.

Gender vs salary



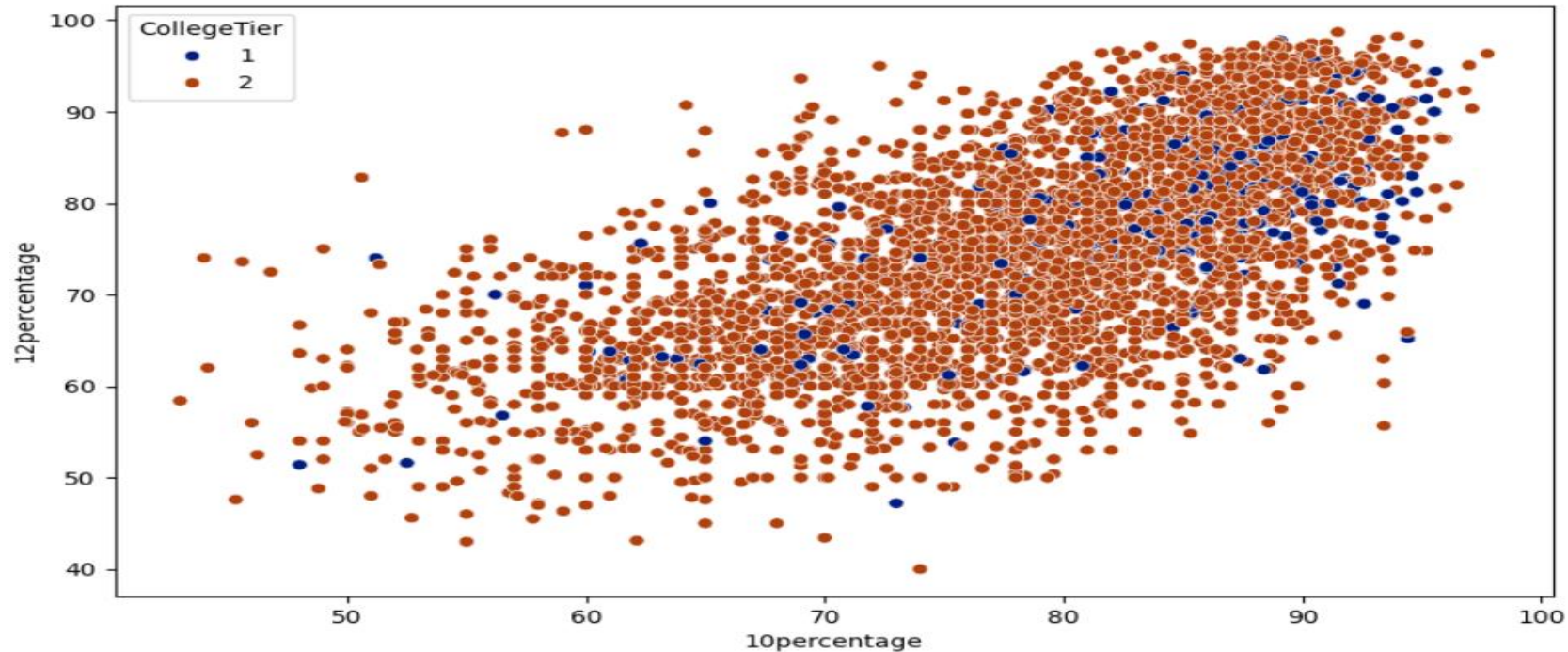
- It is clearly visible from the plot that the Average salary for man and woman are looking almost same.
- There are outliers in salary column.

Salary vs Degree Analysis



Average salary is highest for BE/B.tech graduates as compared to any other degree graduates.

10th and 12th percentage by college tier



- According to scatterplot, we can see that 10th and 12th are positively correlated, this is the case of multicollinearity so I have decided to keep only one.

Research Question 1:

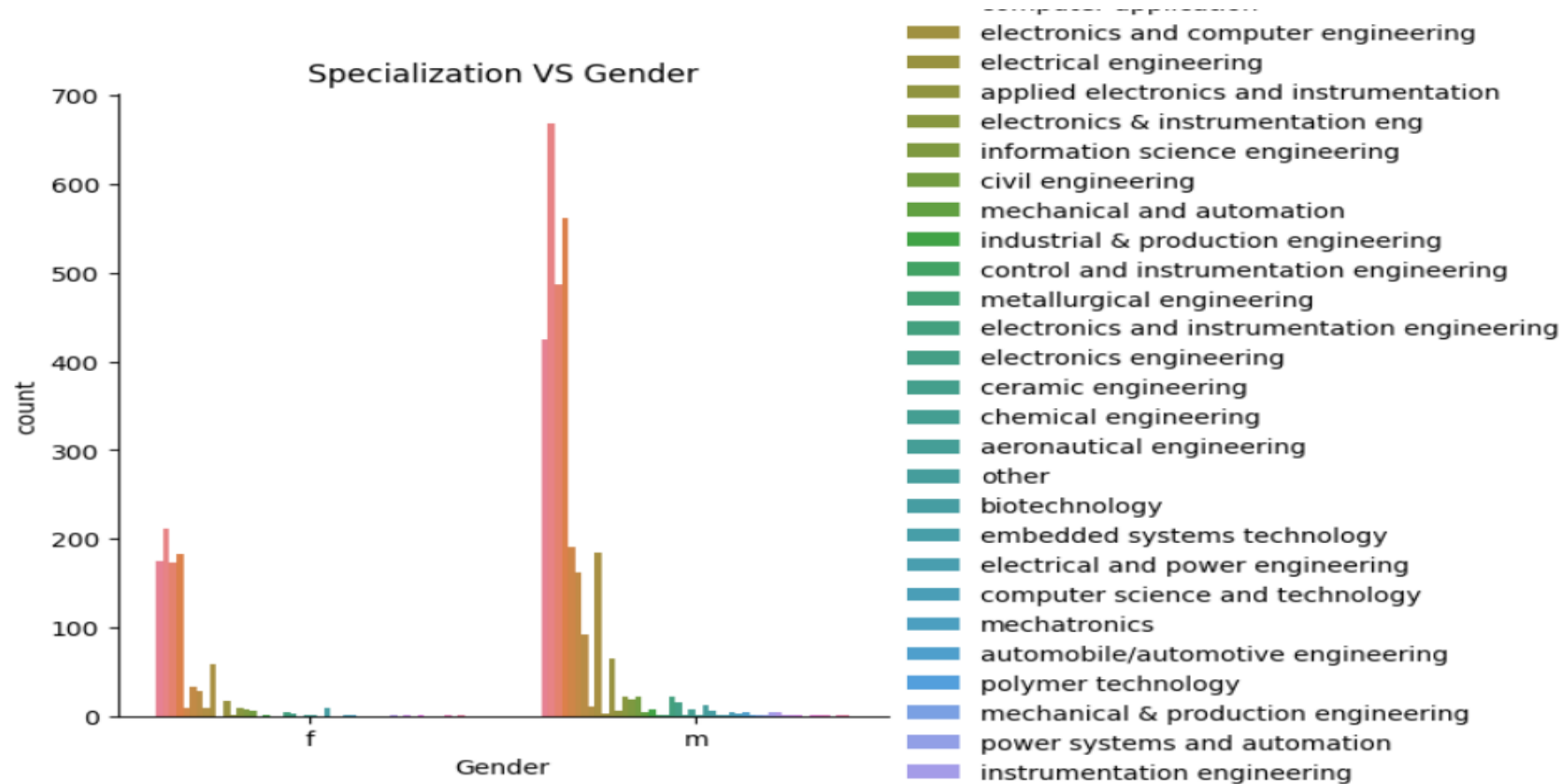
Determine whether fresh graduates earn 2.5-3 lakhs annually as stated in the article.

	Designation	Specialization	Salary
0	.net developer	computer application	231250.000000
1	.net developer	computer engineering	309285.714286
2	.net developer	computer science & engineering	160000.000000
3	.net developer	electronics and communication engineering	232500.000000
4	.net developer	information science engineering	240000.000000

The claim is not correct for the given data.

Solution: We found that the average salary is not within the given range(i.e, 2.5 lakhs to 3 lakhs)

Research Question 2: Determine if gender influences the choice of specialization.



Insights of question 2

- The most popular specializations for both males and females are Computer Science and Electronics and Telecommunications (ETRX), which make up more than 60% and 55% of their respective proportions
- The proportion of females in Aeronautical Engineering, Biomedical Engineering, and Biotechnology is higher than the proportion of males in these fields.
- The proportion of males in Chemical Engineering, Civil Engineering, and Mechanical Engineering is higher than the proportion of females in these fields.
- Other specializations have a higher proportion of males than females.

Conclusion

- Average salary is highest for BE/B.tech graduates as compared to any other degree graduates
- The proportion of females in Aeronautical Engineering, Biomedical Engineering, and Biotechnology is higher than the proportion of males in these fields.
- .The most popular specializations for both males and females are Computer Science and Electronics and Telecommunications (ETRX), which make up more than 60% and 55% of their respective proportions
- These are the 5 most cities which are popular Bangalore,Noida,Hyderabad,Pune,Chennai
- Most of people are software Developer and less people are senior software engineer.

THANK YOU

