

Seaborn-1.ipynb

1. univariate analysis
2. Bivariate Analysis
3. Categorical plot
4. Distribution Observation

[Skip to Main](#)

Seaborn-1

Last Checkpoint: yesterday



```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
# univariate analysis
```

univariate analysis¶

[2]:

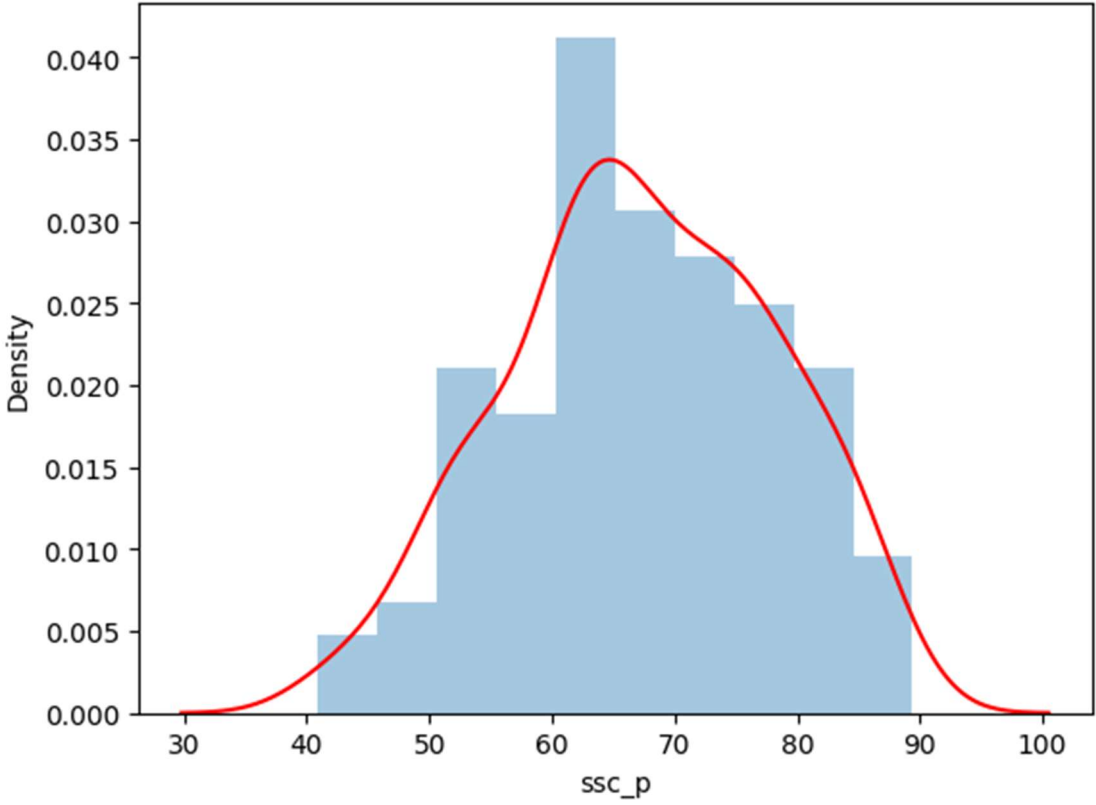
```
import seaborn as sb
dataset=pd.read_csv("Preplacement.csv")
sb.distplot(dataset['ssc_p'],hist=True,kde=True,kde_kws={'color':'red'})
C:\Users\Prabakaran J\AppData\Local\Temp\ipykernel_23996\1975791051.py:3: UserWarning:
```

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

Please adapt your code to use either `displot` (a figure-level function with 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>

```
sb.distplot(dataset['ssc_p'],hist=True,kde=True,kde_kws={'color':'red'})
[2]:
<Axes: xlabel='ssc_p', ylabel='Density'>
```

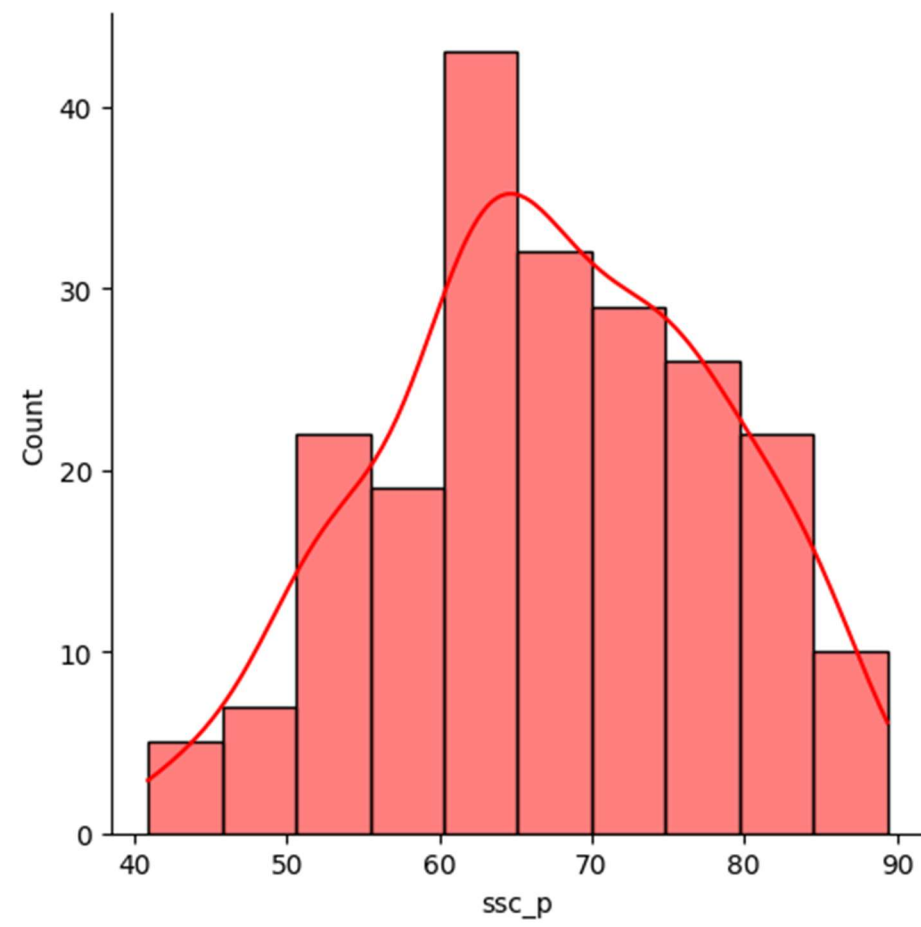


```
[21]:
dataset
[21]:
```

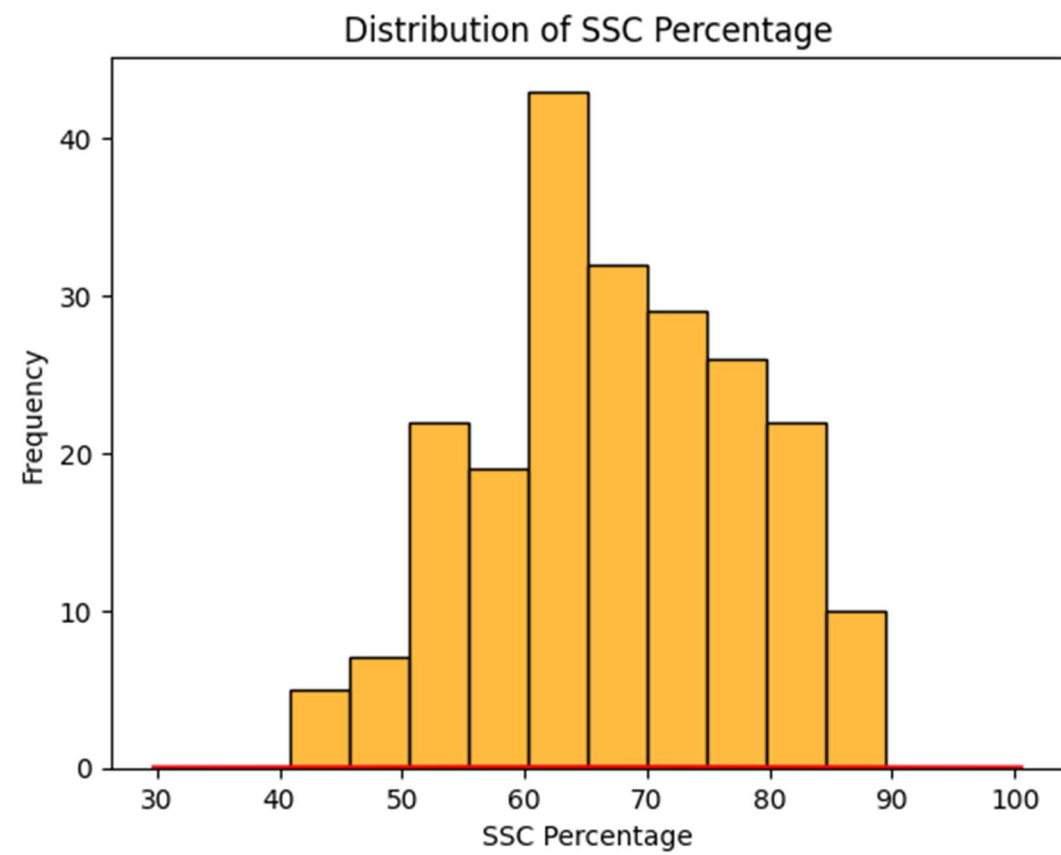
	sl_no	ssc_p	hsc_p	degree_p	etest_p	mba_p	salary	gender	ssc_b	hsc_b	hsc_s	degree_t	workex	specialisation	status
0	1.0	67.00	91.00	58.00	55.0	58.80	270000.000000	M	Others	Others	Commerce	Sci&Tech	No	Mkt&HR	Placed
1	2.0	79.33	78.33	77.48	86.5	66.28	200000.000000	M	Central	Others	Science	Sci&Tech	Yes	Mkt&Fin	Placed
2	3.0	65.00	68.00	64.00	75.0	57.80	250000.000000	M	Central	Central	Arts	Comm&Mgmt	No	Mkt&Fin	Placed
3	4.0	56.00	52.00	52.00	66.0	59.43	288655.405405	M	Central	Central	Science	Sci&Tech	No	Mkt&HR	Not Placed
4	5.0	85.80	73.60	73.30	96.8	55.50	425000.000000	M	Central	Central	Commerce	Comm&Mgmt	No	Mkt&Fin	Placed
...
210	211.0	80.60	82.00	77.60	91.0	74.49	400000.000000	M	Others	Others	Commerce	Comm&Mgmt	No	Mkt&Fin	Placed
211	212.0	58.00	60.00	72.00	74.0	53.62	275000.000000	M	Others	Others	Science	Sci&Tech	No	Mkt&Fin	Placed
212	213.0	67.00	67.00	73.00	59.0	69.72	295000.000000	M	Others	Others	Commerce	Comm&Mgmt	Yes	Mkt&Fin	Placed
213	214.0	74.00	66.00	58.00	70.0	60.23	204000.000000	F	Others	Others	Commerce	Comm&Mgmt	No	Mkt&HR	Placed
214	215.0	62.00	58.00	53.00	89.0	60.22	288655.405405	M	Central	Others	Science	Comm&Mgmt	No	Mkt&HR	Not Placed

215 rows × 15 columns

```
[19]:
sb.displot(dataset['ssc_p'], kde=True, color='red')
[19]:
<seaborn.axisgrid.FacetGrid at 0x185dcae3990>
```



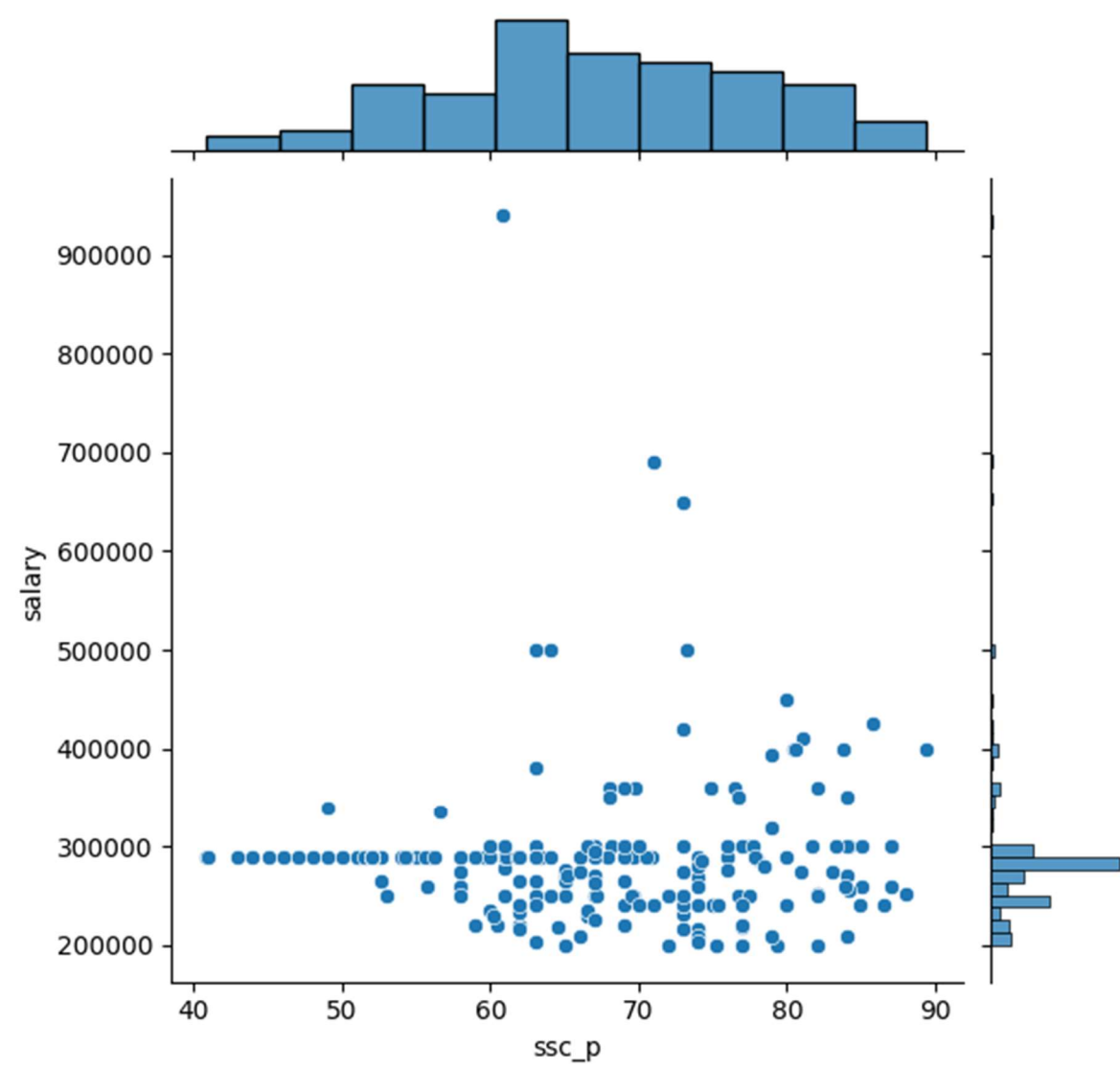
```
[20]:  
# Create the histogram  
sb.histplot(dataset['ssc_p'], color='orange')  
  
# Overlay the KDE plot  
sb.kdeplot(dataset['ssc_p'], color='red', lw=2)  
  
# Add titles and labels  
plt.title('Distribution of SSC Percentage')  
plt.xlabel('SSC Percentage')  
plt.ylabel('Frequency')  
[20]:  
Text(0, 0.5, 'Frequency')
```



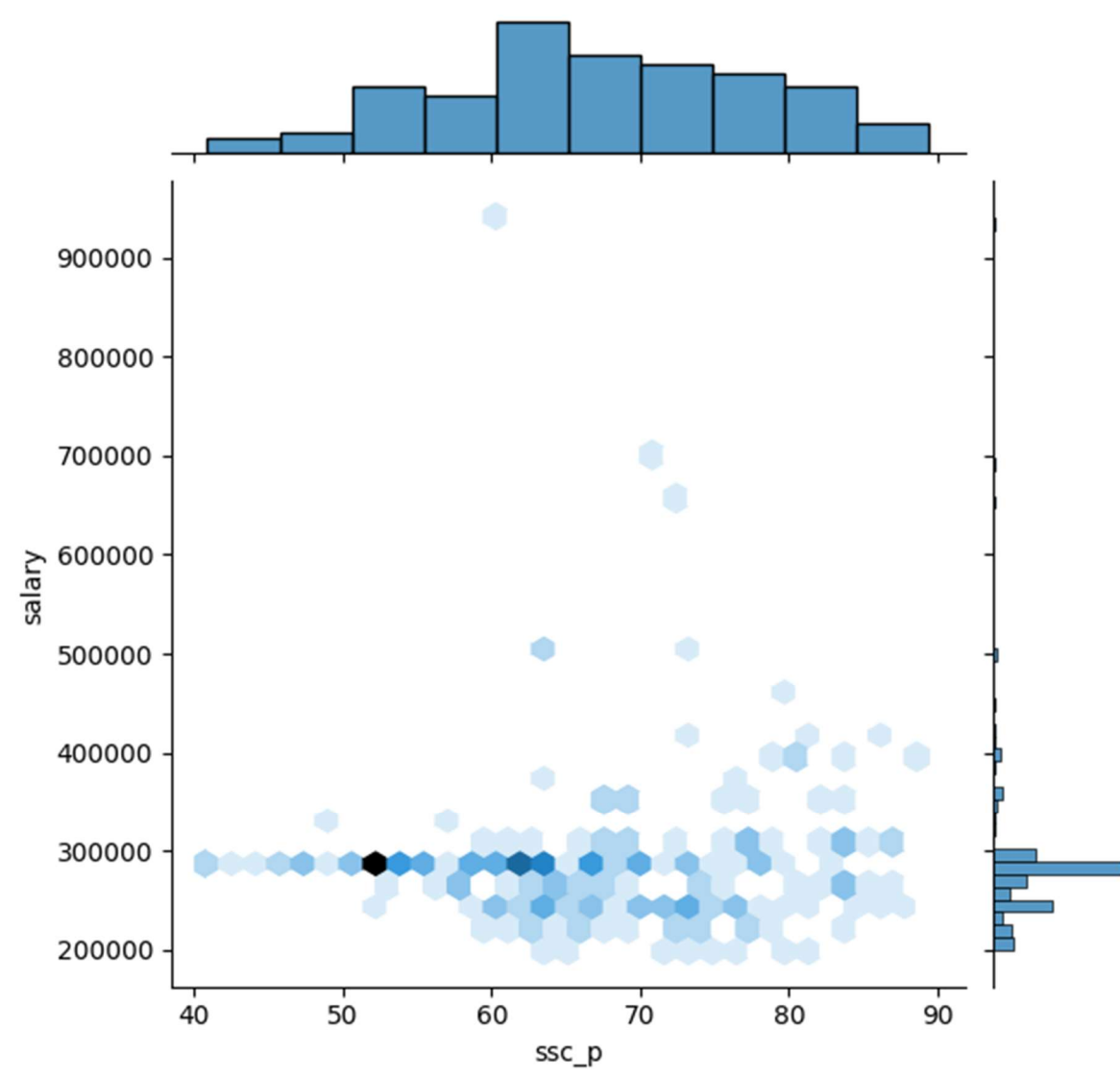
Bivariate Analysis

Bivariate Analysis¶

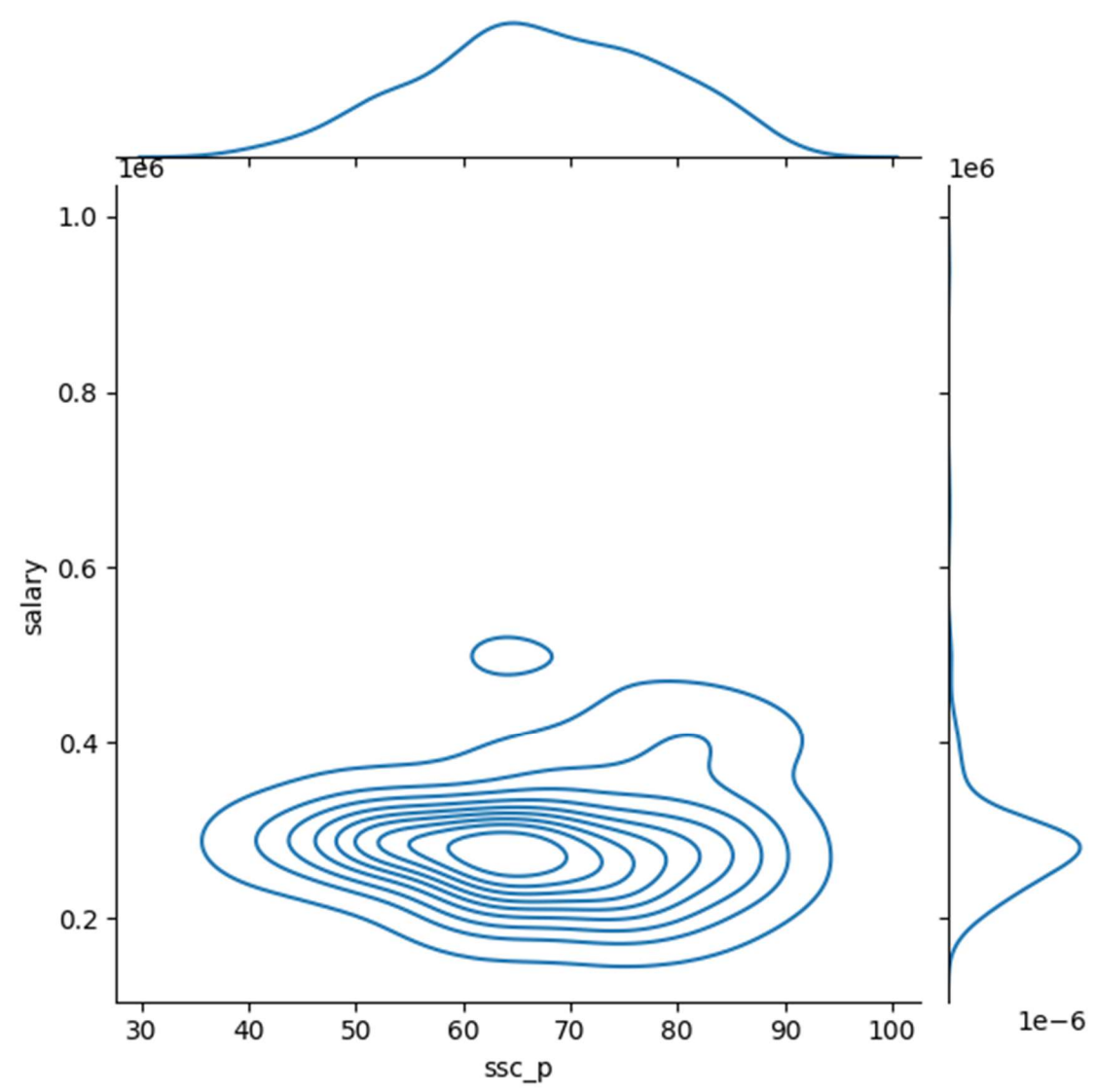
```
[7]:  
sb.jointplot(x='ssc_p',y='salary',data=dataset)  
plt.show()
```



```
[8]:  
sb.jointplot(x='ssc_p',y='salary',data=dataset,kind='hex')  
plt.show()
```



```
[9]:  
sb.jointplot(x='ssc_p',y='salary',data=dataset,kind='kde')  
plt.show()
```



```
[10]:  
sb.pairplot(hue='gender',data=dataset,kind='scatter',palette='husl',diag_kind='kde')  
plt.show()
```

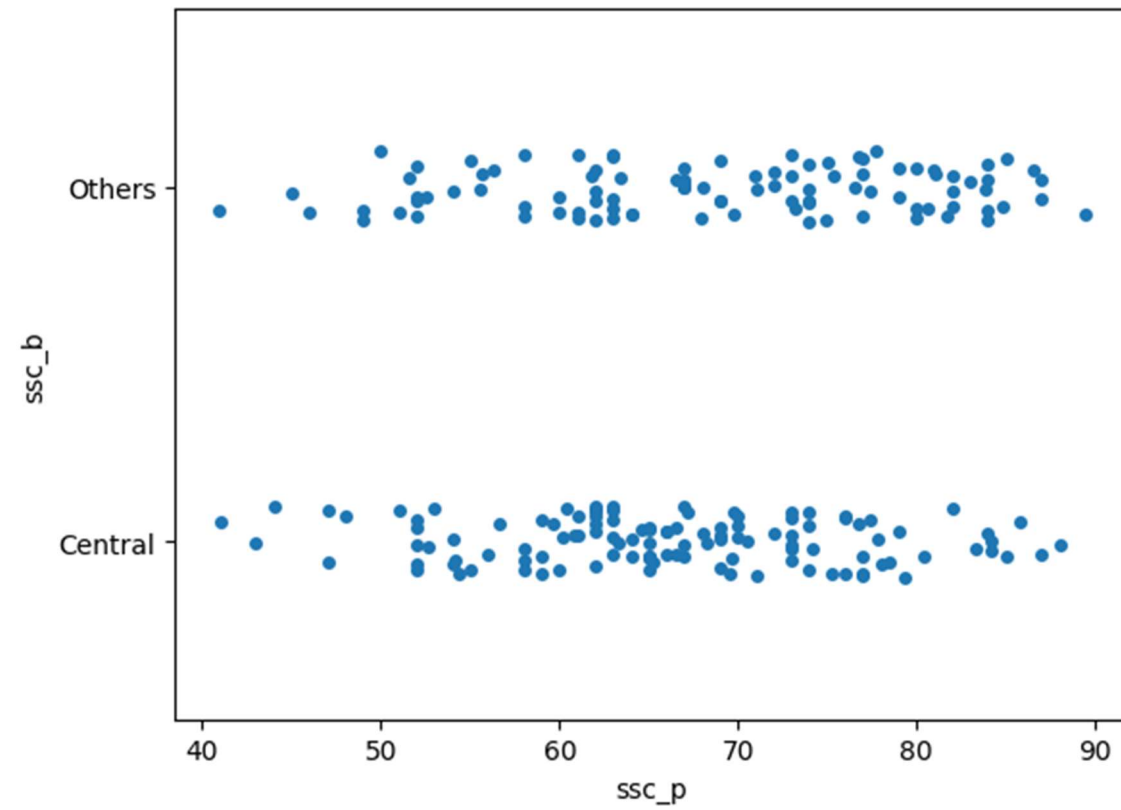




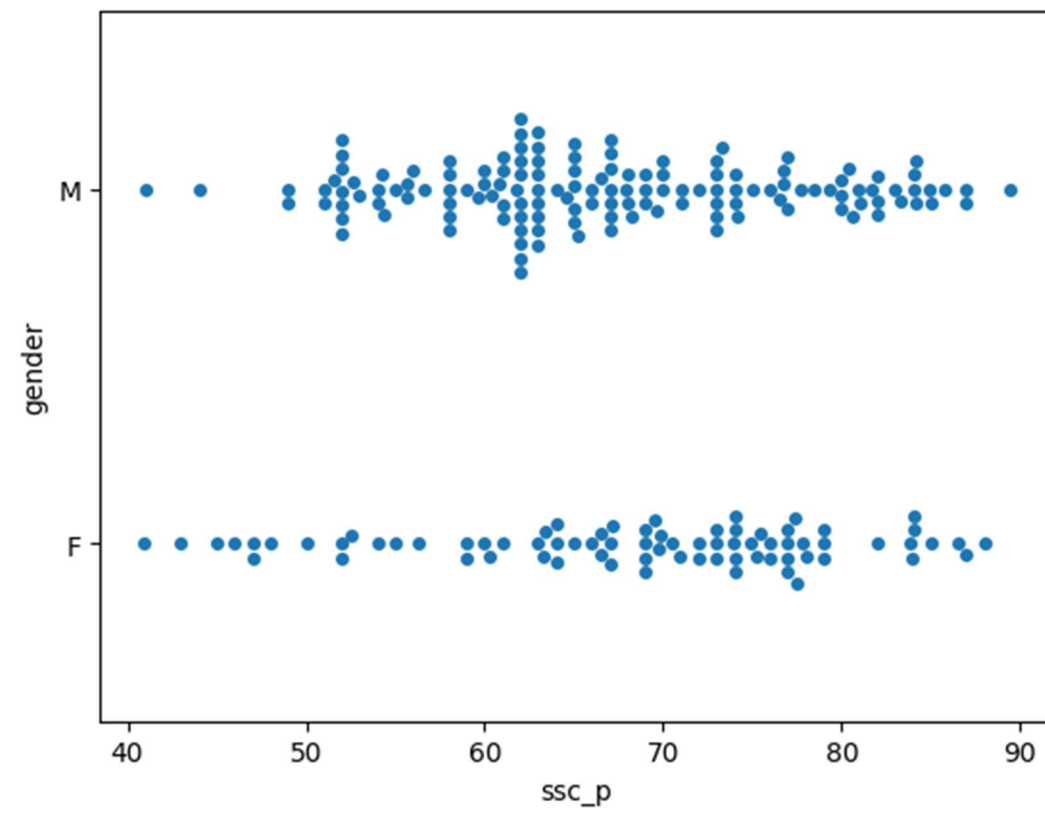
```
# Categorical plot
```

Categorical plot ¶

```
[11]:  
sb.stripplot(x='ssc_p',y='ssc_b',data=dataset)  
plt.show()
```



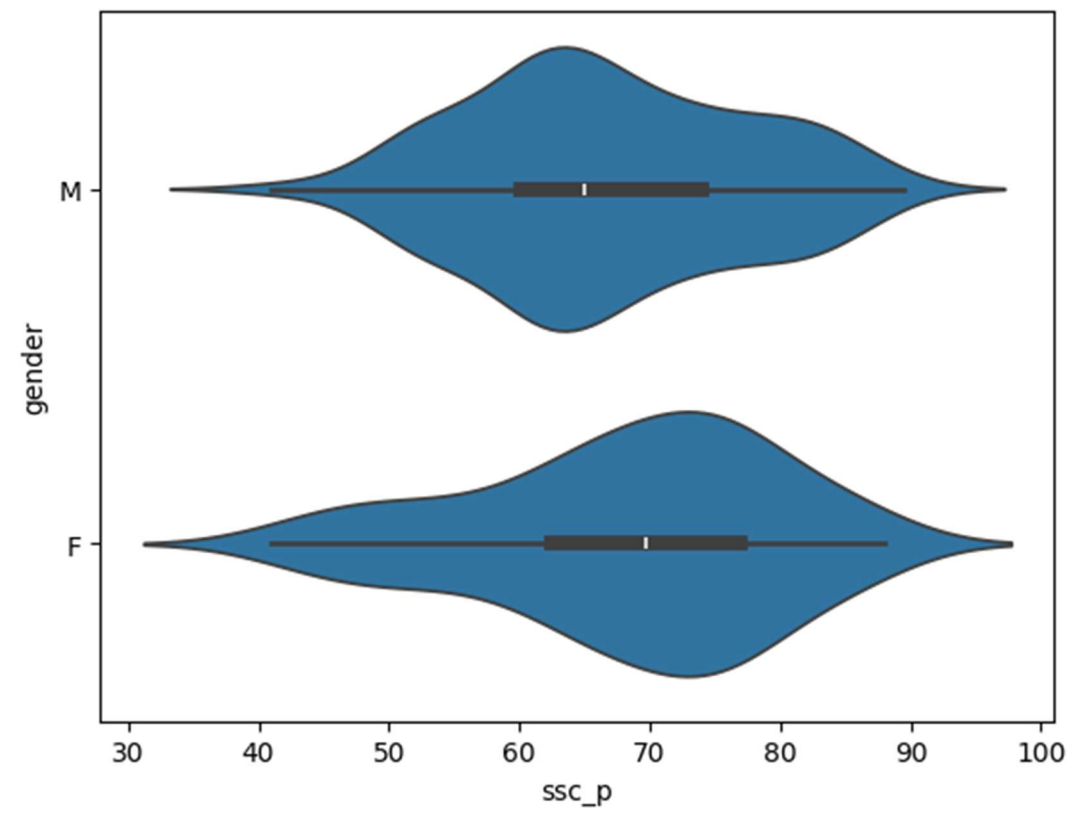
```
[12]:  
sb.swarmplot(x='ssc_p',y='gender',data=dataset)  
plt.show()
```



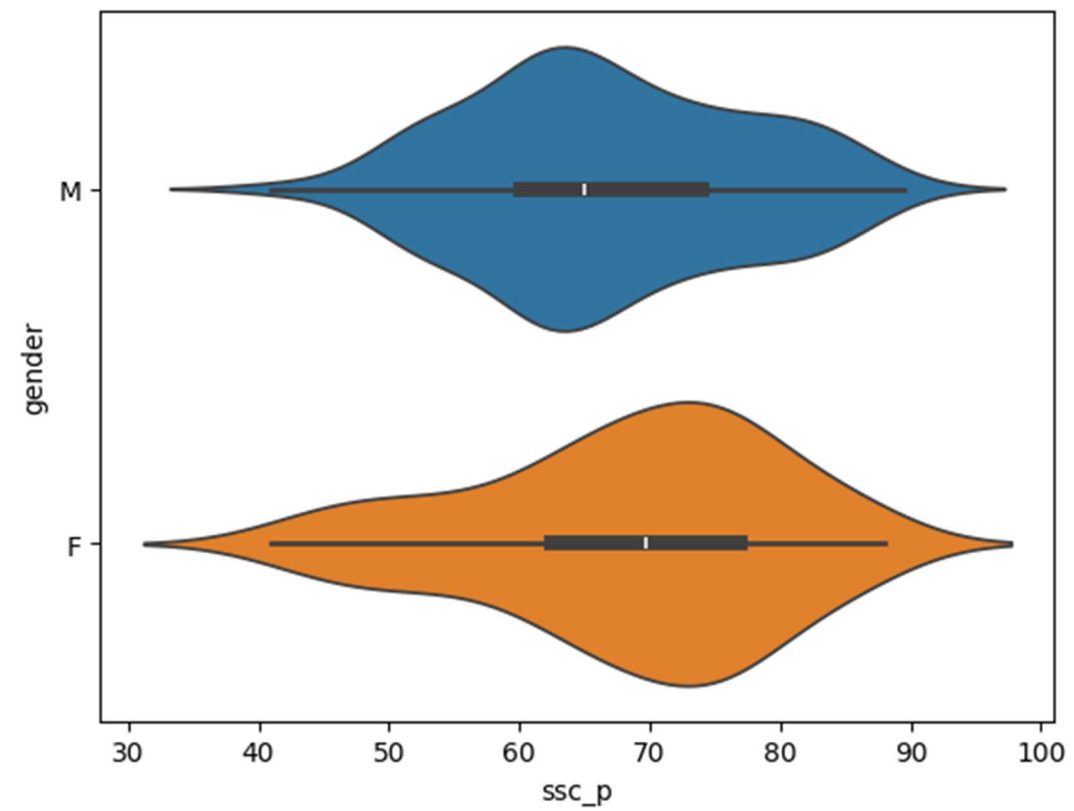
Distribution Observation

Distribution Observation¶

```
[13]:  
sb.violinplot(x='ssc_p',y='gender',data=dataset)  
plt.show()
```

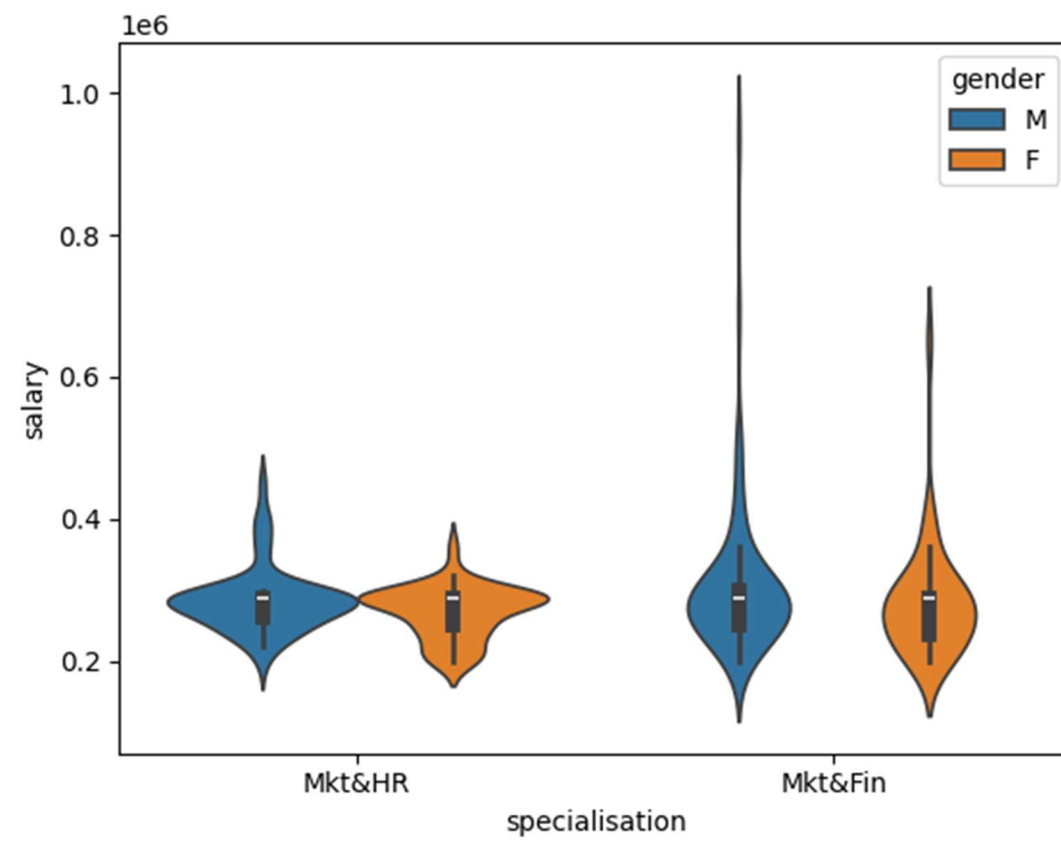


```
[14]:  
sb.violinplot(x='ssc_p',y='gender',data=dataset,hue='gender')  
plt.show()
```

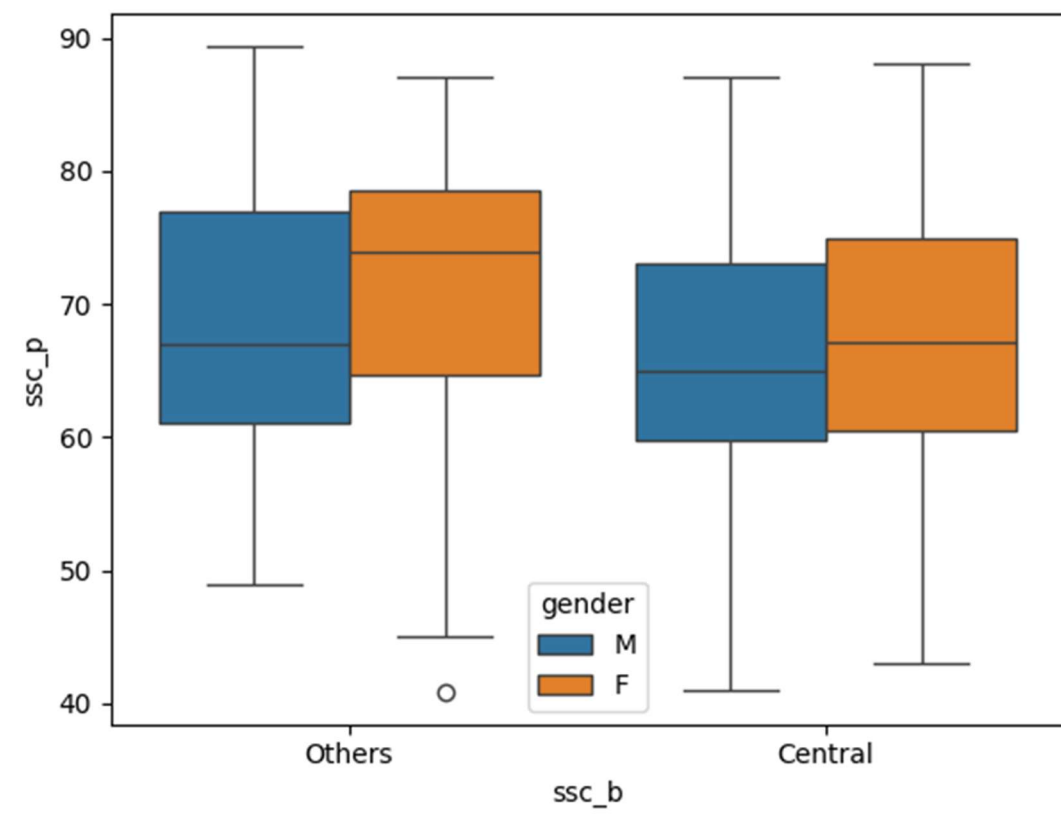


```
[22]:  
sb.violinplot(x='specialisation',y='salary',data=dataset,hue='gender')
```

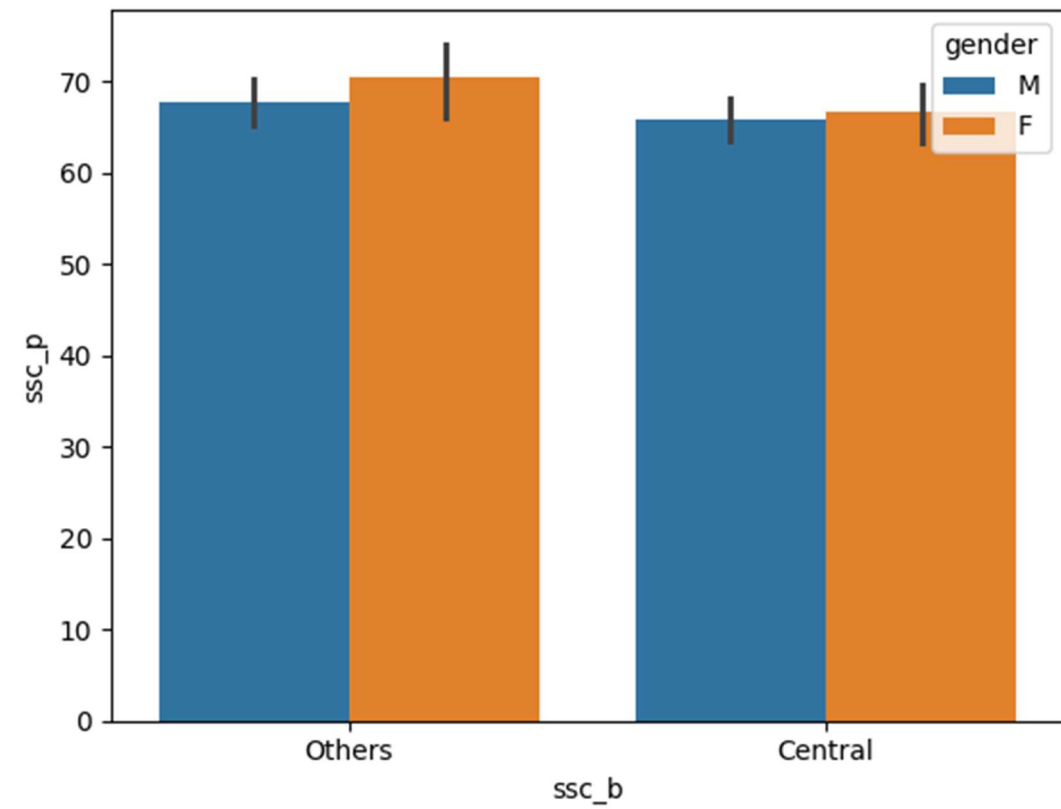
```
plt.show()
```



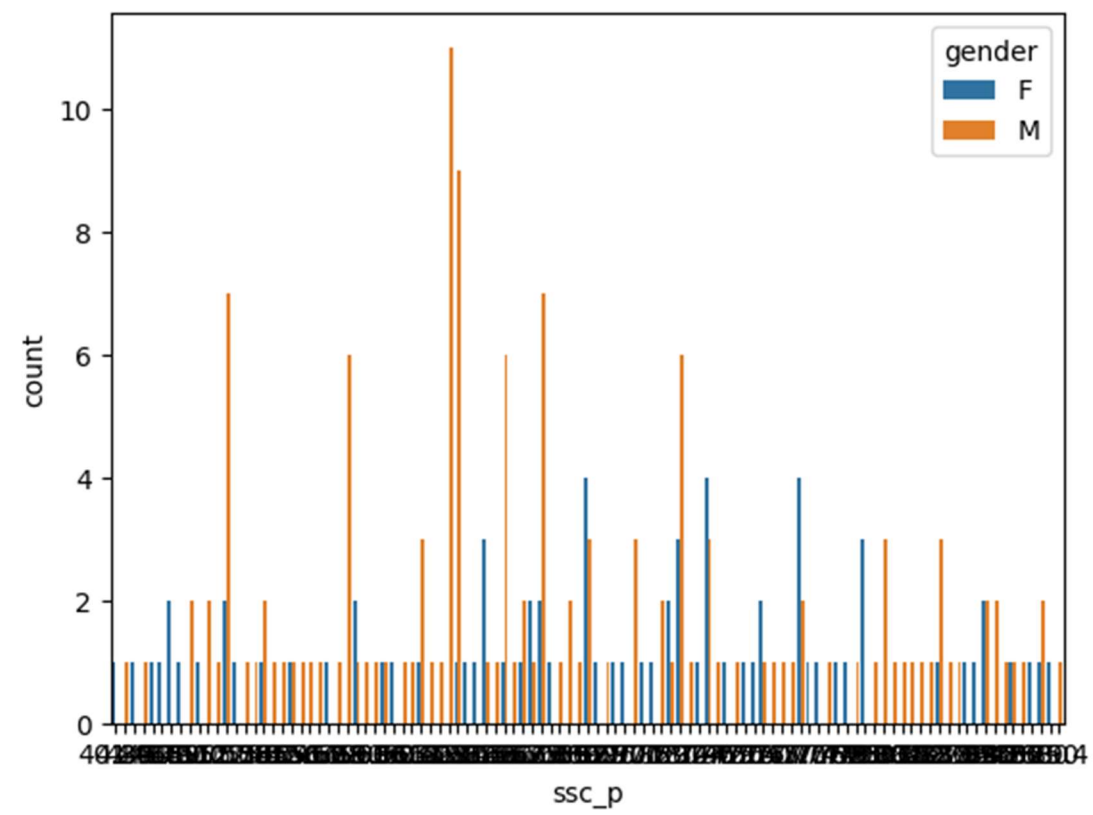
```
[24]:  
sb.boxplot(x='ssc_b',y='ssc_p',data=dataset,hue='gender')  
plt.show()
```



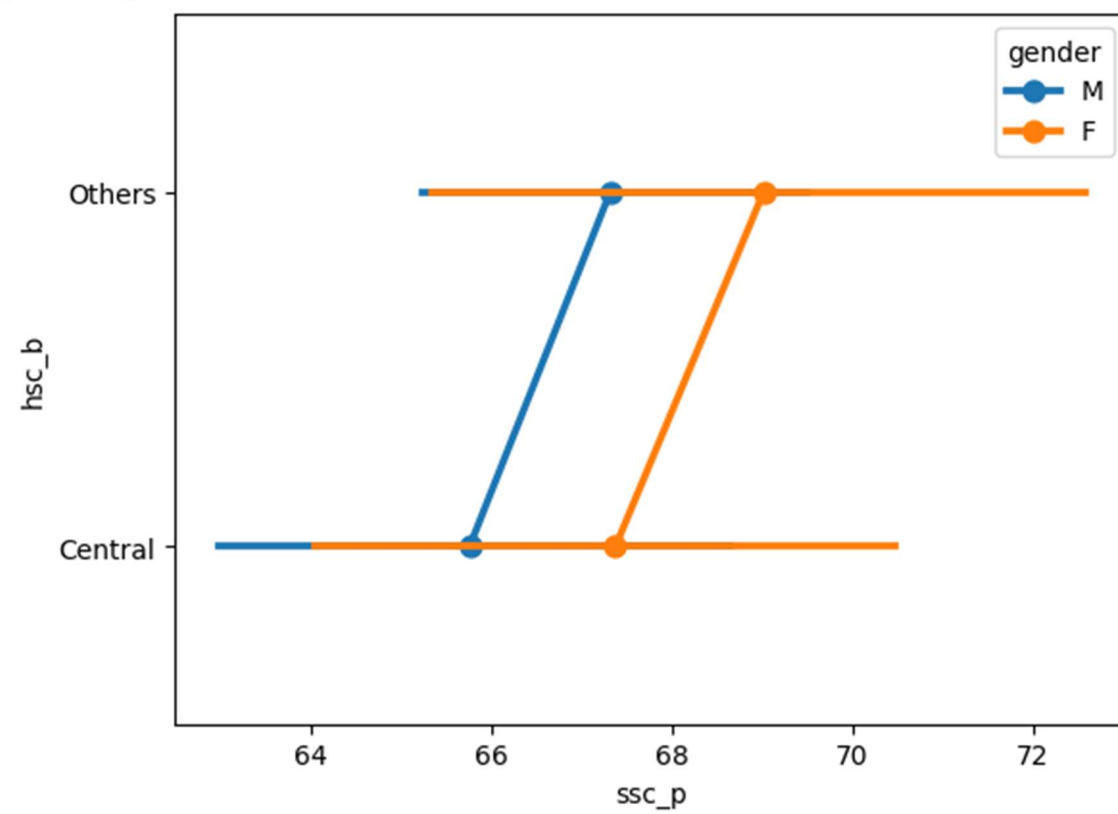
```
[25]:  
sb.barplot(x='ssc_b',y='ssc_p',data=dataset,hue='gender')  
plt.show()
```



```
[17]:  
sb.countplot(x='ssc_p',data=dataset,hue='gender')  
plt.show()
```



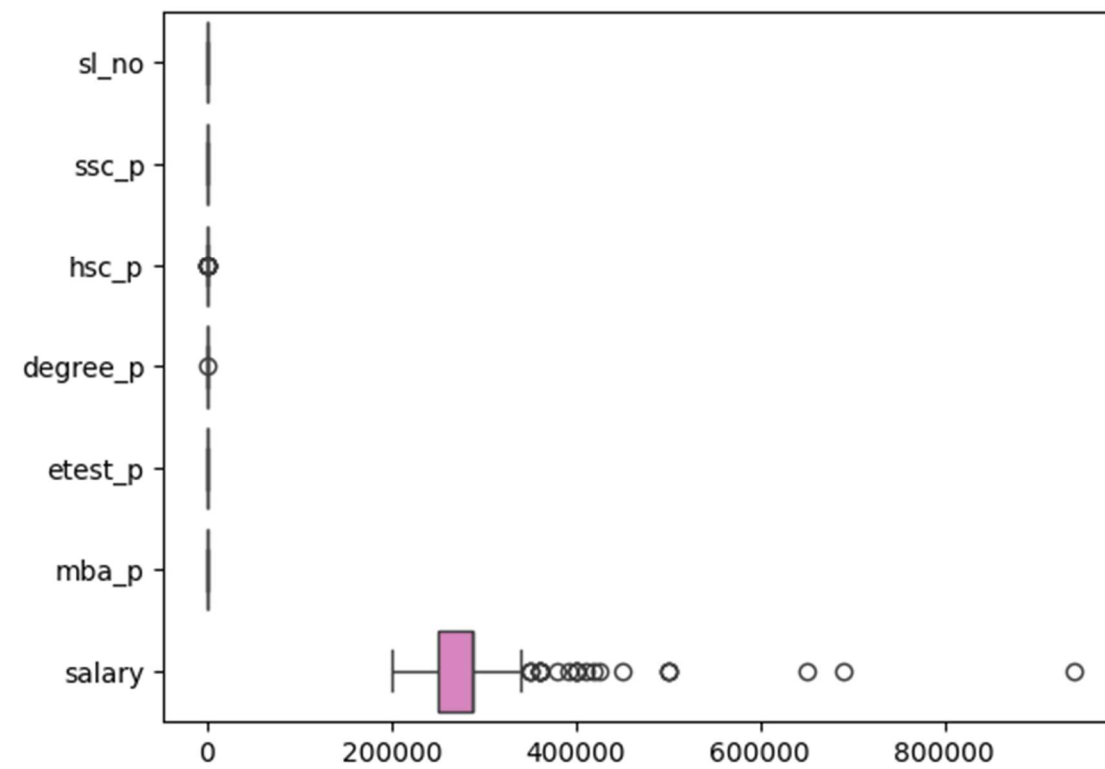
```
[26]:
sb.pointplot(x='ssc_p',y='hsc_b',data=dataset,hue='gender')
plt.show()
```



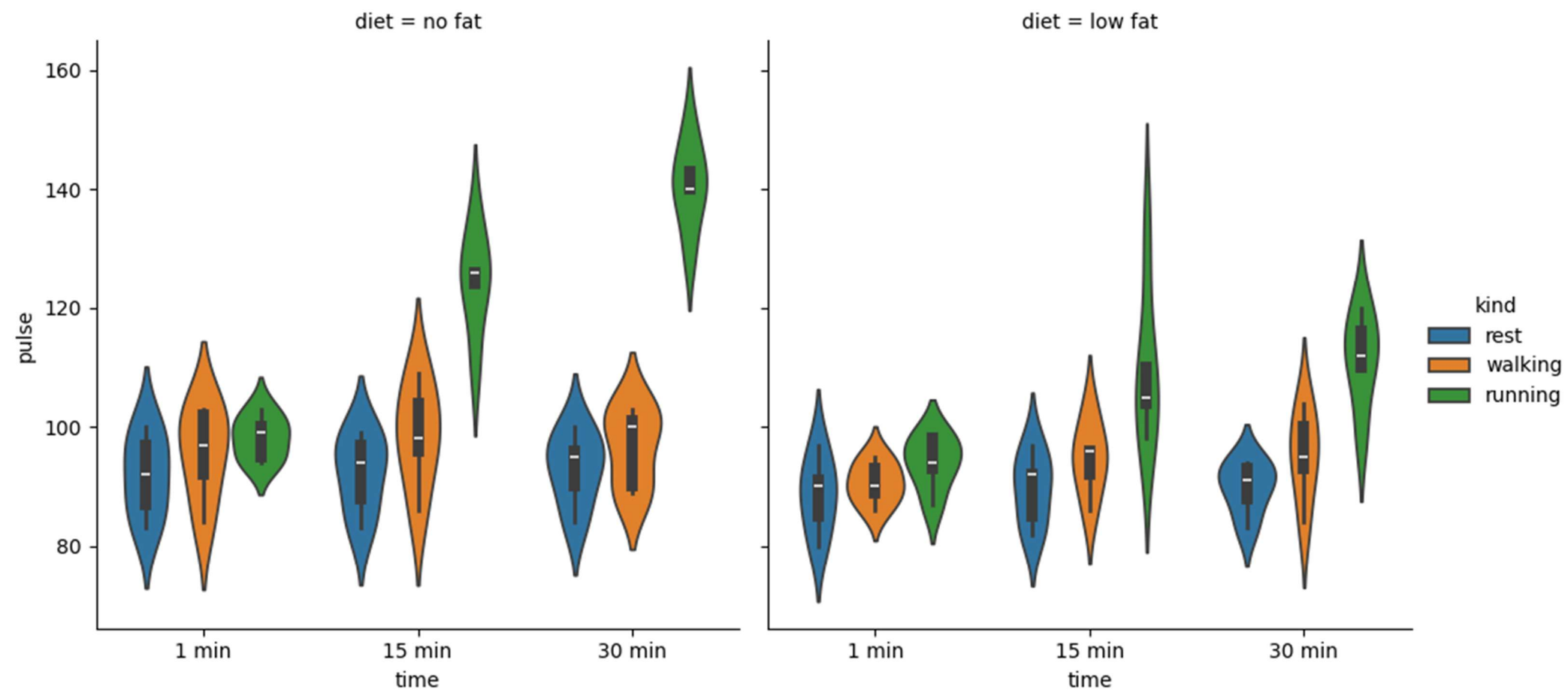
```
[27]:
sb.boxplot(data=dataset,orient='h')
```



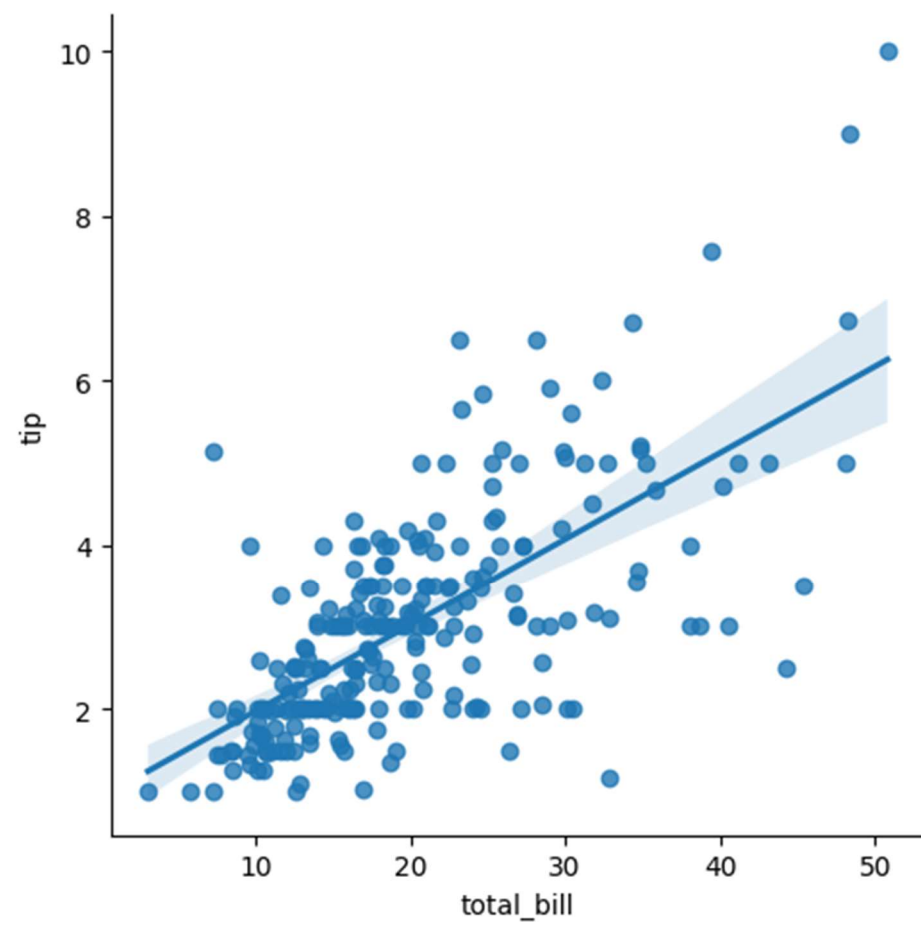
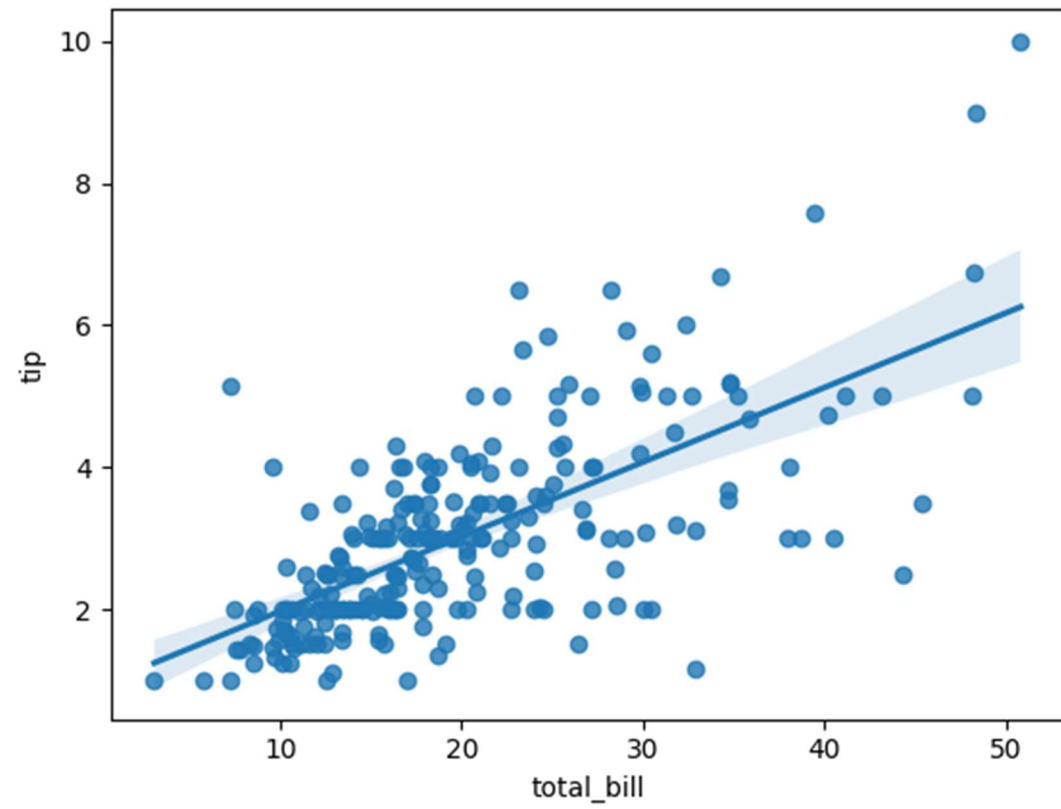
```
plt.show()
```



```
[30]:
import pandas as pd
import seaborn as sb
from matplotlib import pyplot as plt
df=sb.load_dataset('exercise')
sb.catplot(x='time',y='pulse',hue='kind',kind='violin',col='diet',data=df)
[30]:
<seaborn.axisgrid.FacetGrid at 0x185dcffd190>
```



```
[31]:
import pandas as pd
import seaborn as sb
from matplotlib import pyplot as plt
df=sb.load_dataset('tips')
sb.regplot(x='total_bill',y='tip',data=df)
sb.lmplot(x='total_bill',y='tip',data=df)
plt.show()
```



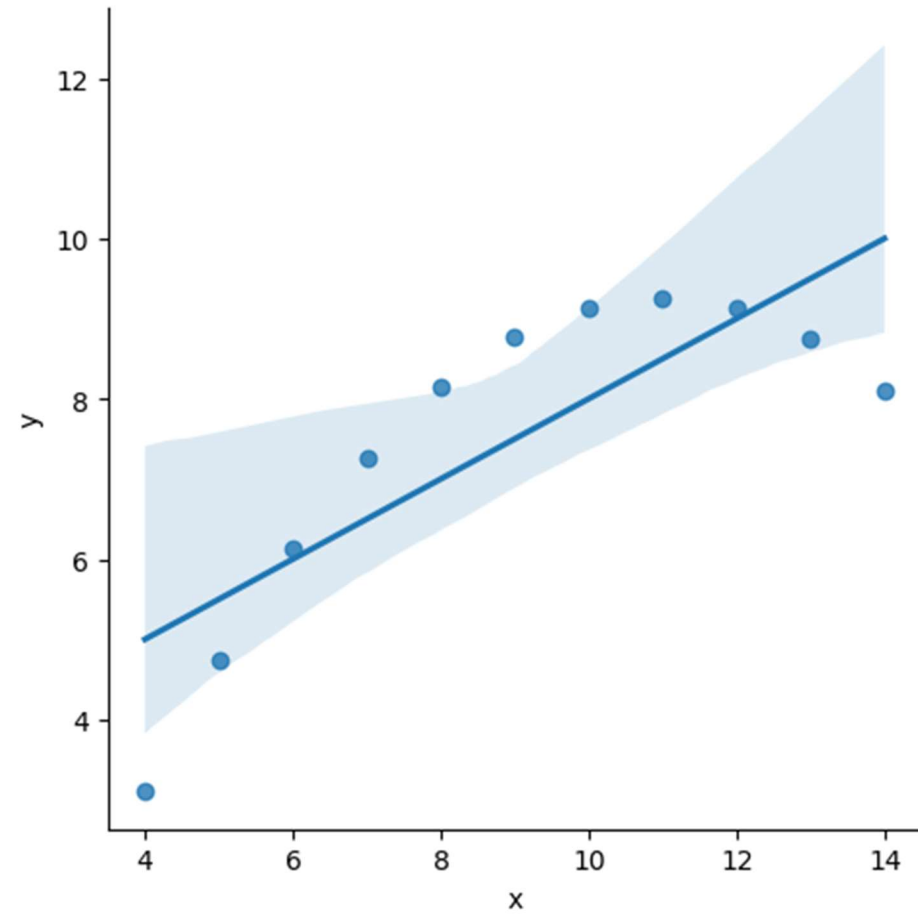
[34]:
import pandas **as** pd

```

import seaborn as sb
from matplotlib import pyplot as plt
df=sb.load_dataset('anscombe')

sb.lmplot(x='x',y='y',data=df.query("dataset=='II'"))
plt.show()

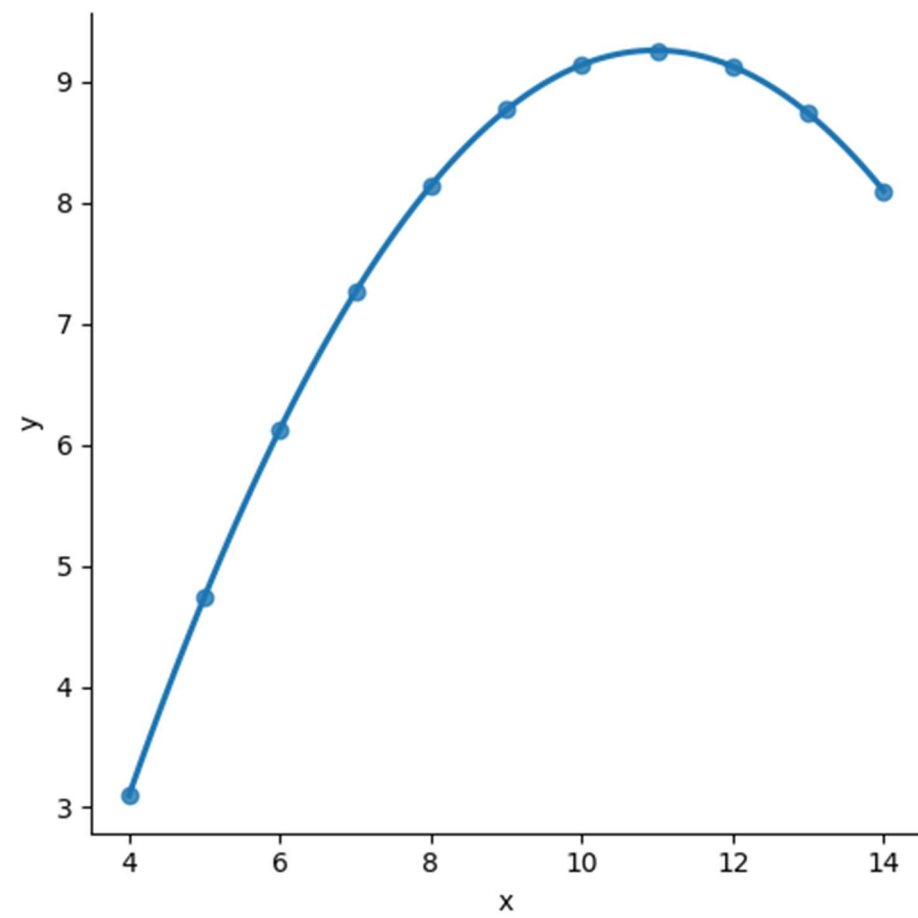
```



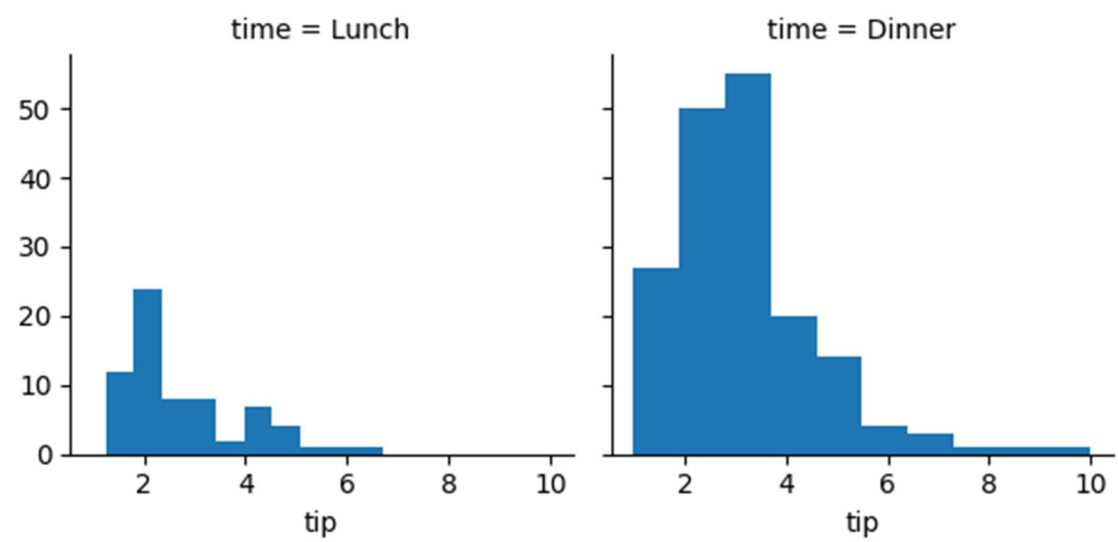
```

[35]:
import pandas as pd
import seaborn as sb
from matplotlib import pyplot as plt
df=sb.load_dataset('anscombe')
sb.lmplot(x='x',y='y',data=df.query("dataset=='II'"),order=2)
plt.show()

```



```
[36]:
import pandas as pd
import seaborn as sb
from matplotlib import pyplot as plt
df=sb.load_dataset('tips')
g=sb.FacetGrid(df,col='time')
g.map(plt.hist,'tip')
plt.show()
```



```
[37]:  
import pandas as pd  
import seaborn as sb  
from matplotlib import pyplot as plt  
df=sb.load_dataset('iris')  
g=sb.PairGrid(df)  
g.map(plt.scatter)  
plt.show()
```