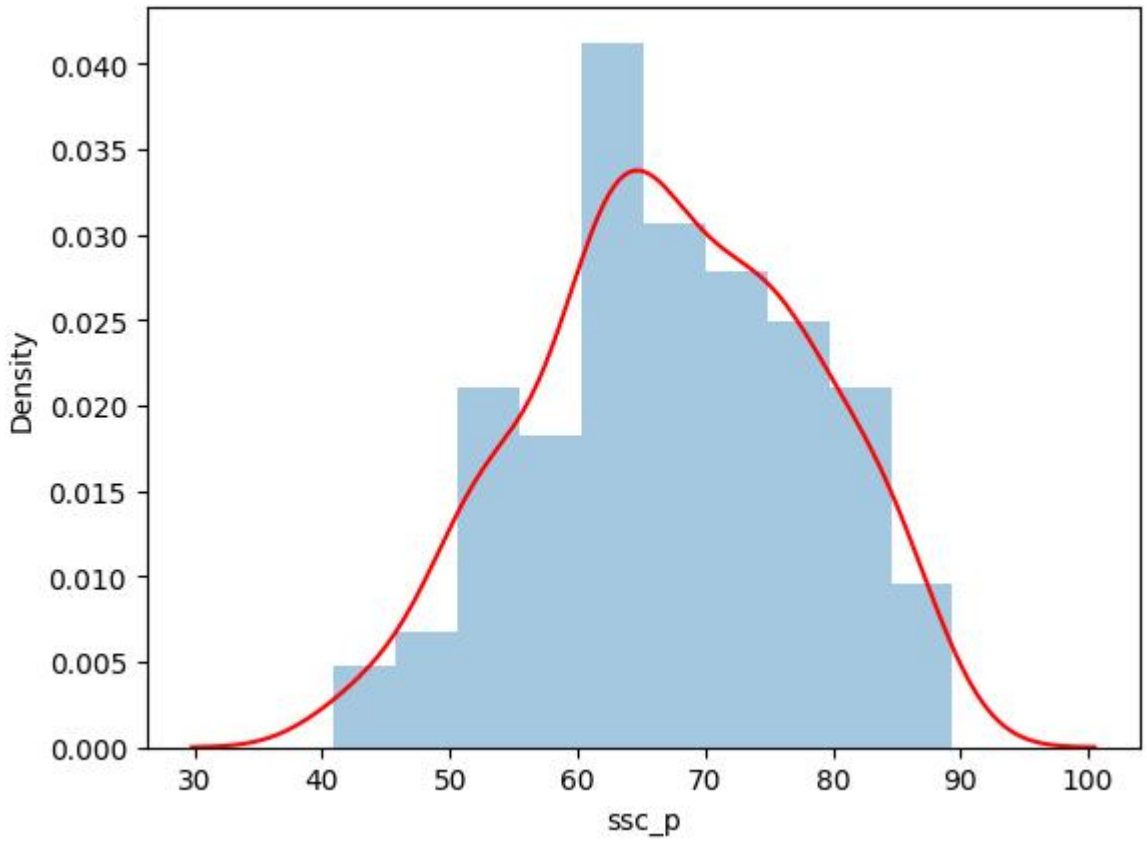
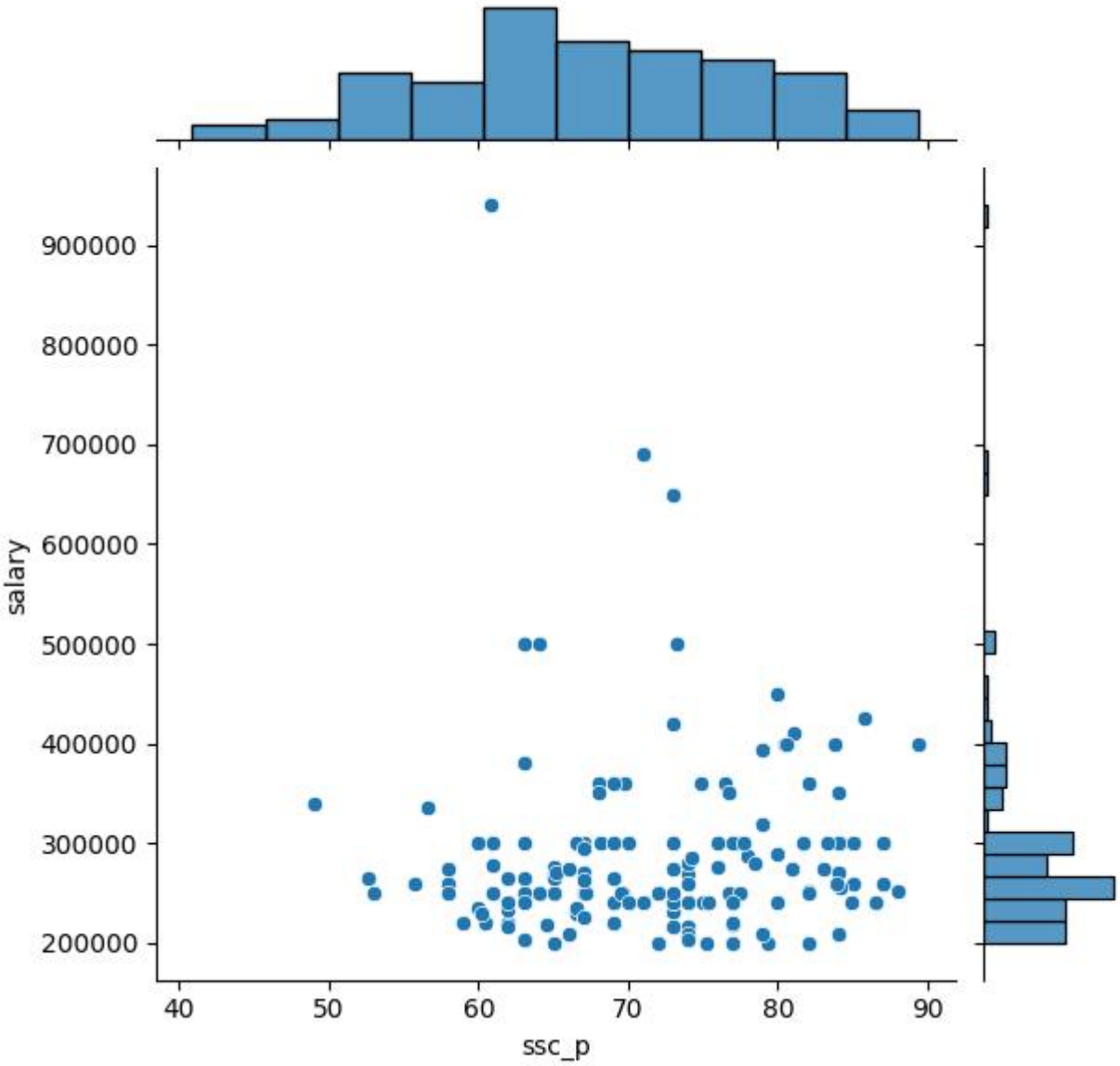
GRAPHS AND EXPLANATION

1) UNIVARIATE ANALYSIS- Dist plot



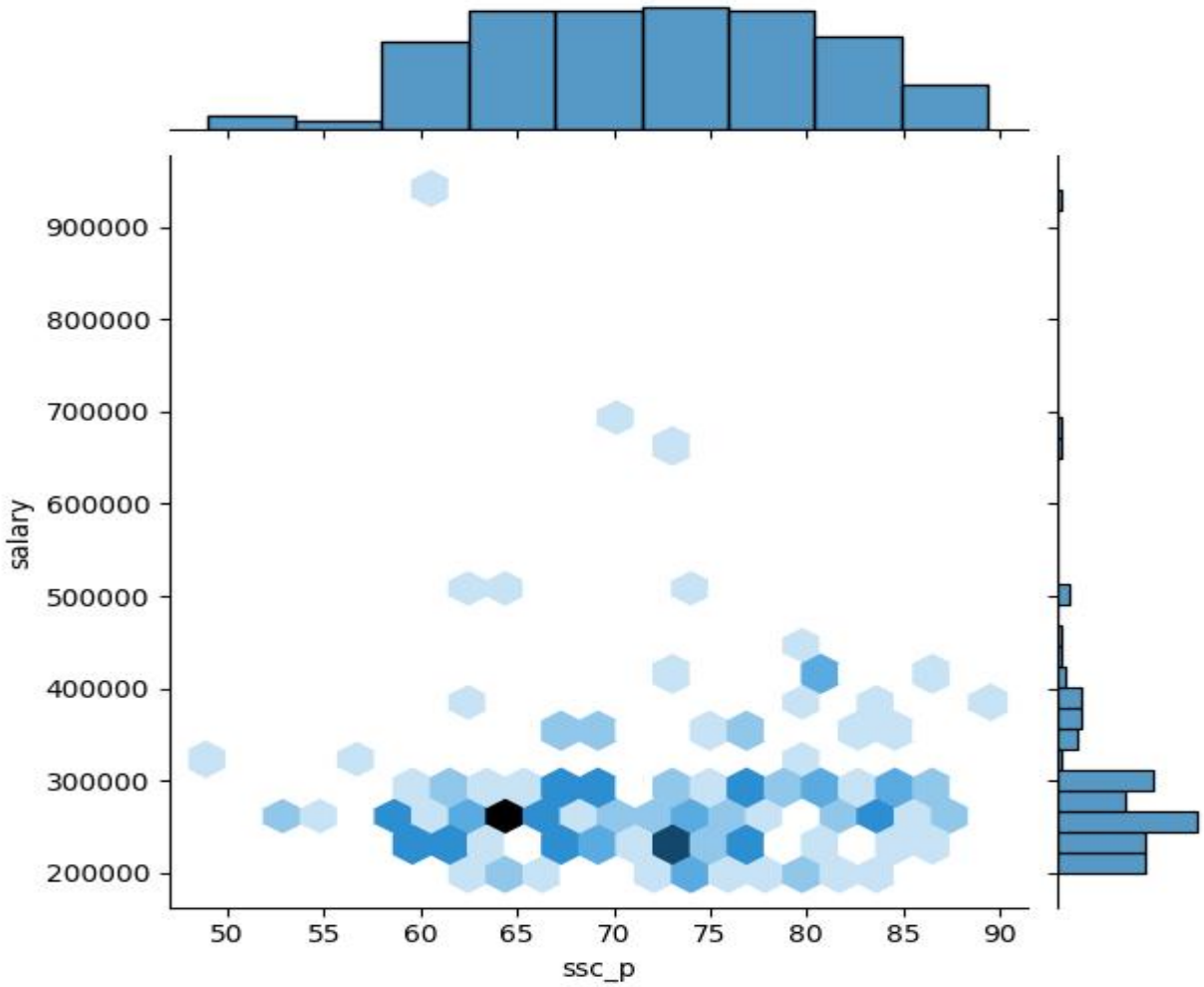
- Dist plot (or hist plot + KDE in modern Seaborn) helps to analyze the distribution of SSC percentages. It gives the shape, spread, and central tendency of the data,Concentration of marks.
- Most students score between 50% to 80% ranges
- Students with very high or very low SSC marks are rare.
- Presence of outliers-A dist plot visually shows there is no Very low SSC scores (outliers) and Very high SSC scores .

2)BIVARIATE ANALYSIS- Joint plot



- A joint plot between ssc_p and salary shows the relationship between a student's Secondary School Percentage (SSC%) and their salary after placement.
- It indicates that SSC percentage does not strongly influence salary. Salary is mostly determined by other factors (like MBA %,specialization, work experience)."

3)BIVARIATE ANALYSIS-HEAT MAP

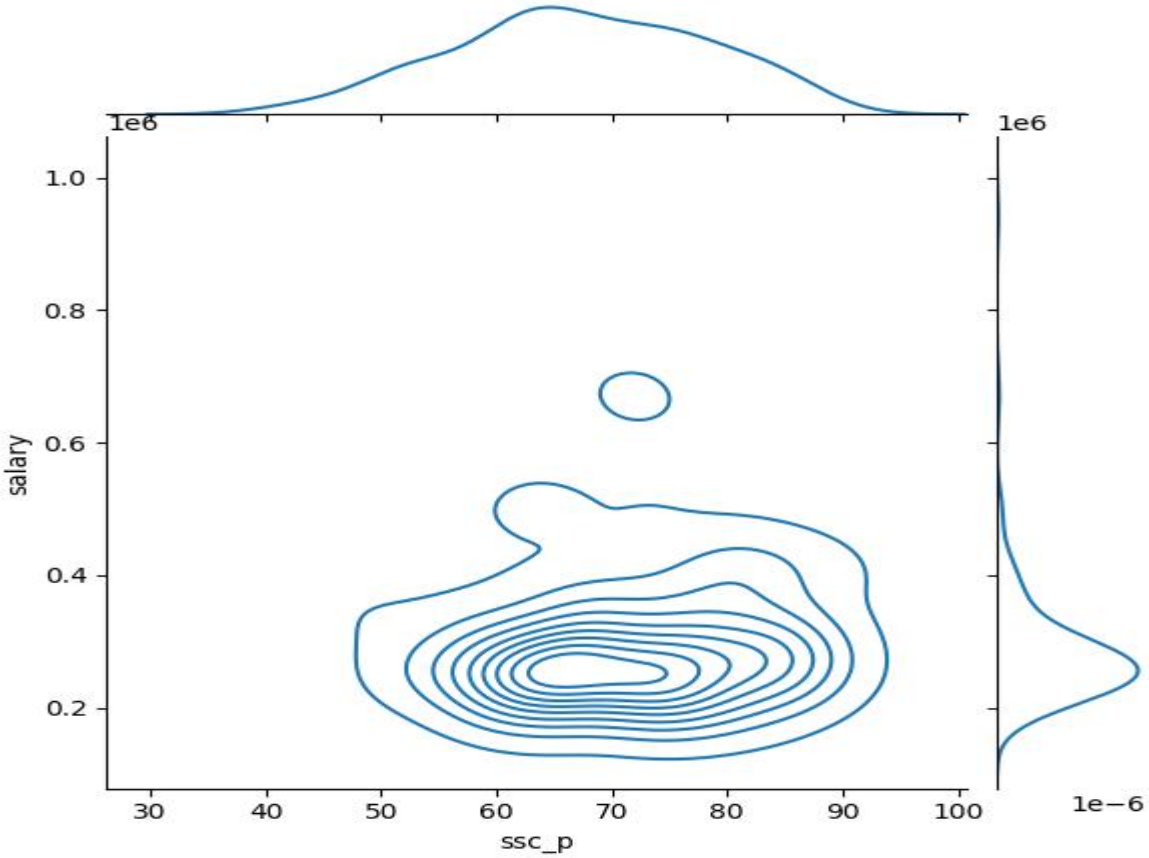


- A heat map between ssc_p and salary is used to show the strength and direction of correlation between the two variables.
- The heat map uses colors:

  Darker colors = stronger correlation

  Lighter colors = weaker correlation
- Here correlation of students who scored 65% of marks and getting around 2.5L salary is strong.and also students who scored 70-75%of marks and getting 2L-2.5L salary is strong, other ssc_p and salary correlations are weak.

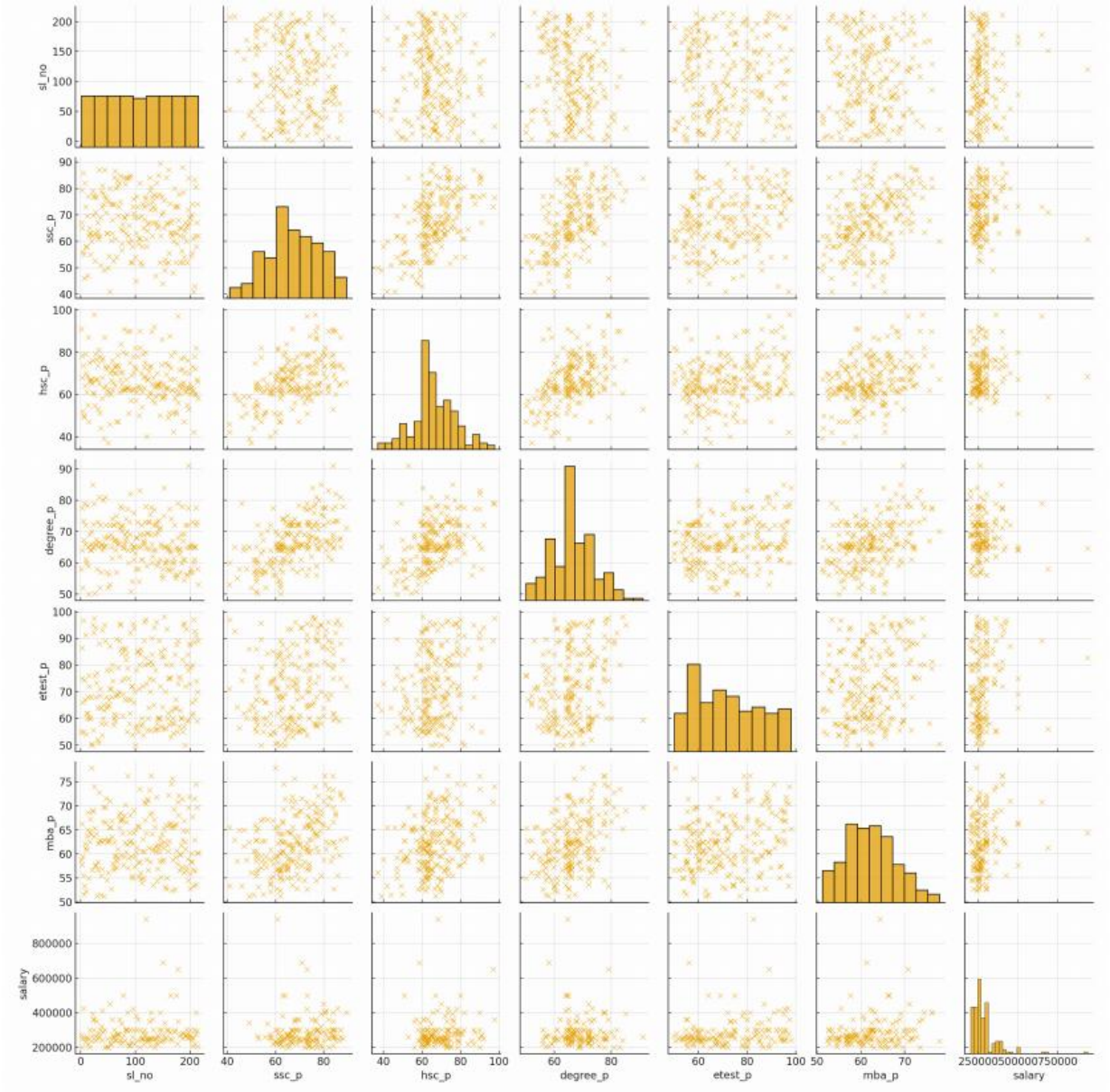
4) BIVARIATE ANALYSIS - Kernel Density Estimate plot



- It is used to show the **density** (concentration) of data points — like a smooth version of a histogram or 2-D heat map.
- It shows where most data points are concentrated. In the placement data set, the KDE map of SSC% and salary shows the density of students in different score and salary ranges, helping identify patterns or clusters." It shows
- Most SSC% lies between 60-90% and Salary lies between 2-3L
- Whether high SSC% students tend to get higher salaries -NO
- It is better than scatter plot when Graph have many overlapping points and want a smooth pattern instead of raw points

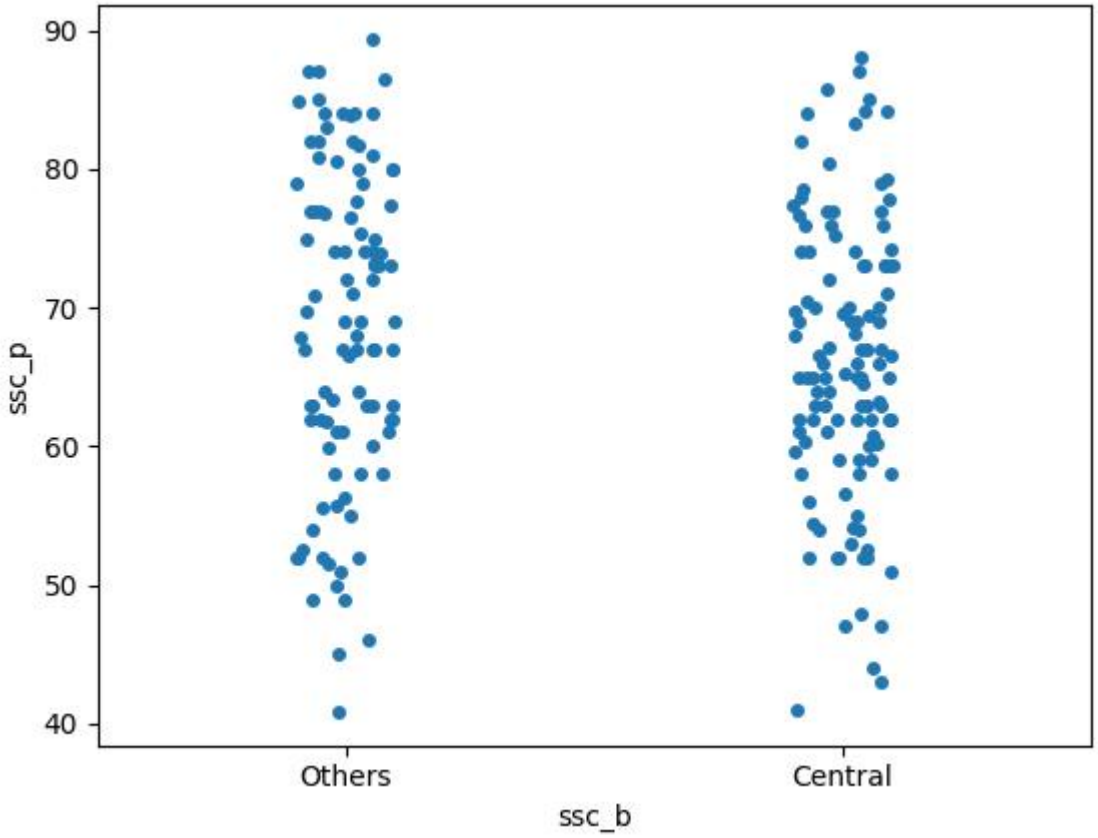
5) BIVARIATE ANALYSIS -Pair plot

- A pair plot is useful for exploratory data analysis because it visualizes the pairwise relationships between multiple variables in a datasets.
- It creates a grid of scatter plots for every pair of numerical variables, showing their joint distributions, and includes histograms or density plots on the diagonal to display the distribution of each individual variable. This allows to quickly identify patterns, trends, correlations, and outliers across different features, which is crucial for gaining insights and making informed decisions early in the analysis.



6) CATEGORICAL PLOT- Strip plot

- A strip plot is useful for visualizing the distribution of individual data points, especially for small to moderate datasets, because it clearly shows the location of every point, making it easy to identify gaps, compare distributions across categories, and spot outliers. It is particularly effective for showing the relationship between a categorical and a continuous variable.



- Both boards have a wide range of SSC percentages
- Students from both boards score between low 40% to high 90%.
- Neither board consistently scores higher than the other.
- Slight spread differences -Central board shows slightly more spread between 60%-70% marks than other board, other difference is not major.

7) CATEGORICAL PLOT- Swarm plot

- A swarm plot is used in data visualization to display the distribution of a continuous variable across different categories, with the key advantage of preventing overlapping data points. This non-overlapping arrangement provides a clear and accurate representation of the data density and individual values within each category.

Key reasons for using a swarmplot:

- **Visualizing individual data points:**

  Unlike box plots or violin plots that summarize distributions, swarm plots show every single data point, making them ideal for understanding the granular structure of the data.

- **Revealing data density and clustering:**

  By arranging points without overlap, swarm plots effectively illustrate where data points are concentrated, revealing patterns, clusters, and the shape of the distribution within each category.

- **Comparing distributions across categories:**

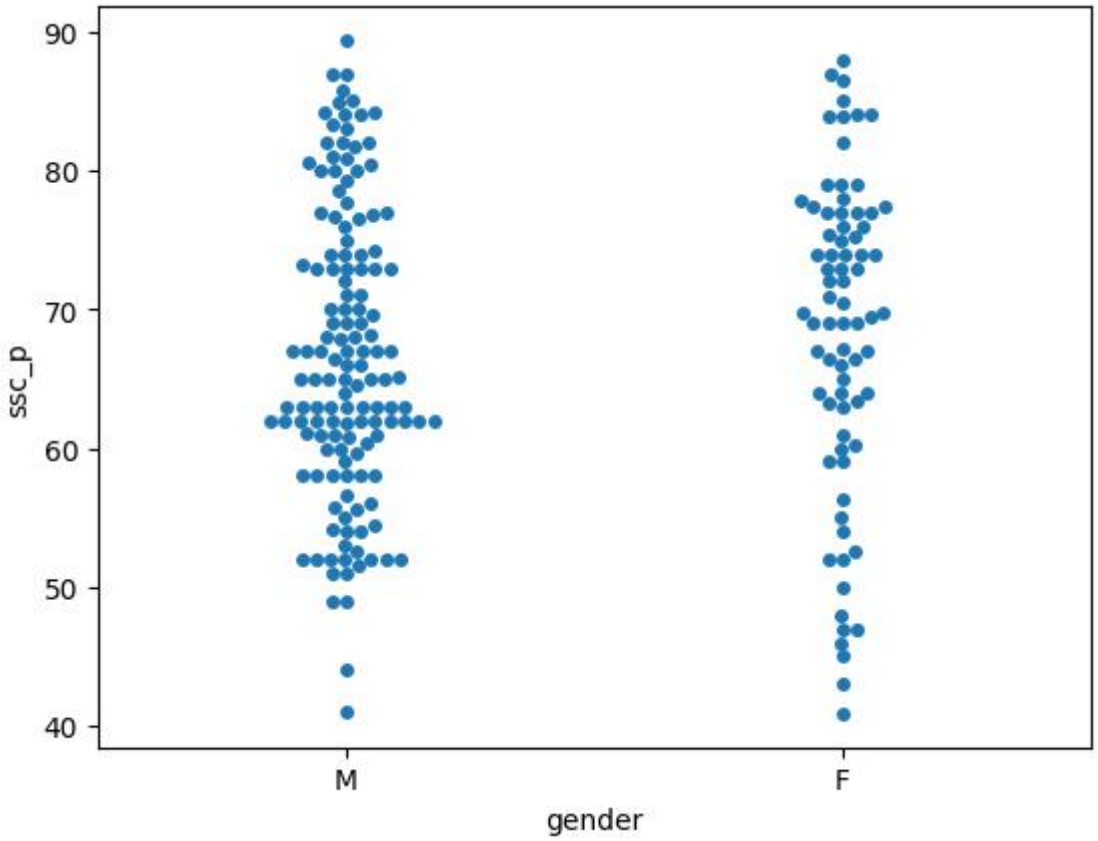
  They allow for a direct visual comparison of the distribution of a continuous variable across different categorical groups.

- **Complementing other plots:**

  Swarm plots can be effectively combined with other statistical summary plots like box plots or violin plots to provide both a detailed view of individual points and an overall summary of the distribution.

- **Understanding the spread of values:**

  The non-overlapping nature helps in understanding the spread and range of values within each category, particularly for small to medium-sized datasets where individual observations are important.

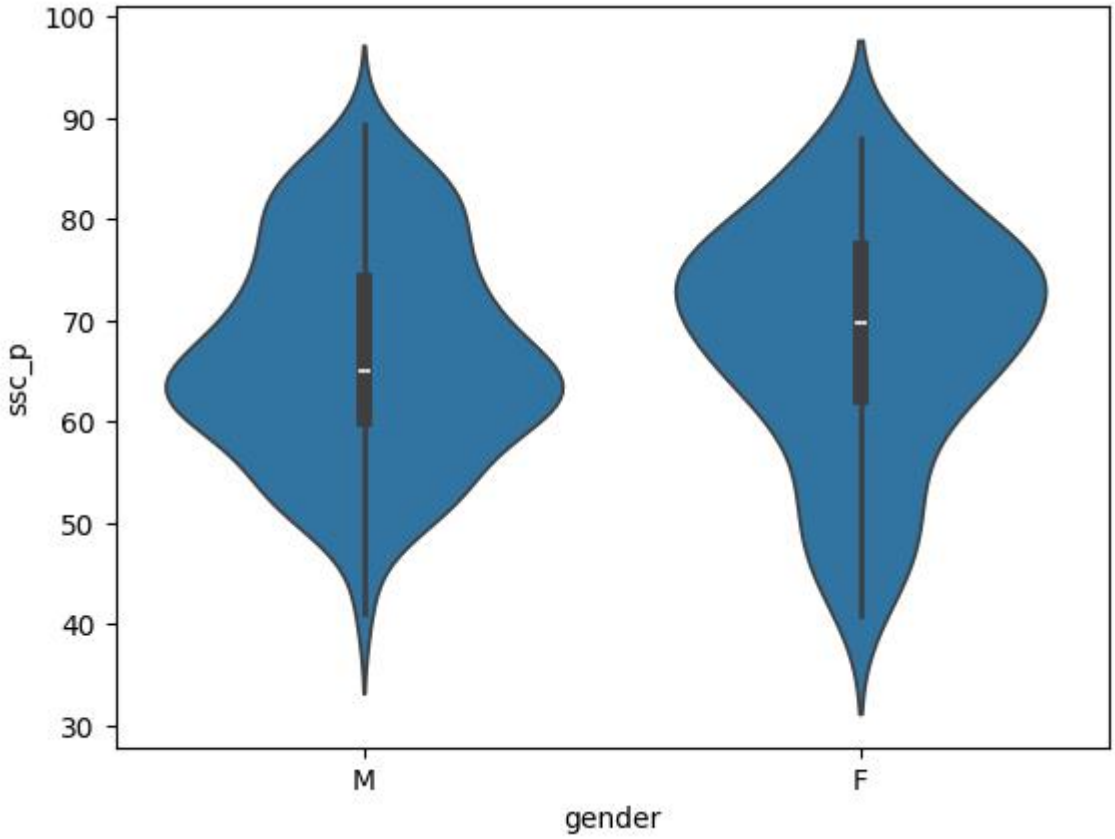


Distribution of SSC Marks by Gender

- Male candidates scored more in the range of 60-70%
- Female candidates concentration is more between 40-50% marks
- Male candidates have higher concentration at 80% marks
- A male candidate scored higher mark around 90%

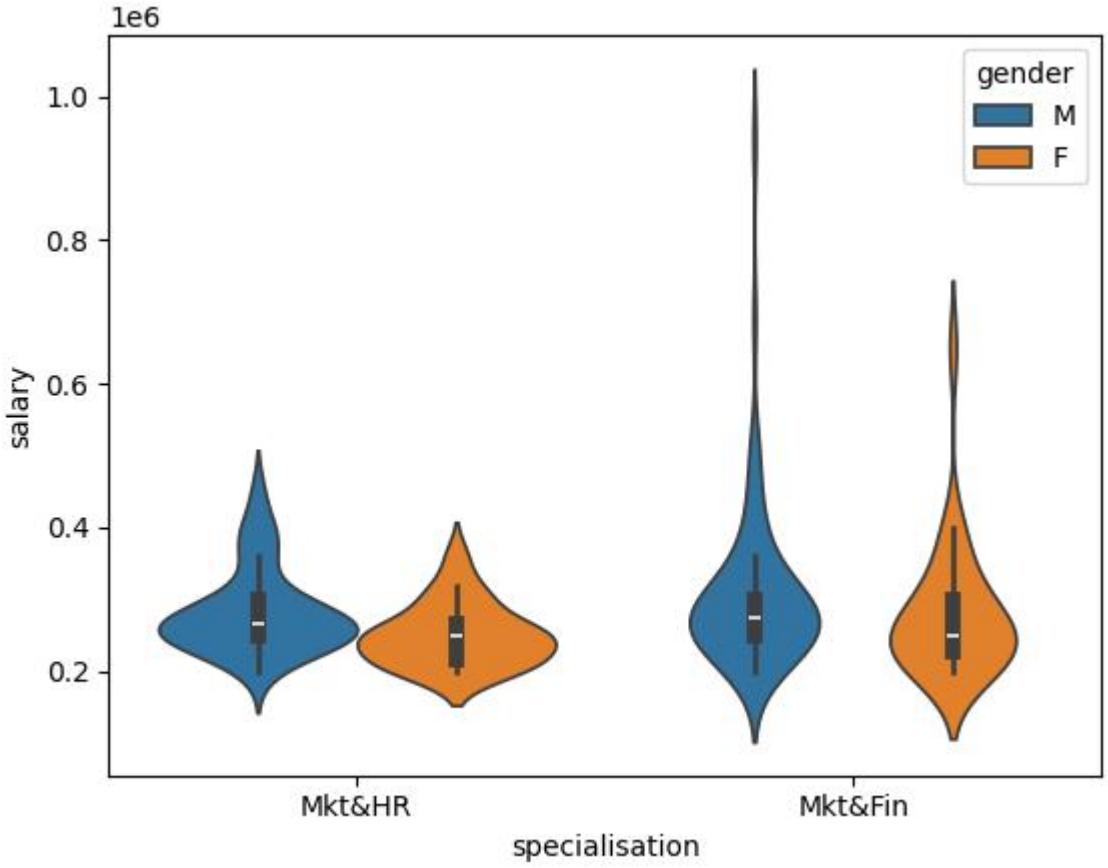
8) Univariate or Bivariate Distribution Plot-Violin Plot

- A violin plot is a distribution plot used in Exploratory Data Analysis.. It can be uni-variate or bi-variate depending on how many variables we use.



- Both males and females show a wide spread of scores
- Male have central peak around 60–70%, this is the range most of the male students scored .
- Female have central peak around 70–80%,this is the range most of the female students scored .
- The white dot in each violin represents the median SSC%, Male have median 65% , female have median of 70%
- Major difference in the way marks are distributed

This violin plot shows the distribution of salary for each specialization, split by gender.



- Wider parts of Salary Levels Between Specializations are same,more students earning salary in same range.
- Mkt&Fin specialization students has wider spread , there is more variation in salary
- Average salary of both specialization are not varied much.
- Mkt&Fin has low and high salary clusters,top long tail shows extreme salary value.
- Salary Levels Between Gender are,
- Some of the Male students get higher salary above 4L but female highest salary is around 4L in Mkt&Fin
- Some of the Male students get higher salary above 7L but female highest salary is around 7L in Mkt&HR
- Male gender has more variation in salary.