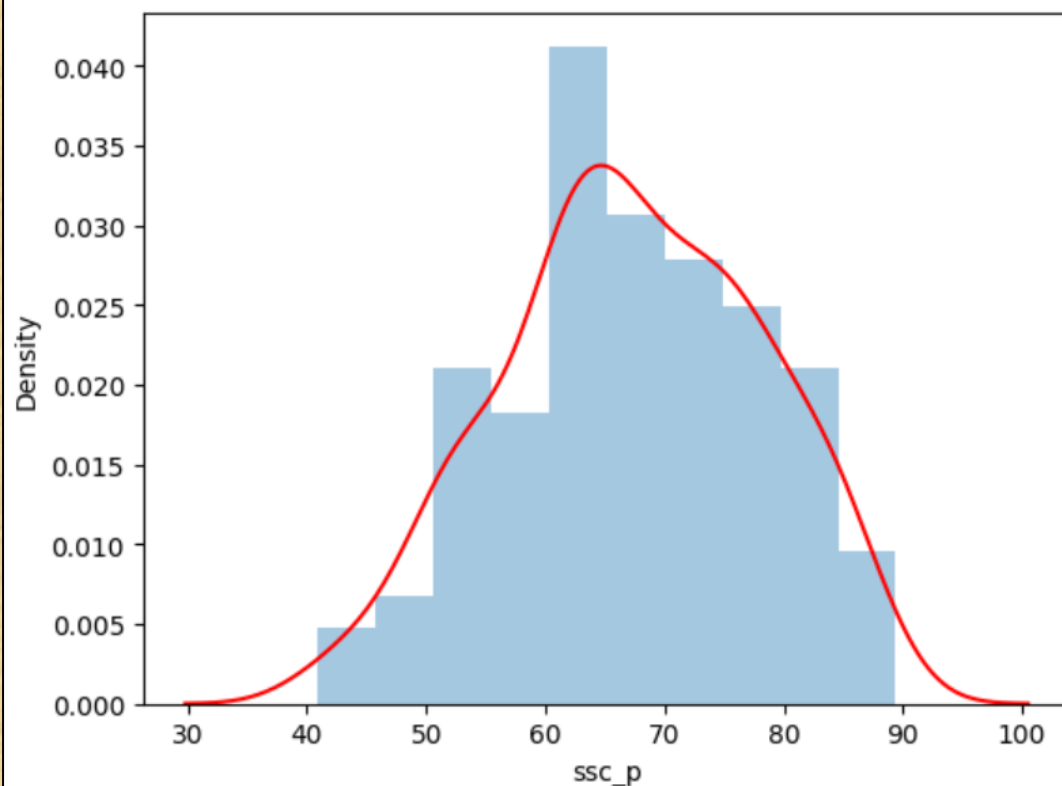# Different Plots of Seaborn with Observations

Dist Plot:

```
similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

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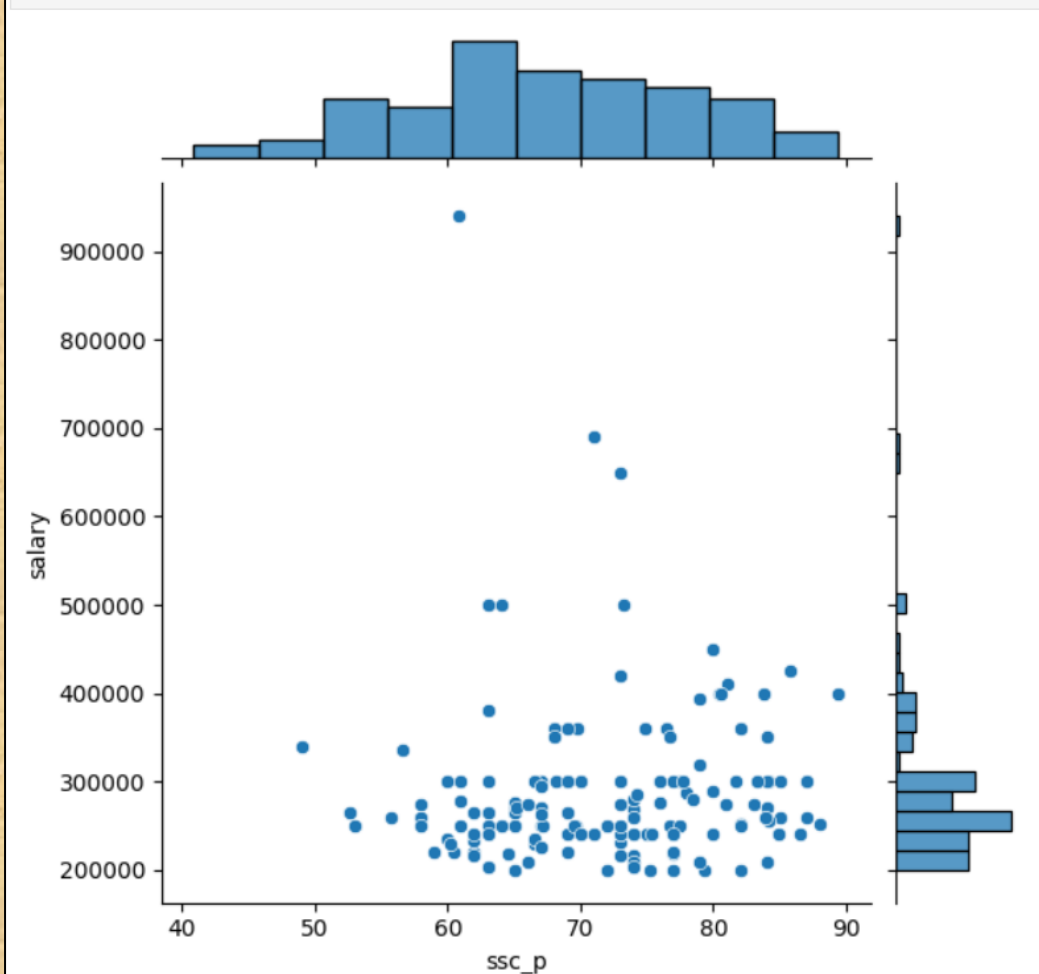

Observations:

- 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:

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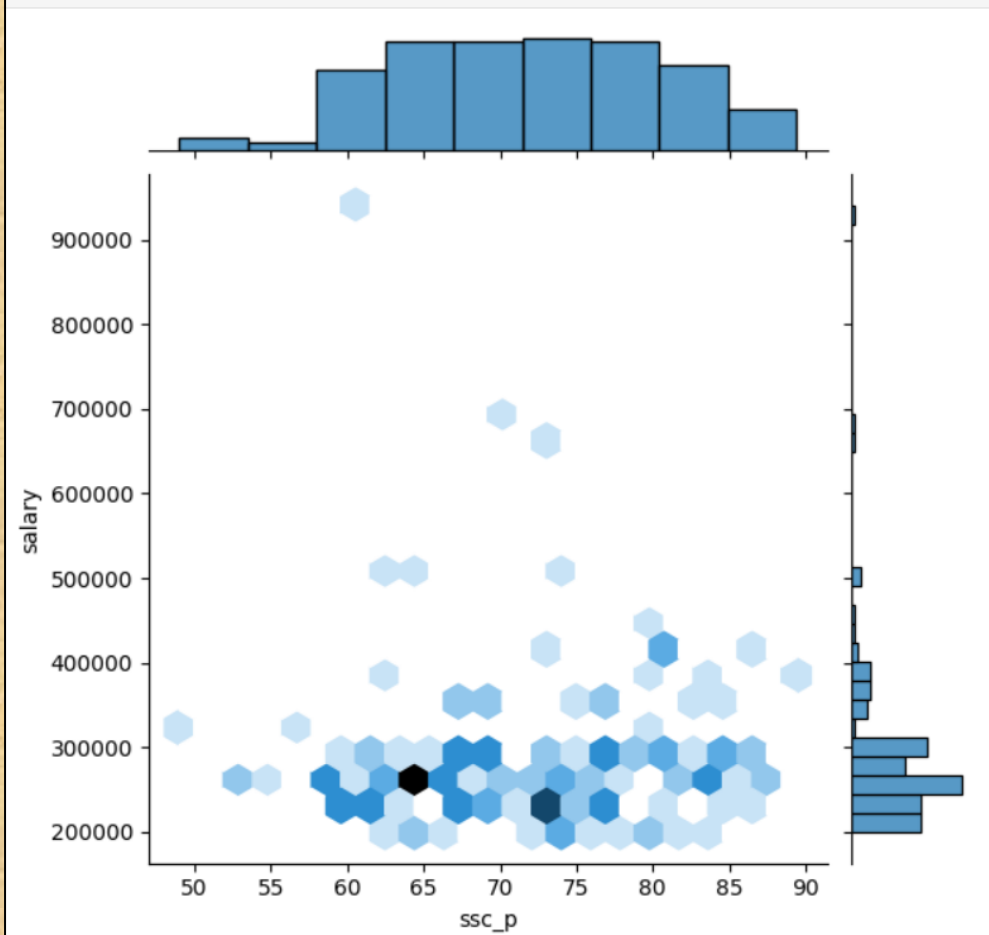

Observations:

- Students with ssc_p marks ranging from *50 to 90* are getting 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):

```python
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

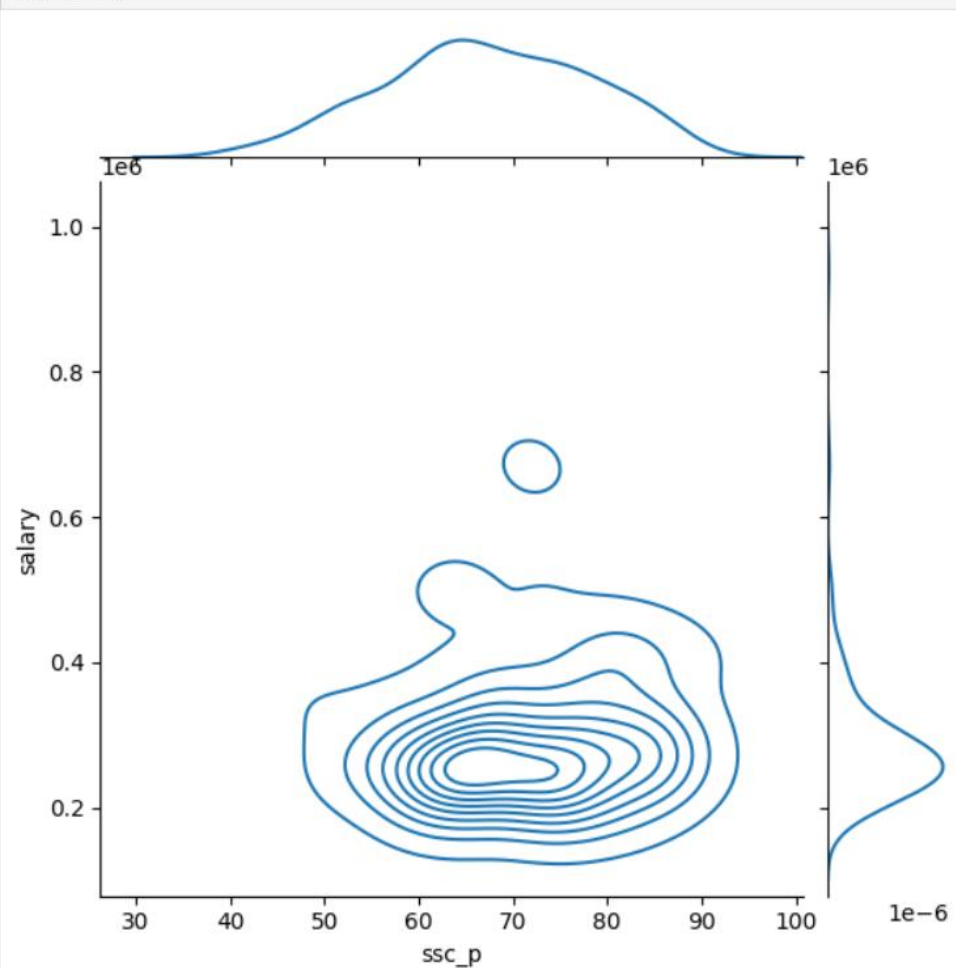*Second top-most frequency(light colour compared to top one-⬢)*

- 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()
```
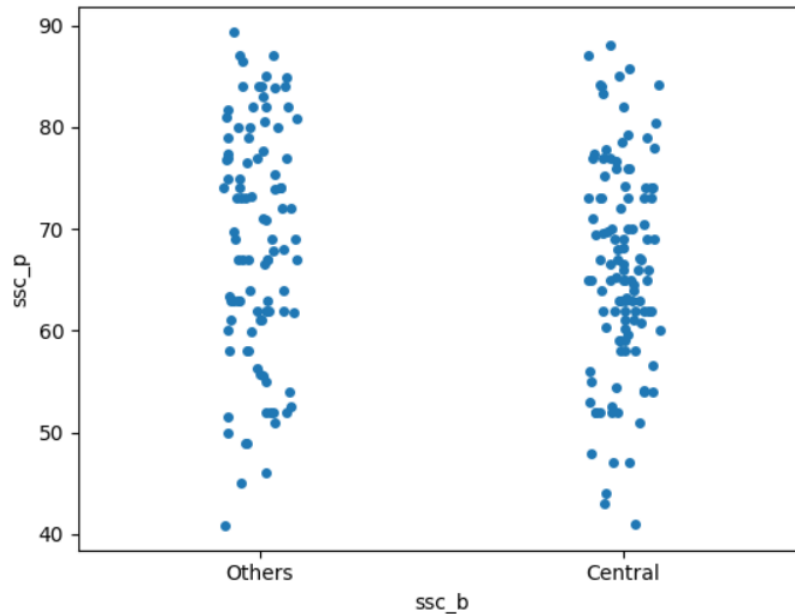

Observations:

- 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**
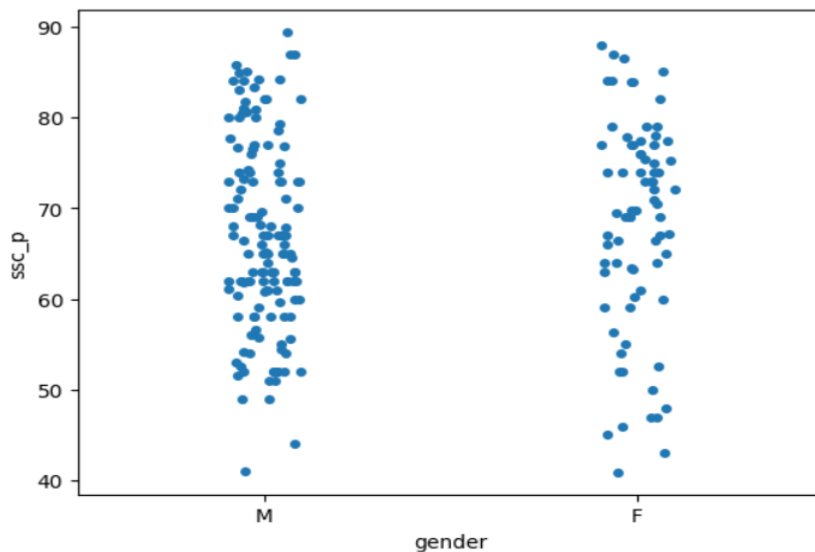
Strip Plot:

```python
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**

```python
sns.stripplot(x='gender',y='ssc_p',data=dataset,jitter=True)
plt.show()
```
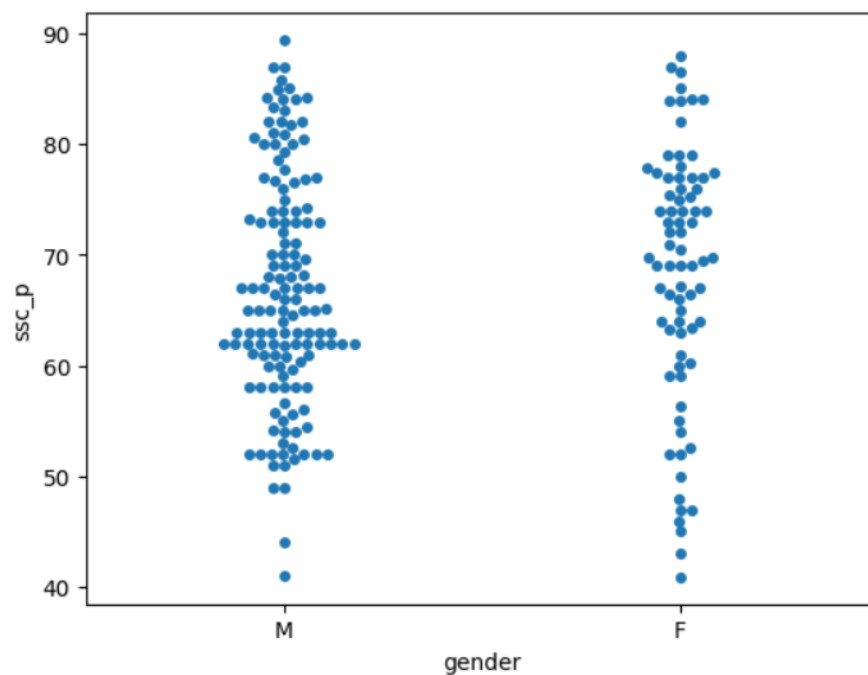
- Most of the male students scored ssc_p marks ranging from *50 to 85*
- Most of the female students scored ssc_p marks ranging from *58 to 80*
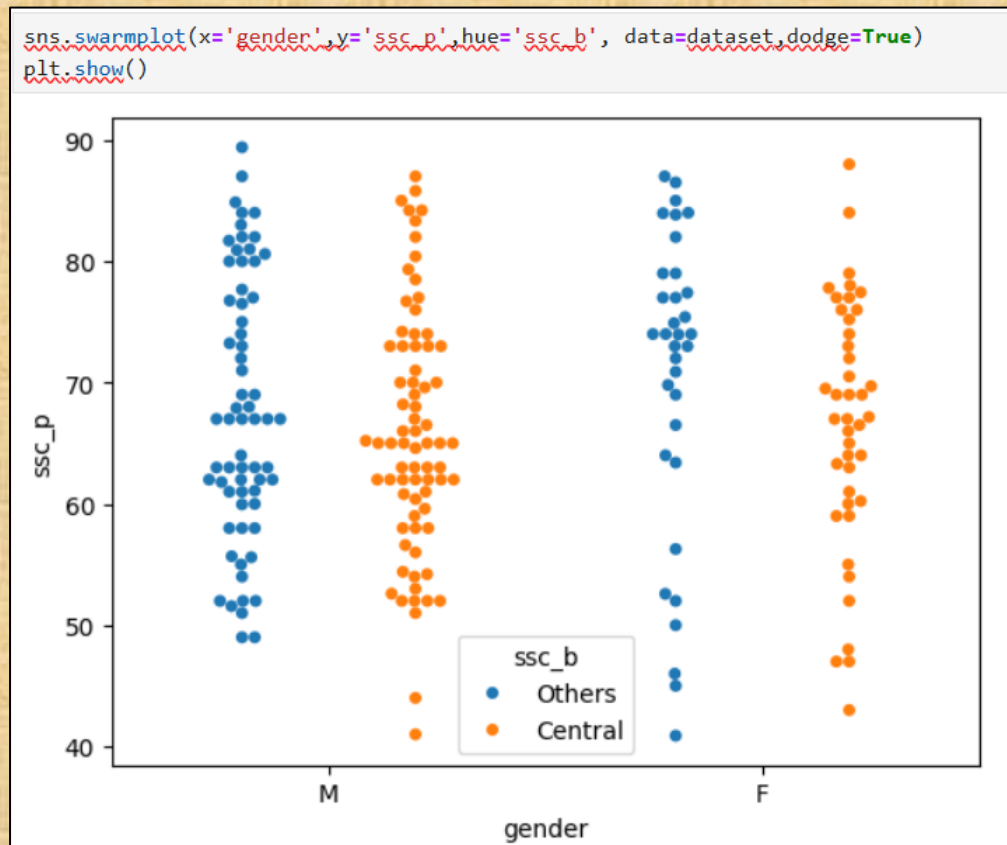
Swarm Plot

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


Observations:

- Most of the male students get *62 marks* in SSC
- Most of the female students get *75 marks* in SSC

Once we give dodge=True in the swarmplot, both the hue (Central and others) is plotted separately in the plot for better understanding
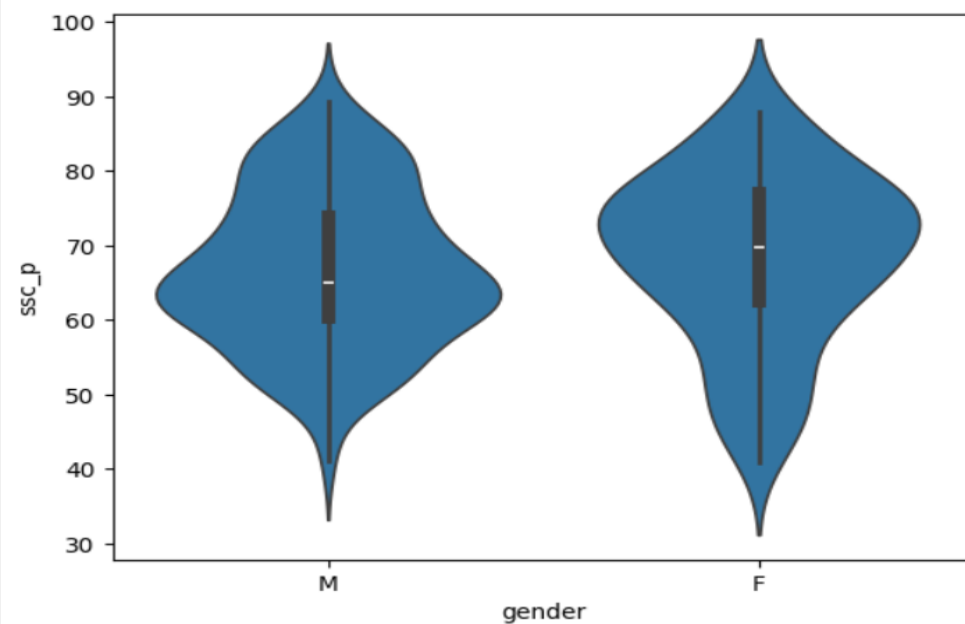
```
sns.swarmplot(x='gender',y='ssc_p',hue='ssc_b', data=dataset,dodge=True)
plt.show()
```


Observations:

- Most of the male students in ssc_b(Central) scored 62 marks in SSC
- Most of the male students in ssc_b(others) scored 68 marks in SSC
- Most of the female students in ssc_b(Central) scored 70 marks in SSC
- Most of the female students in ssc_b(others) scored 75 marks in SSC

Violin Plot:

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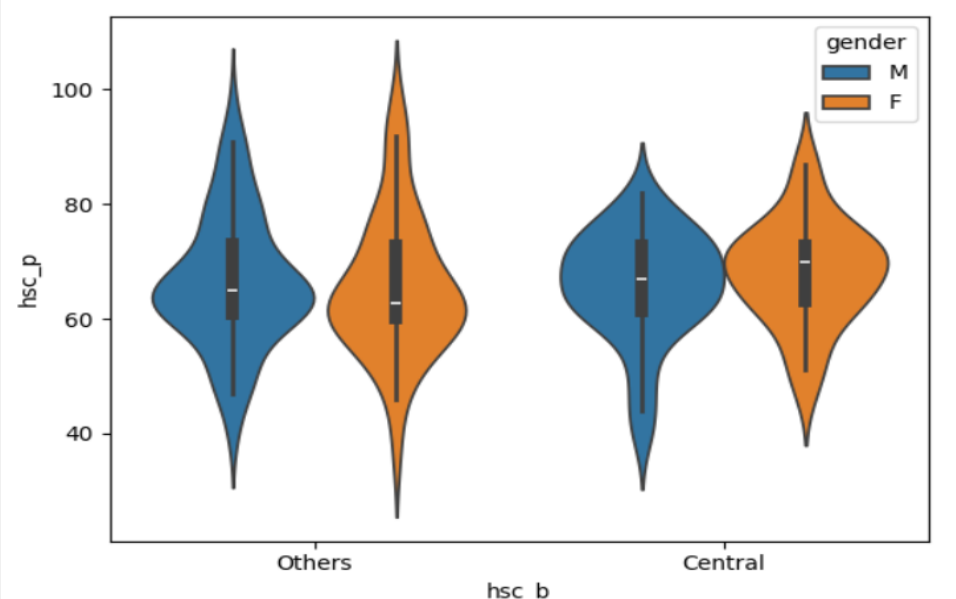

Observations:

- Most of the male students scored from 60 to 74 in SSC
- Most of the female students scored from 62 to 78 in SSC
- Minimum and maximum marks of male students in SSC are 34 and 97 respectively
- Minimum and maximum marks of female students in SSC are 32 and 97 respectively

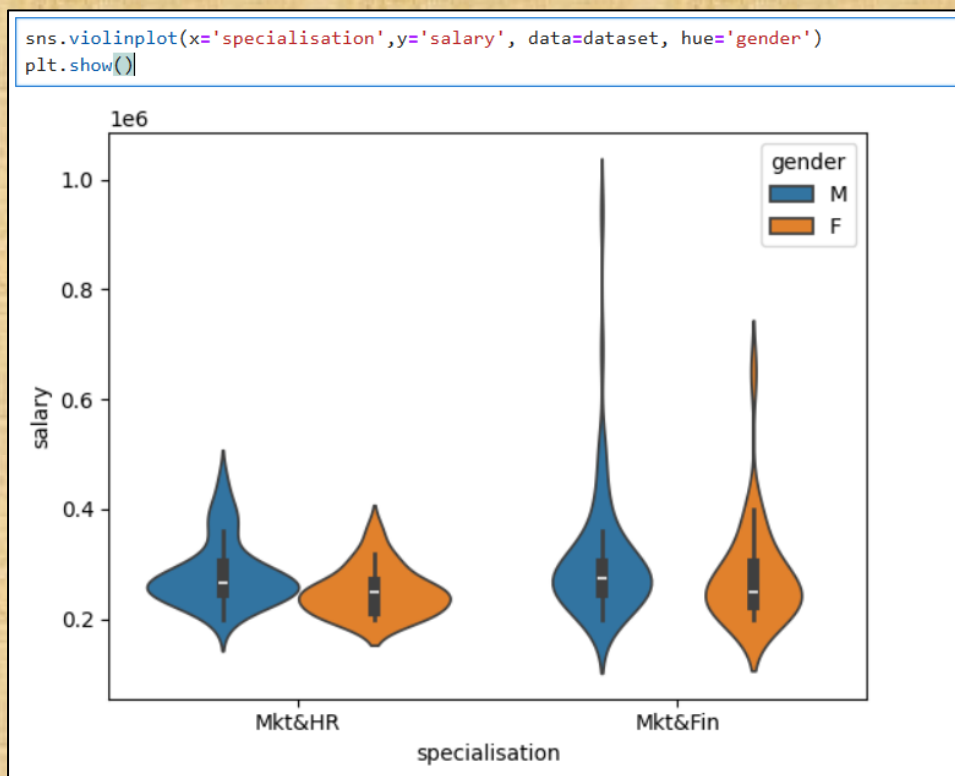Defining hue as 'gender' with x axis as 'hsc_b', y as 'hsc_p' displays the below plot

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

Observations:

- Male students in hsc_b(Others), scored marks from *60 to 78* in hsc_p
- Female students in hsc_b(Others), scored marks from *58 to 78* in hsc_p
- Male students in hsc_b(Central), scored marks from *60 to 78* in hsc_p
- Female students in hsc_b(Central), scored marks from *62 to 78* in hsc_p
- Maximum marks and minimum marks obtained by male students in hsc_b(Others) are *150 and 30* respectively
- Maximum marks obtained by female students in hsc_b(Others) are *160 and 20* respectively
- Maximum marks and minimum marks obtained by male students in hsc_b(Others) are *90 and 20* respectively
- Maximum marks and minimum marks obtained by female students in hsc_b(Central) are *95 and 30* respectively

Below is the violin plot for specialisation and salary with respect to gender:

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


Observations:

- Male students with Mkt and HR specialisation received salary from *2,10,000 to 3,00,000*
- Female students with Mkt and HR specialisation received salary from *2,00,000 to 2,50,000*
- Male students with Mkt and fin specialisation received salary from *2,40,000 to 3,00,000*

- Female students with Mkt and fin specialisation received salary from *2,10,000 to 3,00,000*
- Minimum and maximum marks obtained by male students (Mkt & HR specialisation) are *1,50,000 and 5,00,000* respectively
- Minimum and maximum marks obtained by male students (Mkt & fin specialisation) are *50,000 and 10,00,000* respectively
- Minimum and maximum marks obtained by female students (Mkt & HR specialisation) are *1,60,000 and 4,00,000* respectively
- Minimum and maximum marks obtained by male students (Mkt & fin specialisation) are *1,00,000 and 7,50,000* respectively