Analysis of Campus Recruitment Data

By Manthan Bagade

This is my first project in the field of Data Science. I want to learn and apply my learnings on real life datasets. I am starting my data analysis with the help of Python. Python has extensive libraries for data analysis, and which will make my analysis presentable. I have used the dataset available on **Kaggle** with the name **“Campus Recruitment”** by “**Ben Roshan”**. I acknowledge the efforts of the creator of this dataset who collected this data. It is of great help for students like me. Lastly, I would also like to thank Pushpendra Singh for his collaboration.

# Objective

The objective of this project is to determine and analyze the **Academic and Employability factors** influencing college placements. The project will present relevant statistics and graphs which will be sufficient to draw out comprehensible conclusions for this dataset. I also aim to answer these questions based on my analysis of the data.

1. Which factor influenced a candidate in getting placed?
2. Do percentage matters for one to get placed?
3. Which degree specialization is much demanded by corporate?
4. Play with the data conducting all statistical tests.

Not to forget, the final aim of this project is to learn and apply the skills which I have developed while starting out my journey to master Data Science.

“The best way to learn Data Science is to do Data Science”

# Dataset Source and Details

The datasets are downloaded from Kaggle. The name of the dataset is **“Campus Recruitment”** by “**Ben Roshan”**. The link to the dataset is attached for the reference.

<https://www.kaggle.com/datasets/benroshan/factors-affecting-campus-placement/>

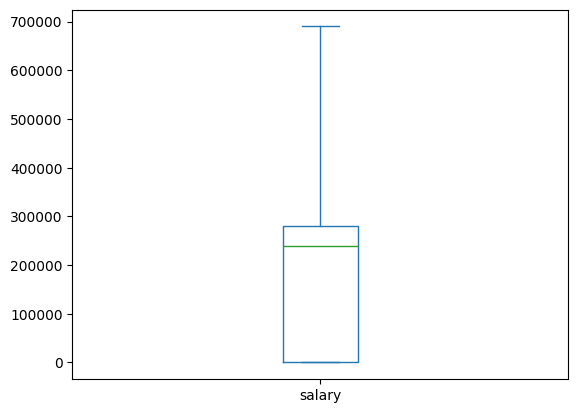
The Dataset is a csv file. The file contains a total of 215 entries and 15 columns.

| Column Name | Datatype | Description |
| --- | --- | --- |
| sl\_no | int | Serial Number |
| gender | categorical(char) | Gender - Male=‘M’, Female=‘F’ |
| ssc\_p | float | Secondary Education percentage- 10th Grade |
| ssc\_b | categorical(string) | Board of Education- Central/ Others |
| hsc\_p | float | Higher Secondary Education percentage- 12th Grade |
| hsc\_b | categorical(string) | Board of Education- Central/ Others |
| hsc\_s | categorical(string) | Specialization in Higher Secondary Education |
| degree\_p | float | Degree Percentage |
| degree\_t | categorical(string) | Under Graduation(Degree type)- Field of degree education |
| workex | categorical(string) | Work Experience |
| etest\_p | float | Employability test percentage ( conducted by college) |
| specialisation | categorical(string) | Post Graduation(MBA)- Specialization |
| mba\_p | float | MBA percentage |
| status | categorical(string) | Status of placement- Placed/Not placed |
| salary | int | Salary offered by corporate to candidates |

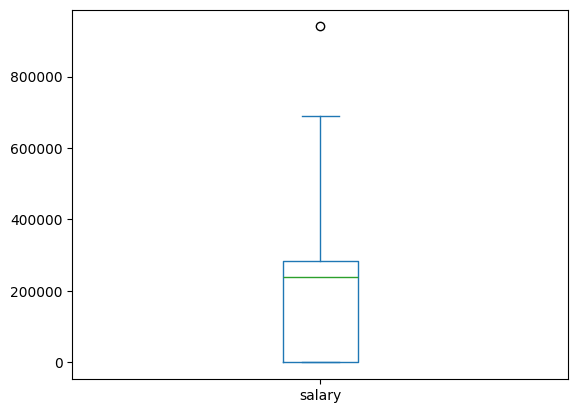
# Dataset Cleaning and Filling Missing Values

All the columns do not have any missing values except the column ‘salary’. It has 67 missing values. I filled all the missing values with 0 since the students who were unplaced did not have any salary.

I also used box plot to identify any outliers using the IQR (Interquartile range) method on the ‘salary’ column. The data was so well collected that I only found one outlier. Dropping the row was the best choice because such high salary as compared to other students cannot be accounted to academic factors. It is a result of personal hard work and can impact the statistics of the data.



Data After Dropping Outlier

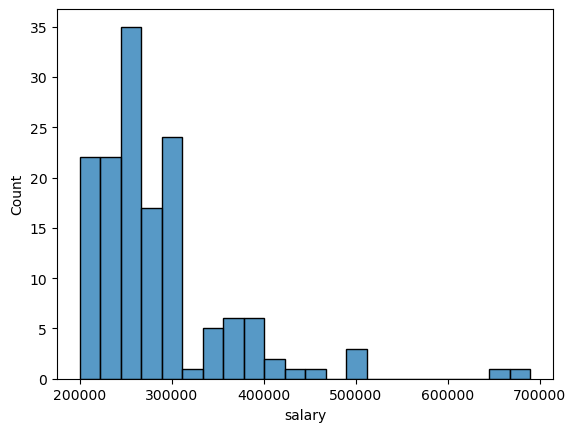


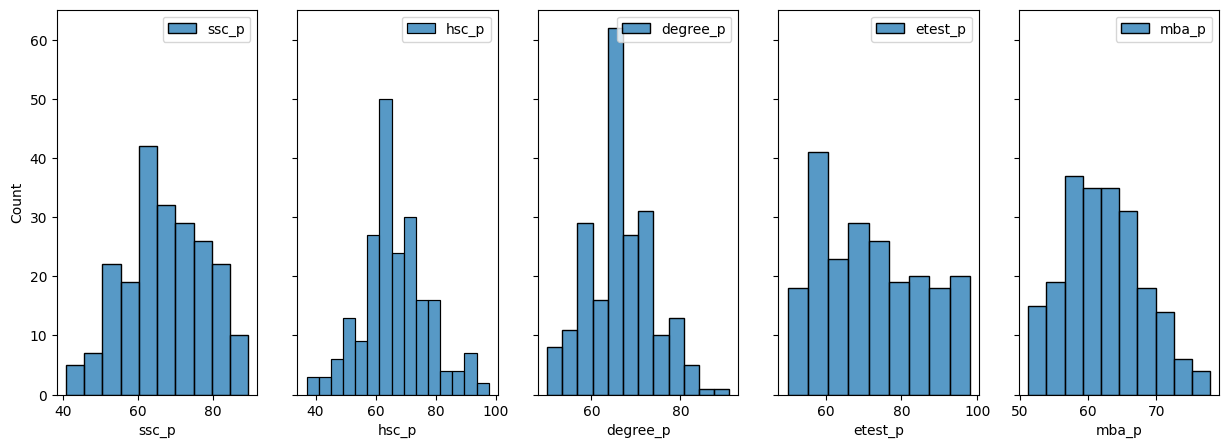
Single Outlier Point

# Data Description

|  | ssc\_p | hsc\_p | degree\_p | etest\_p | mba\_p | salary |
| --- | --- | --- | --- | --- | --- | --- |
| mean | 67.33 | 66.32 | 66.38 | 72.05 | 62.27 | 1,95,238 |
| std | 10.84 | 10.92 | 7.37 | 13.29 | 5.85 | 1,46,553 |
| min | 40.89 | 37.00 | 50.00 | 50.00 | 51.21 | 0 |
| 25% | 60.55 | 60.85 | 61.00 | 60.00 | 57.92 | 0 |
| 50% | 67.00 | 65.00 | 66.00 | 70.50 | 61.95 | 2,40,000 |
| 75% | 75.85 | 73.00 | 72.00 | 83.75 | 66.27 | 2,80,000 |
| Max | 89.40 | 97.70 | 91.00 | 98.00 | 77.89 | 6,90,000 |

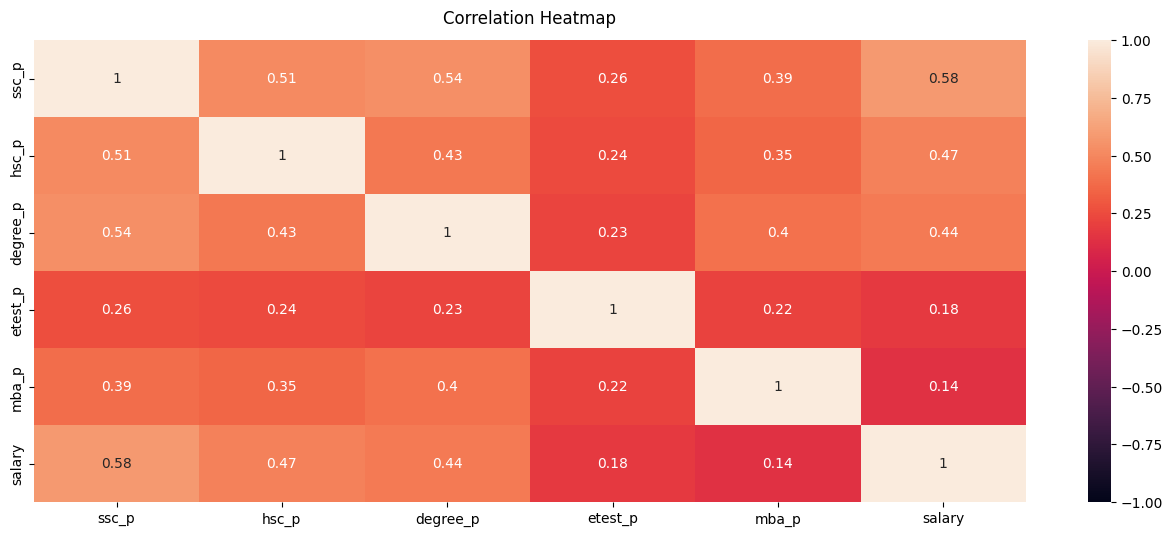
It is clear from the following histogram that maximum number of students had a salary between 2,00,000 to 3,00,000. Thus, we can say that the companies which are coming for the placements are able to pay only that much of the amount. The median salary is 2,40,000 which is verified by histogram.





Histogram of Academic attributes

The other histogram plots show the frequency of the percentages of different examinations.



The result of the Correlation Heatmap is quite surprising. Salary least relates to employability test and MBA percentage. Further, employability test and MBA scores are not related so strongly to any of the other academic factors. The reason for these unexpected results can be the irrelevance of the questions and the syllabus which are asked in tests. The questions asked in employability test could only assess the whether the person is fit for any job rather than evaluating him quantitatively.

Higher Secondary (12th grade) is percentage is quite related to Senior Secondary (10th grade) percentage which is self-explanatory and similarly the degree percentage is related to 10th and 12th grade scores.

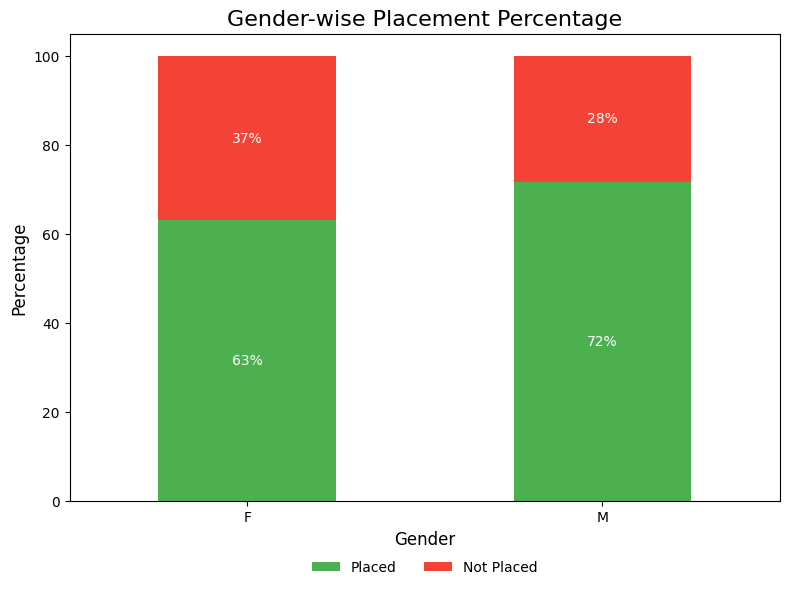
The salary is related to hsc\_p, ssc\_p and degree\_p which points that early development plays a great role in placement season.

# Data Visualization and Interpretation

Let us visualize the answers to the questions in the objectives through charts.

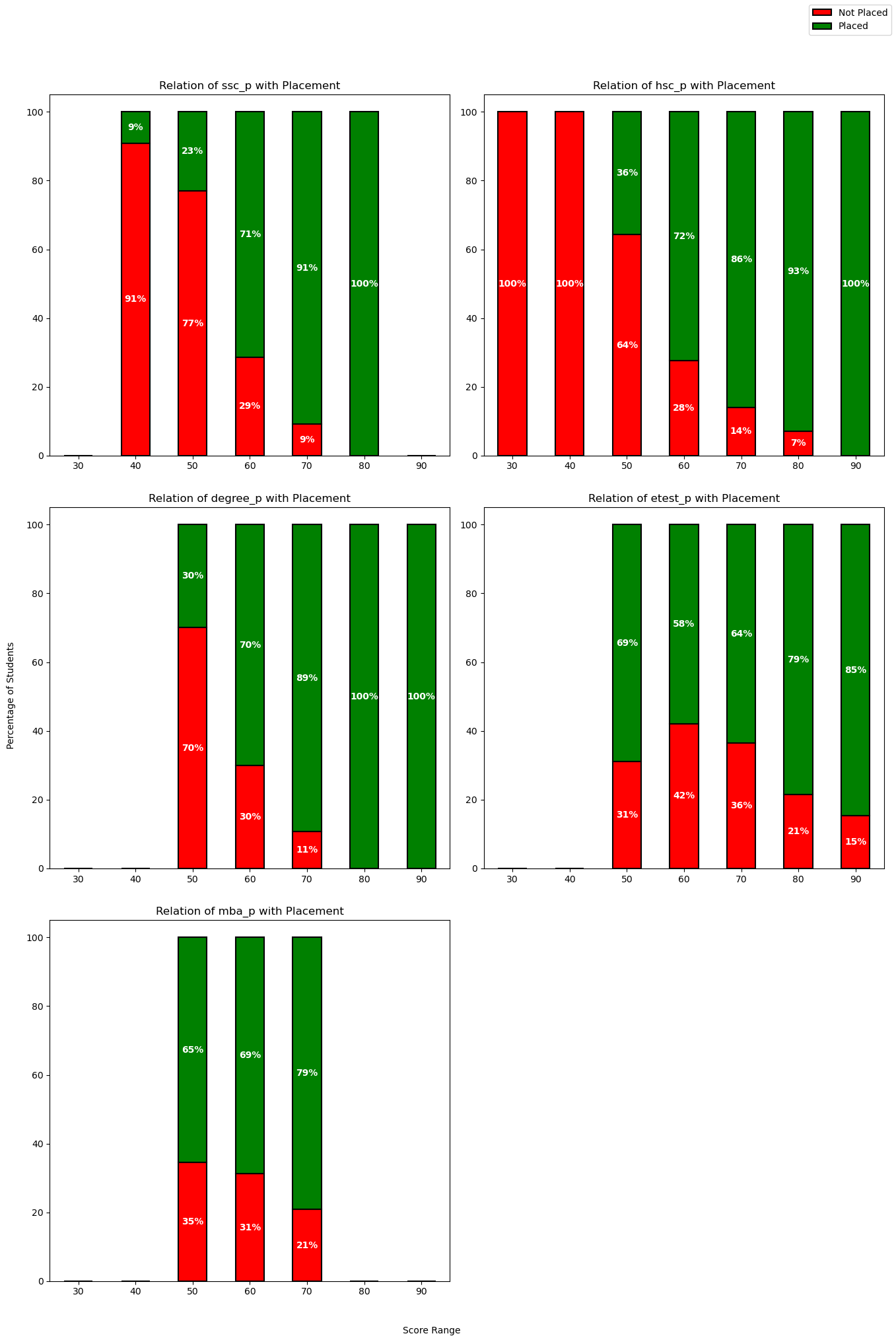
**Gender wise Analysis**

| gender | status | count | mean salary | percentage placed |
| --- | --- | --- | --- | --- |
| F | Not Placed | 28 | 0 | 36.84% |
| F | Placed | 48 | 2,67,292 | 63.16% |
| M | Not Placed | 39 | 0 | 28.26% |
| M | Placed | 99 | 2,92,434 | 71.74% |

****

From the graphs male students have greater percentage of placements than Female students. They also have slightly higher salary than their counterparts. Generally, male students are more flexible, and this can be the reason they are paid higher average salary.

**Relation between ssc\_p, hsc\_p and salary**



Stacked bar graphs reflect a positive trend between test/academic and placement performance. For all the five measures—SSC% (Secondary School Certificate), HSC% (Higher Secondary Certificate), Degree%, E-test%, and MBA%—students who obtain higher marks are far more likely to get placed. The trend is similar, with lower score ranges (30–50%) having a larger number of unplaced students, and upper score ranges (70–100%) with placed candidates being in majority. Companies are screening on academic performance as a parameter, seeking students with a good academic background. Students with good academic performance consistently at different levels of studies may also have skills such as discipline, good study habits, and command over the subject, and thus be more employable.

While the trend is uniform for all the parameters, that of growth in the percentages of placements is marginally different. Some of the parameters may be directly influencing placement opportunities in a different way than others, and that may imply that recruiters are assigning more importance to some academic phases (such as degree performance) compared to others. But let's not forget, correlation is not causation. Communication skills, internships, and co-curricular activities may be contributing significantly to placements too.

Despite these subtleties, the implication becomes clear: higher grades have a direct correlation with better placement. Another incentive to keep high grades throughout one's academic life to gain the rewards of one's professional prowess.

**Undergraduate Degree and Specialization Relation with Salary**

| degree\_t | mean salary |
| --- | --- |
| Comm&Mgmt | 1,90,833 |
| Others | 1,27,454 |
| Sci&Tech | 2,18,627 |

* Students with **Science and Tech** background have higher salaries because they have large skill set. They have pretty good knowledge about technical stuff and their specialization in marketing and finance/HR gives them a better edge. This makes them an ideal candidate for companies.
* Those from a **Commerce & Management** background already have business exposure, making their master's degree less of a specialization leap. They likely get placed in general business roles (marketing, sales, or finance) with moderate starting salaries. These roles may not require advanced analytical or technical skills, leading to slightly lower pay than Sci&Tech graduates.
* The **"Others"** category likely includes students from arts, humanities, or other general disciplines, who may not have had strong exposure to either business or technical skills. Even with a master's in business, they might struggle to compete with Science & Tech or Commerce & Management graduates in the job market. They may get placed in supportive, non-specialized roles with relatively lower pay.

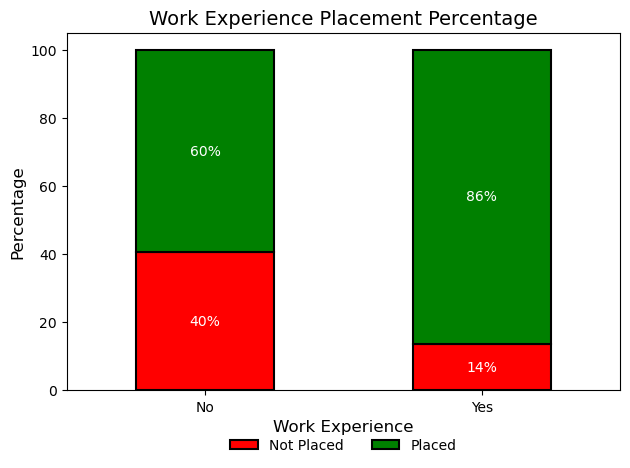
| degree**\_**t | Mkt&Fin | Mkt&HR |
| --- | --- | --- |
| Comm&Mgmt | 220976 | 147407 |
| Others | 150000 | 114571 |
| Sci&Tech | 268933 | 166586 |

Sci&Tech graduates are paid highest in marketing–finance (Mkt&Fin) and marketing–HR (Mkt&HR) jobs. It shows that the technical and analytical ability set with which a Sci&Tech degree equips the graduate is highly prized in those jobs where a finance interface is present.

Moreover, mean salaries for Mkt&Fin jobs are greater than for Mkt&HR jobs at all levels. This may be due to higher financial risk and higher analytical demands traditionally associated with finance-type jobs. This can be seen as evidence that the market is paying a premium for technical skill in situations where financial skill is most important.

Overall, these results provide strategic advice to different stakeholders. To graduates, and especially Sci&Tech graduates, these results provide evidence of the profitability of seeking career prospects in marketing–finance positions. To employers, it suggests the need to utilize technical capability to provide financial performance, which can include restructuring remuneration packages as well as recruitment strategies for employees.

**Relation Of Work Experience With Placements**



We can observe from this graph that there is a distinct difference in placement rate depending on whether the candidate has or has not had work experience. Of the candidates with no work experience, 60% were placed and 40% were not placed. However, the candidates with work experience had an 86% placement rate, with a mere 14% not being placed. This shows that work experience greatly contributes to the probability of obtaining placement. It is a good sign that work experience is a good thing to possess when obtaining placement.

**Answers To Some Questions**

1. Which factor influenced a candidate in getting placed?

Several factors come into play when a candidate is placed, but the analysis identifies that **test performances (higher secondary, secondary, undergraduate, and MBA percentages), the quality of one's bachelor's degree, specialization, and work experience** are more significant. More specifically, better performance on these levels is strongly related to placement success, and the additional increment of work experience further increases a candidate's chances of placement.

1. Do percentage matters for one to get placed?

Yes, academic percentages do matter a lot. The figures show a good trend: higher percentage students at various academic levels—such as SSC, HSC, and degree exams—have much better placement rates. This trend is also supported by the fact that lower score bands have a higher percentage of unplaced candidates and higher score bands have predominantly placed candidates, highlighting the point that maintaining high academic percentages might be crucial to secure a placement.

1. Which degree specialization is much demanded by corporate?

The analysis discovers that Sci&Tech graduates are in great demand by corporates. Not only do they command higher mean salaries in marketing–finance and marketing–HR functions, but they also seem to have an advantage grounded on technical and analytical capability. This implies that a Sci&Tech major is more sought after than Commerce & Management or other undergraduate streams. The analysis also points to the fact that students who have post graduated in Mkt&Fin tend to have better salaries. Overall, we can say that the most sought degree and specialization is Sci&Tech with specialization in Mkt&Fin.