

Swarm Plot

Graph assignment
By
MuthuKumaranM

Why Swarm Plot is Important in Seaborn (Easy Explanation)

A Swarm Plot in Seaborn is used to visualize the distribution of data points while avoiding overlap. It is very useful in exploratory data analysis (EDA)

What is a Swarm Plot?

A swarm plot displays all individual data points for a variable, spread out along a categorical axis so that no two points overlap. It is an improved version of a strip plot.

1 Shows Actual Data Points

Unlike boxplots or violin plots, swarm plots show every single observation

Helps understand true data distribution

2 Avoids Overlapping

Automatically adjusts points so they don't overlap

Makes dense data clearly visible

3 Best for Small to Medium Datasets

Ideal when you want to see all values

Perfect for academic & project analysis

4. Great for Comparing Categories

Useful in bivariate analysis Example: salary vs specialisation, marks vs placement status

5.Detects Outliers Clearly

Outliers stand out visually easier than boxplots in some cases

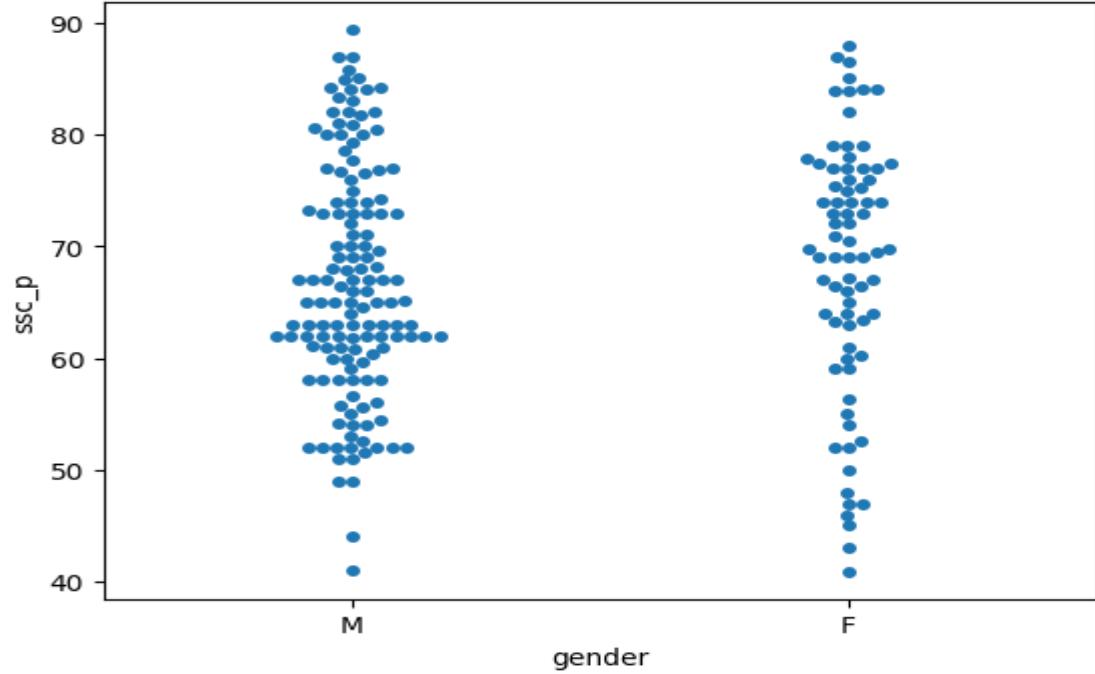
Code to make swarmPlot:

```
import seaborn as sns  
import matplotlib.pyplot as plt  
sns.swarmplot(x='specialisation', y='salary', data=dataset)  
plt.show()
```

Swarm Plot vs Others

Plot	Shows points	Handles overlap	Best for
Strip Plot	<input checked="" type="checkbox"/>	<input type="checkbox"/>	Simple view
Swarm Plot	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	Clean comparison
Box Plot	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Summary stats
Violin Plot	<input type="checkbox"/>	<input checked="" type="checkbox"/>	Distribution

```
sns.swarmplot(x='gender',y='ssc_p',data=df)
plt.show()
```



Information In Violin Plot

Most salaries for all groups are concentrated between ₹200,000 and ₹300,000.

Mkt&Fin shows a wider salary spread compared to Mkt&HR.

Mkt&HR salaries are more consistent and compact.

Median salary of Mkt&Fin students is higher than Mkt&HR.

In both specialisations:

Male median salary is slightly higher than female.

Difference is small and overlapping.

Mkt&Fin (Male) has the highest variability in salary.

Female salaries in both specialisations show less variation.

One or two very high salary outliers (around ₹1,000,000) exist in Mkt&Fin (Male).

These outliers stretch the violin upward but do not affect most observations

Wider sections of the violin indicate more students at that salary range.

Narrow ends indicate fewer students.

Mkt&Fin → higher salary potential

Mkt&HR → more stable salaries

Males show slightly higher salary spread

Extreme high salaries are rare

Boxplot

Variables Compared

Y-axis: ssc_p (Secondary School Percentage)

X-axis: ssc_b (Board type)

Others

Central

Grouping: Gender

M (Male)

F (Female)

Spread (Interquartile Range – IQR)

Others board has a wider IQR, meaning more variation in scores.

Central board scores are more tightly clustered, showing consistency.

Male scores show slightly more variability than females.

Maximum ssc_p values for all groups reach around 88–90%.

Minimum values:

Others (Male): around 49%

Others (Female): around 45%

Central (Male): around 41%

Central (Female): around 43%

Central Tendency (Median Line)

Female students consistently have a higher median SSC percentage than males in both board types.

The highest median is observed for Females from “Others” board.

Central board males show the lowest median among all four groups

Performance Distribution (IQR – Box Width)

Others board:

Shows larger IQR, meaning student performance varies more.

Indicates presence of both average and high scorers.

Central board:

Narrower IQR suggests more uniform academic performance.

Females have a slightly tighter IQR than males → more consistent results.

Score Extremes:

Upper whiskers:

Reach close to 90% for all groups → top performers exist across boards and genders.

Lower whiskers:

Central board shows lower minimum scores than Others board.

Indicates higher risk of low performance in Central board students.

Outlier Behavior:

A single low-end outlier appears for Others board – Female, indicating:

One female student scored significantly lower than peers.

No extreme high outliers → top scores are stable and not anomalies.

Gender Gap Insights (from Graph Only)

Gender-based gap is more visible in Others board.

In Central board, male and female medians are closer, indicating reduced gender difference.

Overall Visual Pattern:

Board type influences performance more strongly than gender.

Others board dominates in terms of:

Higher medians

Higher maximum scores

Central board emphasizes consistency over peak performance.