# INNOMATICS®
## RESEARCH LABS

### INNOVATION. AUTOMATION. ANALYTICS

## PROJECT ON

**AMCAT data - An Exploratory Data Analysis**

# About me

- Hi there! I'm UmaMahesh Yenamala [IN1240703], a 22-year-old data enthusiast, Computer Science and Engineering graduate student at CMR Institute Of Technology. Sure, I might be fresh, but I'm your one-stop shop for unlocking insights from numbers.

- Speaking of Experience, my internship at Feynn Labs was a rocket ship ride! For 60 days, I worked on real-time dataset projects, tamed them with machine learning algorithms, and generated insights from the dataset using EDA.

- Feel free to reach out! I'm always down to connect with you all. ☺
    Github: https://github.com/UMAMAHESHWARRAO302001/Amcat_EDA

    LinkedIn: https://www.linkedin.com/in/umamahesh-yenamala-a20423197/

# Objective

- **To find the insights and decode the patterns on how the salary pay scale is impacted based on various parameters like academic achievements, AMCAT scores, Age, Years of Work Experiences etc..**

# Summary of the Dataset

- **There are total of <u>38 columns</u> that are used to find their individual impacts on salary.**

- **With <u>4000 data points</u> that make our analysis to the optimal insights with all the required information.**
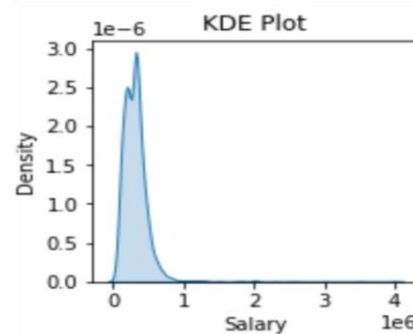
# What is our Target variable 'Salary' informs to us?


Box Plot
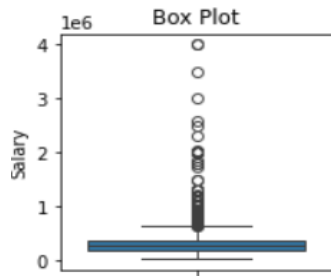
- **count** 3998
- **min** 35000
- **max** 4000000
- **mean** 307699
- **median** 300000
- **var** 4.52e+10
- **std** 2.12e+05
- **skew** 6.45
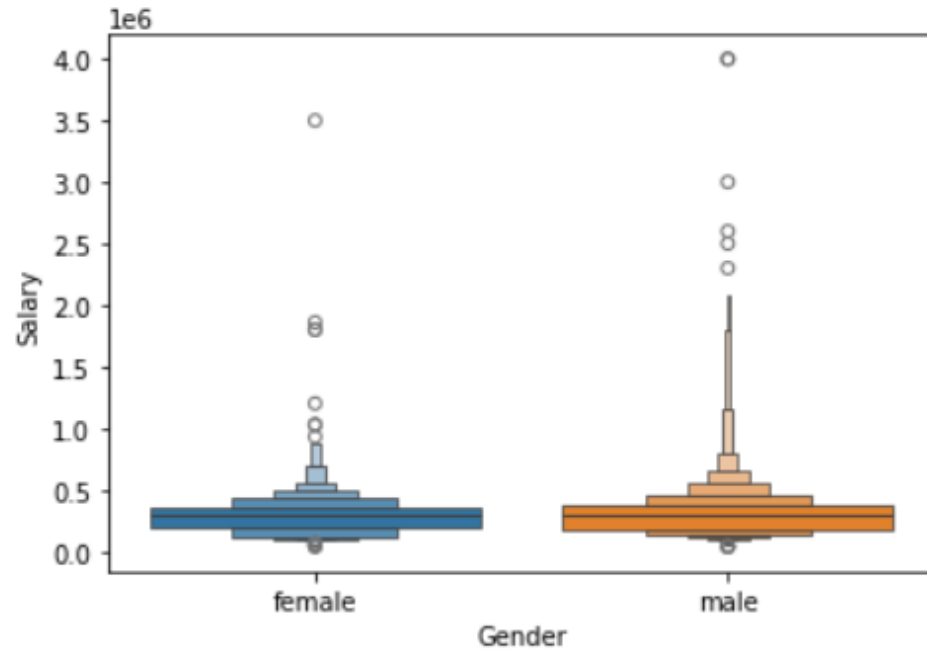- **kurt** 80.9


Histogram Plot
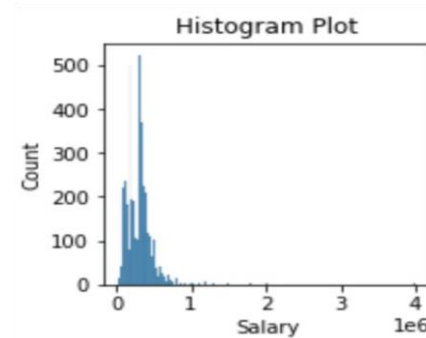
Salary ranges from 35k to 40L with higher density of 3L
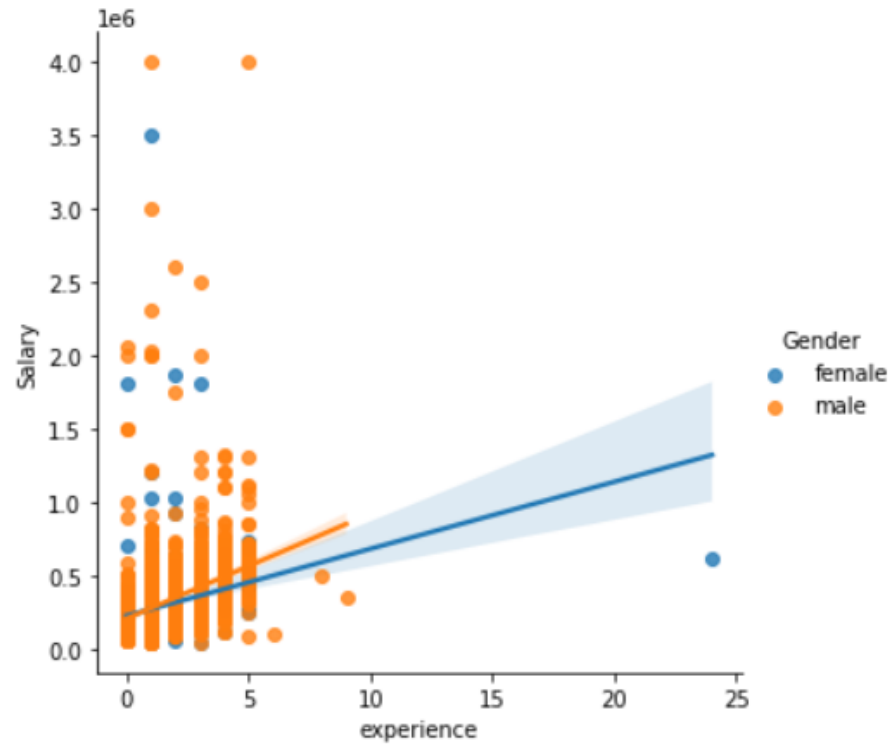

KDE Plot

It is right skewed

# Do you think there is salary difference for genders?



- **Yes, <u>Male</u> has higher median salary than Female.**

# Experience Pays Off? Is Salary increases with Experience?



There is a **slight increasing slope** for experience and salary. However we can observe that there are some good pay scale for the fresher's and we see that most of the employees are male when we compared with females, but still there is increasing proportion salary based on experience

# Let's Explore Board Transitions Across 10th and 12th using Chi Square test:

- Chi-square value: 4503.58

```python
from scipy.stats import chi2_contingency

contingency_table = pd.crosstab(df['10board'], df['12board'])
chi2, p, _, _ = chi2_contingency(contingency_table)

print(f"Chi-square value: {chi2}")
print(f"P-value: {p}")
```

- P-value: 0.0

- Since, P-Value equals to zero, we have enough evidence to reject H0 ( Null Hypothesis ).

- **Your 10th board choice might significantly impact your 12th board decision. This analysis reveals a hidden link between the board you choose in 10th grade and where your academic journey takes you in 12th.**

# Do Students Maintain Consistent Performance Across School and College?


Scatter Plot of School Marks vs College Marks
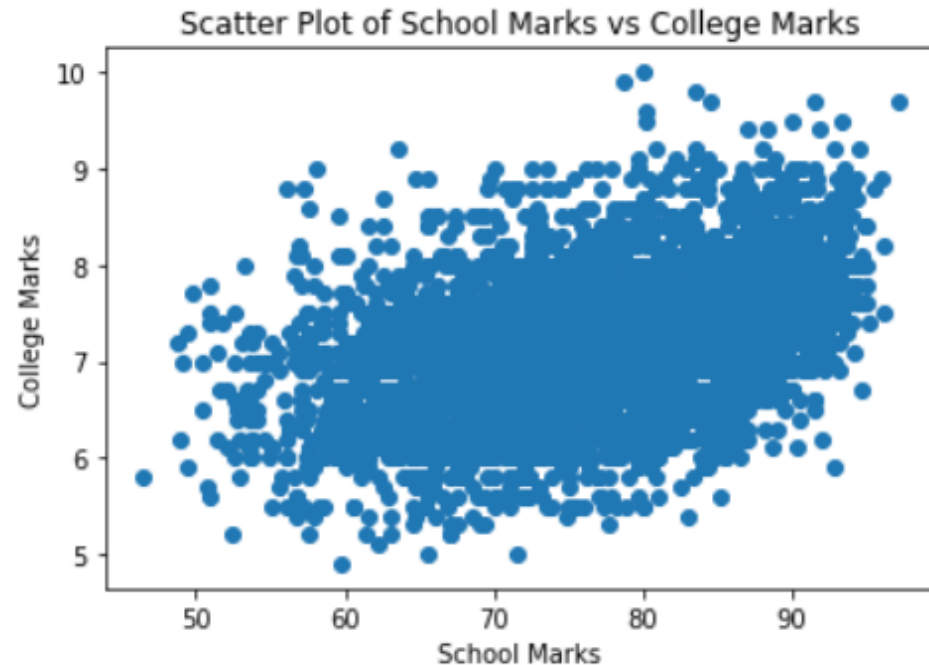
- **Pearson correlation coefficient: 0.36**

- **P-value: 1.77e-125**

- **Since, P-Value equals to zero, we have enough evidence to reject H0 ( Null Hypothesis ).**

- **We could see there is a correlation among school and college**

| col_0 | Salary |
| --- | --- |
| **Designation** | |
| technology lead | 645000.0 |
| sr. database engineer | 600000.0 |
| technical analyst | 600000.0 |
| asst. manager | 600000.0 |
| technology analyst | 552500.0 |
| ... | ... |
| documentation specialist | 80000.0 |
| visiting faculty | 70000.0 |
| web designer and joomla administrator | 60000.0 |
| trainee software developer | 60000.0 |
| secretary | 45000.0 |

413 rows × 1 columns

| col_0 | Salary |
| --- | --- |
| **Specialization** | |
| internal combustion engine | 360000.0 |
| metallurgical engineering | 337500.0 |
| ceramic engineering | 335000.0 |
| chemical engineering | 325000.0 |
| control and instrumentation engineering | 305000.0 |
| instrumentation and control engineering | 302500.0 |
| CSE | 300000.0 |
| CIVIL | 300000.0 |
| IT | 300000.0 |
| ECE | 300000.0 |

# Average salaries based on Specialization

- **Observation :** "CSE","information & CIVIL" specialization is having highest average salary among all specializations followed by "industrial engineering" and "telecommunication engineering".

- "aeronautical engineering" commanding the least average salary.
- Average salaries based on  Designation

- **Observation :** technology lead, sr. database engineer, and technical analyst positions offer the highest average salaries, each at

  60k – 64k

- technology lead and senior database engineer roles also provide competitive average salaries, with earnings reaching 645,000 and 600,000  respectively.

conversely, roles such as documentation specialist, visiting faculty, and web designer and joomla administrator offer comparatively lower average salaries, ranging from 80,000  to 60,000 . • trainee software developer and secretary positions

# Do companies still look for College Tier?
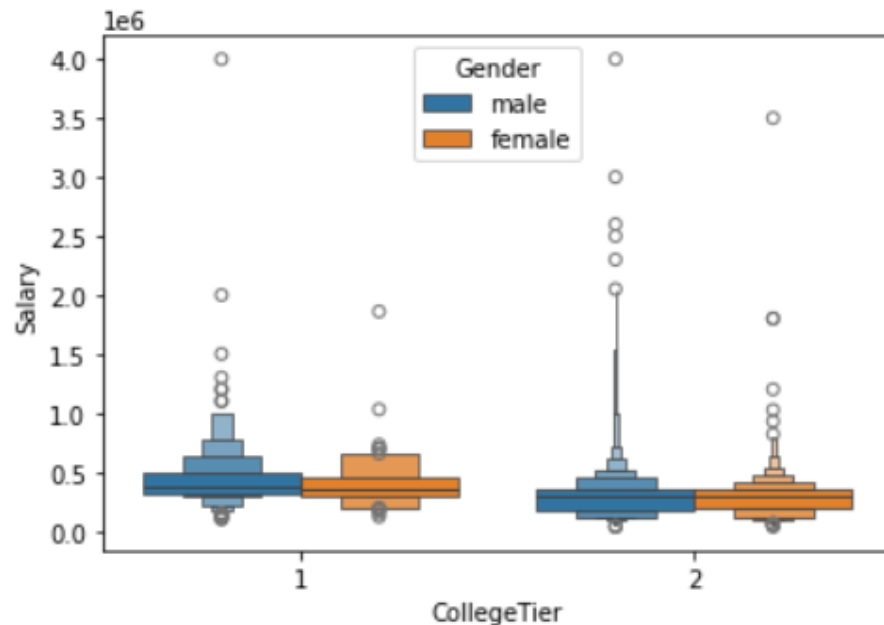


- This boxplot suggests attending a tier 1 college could be the key to unlocking a higher salary, but remember, **SKILLSET** matters!

- And companies offering big packages to Tier-1 colleges when we compared with Tier-2.

- When it comes to campus recruitment top companies will give first priority to Tier-1 colleges.

# Are Students Following the Crowd or Charting Their Own Path?
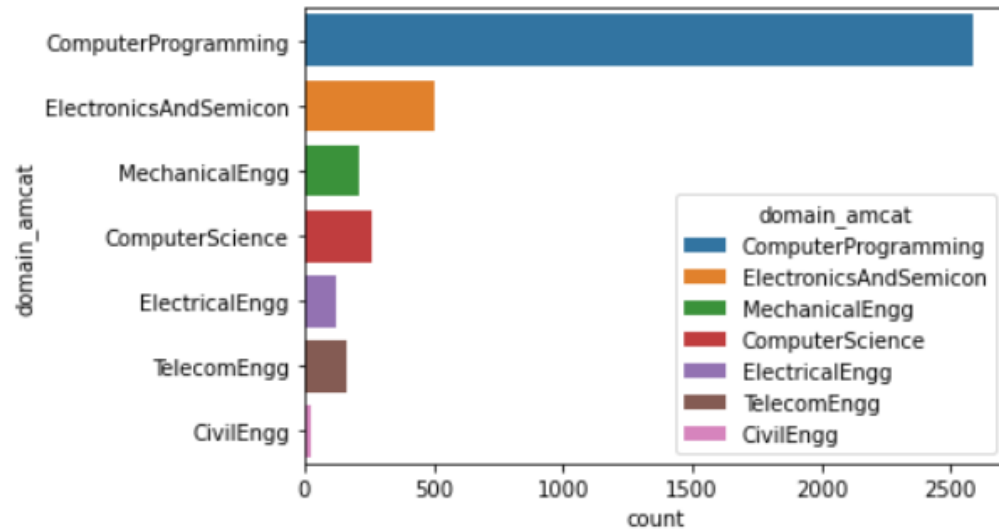


- From the plot we can conclude, Because of demand in software industry, most of the students are graduated from computers field. By their will be more competition in IT field regarding jobs.

- We can predict that there will be more demand for mechanical and civil in future as manufacturing industry increases.

# RESEARCH QUESTION - 1

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." Test this claim with the data given to you.

- **Test used – <span style="color:red">1 Sample T-Test</span>**

- **Level of Significance – <span style="color:red">0.05</span>**

| Designation | Test Statistic | P - Value | Conclusion |
|---|---|---|---|
| Software Engineer | 13.02 | 1.85e-36 | Reject H0 |
| Hardware Engineer | 0.68 | 0.513 | Do not reject H0 |
| Associate Engineer | 1.06 | 0.29 | Do not reject H0 |
| Programming Analyst | 5.40 | 1.99e-07 | Reject H0 |

# Conclusion:

- Only Software Engineer and Programming Analysts are having higher income range.

- Hardware and Associate Engineers do not have proper evidence to prove against.

# RESEARCH QUESTION - 2

According to (AMEO-2015) dataset. It has been observed from the analysis that B.Tech is the most preferred course in India with Electronics and Communication engineering stream as the most preferred branch with highest package around 13 lakhs per annum and average package around 5 lakhs
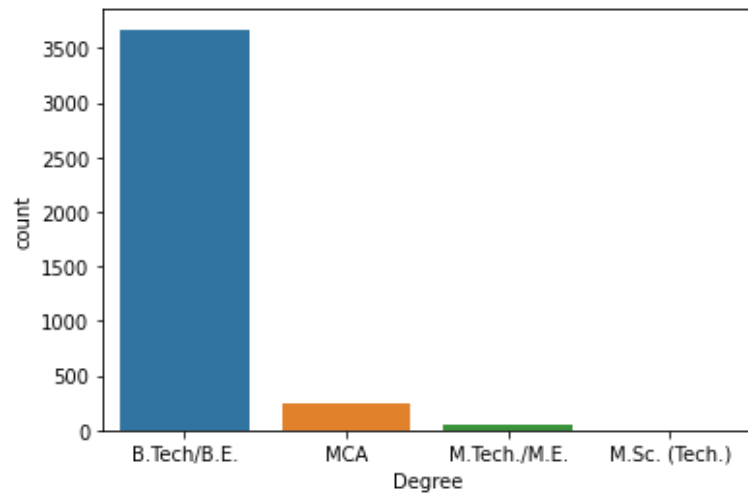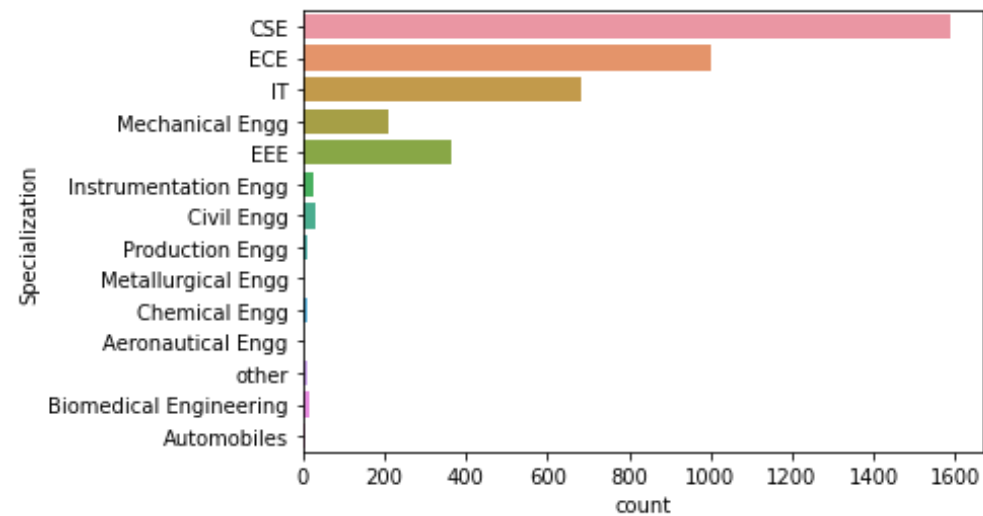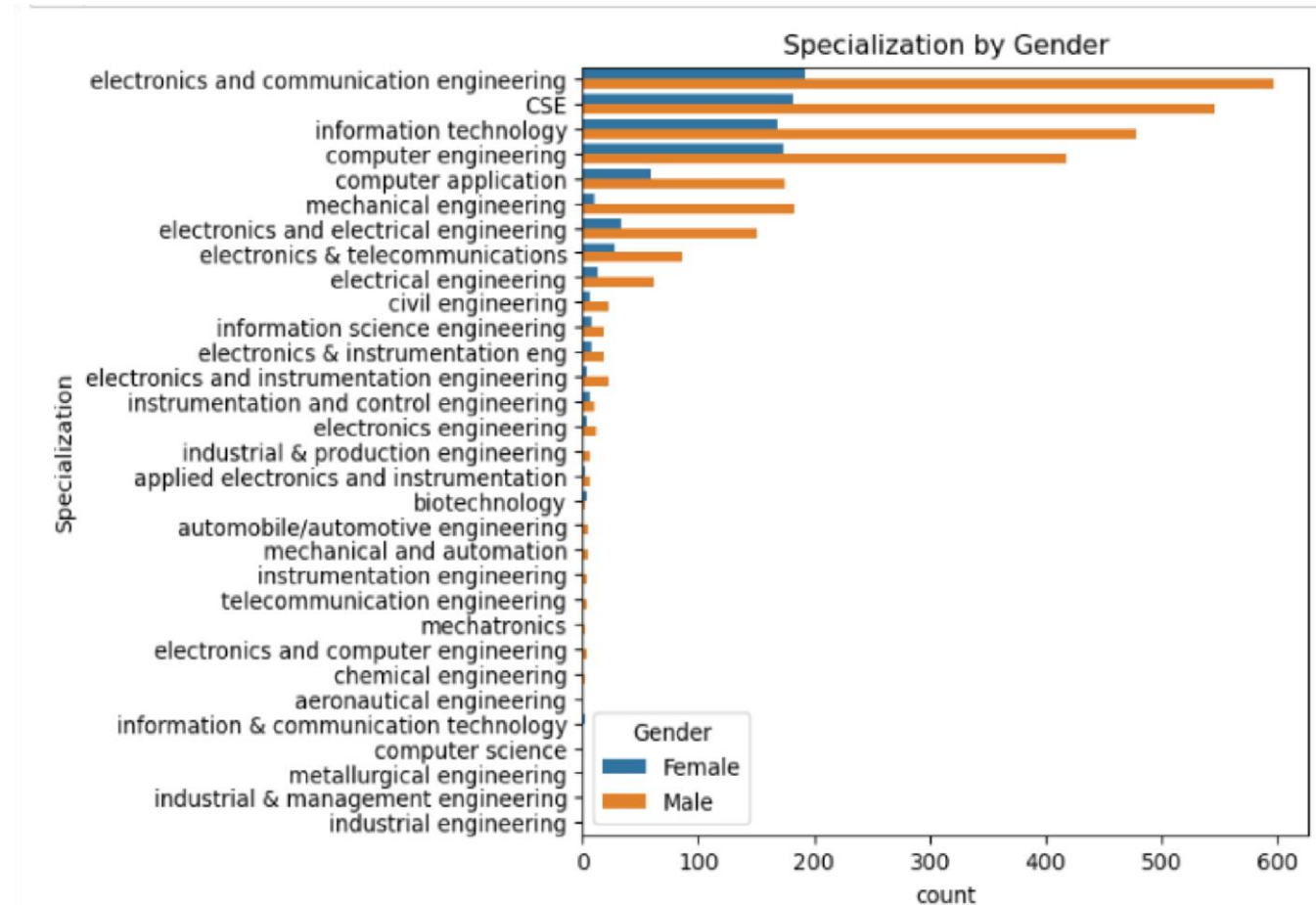
```
ece_salary['Salary'].max()
```
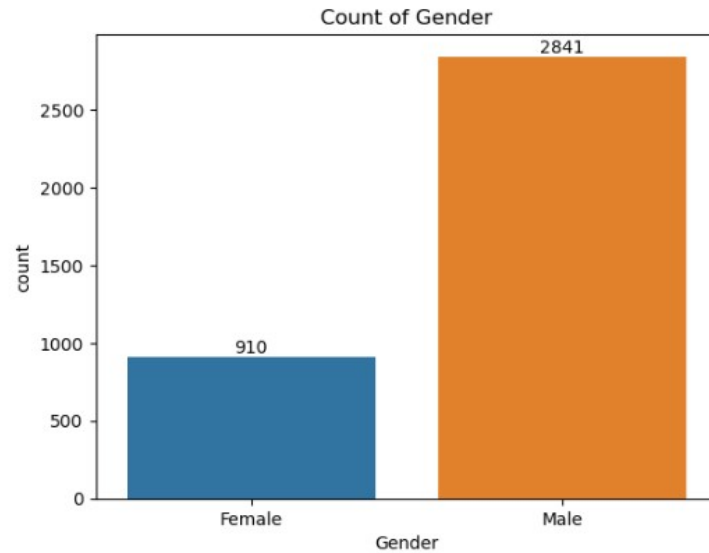
3000000.0

```
ece_salary['Salary'].mean()
```

296869.3918245264

# Research question 3: Is there a relationship between Gender and Specialization

# Observations

1. The dataset indicates a significant gender disparity, with a smaller number of females compared to males.

2. Despite the gender imbalance, the selection of specializations appears to be more diverse and less influenced by gender.

3. This means that both males and females in the dataset are opting for a wide range of specializations, suggesting that individual preference s, interests, or career goals play a more significant role in specialization choice than gender.

4. Various factors may influence specialization selection, including personal interests, academic background, career aspirations, opportunities for advancement, salary prospects, and societal trends.

5. Cannot conclude that there exists a relationship between "Gender" and "Specialization"

# Conclusion:

- **Male** earns more salary

- **Experience** pays you higher salary

- 10$^{th}$ & 12$^{th}$ board are **same** (Chi-sq test)

- Students maintain **consistence performance** in school and college as well

- **Top Tier college** graduates have chance of getting **more salary** pay

- Students wish to **relocate** themselves for works

- **AMCAT Scores** is one of the key feature that **determines your pay**

THANK YOU

INNOMATICS
RESEARCH LABS