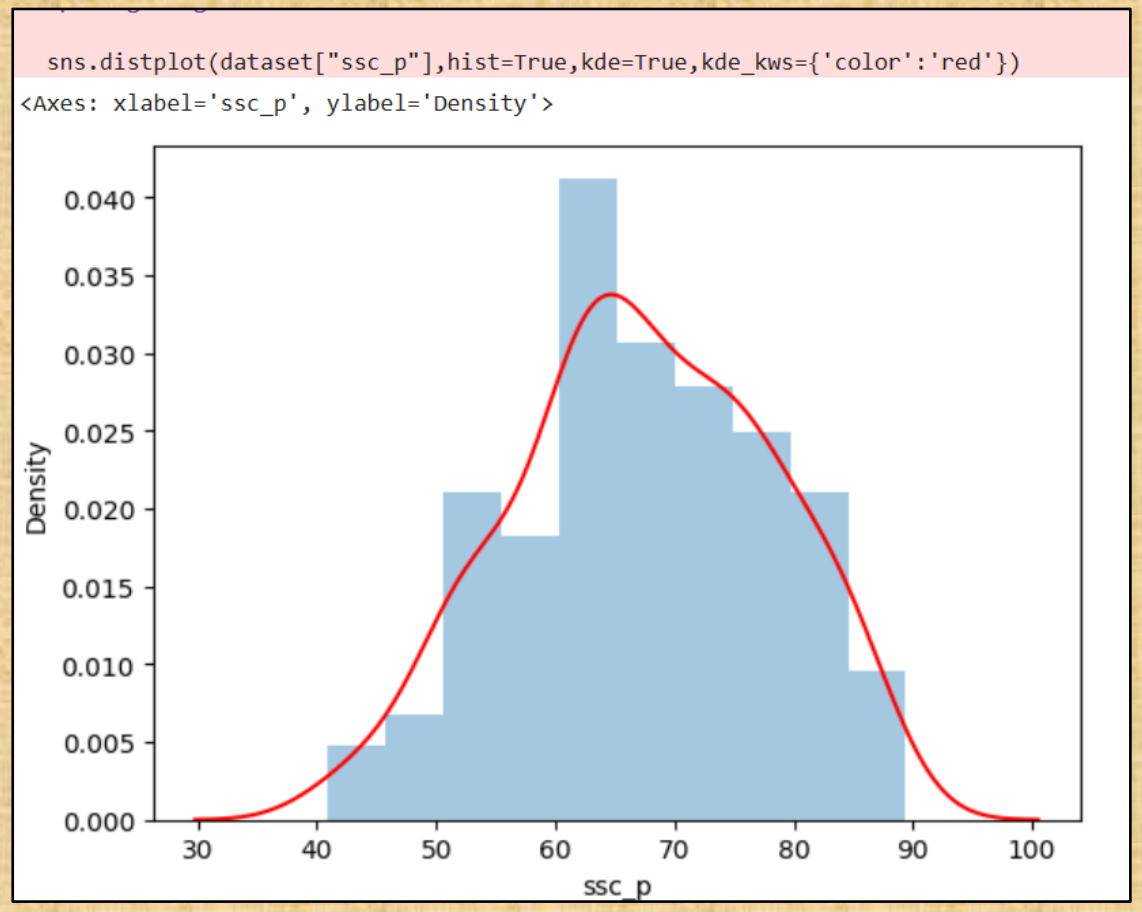
Dist Plot: Combine histograms and smooth curves to visualize a variable's distribution.

```
sns.distplot(dataset["ssc_p"],hist=True,kde=True,kde_kws={'color':'red'})
<Axes: xlabel='ssc_p', ylabel='Density'>
```

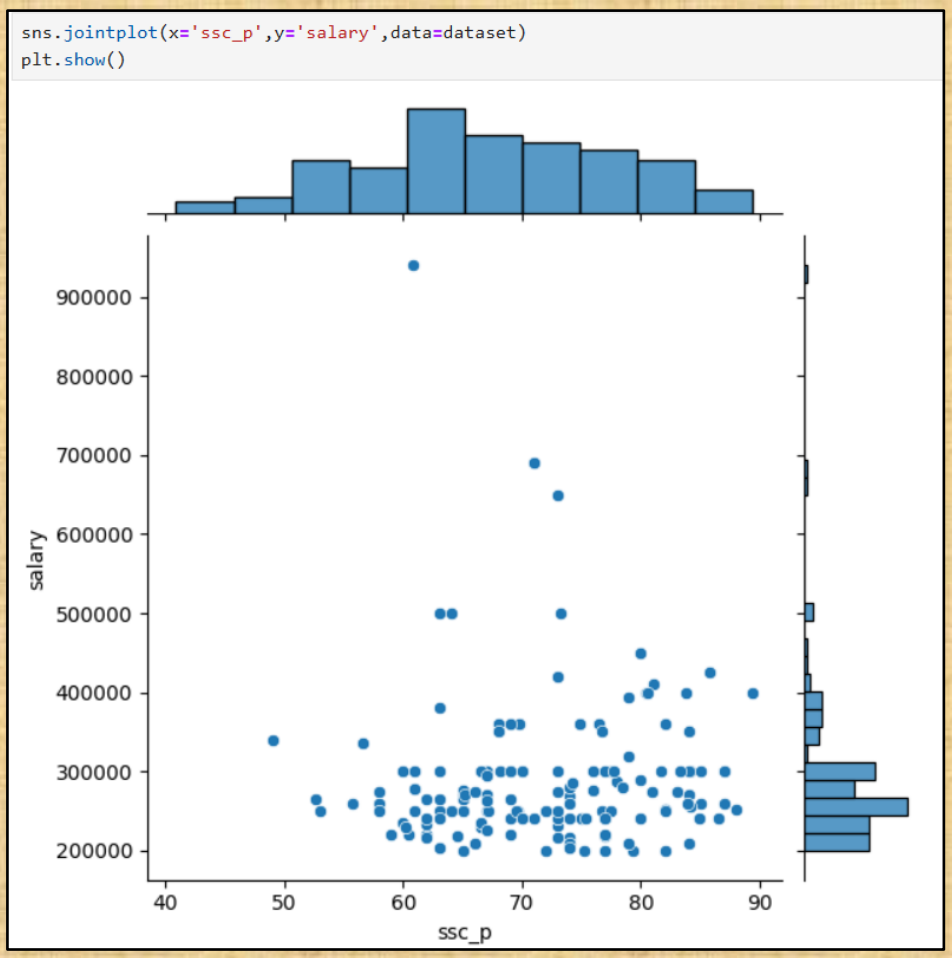


Observations:

- There is no repetition in the SSC_P marks from the **30 - 40 and 90 - 100** range
- SSC_P marks from **60 to 65** have a higher density as per the KDE curve

Joint Plot: Shows the relationship between two variables along with their individual distributions (scatter or hex or kde)

```
sns.jointplot(x='ssc_p',y='salary',data=dataset)
plt.show()
```

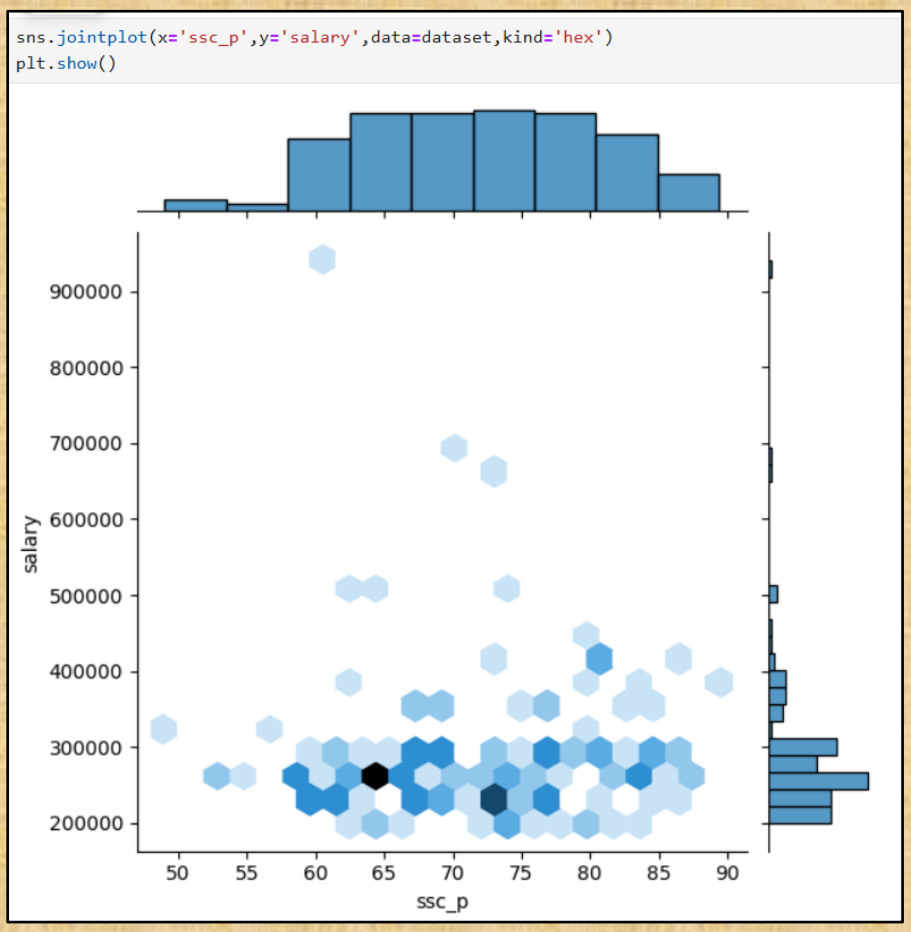


Observations:

- Students with ssc_p marks ranging from **50 to 90** are getting a salary from **2,00,000 to 4,00,000**
- Few students (5 of them) with ssc_p marks ranging from **65 to 75** are getting salary from **5,00,000 to 7,00,000**
- One student with ssc_p marks as 62 is getting salary greater than **9,00,000**(this might be an outlier for the dataset)

Joint Plot(kind:hex):

```
sns.jointplot(x='ssc_p',y='salary',data=dataset,kind='hex')
plt.show()
```



Observations:

Topmost frequency (dark color hex-⬢) `

- There are more repetitions for ssc_p marks from **63 to 65**
- There are more repetitions for salary from **2,30,000 to 2,70,000** approximately

Second top-most frequency(light colour compared to top one-⬢)

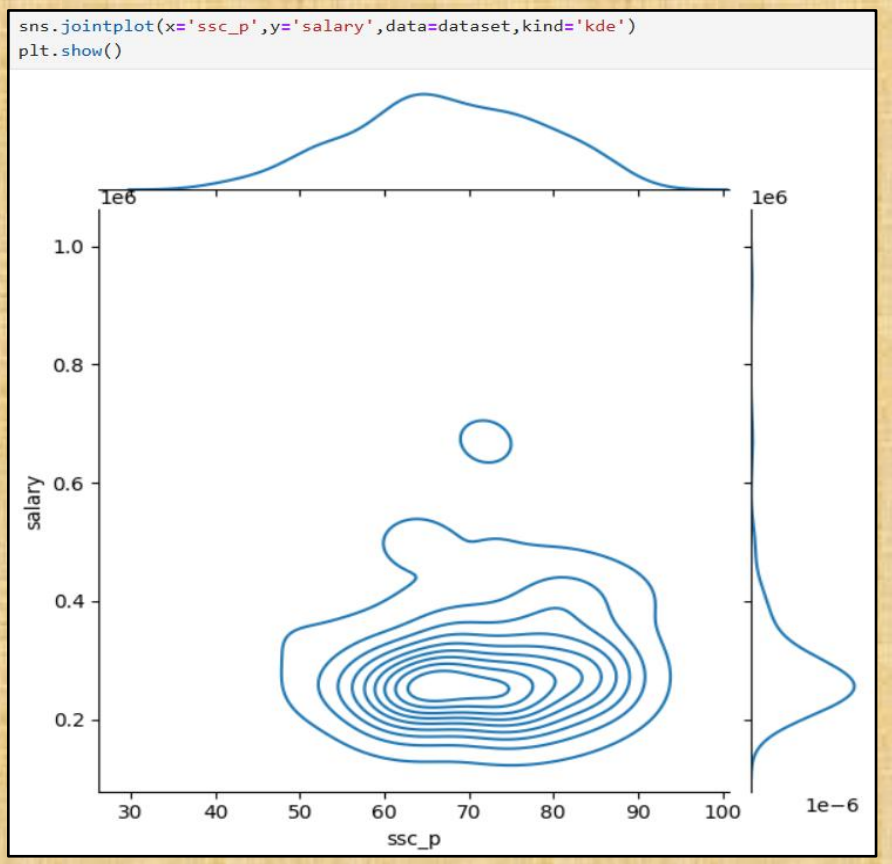
- There are repetitions for ssc_p marks from **72 to 74**
- There are repetitions for salary from **2,10,000 to 2,40,000** approximately

Third top-most frequency light colour compared to second top-most:⬢)

- There are repetitions for ssc_p marks from **57-62, 66-70,73-77,80-86**
- There are repetitions for salary from **2,10,000 to 3,10,000**

Joint Plot (kind: kde)

```
sns.jointplot(x='ssc_p',y='salary',data=dataset,kind='kde')
plt.show()
```

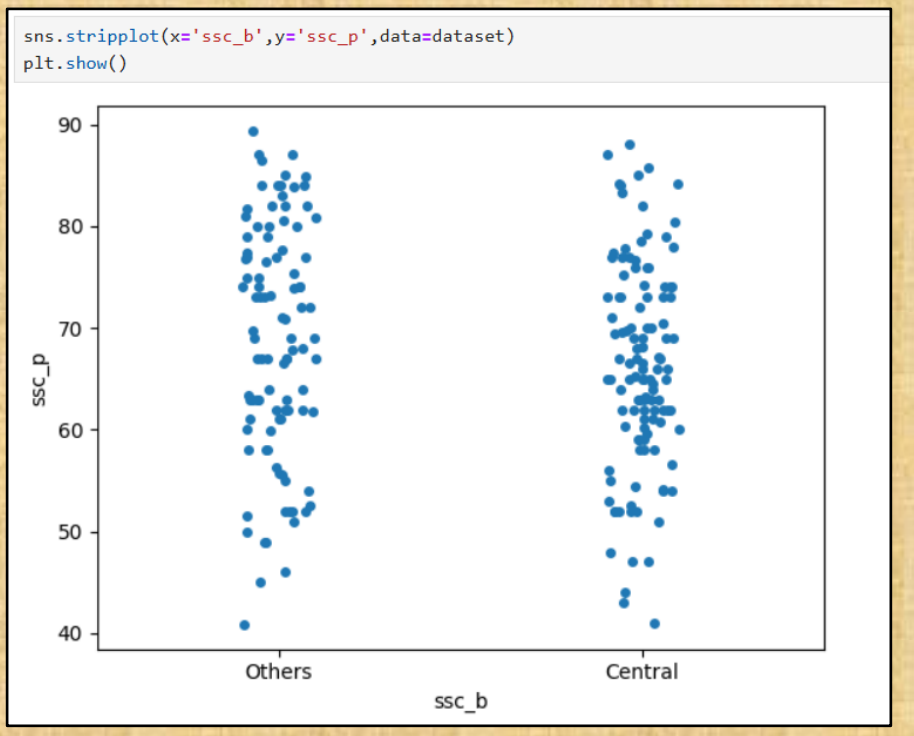


Observations:

- There are more repetitions in the SSC_P marks, which range from **48 to 94** approximately
- There are more repetitions in the Salary range from **1,00,000 to 5,00,000** approximately
- There are a few repetitions in the salary range from **6,50,000 to 7,50,000**

Strip Plot: Individual data points along a categorical axis to reveal distribution, density and outliers

```
sns.stripplot(x='ssc_b',y='ssc_p',data=dataset)
plt.show()
```

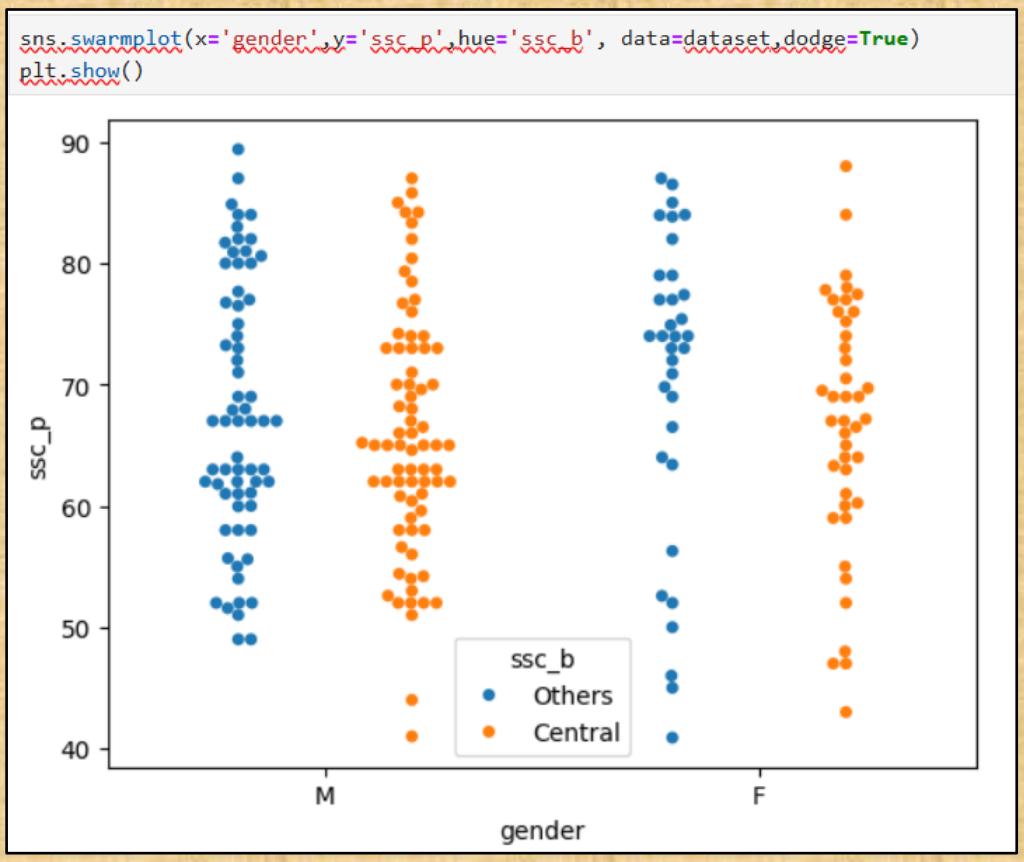


Observations:

- Students of ssc_b(central), scored marks from **41 to 88**
- Students of ssc_b(Others), scored marks from **41 to 89**
- Most of the students in ssc_b(central) scored marks from **52 to 80**
- Most of the students in ssc_b(others) scored marks from **50 to 85**

Swarm Plot: Display individual values along a categorical axis, spaced to prevent overlap and highlight clusters and outliers.

```
sns.swarmplot(x='gender',y='ssc_p',hue='ssc_b', data=dataset,dodge=True)
plt.show()
```

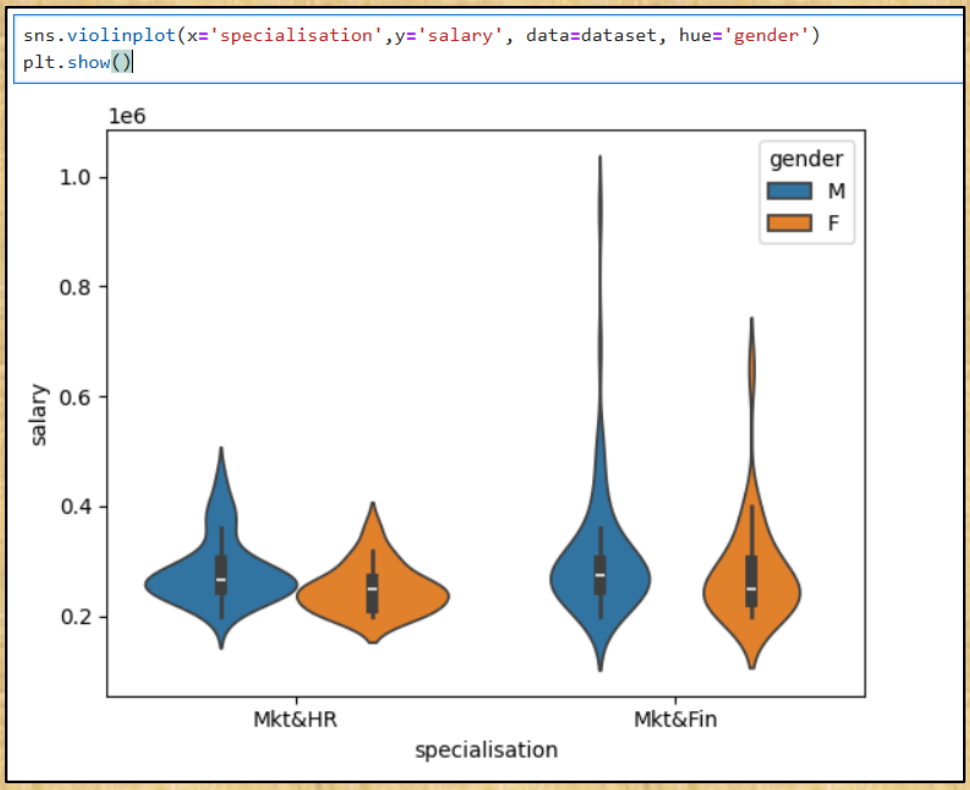


Observations:

- ***Female*** students from ***SSC_B(Central)*** have slightly tighter clusters of higher scores compared to ***male*** students in ***SSC(Central)***
- ***Female*** students from ***SSC_B(Central)*** have better performance compared to male students from ***SSC_B(Central)***
- Students from ***SSC(Central)*** have better performance compared to Students from ***SSC_B(Others)***

Violin Plot: A Visual symphony of data spread and central Tendency

```
sns.violinplot(x='specialisation',y='salary', data=dataset, hue='gender')
plt.show()
```

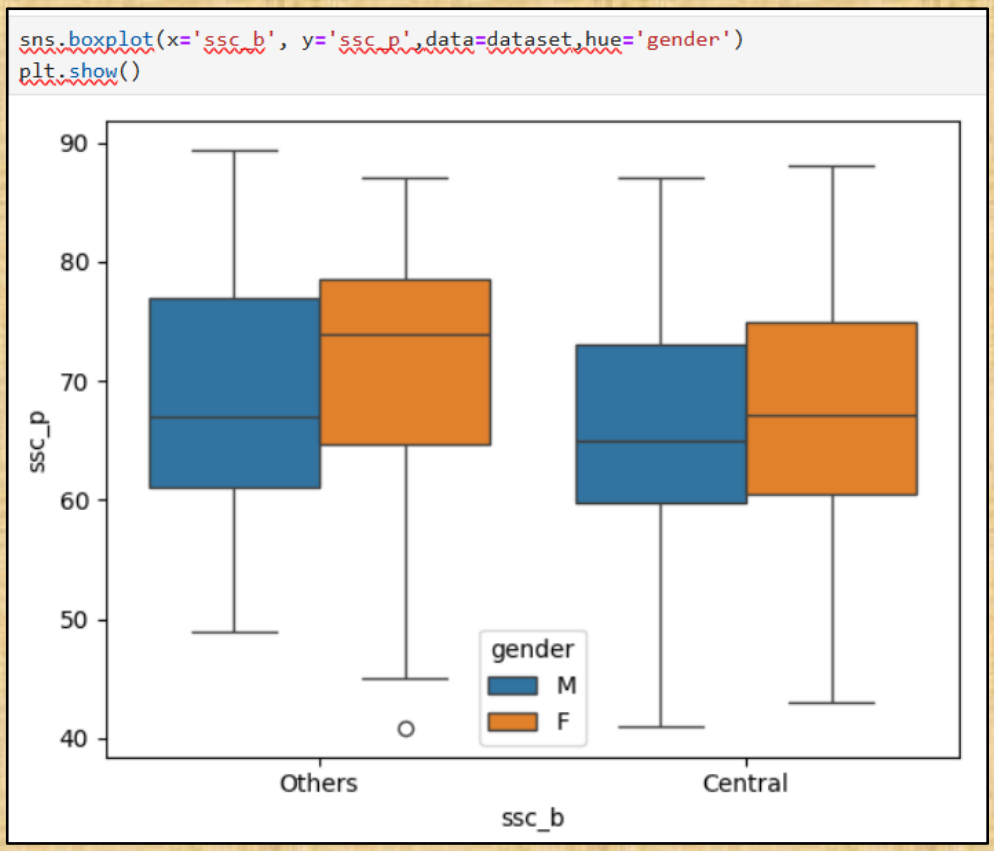


Observations:

- Most of the Male students with Mkt and HR specialisation received salary from **2,10,000 to 3,00,000**
- Most of the Female students with Mkt and HR specialisation received salary from **2,00,000 to 2,50,000**
- Most of the Male students with Mkt and fin specialisation received salary from **2,40,000 to 3,00,000**
- Most of the Female students with Mkt and fin specialisation received salary from **2,10,000 to 3,00,000**
- Minimum and maximum salary received by male students (Mkt & HR specialisation) are **1,50,000 and 5,00,000** respectively
- Minimum and maximum salary received by male students (Mkt & fin specialisation) are **50,000 and 10,00,000** respectively
- Minimum and maximum salary received by female students (Mkt & HR specialisation) are **1,60,000 and 4,00,000** respectively
- Minimum and maximum salary received by male students (Mkt & fin specialisation) are **1,00,000 and 7,50,000** respectively
- ***Male*** students are receiving ***higher salary*** in both specializations (Mkt & HR and Mkt & Fin) compared to ***female*** students.

Box Plot: A compact way to reveal the center, spread and outliers of a dataset

```
sns.boxplot(x='ssc_b', y='ssc_p',data=dataset,hue='gender')
plt.show()
```



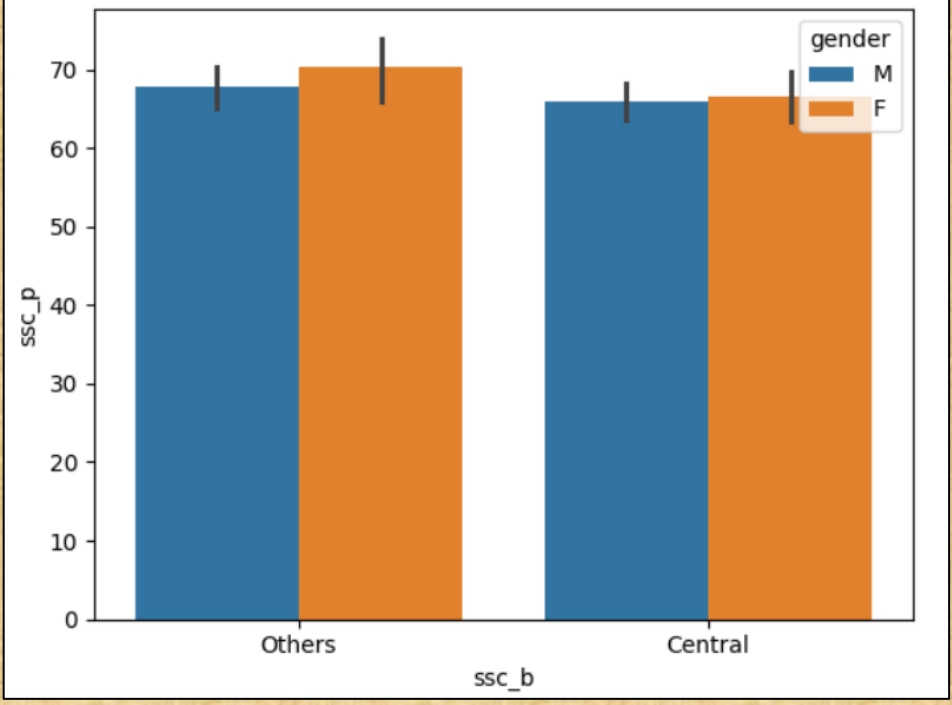
Observations:

- Lower bound outlier value for ***SSC_B(Others)*** is ***42*** marks
- ***Minimum*** and ***maximum*** marks of ***male*** students in SSC_B(Others) is ***49*** and ***89*** respectively
- ***25%(Q1)*** of male students in SSC_B(Others) scored ***62*** marks
- ***50%(Q2)*** of male students in SSC_B(Others) scored ***67*** marks
- ***75%(Q3)*** of male students in SSC_B(others) scored ***76*** marks
- ***Minimum*** and ***maximum*** marks of ***female*** students in ***SSC_B(others)*** is ***45*** and ***86*** respectively
- ***25%(Q1)*** of female students in SSC_B(Others) scored 65 marks
- ***50%(Q2)*** of female students in SSC_B(Others) scored 74 marks
- ***75%(Q3***) of female students in SSC_B(others) scored 78 marks
- ***Minimum*** and ***maximum*** marks of ***male*** students in ***SSC_B(Central)*** is ***42*** and ***87*** respectively
- ***25%(Q1)*** of male students in SSC_B(Others) scored 60 marks
- ***50%(Q2)*** of male students in SSC_B(Others) scored 65 marks

- ***75%(Q3)*** of male students in SSC_B(others) scored 73 marks
- ***Minimum*** and ***maximum*** marks of ***female*** students in ***SSC_B(Central)*** is ***43*** and ***88*** respectively
- ***25%(Q1)*** of female students in SSC_B(Others) scored ***61*** marks
- ***50%(Q2)*** of female students in SSC_B(Others) scored ***67*** marks
- ***75%(Q3***) of female students in SSC_B(others) scored ***75*** marks
- Students from ***SSC_B(Other)*** performed well compared to Students from ***SSC_B(Central)***
- In SSC_B(Others), ***Male*** students performed well compared to ***Female*** students
- In SSC_B(Others), ***Female*** students perform well compared to ***male students(percentile-wise).***

Bar Plot: Make it easy to compare quantities across categories briefly

```
sns.barplot(x='ssc_b', y='ssc_p', data=dataset, hue='gender')
plt.show()
```

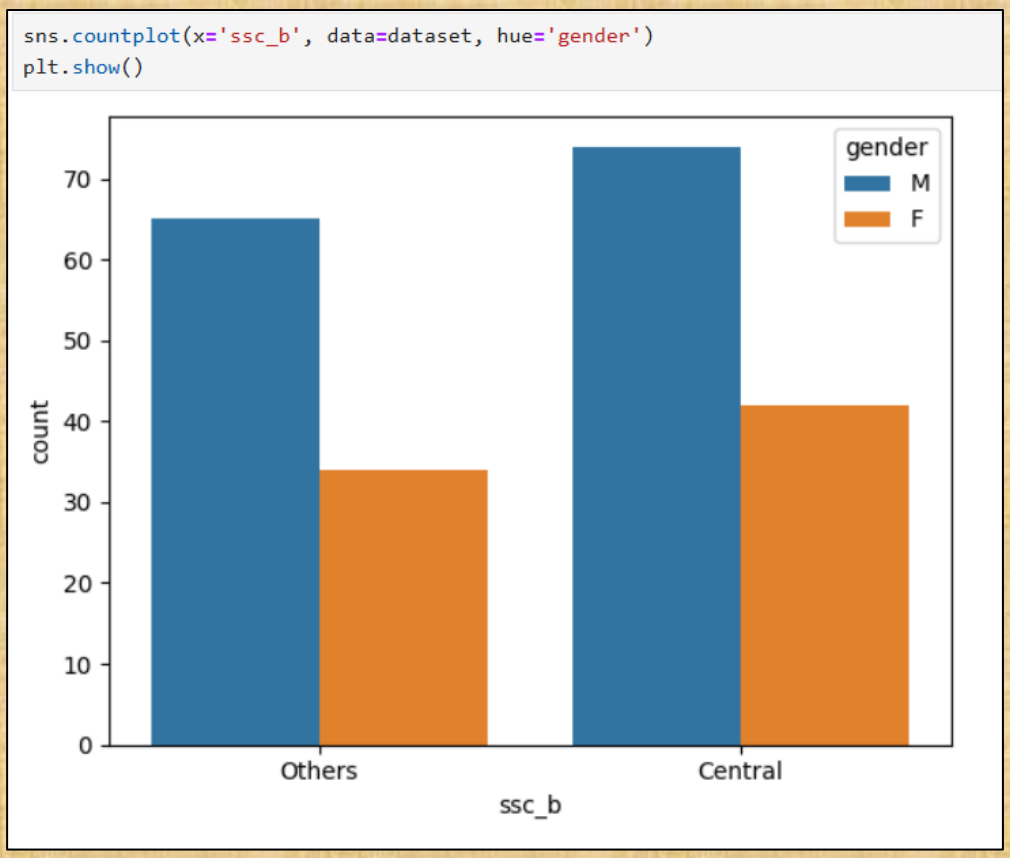


Observations:

- Students in ***SSC_B(Others)*** performed well compared to ***SSC_B(Central)***
- Female students performed well in ***SSC_B(Others)*** and ***SSC_B(central)*** compared to Male students

Count Plot: Visual Headcount of Categorical Values

```
sns.countplot(x='ssc_b', data=dataset, hue='gender')
plt.show()
```

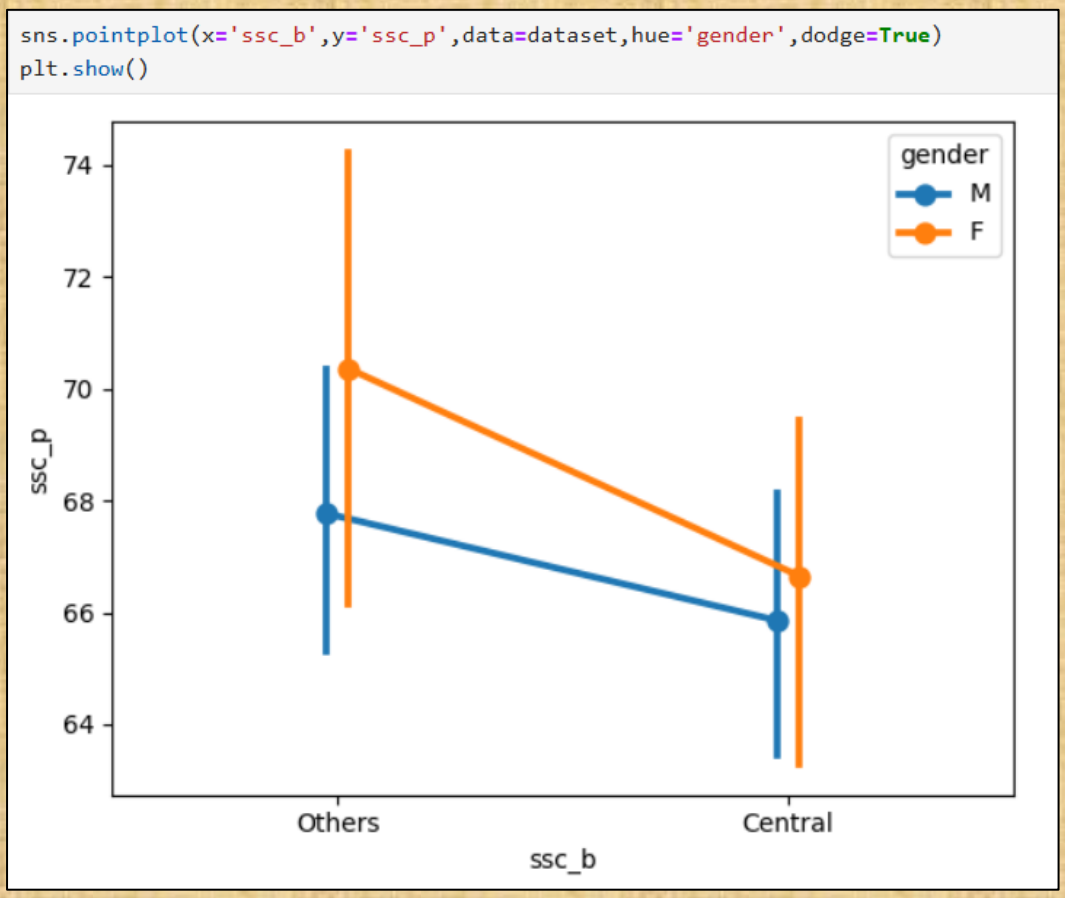


Observations:

- Most of the students belongs to **ssc_b(Central)** compared to **ssc_b(Others)**
- Higher percentage of students enrolled on the **Central board** compared to **Others** in SSC
- Number of **male** students in both the boards (Central and others) is **higher** than that of **female** students
- Higher percentage of **male** students are enrolled for **SSC** compared to **female** students.

Point Plot: Visualize trends by plotting means or other estimates with confidence intervals across categorical groups

```
sns.pointplot(x='ssc_b',y='ssc_p',data=dataset,hue='gender',dodge=True)
plt.show()
```

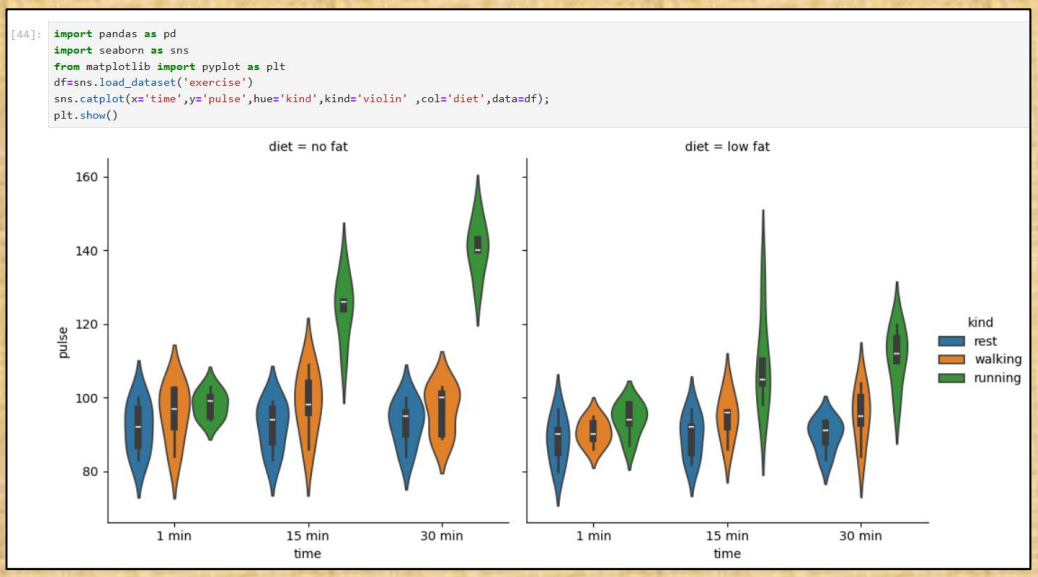


Observations:

- Average score of **male** students from ssc_b(Others) is **67.8**
- Average score of **female** students from ssc_b(Others) is **70.5**
- Average score of **male** students from ssc_b(Central) is **66.25**
- Average score of **female** students from ssc_b(Central) is **67**
- Based on the **average** score, **female** students perform well in both **Central and other** boards compared to **male** students
- **Female** students from **ssc_b(Others)** are consistent (less variance as it has a high error bar) compared to **male** students
- **Male** students from **ssc_b(Central)** are consistent (less variance as it has a high error bar) compared to **female** students
- Based on **error bar**, **Female** students perform well in **ssc_b(others)** compared to **male** students
- **Male** students perform well in **ssc_b(Central)** compared to **female** students

Factor Plot or Cat Plot: Combine multiple plot types to visualize category-based distributions and comparisons.

```
[44]: import pandas as pd
      import seaborn as sns
      from matplotlib import pyplot as plt
      df=sns.load_dataset('exercise')
      sns.catplot(x='time',y='pulse',hue='kind',kind='violin' ,col='diet',data=df);
      plt.show()
```



Observations:

Left plot: diet-no fat

Overall Observation - 1 min:

- People who rest, walk and run for a ***minute*** have a normal pulse rate (***within normal range<120***)
- People who ***walk, rest*** for a minute have a lower pulse rate compared to those who ***run*** for a minute (based on min value)
- People who ***walk*** for a minute have a higher pulse rate (***114***) compared to those who run and ***rest*** for a minute (based on max value)
- People who run for a minute have higher pulse rates starting from ***88*** whereas those who walk or rest have lower pulse rates (***40***)

Overall observation-15 min:

- People have a normal pulse rate while ***resting, walking*** for 15 minutes compared to those who ***run*** for 15 minutes
- People who run for 15 minutes have a ***higher pulse rate*** (146) compared to those who walk or rest
- People who walk, rest for 15 minutes have a ***lower pulse rate*** (40) compared to those who run

Overall observation-30 min:

- People have a normal pulse rate while ***resting, walking*** for 30 minutes compared to people who ***run*** for 30 minutes
- People who run for 30 minutes have a ***higher pulse rate*** (160) compared to those who walk and rest
- People who rest have a ***lower pulse rate*** (60) compared to those who walk and run

Observations for exercising with a no-fat diet for different time intervals:

- Pulse rate of people who ***rest, walk and run*** for 1 min with a ***no-fat diet*** is normal (***within 120***)
- Pulse rate of some people who run for ***15 minutes*** and ***30 minutes*** with a ***no-fat diet*** is abnormal (i.e. ***> 120***)
- Pulse rate of people who rest and walk for ***15 minutes*** and ***30 minutes*** with a ***no-fat diet*** is within the normal range (***within 120***)
- When the time duration increases (***15 min,30 min***) for running, the pulse rate of people following a no-fat diet plan becomes abnormal (i.e***.>120***)
- ***No-fat diet*** is not suitable for some people to run for more than one minute
- Further observations to be done on the people with respect to their health conditions for more details.

Right plot: diet-low fat

Overall observation-1 min:

- People have a normal pulse rate while resting, walking and running for a minute
- People who rest for a minute have a much lower pulse rate (***30***) compared to those who run for a minute (based on min value)
- People who rest for a minute have a higher pulse rate ***(106)*** compared to those who ***run and walk*** for a minute (based on max value)

Overall observations-15 min:

- People have a normal pulse rate while ***resting and walking*** for 15 minutes compared to people who **run** for 15 minutes
- People who run for 15 minutes have a higher pulse rate (***150***) compared to those who walk and rest (***based on max value***)
- People who rest for ***15 minutes*** have a lower pulse rate ***(40)*** compared to those who run and walk

Overall Observations-30 min:

- People have a normal pulse rate while ***resting, walking*** for ***30 minutes*** compared to people who ***run for 30 minutes***
- People who ***run for 30 minutes*** have a higher pulse rate (***130)*** compared to those who walk and rest
- People who ***walk for 30 minutes*** have a lower pulse rate (***40)*** compared to those who rest and run

Observations for exercising with a low-fat diet for different time intervals:

- Pulse rate of some people who ***run*** for 15 minutes and 30 minutes with a ***low-fat diet*** is abnormal (***i.e. > 120***)
- Pulse rate of people who ***rest, walk and run*** for ***1 min*** with a ***low-fat diet*** is normal (***within 120***)
- Pulse rate of people who rest and walk for ***15 minutes*** and 30 minutes with a ***low-fat diet*** is normal (***within 120***)
- Pulse rate of some people who walk for ***30 minutes*** is lower (***40***) compared to those who rest and run
- Pulse rate of some people who rest for ***a minute*** is lower (***30***) compared to those who ***run and walk***
- When the time duration ***increases*** (15 min,30 min) for running, the pulse rate of people following a low-fat diet plan becomes ***abnormal*** (i.e. >120)

Observations for diet-no fat and diet-low fat:

- Both categories (***diet-no fat & diet-low fat***) people have a normal pulse while resting, walking and running for a ***minute***
- People who follow a ***low-fat diet*** have a lower pulse spike rate (***130***) while running for 30 minutes compared to those (***160***) who follow a no-fat diet.
- People who follow a low-fat diet have a higher pulse spike rate (***150***) while running for 15 minutes compared to those (***146***) who follow a no-fat diet.
- ***Low-fat diet*** is quite suitable for people who run for ***30 minutes*** as it has a lower pulse spike rate
- ***No-fat diet*** is quite suitable for people who run for ***15 minutes*** as it has a lower pulse spike rate
- People who follow a ***low-fat diet*** have a lower pulse rate (***30***) while resting for a minute compared to those (***40***) who follow a ***no-fat diet***
- People who follow a ***low-fat diet*** have a lower pulse rate (***40***) while walking for 30 minutes compared to those who follow a ***no-fat diet***
- No-fat diet is suitable for most of them who rest and walk for different durations (as some people who follow a low-fat diet have lower pulse rates ***30,40)***