TOPIC :

ANALYSIS OF FACTORS INFLUENCING CAMPUS PLACEMENTS

[KAIBURR ASSESMENT TASK 6]

Abstract

In today's world with the rapid increase in the number of engineering colleges and students, placements have become everyone's first priority before applying to any of the colleges. But when you search on google it doesn't really show any statistics or in-depth information about the college placements. So, we came up with a campus placement project idea.

This project helps college students in analysing and visualising the campus placement statistics. Graphs or pie charts are always much easier than textual data for understanding the huge dataset. This project will help in understanding the process of placement in a college, how the placement department works and what are the major factors that affect the placements. Python is used for coding.

This project further provides a prediction model which will help students understand their chances of getting placed or not. The applicability of this feature will help students to understand what are the factors they need to improve in order to get a good placement package. The project implementation will be useful for all colleges as the dataset has common fields whose data is easily available.

1. INTRODUCTION

When a student joins a college or any institute the first thing that they have in mind is that they will get a good placement from this college after graduation. Campus placements help the students to get a platform for themselves and they don't have to struggle themselves in the search for a job. Hence it is important to properly analyse the whole process of placements. In order to perform thorough data analysis, we chose a data set that consists of Placement data of students on the campus. It includes secondary and higher secondary school percentages and specialisation. It also includes degree specialisation, type and Work experience and salary offers to the placed students. This analysis will help students understand the process and will help them in getting better results in the future.

1.1. How did the project evolve?

Campus recruitment helps students in their career exploration and career path considerations just before they enter the corporate world. The recruitment process gives companies a chance to select employers and the jobs of their choice. However, it is important for students to be increasingly careful with regard to employers' hiring and selection decisions, due to an increase in the competition between students. Today, most students achieve decent scores and appropriate subject knowledge; however, employers are now looking for additional basic and soft skills that they consider essential for effective performance in the workplace. Very few studies had been conducted by researchers on a similar basis to ascertain students' perceptions about campus placement and the factors affecting their selection or rejection by employers visiting the campus. This motivated us to explore the perceptions of students with regard to campus placements and the corporate world in order to ensure a smooth transition from academics to employment.

As our project focuses on college data and campus placement statistics, its scope extends to college students or freshers ranging between age group 18-25. Along with analysis and visualisation of data, our project will also provide a prediction model which will help students understand their chances of getting placed or not. The applicability of this feature will help students to understand what are the factors they need to improve in order to get a good placement package. The project implementation will be useful for all colleges as the dataset has common fields whose data is easily available. Hence this project has a good scope if properly utilised in the campuses by the students

1.2. Why is this helpful?

As explained above it is important to have prior research about the placement trends before applying for any of the colleges. The process of data visualisation holds a lot of significance in doing so. Since there are a lot of students in a college and understanding the data trends manually will cause great difficulty, it is much easier to analyse and understand data if it's in a visual form like a bar graph or pie chart rather than in a textual form like spreadsheets. Understanding data quickly also means that students can make decisions based on that data much more quickly as well. It is sometimes possible to even estimate future trends using Data Visualization. This gives a huge edge to students as they can move ahead of their competitors by analysing future placement trends.

2. Background

[1] Factors Affecting the Placement Prospects of MBA Students: An Exploratory Study

The authors presented the results of a study that attempted to link students' placement prospects, operationalised through student perceptions of recruiting organisations to certain academic parameters such as performance in the entrance examination, group discussion, as well as personal interview, grade point average, internship, marks, and ratings on extra- curricular activities.

A comprehensive data analysis was conducted using various statistical tools to find out the factors having an impact on the campus placement of the students. Results were calculated using concepts of Averages and Chi squared tests of homogeneity on different factors that were initially considered.

The study revealed that students having prior work experience got placement in organisations which are perceived to be better. Also, students with work experience performed better in the personal interview as compared to freshers.

Gaps Identified:

1. Less explanation about the methods and algorithms used to complete the analysis.

2. The tabulated results had no kind of visual representations that made understanding the trends a little difficult.

[2] ABCD Analysis of On-Line Campus Placement Model

The authors studied deeply to discover the practical viability, usefulness, resourcefulness and universal applicability of the on-line campus placement model to Recruiters, Students, Parents, Society, and Institutions.

Determinant issues and affecting factors of the online campus placement model were identified using ABCD analysis framework. ABCD stands A for Advantages, B for Benefits, and C for Constraints and D for Disadvantages. The analysis identified the affecting factors for various determinant issues under four constructs: advantages, benefits, constraints, and disadvantages. The analysis has brought about 158 critical constituent elements which satisfy the success for its existence. Through this research, new insights were also generated about possibilities of building a robust online placement training model as well to cater and adhere to the inevitable application of Online Industry Oriented Placement Model (OIOC) in the near future. With digital India initiatives gathering momentum in the nation, employers are shifting towards OIOC adaptability and approach.

[3] Factors affecting placement and hiring decisions: A study of students' perceptions

The authors presented a paper where they assessed students' perceptions concerning the level of campus placement activities of their higher education institution and determined the order of importance of various factors, as perceived by students, relating to employers' selection criteria.

Primary data that was analysed was based on the responses of 621 students in 29 HEIs operating in the National Capital Region of India. Statistical tools such as average, t-test and Garrett Ranking were used. The t-test results indicated that students' perceptions across the five demographics under study were significantly different. Further, the Garrett Ranking method revealed that students perceived performance in the interview as the most important factor affecting their selection or rejection by employers during campus placements. The major implication of the study is that the gap in perception between students and employers needs to be bridged through collaboration between HEIs and industry.

Gaps identified:

1. Tables were the only visualization method used.

2. The research was based on students' perceptions and not the reality which could be very different.

[4] Campus placements in Kerala-An empirical study at the selected Engineering Colleges in Kerala

The authors presented a paper focusing on what are the various parameters that affect the students in tackling campus placements and how these parameters affect the employability of the students.

The study was conducted at the leading engineering colleges in Kerala. The various elements of research design are;

1. Database Design-The primary data was collected from students of engineering colleges in Kerala. The secondary data was collected from government records, other official records journals, text books and internet portals.

2. Measurement Design- The data was collected by using questionnaire. Nominal, ordinal; interval and ratio scales were used depending upon the data collected.

3. Sampling Design- The simple random sampling is used for the study. The total sample size is 114 and the samples were collected from the student's community of selected engineering colleges in Kerala. The period of the study was from January 02, 2012 to May 25, 2012.

4. Statistical design: Appropriate mathematical and statistical tool were used for analysis.

The data was analysed using a statistical package for social science (SPSS V 12.0). Descriptive statistics such as mean and standard deviation were generated to provide an overview of the data. ANOVA was used for finding significant differences between Educational Qualification

with respect to level of interest in campus placement. Chi-square test were used to find out association between number of placement drives attended and the percentage marks obtained by the students, number of times the student cleared the aptitude test and the percentage marks obtained by the students, number of times the student cleared the Group Discussion and the percentage marks obtained by the students. The Friedman test was used for finding out the various parameters of teaching and extracurricular activities on campus placements.

Gaps identified:

1. Tables were the only visualization method used.

2. Sample size for the study was low.

3. There are chances that the respondent's bias may also reduce the effectiveness of the data collected.

4. The result of the study cannot be generalized.

3. Methodology

3.1. Full description of the project

- Dataset collected from: https://www.kaggle.com/benroshan/factors-affecting-campus-placement

Placement_Data_Full_Class

sl_no	gender	ssc_p	ssc_b	hsc_p	hsc_b	hsc_s	degree_p	degree_t	workex	etest_p	specialisation	mba_p	status	salary
1	M	67.00	Others	91.00	Others	Commerce	58.00	Sci&Tech	No	55	Mkt&HR	58.8	Placed	270000
2	M	79.33	Central	78.33	Others	Science	77.48	Sci&Tech	Yes	86.5	Mkt&Fin	66.28	Placed	200000
3	M	65.00	Central	68.00	Central	Arts	64.00	Comm&Mgmt	No	75	Mkt&Fin	57.8	Placed	250000
4	M	56.00	Central	52.00	Central	Science	52.00	Sci&Tech	No	66	Mkt&HR	59.43	Not Placed	
5	M	85.80	Central	73.60	Central	Commerce	73.30	Comm&Mgmt	No	96.8	Mkt&Fin	55.5	Placed	425000
6	M	55.00	Others	49.80	Others	Science	67.25	Sci&Tech	Yes	55	Mkt&Fin	51.58	Not Placed	
7	F	46.00	Others	49.20	Others	Commerce	79.00	Comm&Mgmt	No	74.28	Mkt&Fin	53.29	Not Placed	
8	M	82.00	Central	64.00	Central	Science	66.00	Sci&Tech	Yes	67	Mkt&Fin	62.14	Placed	252000
9	M	73.00	Central	79.00	Central	Commerce	72.00	Comm&Mgmt	No	91.34	Mkt&Fin	61.29	Placed	231000
10	M	58.00	Central	70.00	Central	Commerce	61.00	Comm&Mgmt	No	54	Mkt&Fin	52.21	Not Placed	
11	M	58.00	Central	61.00	Central	Commerce	60.00	Comm&Mgmt	Yes	62	Mkt&HR	60.85	Placed	260000
12	M	69.60	Central	68.40	Central	Commerce	78.30	Comm&Mgmt	Yes	60	Mkt&Fin	63.7	Placed	250000
13	F	47.00	Central	55.00	Others	Science	65.00	Comm&Mgmt	No	62	Mkt&HR	65.04	Not Placed	
14	F	77.00	Central	87.00	Central	Commerce	59.00	Comm&Mgmt	No	68	Mkt&Fin	68.63	Placed	218000
15	M	62.00	Central	47.00	Central	Commerce	50.00	Comm&Mgmt	No	76	Mkt&HR	54.96	Not Placed	
16	F	65.00	Central	75.00	Central	Commerce	69.00	Comm&Mgmt	Yes	72	Mkt&Fin	64.66	Placed	200000
17	M	63.00	Central	66.20	Central	Commerce	65.60	Comm&Mgmt	Yes	60	Mkt&Fin	62.54	Placed	300000
18	F	55.00	Central	67.00	Central	Commerce	64.00	Comm&Mgmt	No	60	Mkt&Fin	67.28	Not Placed	
19	F	63.00	Central	66.00	Central	Commerce	64.00	Comm&Mgmt	No	68	Mkt&HR	64.08	Not Placed	
20	M	60.00	Others	67.00	Others	Arts	70.00	Comm&Mgmt	Yes	50.48	Mkt&Fin	77.89	Placed	236000
21	M	62.00	Others	65.00	Others	Commerce	66.00	Comm&Mgmt	No	50	Mkt&HR	56.7	Placed	265000
22	F	79.00	Others	76.00	Others	Commerce	85.00	Comm&Mgmt	No	95	Mkt&Fin	69.06	Placed	393000
23	F	69.80	Others	60.80	Others	Science	72.23	Sci&Tech	No	55.53	Mkt&HR	68.81	Placed	360000
24	F	77.40	Others	60.00	Others	Science	64.74	Sci&Tech	Yes	92	Mkt&Fin	63.62	Placed	300000
25	M	76.50	Others	97.70	Others	Science	78.86	Sci&Tech	No	97.4	Mkt&Fin	74.01	Placed	360000

- Data determination: -

| Attribute	Attribute Type
Gender	Categorical (nominal)
ssc_p	Quantitative
hsc_p	Quantitative
Ssc_b, hsc_b, hscc_s	Categorical (nominal)
degree_p	Quantitative
degree_t	Categorical (nominal)
workex	Qualitative-Binary attribute
etest_p	Quantitative
Specialization	Categorical (nominal)
mba_p	Quantitative
status	Qualitative-Binary attribute
salary	Quantitative

The data is stored in **.csv** format

- Data is accessed using python pandas library using **read_csv()** method.
- **Language Used**: Python.
- **Libraries**:
  - a. Pandas
  - b. Numpy
  - c. matplotlib.pyplot
  - d. seaborn
  - e. sklearn
- **Classification Algorithms Used:**
  - a) **Logistic Regression:** Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.
  - b) **Decision Tree:** Decision trees use multiple algorithms to decide to split a node into two or more sub-nodes. The creation of sub-nodes increases the homogeneity of resultant sub-nodes. The decision tree splits the nodes on all available variables and then selects the split which results in most homogeneous sub-nodes.

Implementation - Flow Chart
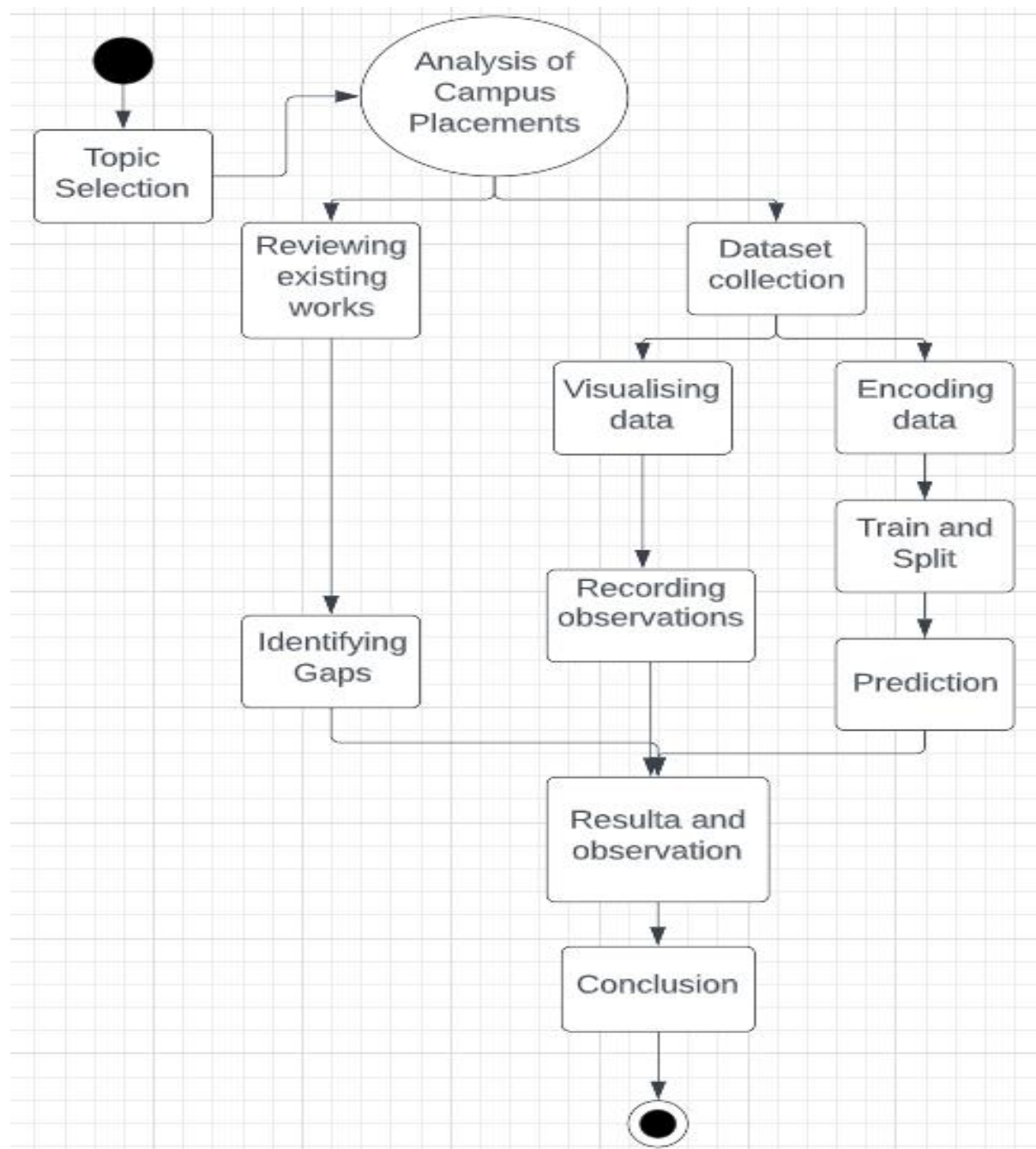
3.2. Full architecture
3.2.1. Low level design
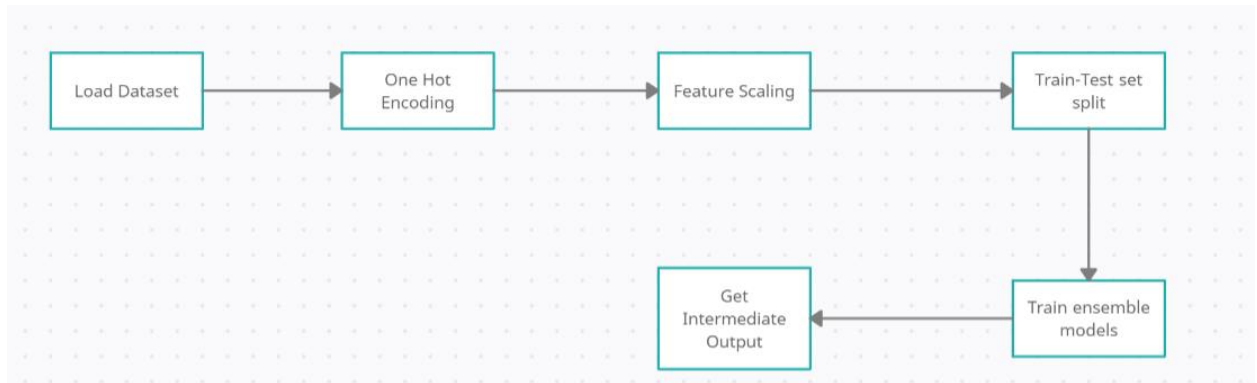
3.2.2. High level design
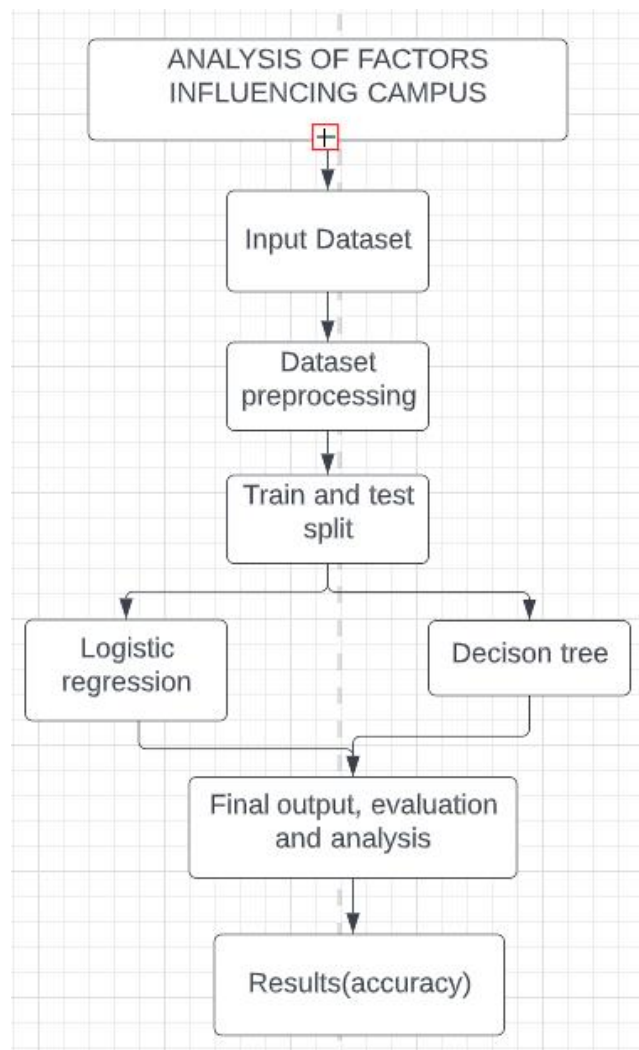
3.3. Obtained Result Analysis

3.3.1. User can compare the Male and Female candidates who applied for Placement & their Placement status

Abstract Level Task:

- Presentation of number of candidates for placement- Gender wise
- Comparing placement status -Gender wise
- Dataset type: Table
- Data type: Items and Attributes
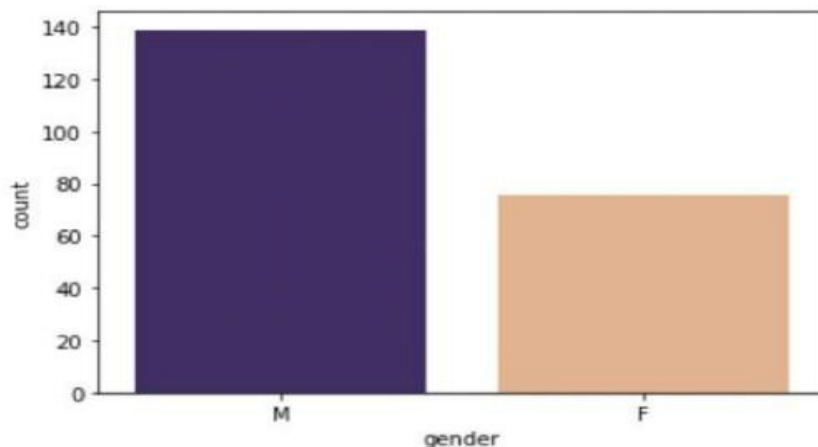
Observations: More number of male candidates applied for the placement process than female candidates.

What is the placement Status of male and Female?
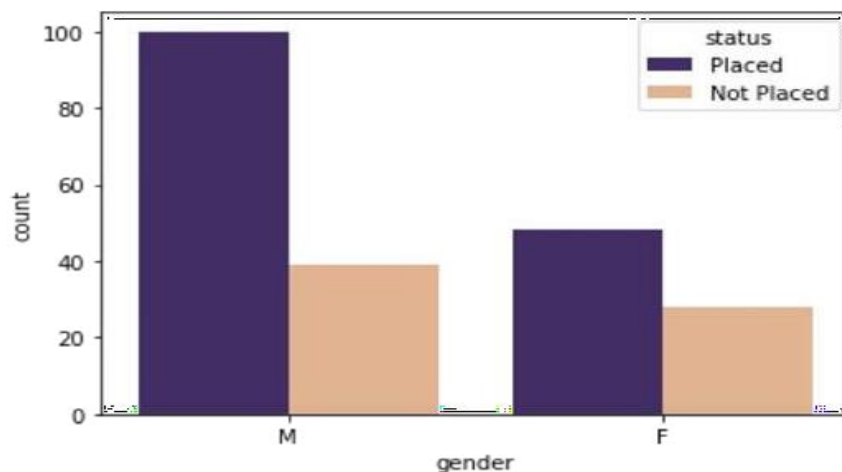
Observations:

- Number of male students is almost double as compared to females.

- Fraction of placed vs not placed for female candidates is significantly low as compared to male candidates.

Hence, we can conclude male candidates are accepted more often than female.

3.3.2. Users will be asked to specify their board in 10th,12th grade.
Abstract Level Task:
- Impact of taking a specific board in 10th grade on placements
- Dataset type: Table
- Data type: Items and A
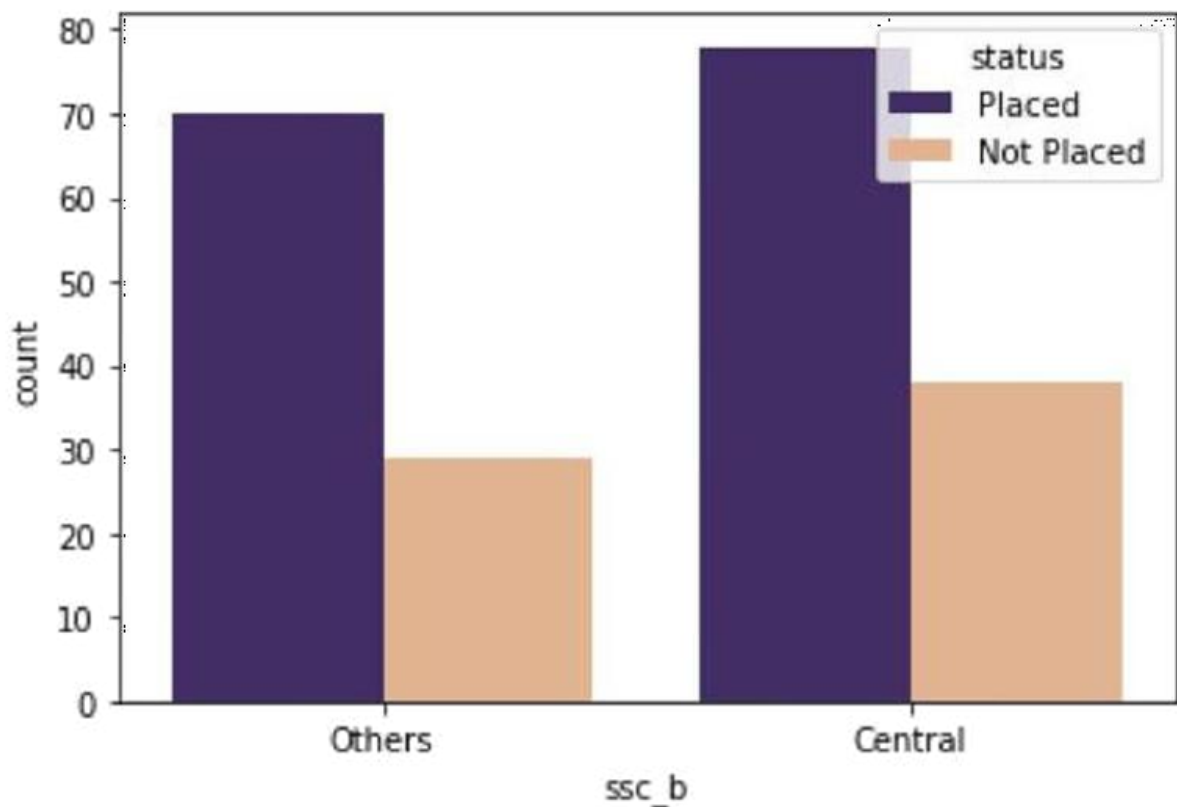
Observations:
- The number of central board students is very high as compared to all other boards.
- The number of placed students from the central board is little more than other categories which doesn't say much.

3.3.3. User will be asked to specify their CGPA and Degree Percentage
Abstract Level Task:
- Analyse whether CGPA and Degree Percentage Matter in Placements
- Understanding placement status on the basis of CGPA in 10th,12th grade and degree.
- Dataset type: Table
- Data type: Items and Attributes

X
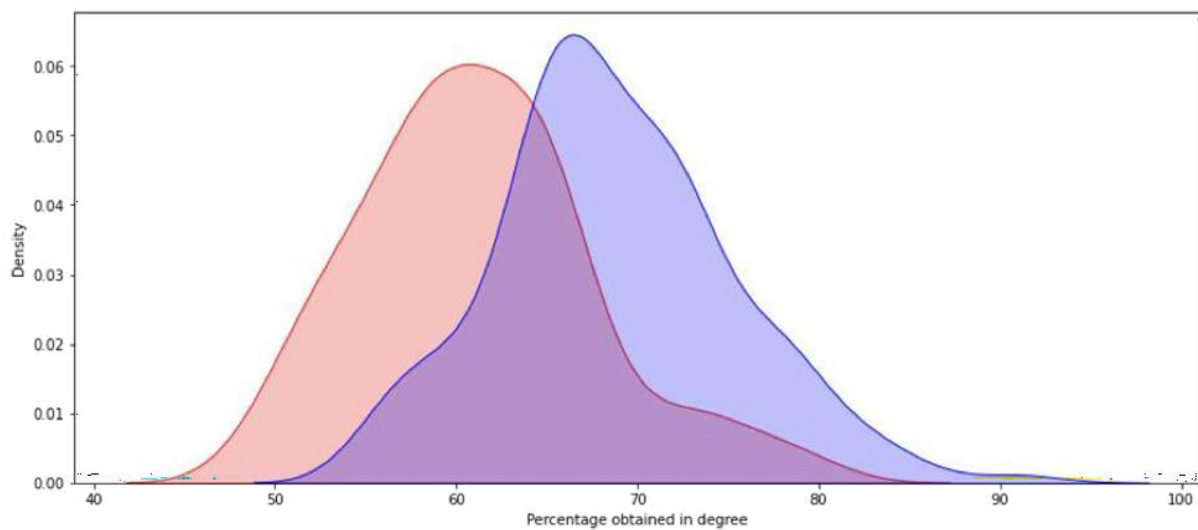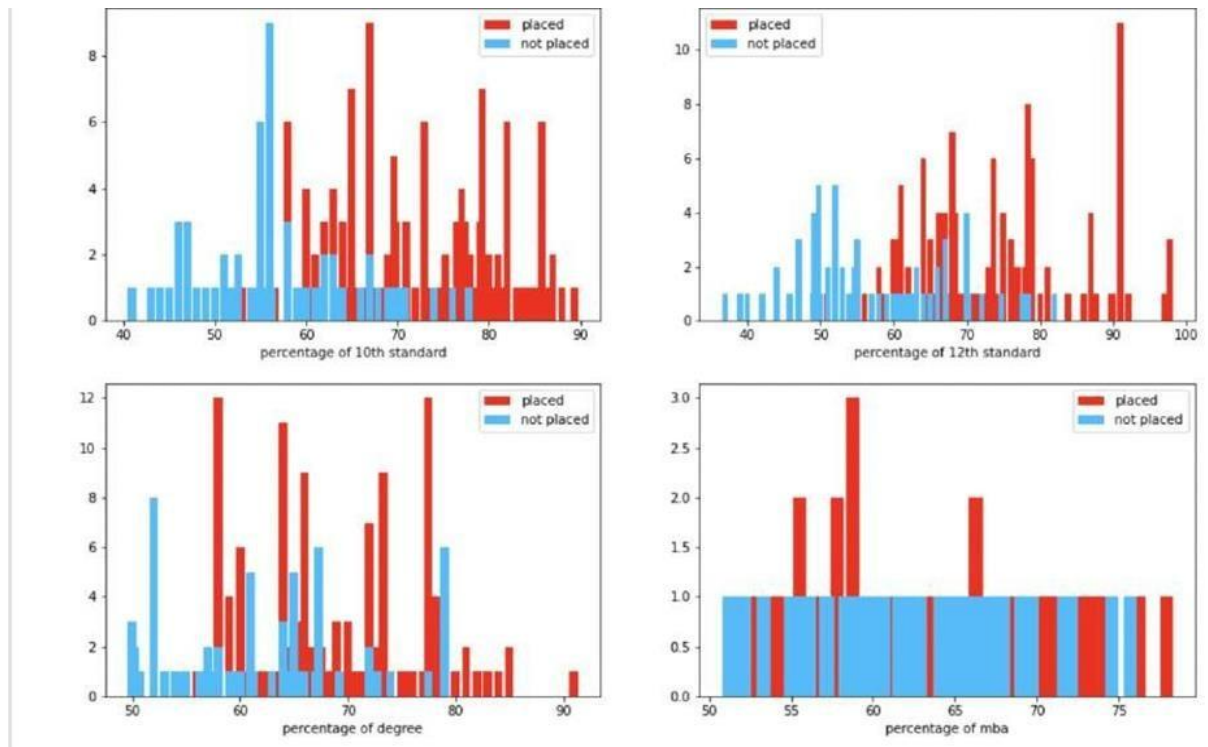
Observations:

- Students with percentages from 90-100 are fully placed.
- Students with percentages from 40-50 are not at all placed.

Do Etest marks Matter in Placements?
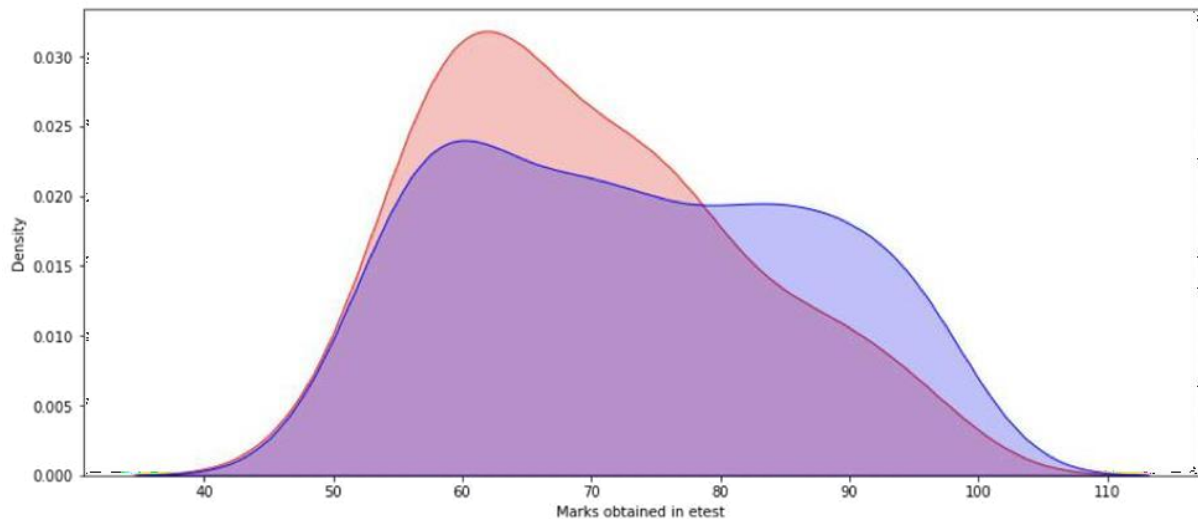
Observations:

- Etest marks cannot be considered as a significant factor as the marks are even distributed along with the placement status.

3.3.4. Users can analyse the impact of hsc specializations in placement.
Abstract Level Task:
- Understanding the placement status trend on the basis of hsc specialization [Science, Commerce, Arts]
- Dataset type: Table
- Data type: Items and Attributes
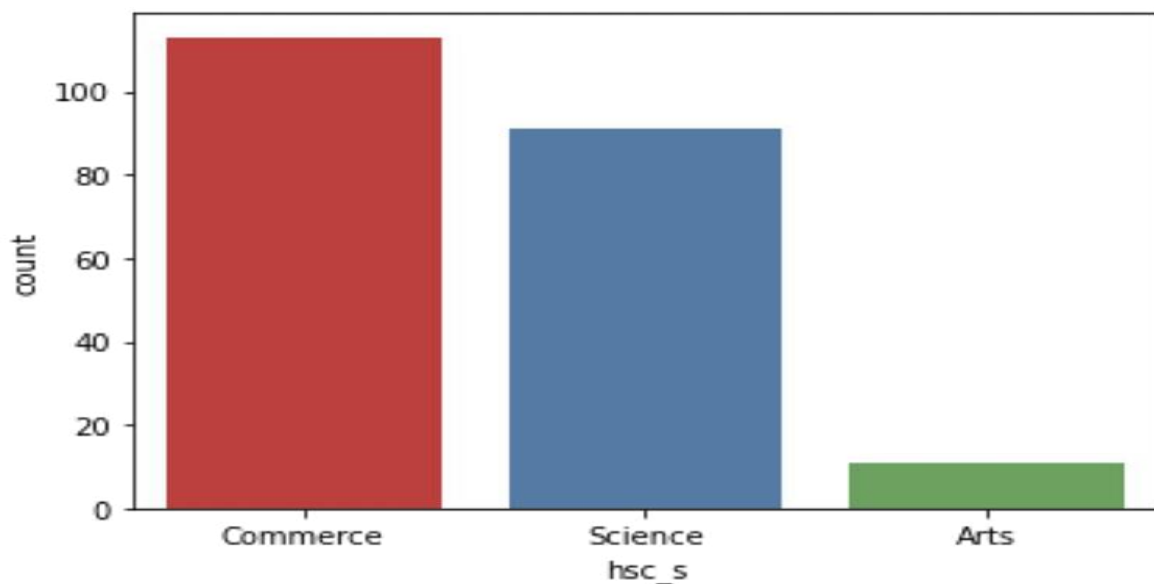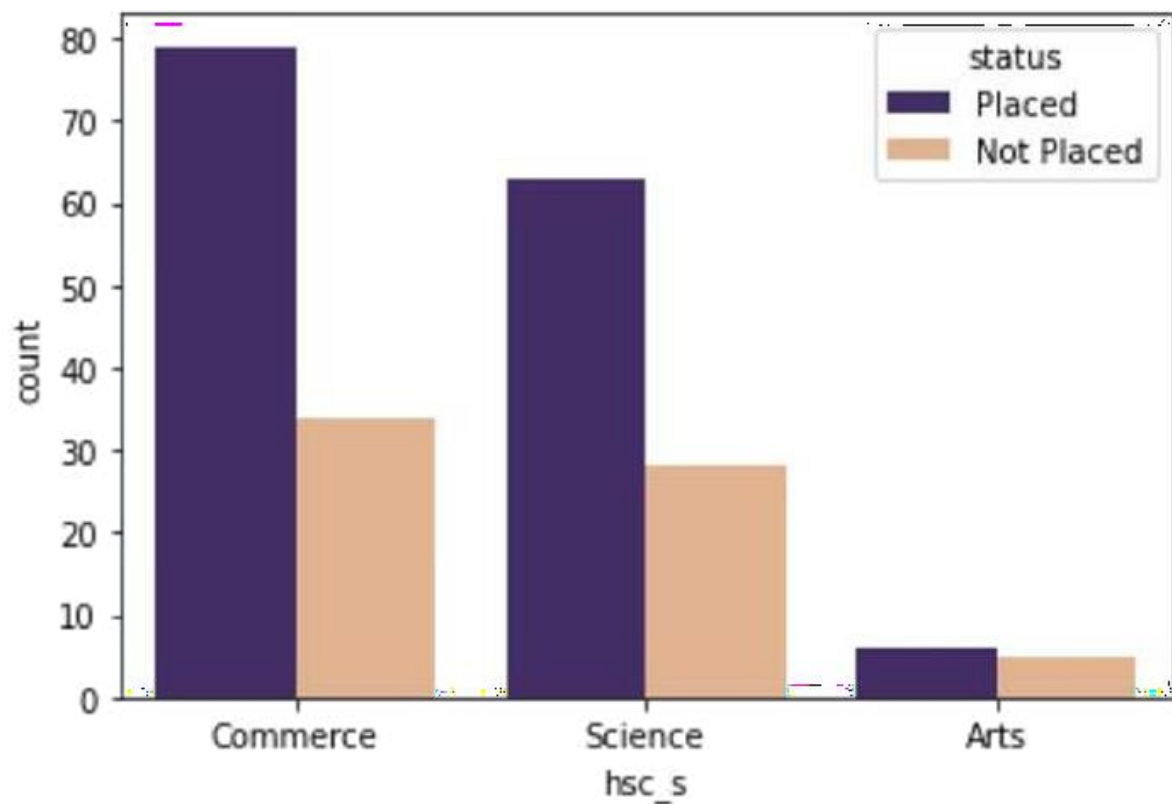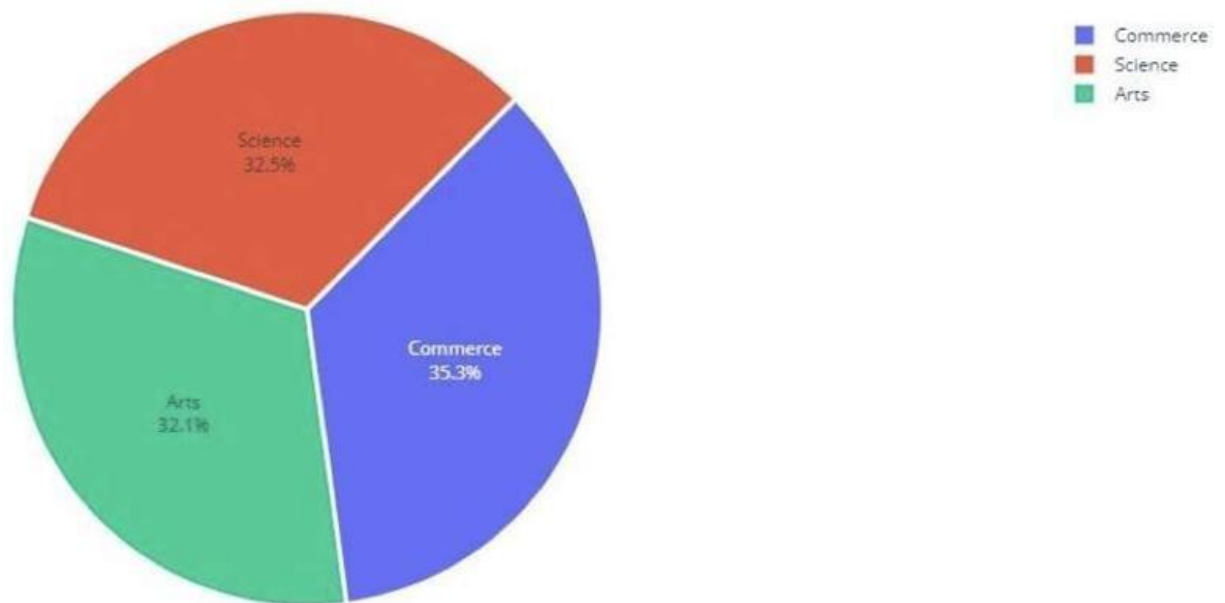
Observations:

- The most popular branch turns out to be commerce or maybe as most of the students get average marks so they were admitted to get commerce based on their marks. Science is the second most popular and the least popular is arts.

- Almost every branch student performed equally but commerce students have slightly better scores than the other two.

- Looking at the fraction of placed and not placed we can say that science branch students have more chance of getting placed than commerce students and around 45% of the students in arts are not placed.

3.3.5. **Users can analyse the impact of Degree type specializations in placement.**
Abstract Level Task: Understanding the placement status trend on the basis of
- Degree type [Science, Commerce, Arts]
- Dataset type: Table
- Data type: Items and Attributes
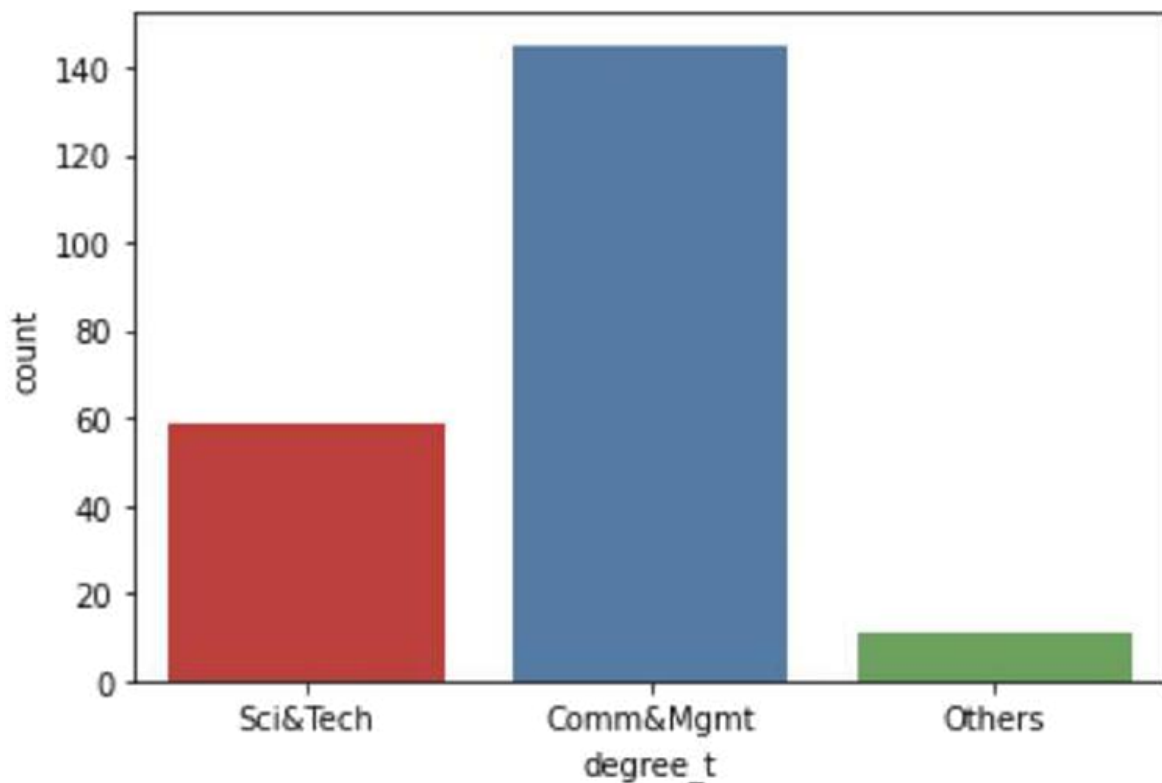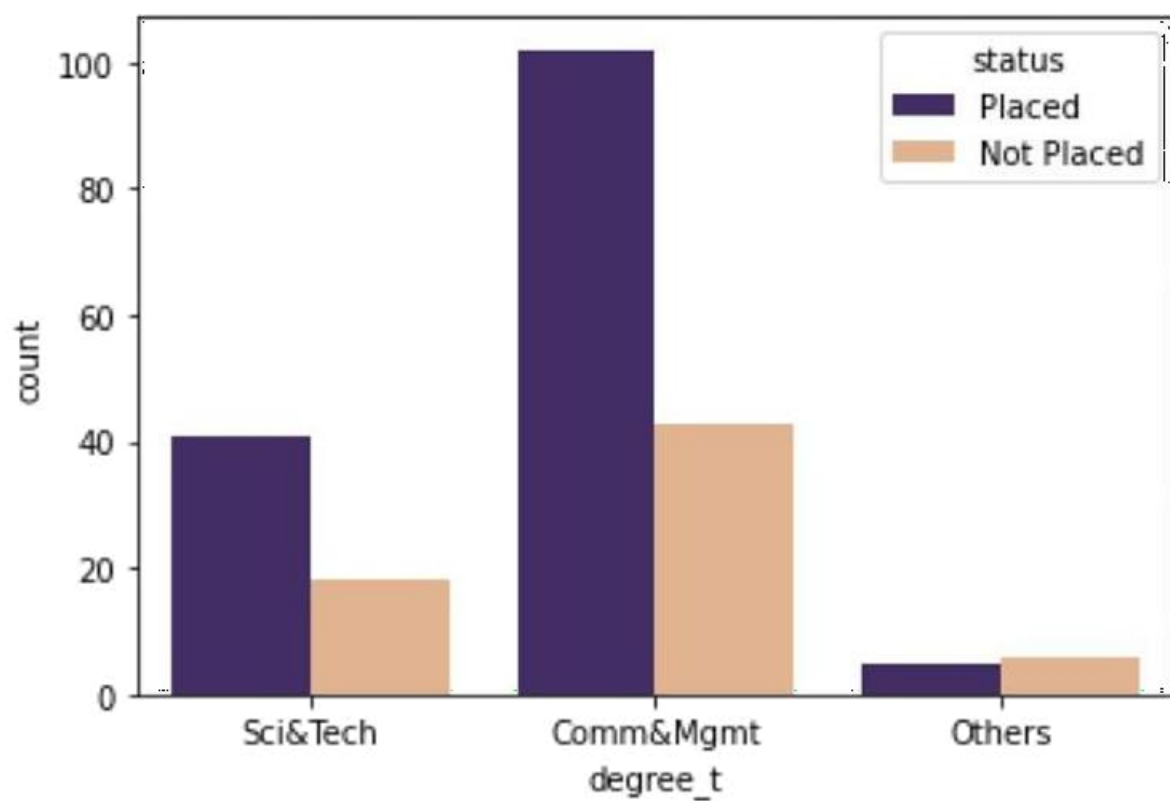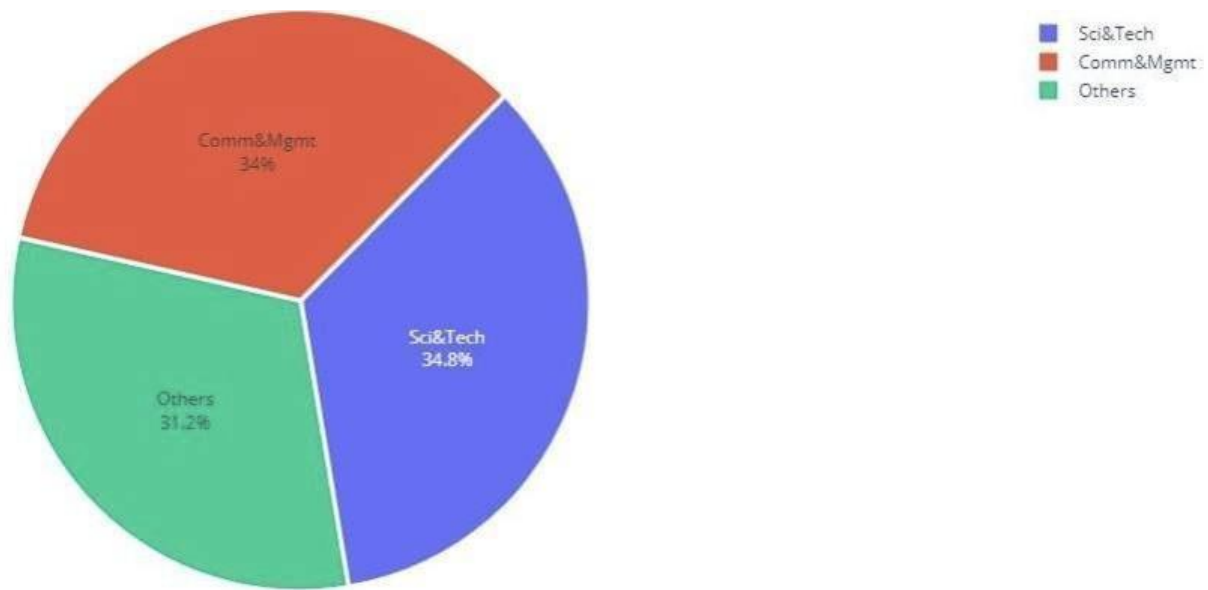
Observations:

- The students opted for following fields:
- Science and Technology (must be science students)
- Commerce and management (might be a mixture of commerce and Arts)
- Others

- There is not much difference in performance of students from Science and Commerce but there but students who opted for "Others" have low performance.
- Looks like Commerce and Science degree students are preferred by companies which is obvious. Students who opted for Others have very low placement chances.

3.3.6. User needs to specify their specialization
Abstract Level Task: Distribution of students based on their specialization -
- Gender wise analysis
- Dataset type: Table
- Data type: Items and Attributes
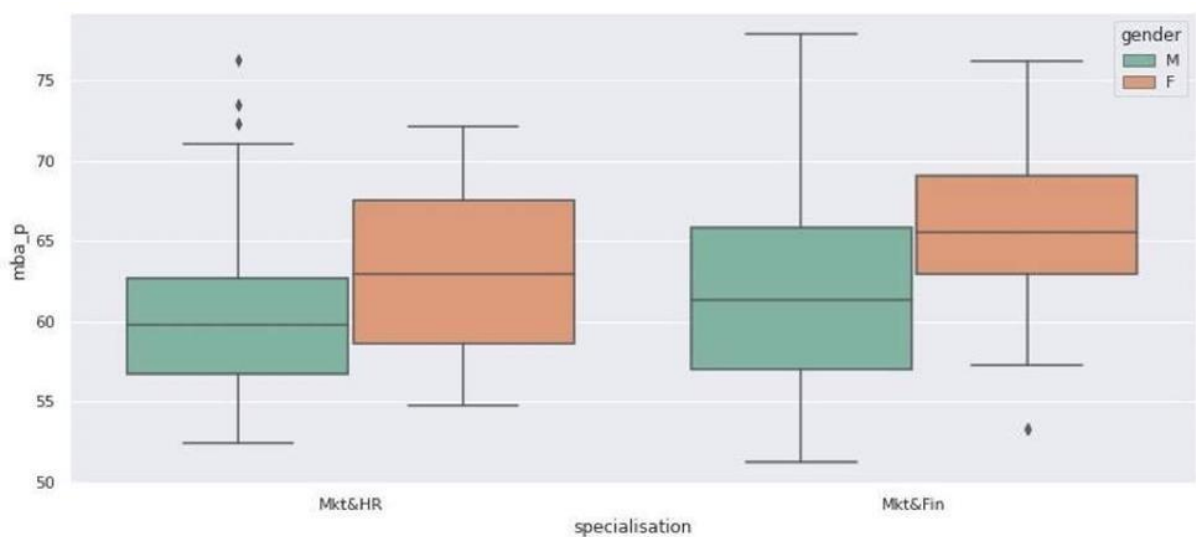
Observations:

- Females of Mkt and Fin are having higher average mba percentages
- Males of Mkt and HR are having lowest average mba percentages.

3.3.7. Users need to mention their work experience.
Abstract Level Task:
- Analysing the work experience trend in Campus Placement
- Dataset type: Table
- Data type: Items and Attributes

XVI
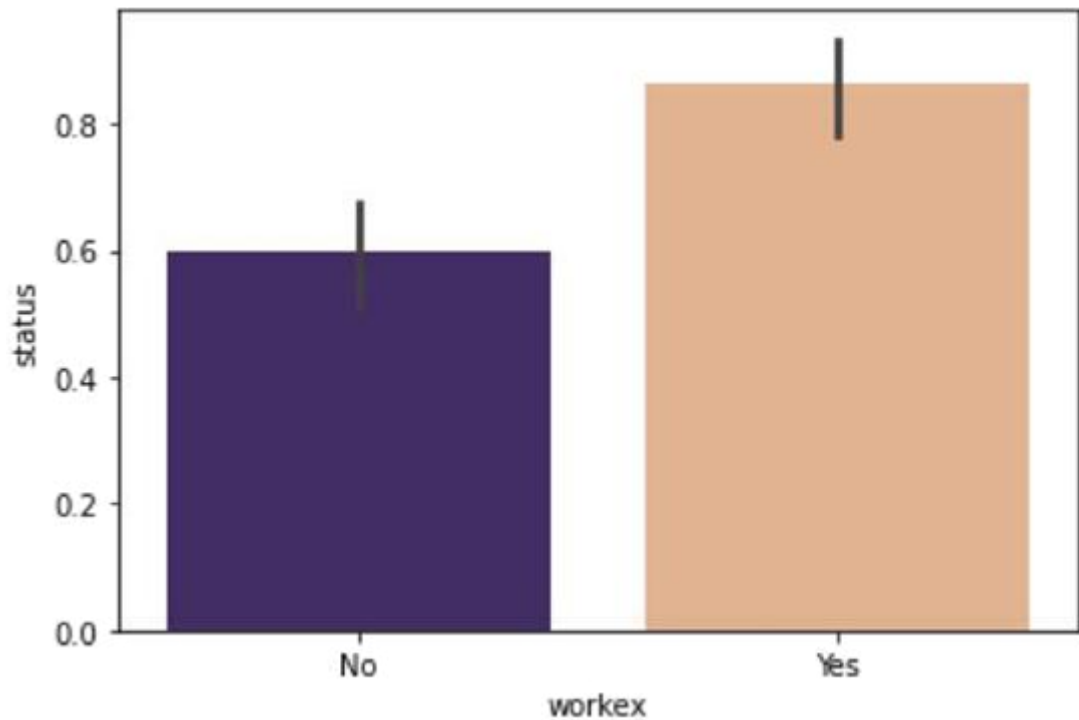
Observations:

- Companies prefer candidates with work experience so the students with internships and past job experience have better chances of being placed.

3.3.8. User needs to specific their MBA percentage

Abstract level Task:

- If I have a high MBA percentage, will I get placed?
- Understanding the importance of MBA placement in college placements
- Dataset type: Table
- Data type: Items and Attributes

XVII
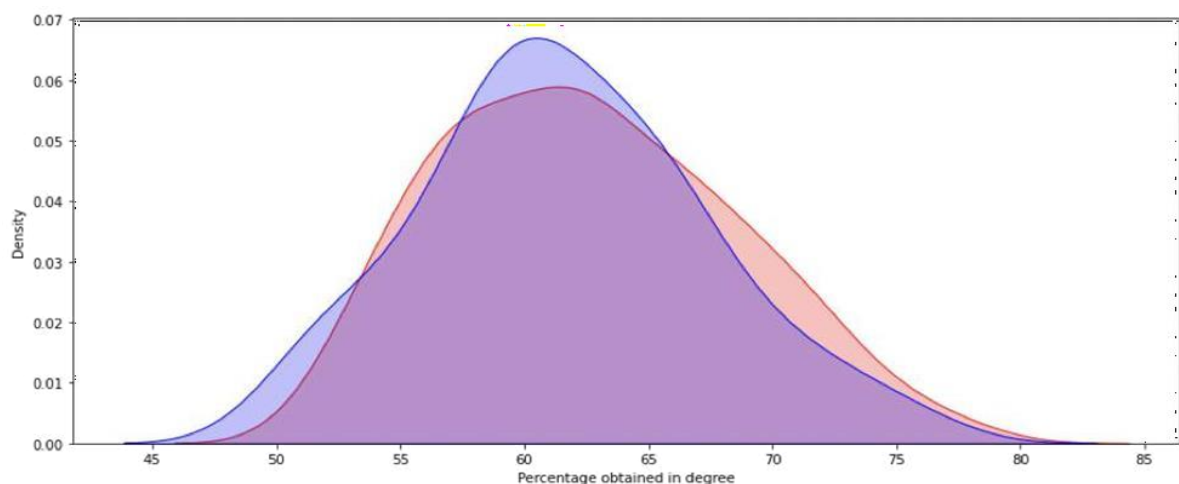
Observations:

- We can see that getting good percentages in MBA does not guarantee placement of the candidate.

3.3.9. Salary Analysis:

Abstract level Task:

- Find the package received by maximum number of students
- Analysis of salary
  - based on specialisation
  -based on work experience
  -based on Board in 10th grade
- Dataset type: Table
- Data type: Items and Attributes

Salary Count Distribution

Salary vs Gender based on specialisation
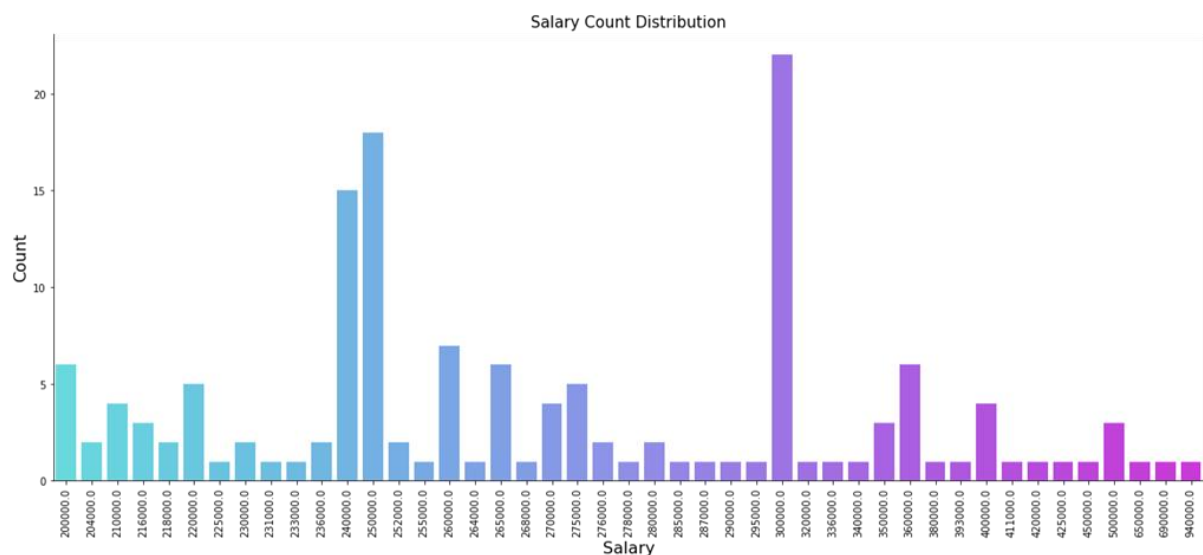
Salary vs Gender based on specialization
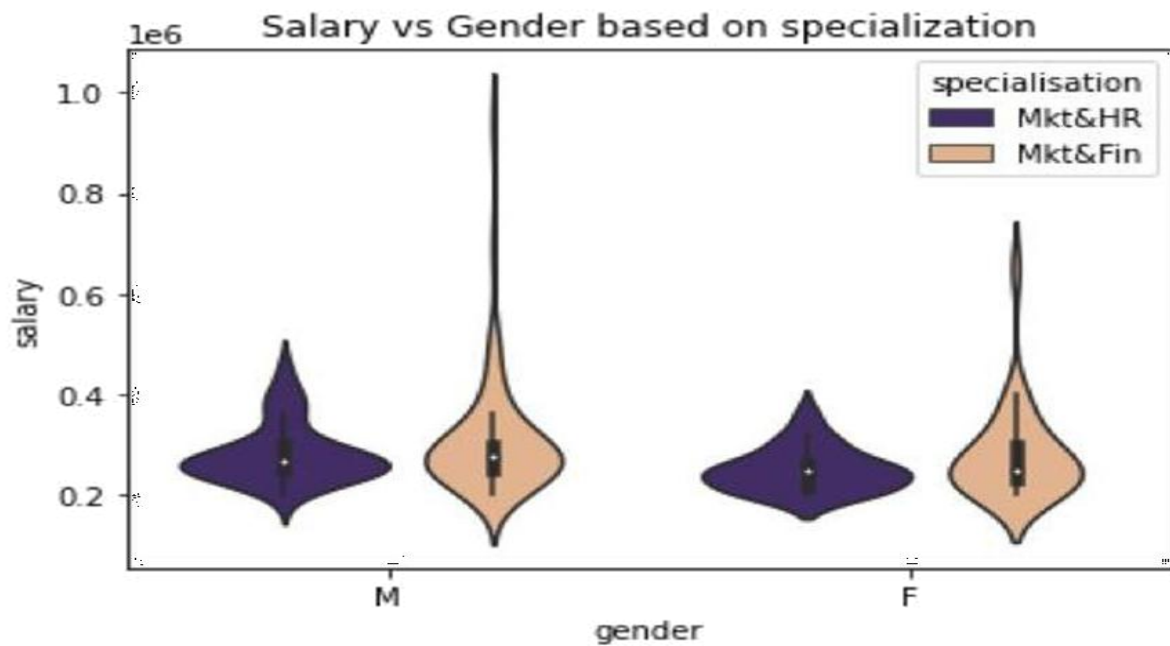
Observations:
- Salary column for male candidates seems to have more outliers than females which means that a lot more male candidates got more than the averageCTC.
- Mean salary is somewhere around 220k.
- Mkt&Fin students are given higher salaries as compared to Mkt&HR.

Gender vs Salary based on work experience

Observations:
- Work Experience is a clear indicator as more work experience results in higher CTC jobs.

XIX

- The maximum salary in male candidates with experience is >1M and for females it is ~700k. The maximum salary in male candidates without experience is ~550k and for females it is ~430k.

**Salary vs Gender based on Board in 10th grade**
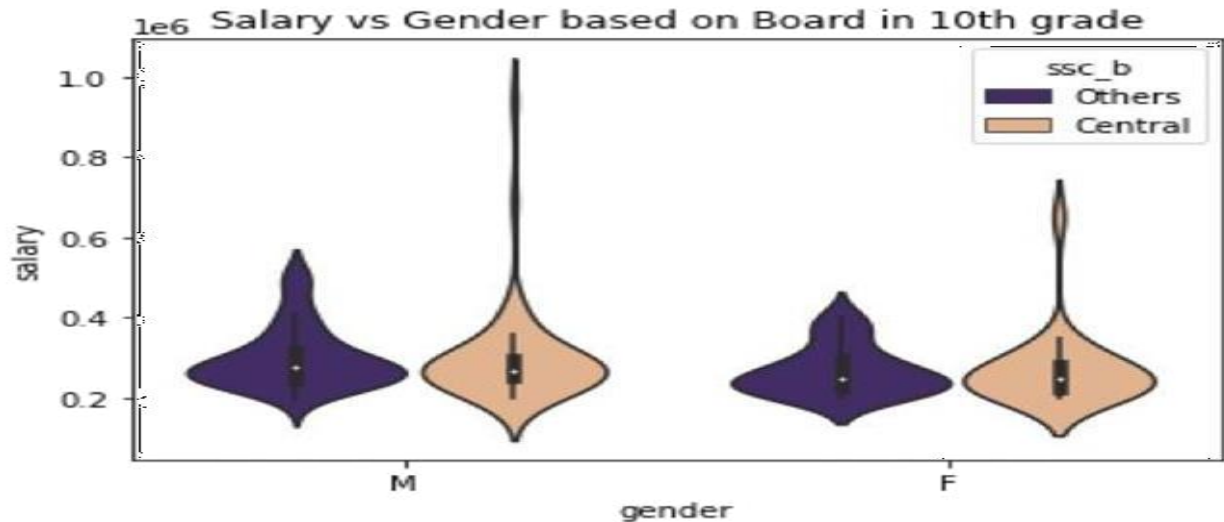
Observations:

- Both Male and Female candidates from Central board got higher CTC as compared to other boards thus we can say that central board in 10th grade might fetch you higher CTCs.
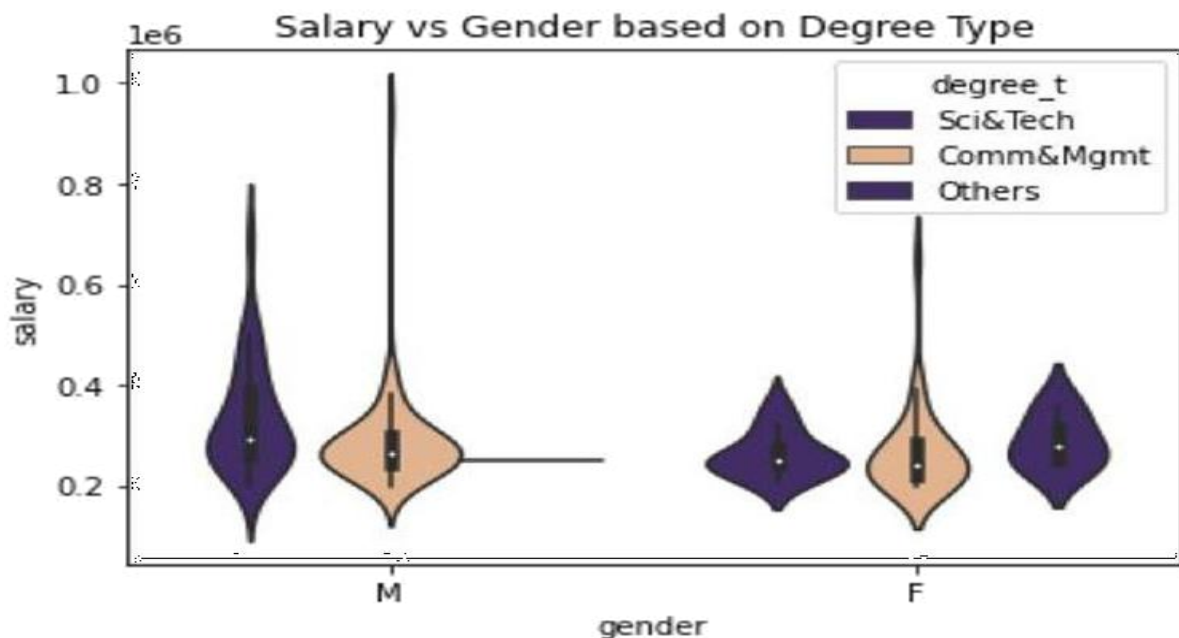
Observations:

- Both male and female candidate got high CTCs choosing Comm&Mgmt as their degree.

- Male candidates from Sci&Tech got high CTCs as compared to Female candidates.
- None of the male candidates got placed from "Others" category whereas for female candidates the package is close to what female Sci&Tech candidates got.

Salary v/s etest_p

Salary v/s Percentage in MBA
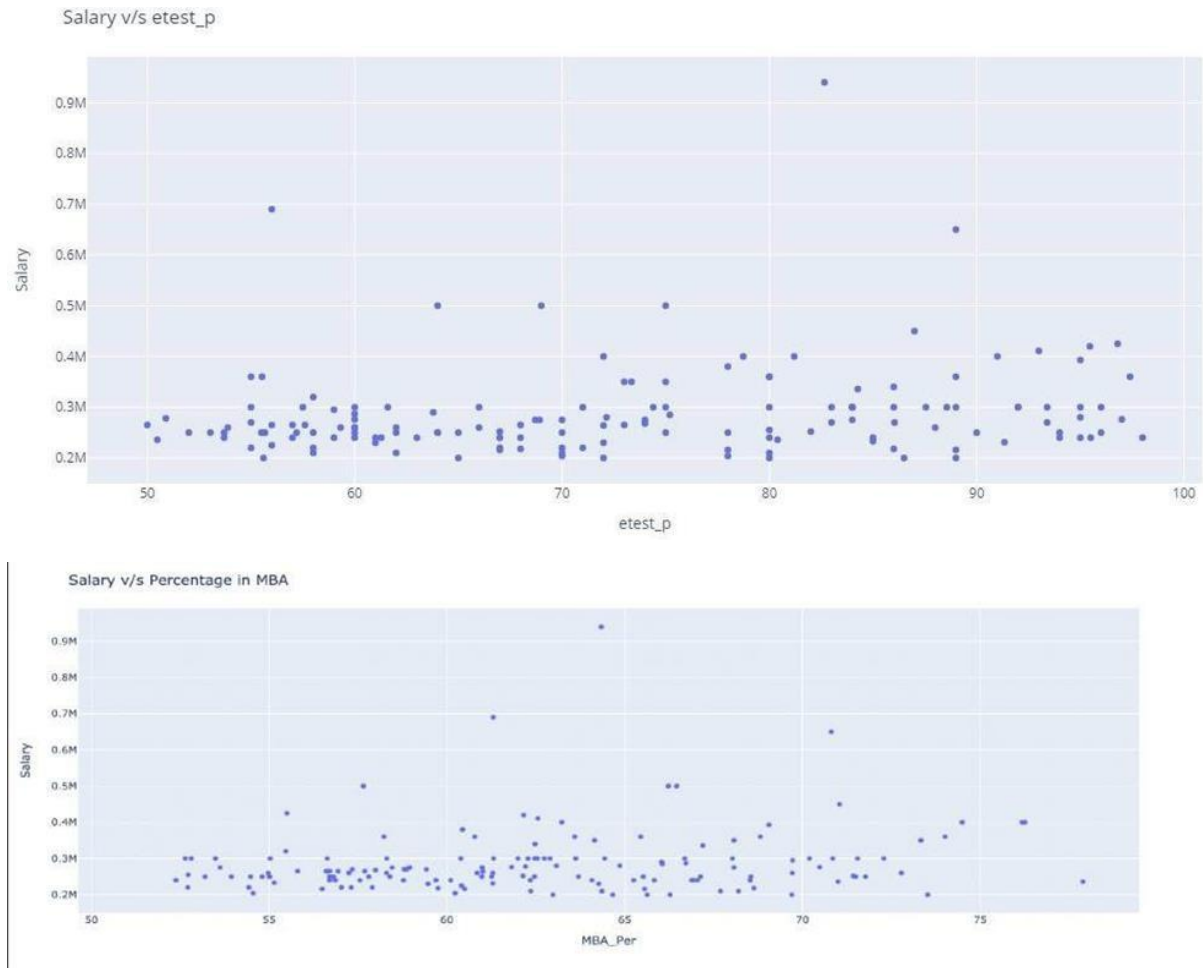
3.3.10. Dataset analysis – Correlation
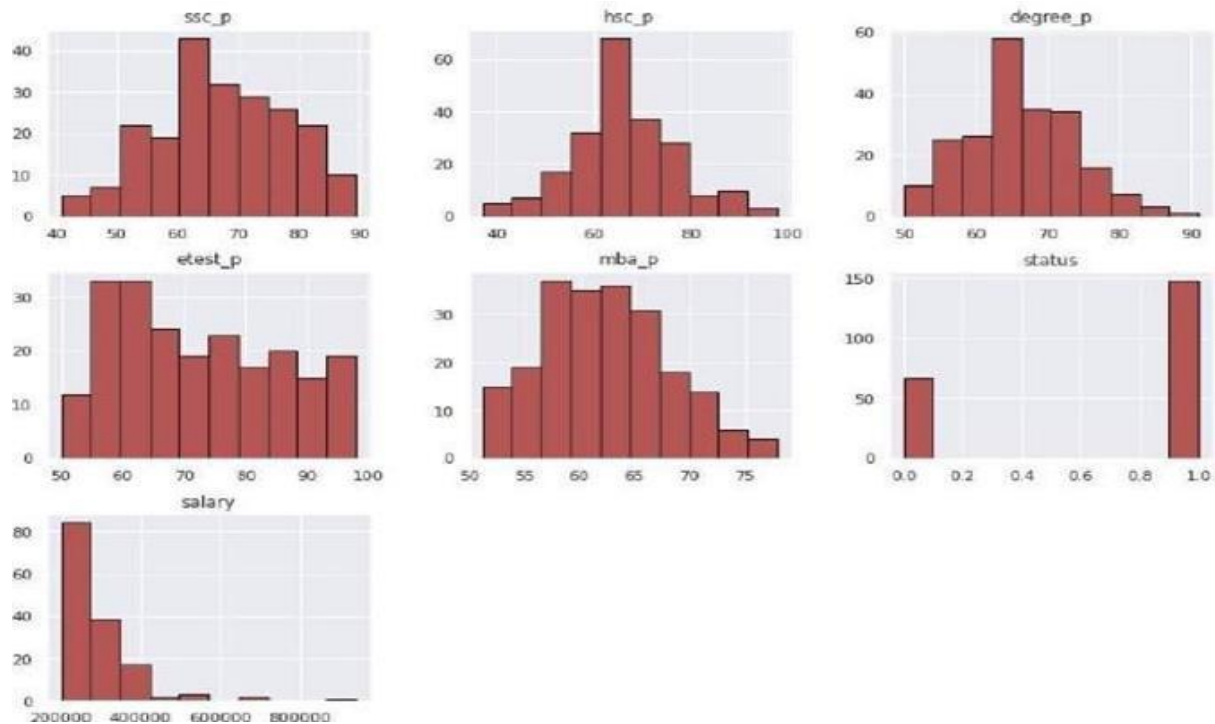
Abstract level Task:

- Relation between numerical values in data set
- Overall impact of mentioned attributes in College Placements
- Dataset type: Table
- Data type: Items and Attributes

Correlation values
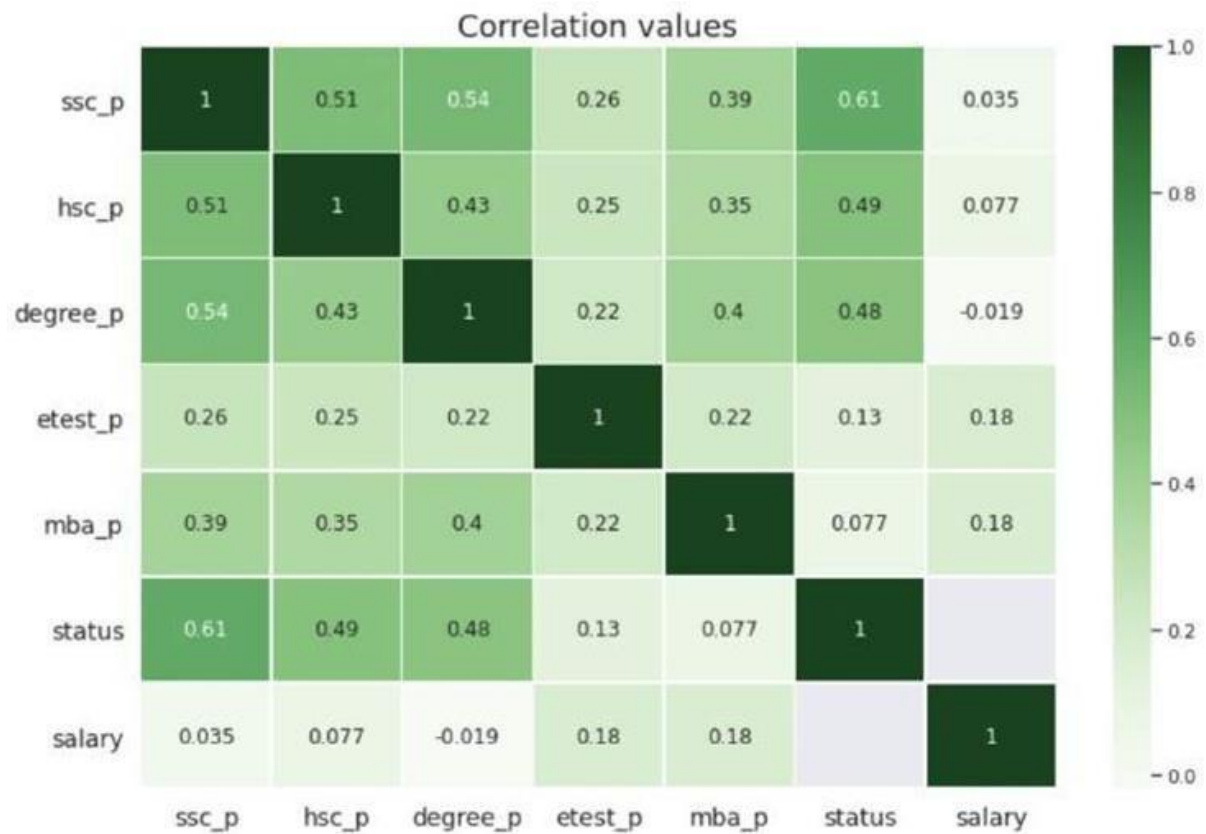
Observations:

- The ssc_p,hsc_p,degree_p have higher correlation with status, hence affect the placement procedure more.

Summary (Histogram Distribution)

4. Conclusion

1. More male candidates got placed as compared to female candidates.
2. Male Candidates got higher CTCs as compared to female candidates.
3. Type of Board chosen does not have any effect on placements thus we can drop in pre-processing steps.
4. Most of the students preferred the Central board in 10th grade whereas other boards in 12th grade.
5. Candidates with higher percentages have better chances of placements.
6. Choosing Science and Commerce as Specialisation seems to have perks when it comes to placements.
7. Maximum package was bagged by male candidate from Mkt&Fin branch which is around 940k.
8. Commerce is the most popular branch among candidates.
9. Mean CTC is around 220k for male and female candidates individually.
10. Choosing Sci & Tech and Comm Mgmt as degrees will fetch you higher CTCs.
11. Mkt&Fin major have higher salaries and more placement chance as compared to Mkt&HR.
12. Employability test percentage and MBA percentage does not affect the placements.
13. Ensemble Modelling gives better accuracy (92%) when predicting the data.

5. References

[1] Shreekumar K. Nair and Sadhana Ghosh, "Factors Affecting the Placement Prospects of MBA Students: An Exploratory Study", 2006 Vision: The journal of business perspective, 2006, doi: 10.1177/097226290601000104

https://journals.sagepub.com/doi/abs/10.1177/097226290601000104

[2] Varun Shenoy and P.S Aithal, "ABCD Analysis of On-Line Campus Placement Model", 2016 International Journal of Management and Social Sciences, 2016, *(ISSN 2455-2267)., Vol. 5, No. 2, pp. 227-244. DOI:10.21013/jmss.v5.n2.p3*

https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2875403

[3] Meenakshi Dhingra and Subhash C Kundu, "Factors affecting placement and hiring decisions: A study of students' perceptions", 2020 Industry and Higher Education, 2020, doi: 10.1177/0950422220950191

https://journals.sagepub.com/doi/abs/10.1177/0950422220950191

[4] Varun Shenoy and P.S Aithal, "Campus placements in Kerala-An empirical study at the selected Engineering Colleges in Kerala", 2016 International Journal of Scientific Research and Publication, 2013, *(ISSN 2250-3153).,* Volume 3, Issue 1, January 2013

https://www.academia.edu/27086150/Campus_placements_in_Kerala-An_empirical_study_at_the_selected_Engineering_Colleges_in_Kerala

[5] Dataset taken from: https://www.kaggle.com/benroshan/factors-affecting-campus-placement