INNOMATICS
RESEARCH LABS

INNOVATION. AUTO**MAT**ION. ANALY**TICS**

PROJECT ON

EXPLORATORY DATA ANALYSIS – AMCAT DATA

About me

- **I am Manju Bhargavi Danalakoti from Hyderabad. I am a B-Tech Mechanical Engineering graduate.**

- **I have about 6 years of working experience in Automobile Industry.**

- **I look forward to become a Data Analyst/Data Science Engineer to take my career to next level, applying the domain knowledge for data driven decisions.**

- **LinkedIn URL -** www.linkedin.com/in/manju-danalakoti

Business Problem :

- Analysis of AMCAT data and understand the employment opportunities .

Objective of the Project:

- Analysis on students data who have attended AMCAT test and their results.

- Output of the Analysis gives an understanding on graduates getting jobs in different domains.

- Comparison of different domains salaries w.r.t specialization.

Web Scraping – Details (Websites, Process you followed):

- **Web scrapping** is a process of extracting data or information from website.

- Website –AMCAT – AMCAT (Aspiring Minds Computer Adaptive Test) is a popular employability assessment test in India, designed to evaluate candidates' skills and job readiness for a variety of roles across industries. Many companies use AMCAT as part of their recruitment process to find candidates who are a good fit for their open positions, especially in fields like IT, engineering, management, and finance.

- **Process followed:**

- Imported libraries for EDA – pandas, numpy, matplotlib and seaborn.

Summary of the Data :

- Data gives the details of the AMCAT test results of the students who have attempted the test.

- The study is primarily limited only to students with engineering disciplines. The dataset contains the employment outcomes of engineering graduates as dependent variables (Salary, Job Titles, and Job Locations) along with the standardized scores from three different areas – cognitive skills, technical skills and personality skills.

- The dataset contains around 40 independent variables and 4000 data points.

- The independent variables are both continuous and categorical in nature.

Exploratory Data Analysis :

- *Data Cleaning Steps :*

1. First level check for total number of null values in the data set for each column.

2. Check for the data type of individual column for numerical and categorical.

3. Check for the data relevance of each column with the data scrapped.

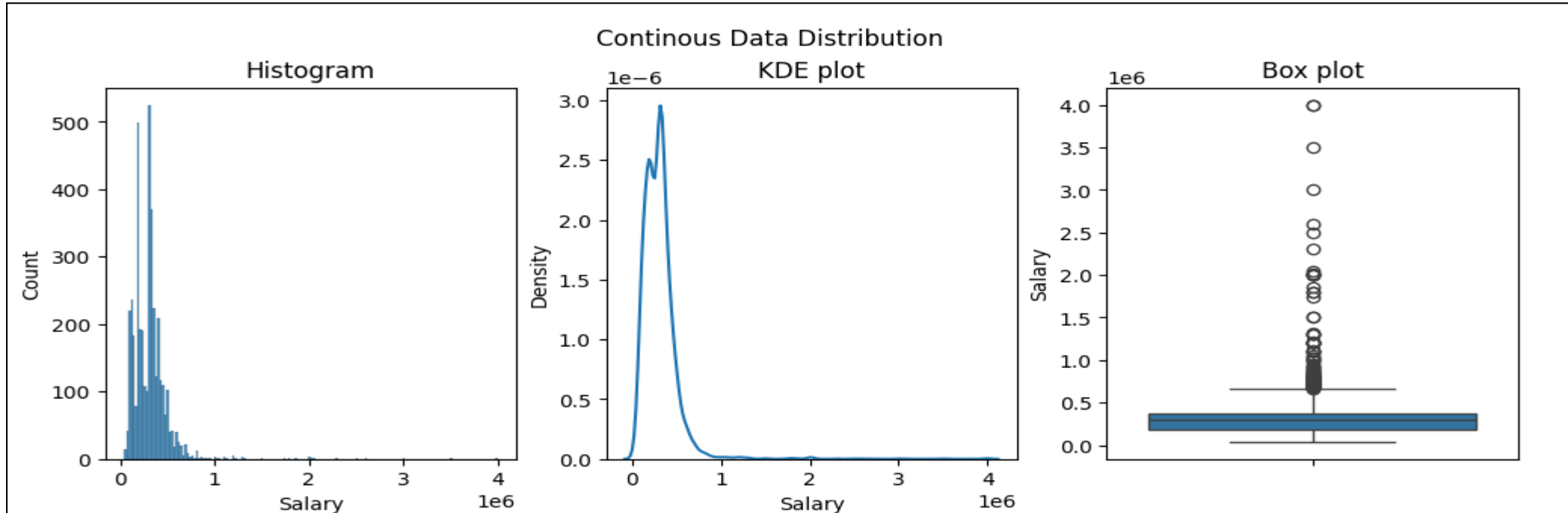4. Check for duplicates in the data and standardize.

- *Data Manipulation Steps:*

1. Check the percentage of null values in each column and decide on the method of data imputation.

2. Filling the null values with valid data/meaningful data or removing non-value added points in data frame.

3. Saving the cleaned data to CSV or Excel file.

INNOMATICS
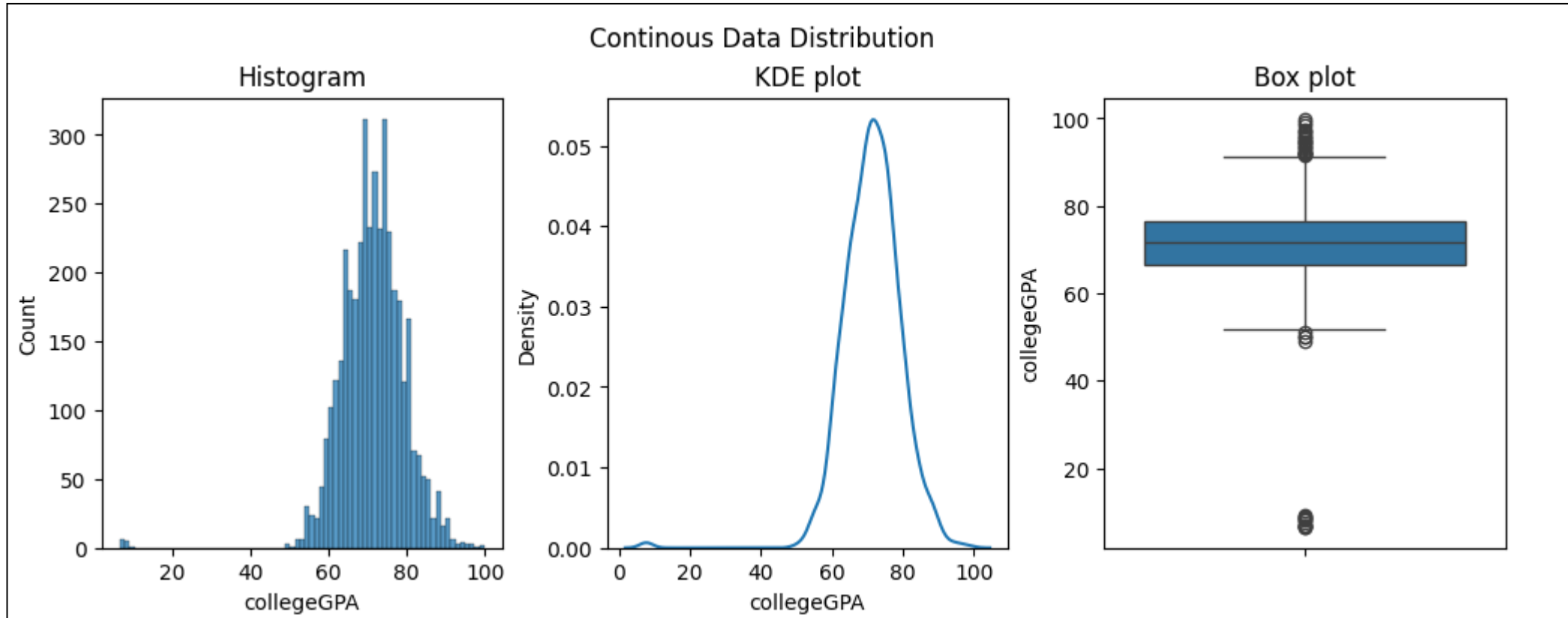RESEARCH LABS

- **Univariate Analysis Steps:**

Check for the columns wise data for numerical or categorical and plot graphs accordingly for meaningful insights.
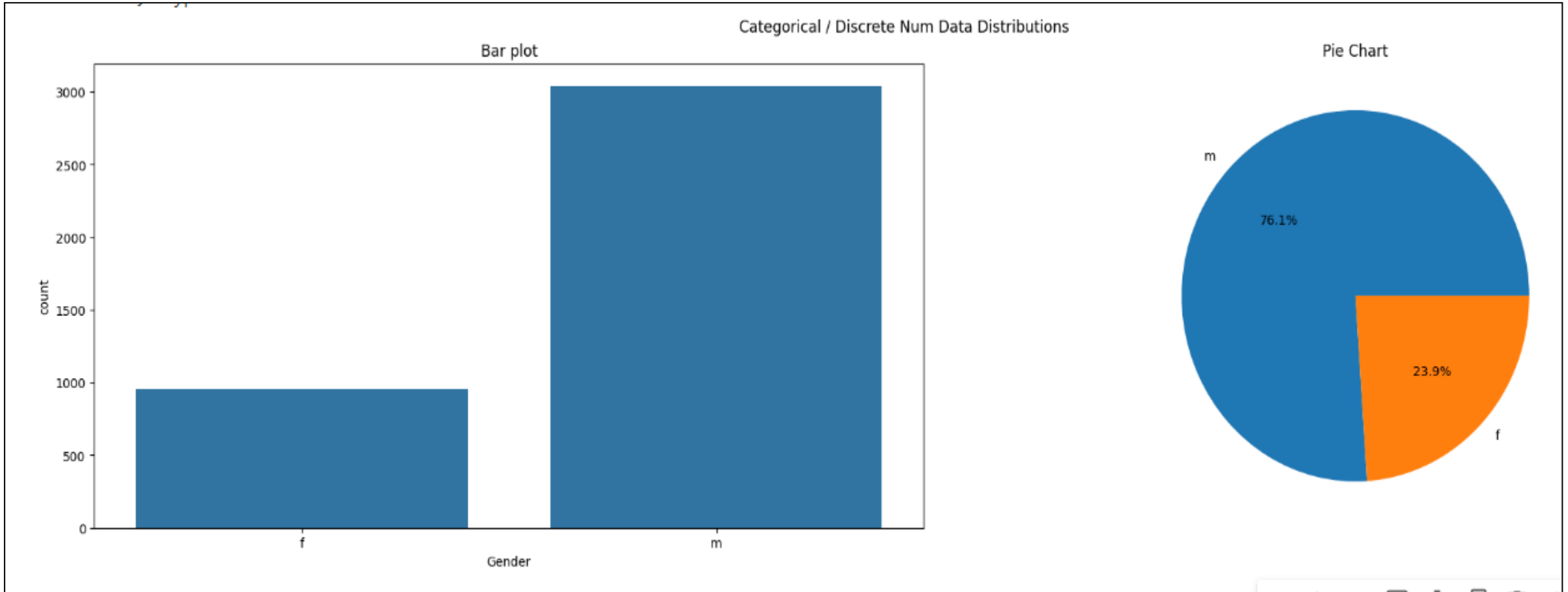
Insight-1 -Based on Salary:

1. **KDE plot shows data is highly right-skewed.**
2. **Kurtosis value depicts the graph is lyptokurtic**
3. **Box plot shows there are so many outliers in the data.**

Insight-2 -Based on College GPA:

Continous Data Distribution

1. **Histogram shows the most of the students got CGPA between 60 to 80**
2. **Kde plot shows data is slightly left skewed**
3. **Box plot shows outliers below 20**
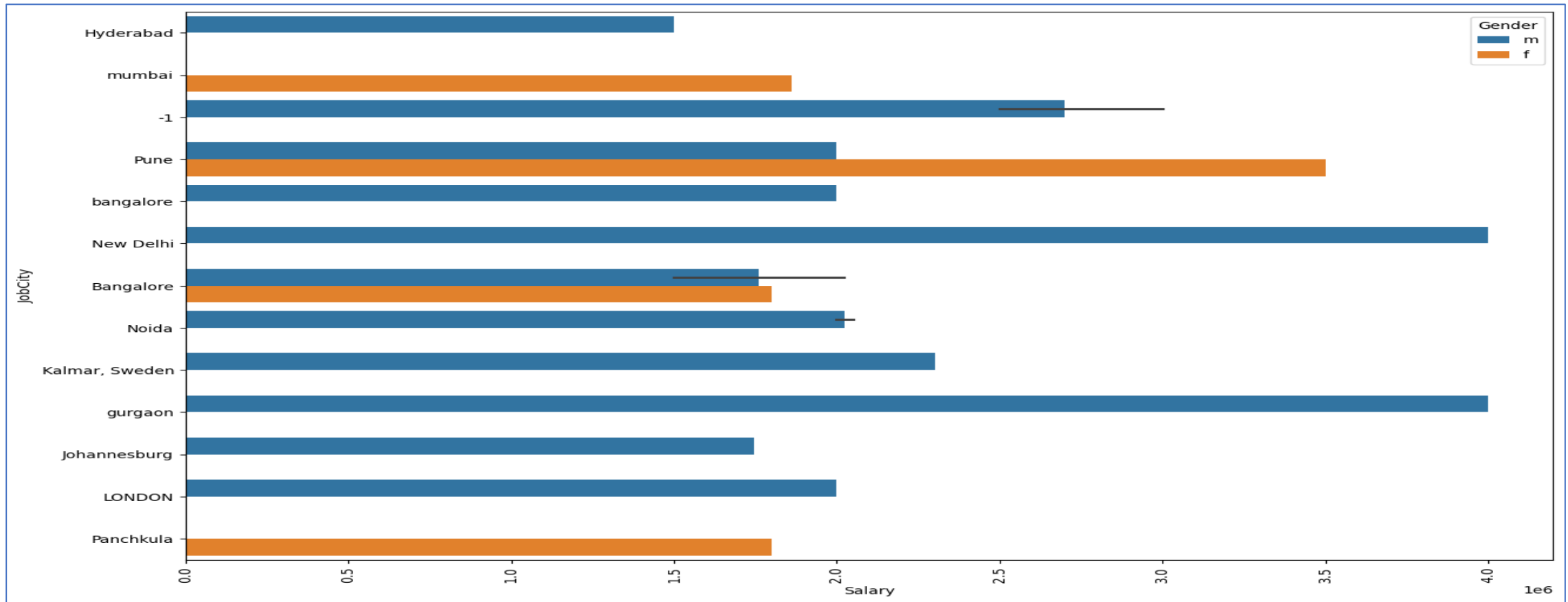
Insight-3 -Based on Gender:

- **Out of 100%, 76% are the male candidates got the job after graduation and 24% of Female candidates got Job after graduation**

- *Bivariate Analysis Steps:*

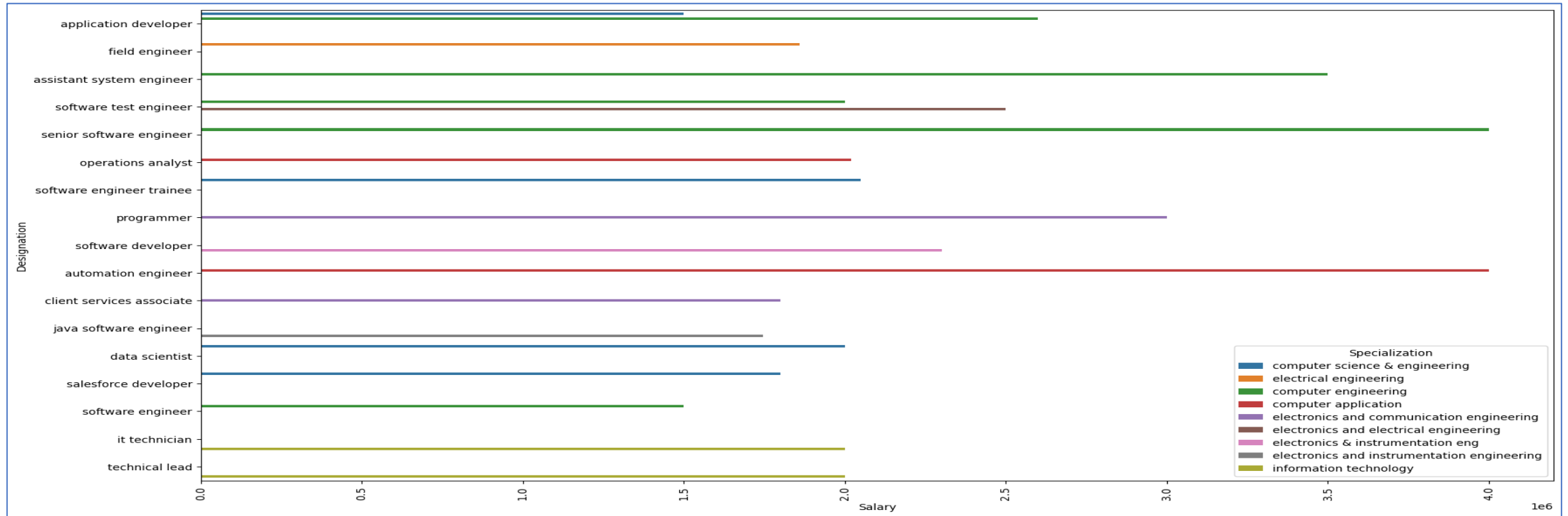Meaningful insights to be drawn by combining any 2 columns data.

Insight-4 -Based on : Job city vs Salary

From the above graph we understand that,
1. Gurgaon and New Delhi stands with highest salary of 35 to 40 LPA with Men Employees
2. Most of the Female associates are working in Pune are having high salary package of 35LPA

INNOMATICS
RESEARCH LABS

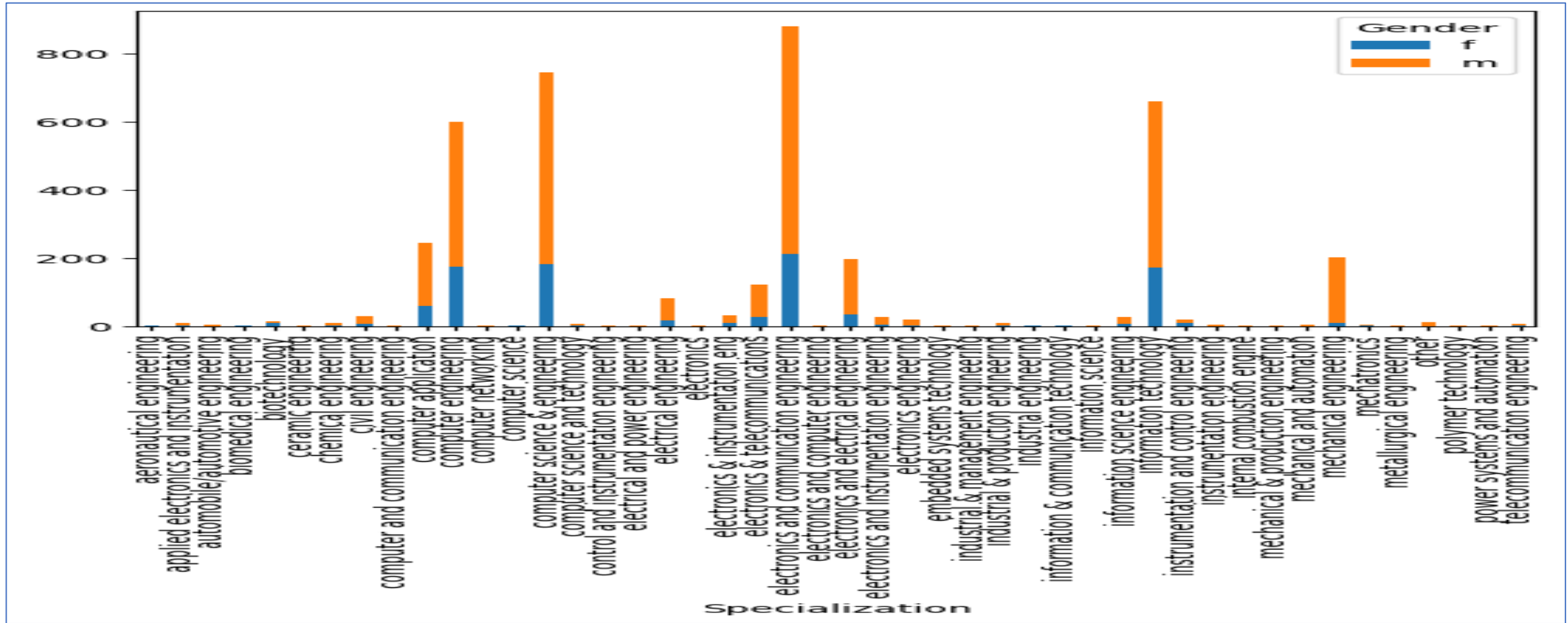**Insight-5 -Based on Specialization and Salary:**

From this data we can get an insight of salaries of students above 15LPA with designation and enggineering stream:
Designations:'application developer', 'field engineer','assistant system engineer', 'software test engineer','senior software engineer', 'operations analyst','software engineer trainee', 'programmer', 'software developer','automation engineer', 'client services associate','java software engineer', 'data scientist', 'salesforce developer','software engineer', 'it technician', 'technical lead'
Specialization: 'computer science & engineering', 'electrical engineering','computer engineering', 'computer application','electronics and communication engineering','electronics and electrical engineering','electronics & instrumentation eng','electronics and instrumentation engineering','information technology'
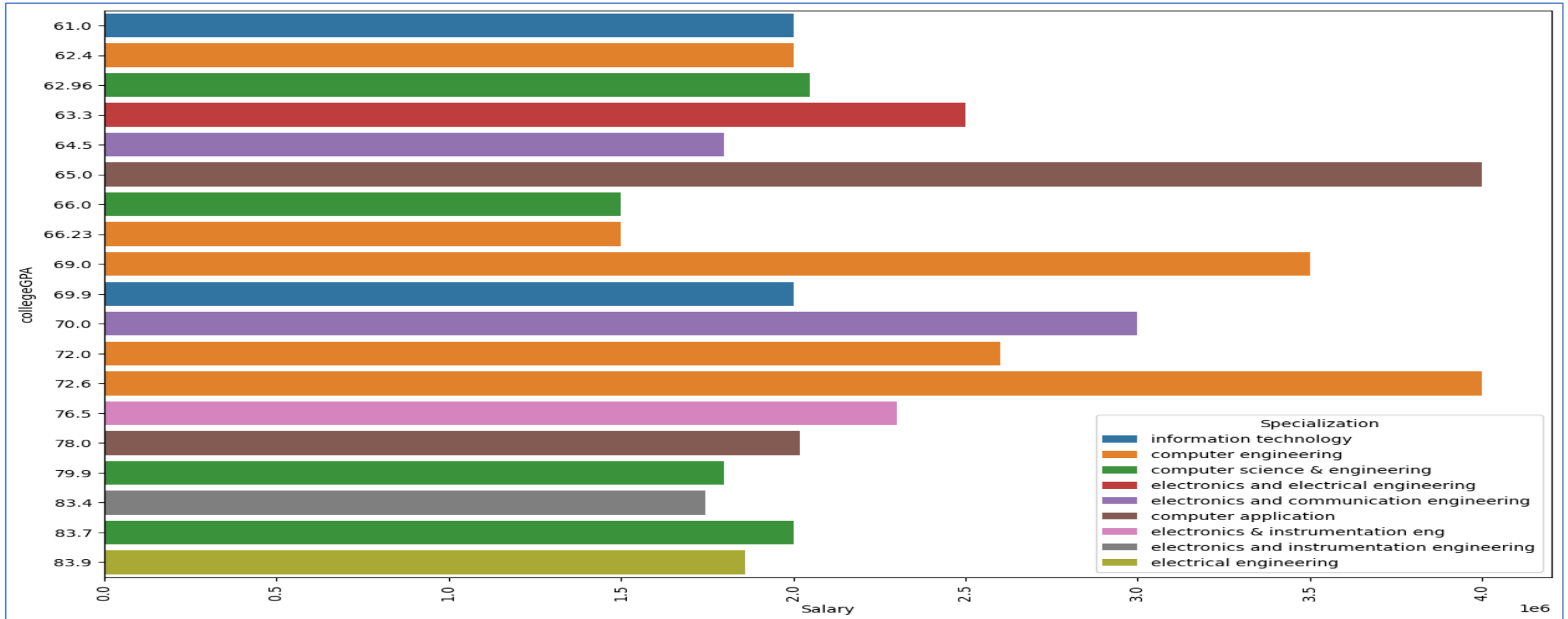
Insight-6 -Based on Gender and Specialization:

From the above graph we can understand that:
1. Based on Gender there is no preference of specialization.
2. Maximum number of students opted for electronics and communication engineering,information technology and electronics and electrical engineering, computer engineering, computer science & engineering have attended the AMCAT test.
3. Some bias can be seen for students opting mechanical engineering as 191 male and only 10 female students.
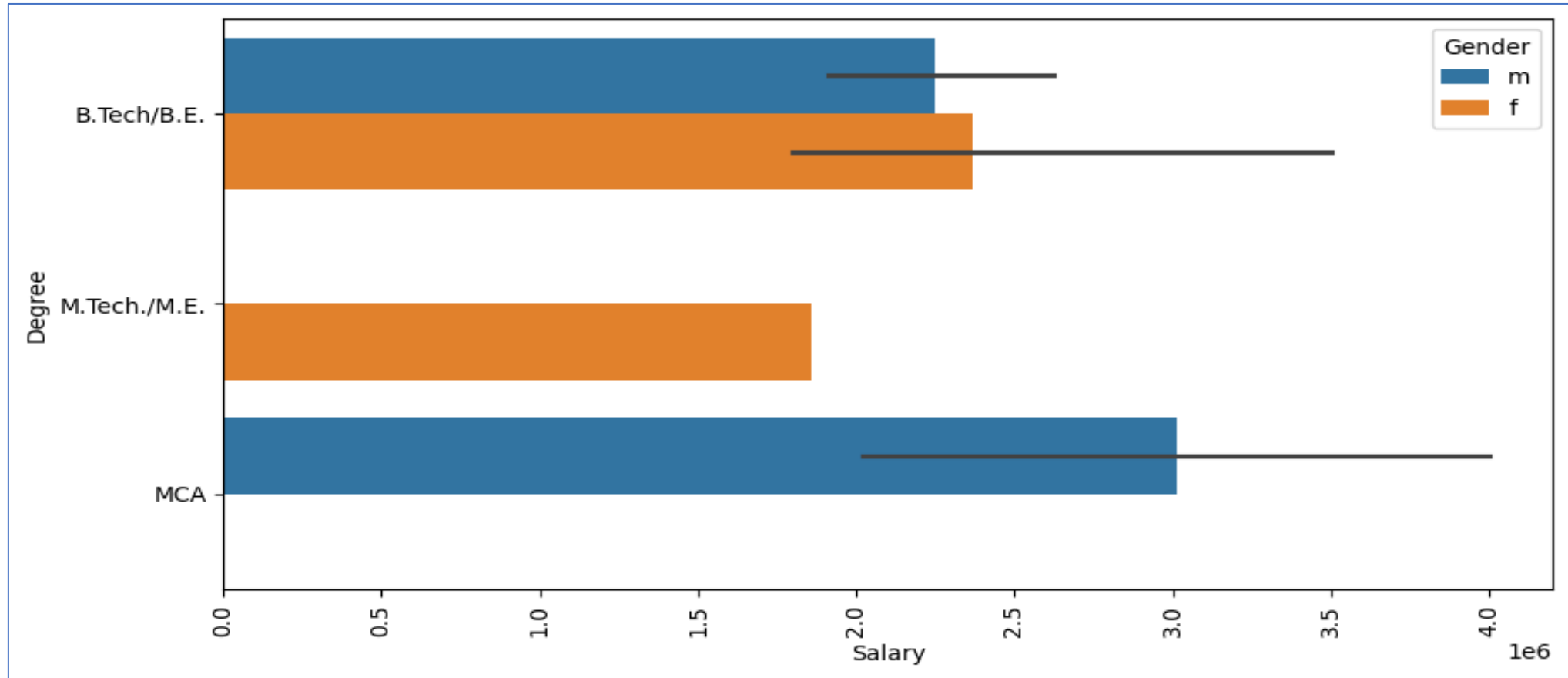
Insight-7 -Based on CollegeGPA and Salary:

From the above graph:
1. Surprisingly students with GPA above 76 have placed in less salary compared to students securing below 75GPA.
2. Computer science students with GPA of 75 have got the highest package of 40LPA

From the above comparative graph:

1. Most of the female employees earning above 15LPA have completed their M.Tech.

2. MCA candidates have got package above 30LPA.

Key Business Questions:

1. *What is the specialization of students who have attended AMCAT test and got salary above 35LPA?*

- Computer application and computer engineering students got salary package above 35LPA.

2. *What is the count of Male and Female candidates attending the test in AMCAT?*

- Out of 4000 candidates – 3071 are male and 957 are Female.

3. *What is the highest degree of students attending AMCAT test?*

- M.Tech /M.E is the highest degree of the candidates opting AMCAT test.

4. *Where are the top salary paying companies located?*

- New Delhi and Gurgaon are the 2 locations with highest salary package of 40LPA.

5. *What is the GPA of students with highest Salary package?*

- Computer science students with GPA of 75 have got the highest package of 40LPA

6. *What is the minimum and maximum salary packages?*

- Minimum salary starting from 3.5LPA to maximum of 40LPA.

Conclusion (Key finding overall):

1. AMCAT helped in understanding the placement opportunities in different companies based on specializations.

2. Data provides an clear idea of salary packages based on different features like College GPA, Job city, Degree and specializations.

3. Most of the students attempting this test are having CGPA of 60 to 80+. And have secured jobs with good salary packages.

Your Experience/Challenges working on – Data Analysis Project.

1. After scraping the data, the challenging part is to clean the data and handling the Null Values.

2. At every stage data to be checked as some irrelevant data may get scrapped from same class.

3. After cleaning the data, driving the insights is challenging as data retrieved to be understood before giving any insight.

4. Analyzing large amount of data at a time to get insights is difficult as graphical representation becomes complex.

QUESTIONS AND ANSWERS

THANK YOU

INNOMATICS
RESEARCH LABS