Author : Sanjoy Biswas

Email :sanjoy.eee32@gmail.com

Topic : Seaborn : Data Visualization

Important Features of Seaborn

Seaborn is built on top of Python's core visualization library Matplotlib. It is meant to serve as a complement, and not a replacement. However, Seaborn comes with some very important features. Let us see a few of them here. The features help in −

Built in themes for styling matplotlib graphics

Visualizing univariate and bivariate data

Fitting in and visualizing linear regression models

Plotting statistical time series data

Seaborn works well with NumPy and Pandas data structures

It comes with built in themes for styling Matplotlib graphics

In most cases, you will still use Matplotlib for simple plotting. The knowledge of Matplotlib is recommended to tweak Seaborn's default plots.

Import Necessary Libraries

In [3]:

```
import seaborn as sns
import matplotlib.pyplot as plt
import numpy as np

%matplotlib inline
```

In [4]:

```
tips = sns.load_dataset('tips')
```

In [5]:

```
tips.head()
```

Out[5]:

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

Numerical Data plot

Distribution plot

In [14]:

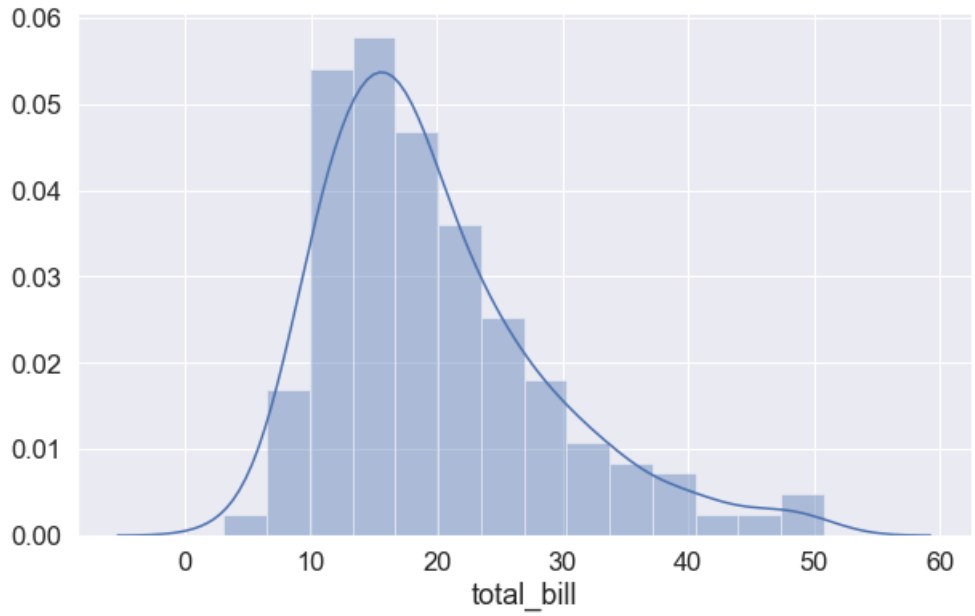
```
sns.set(style = 'darkgrid', rc = {'figure.figsize':(10, 6)}, font_scale = 1.5)
```

In [16]:

```
sns.distplot(tips['total_bill'])
```

Out[16]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754c3e03c8>
```



In [17]:

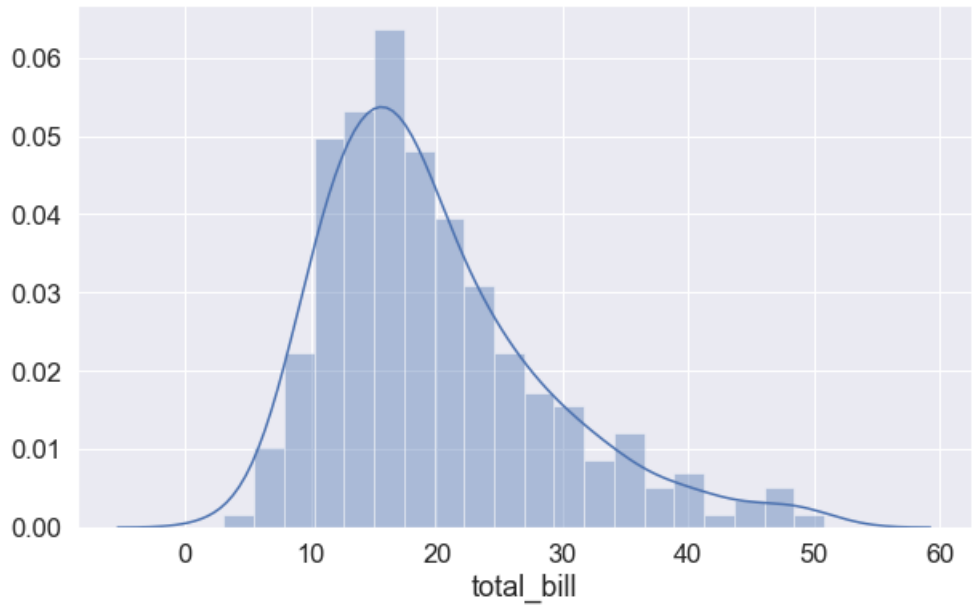
```
sns.distplot??
```

In [19]:

```
sns.distplot(tips['total_bill'], bins = 20)
```

Out[19]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754c679648>
```

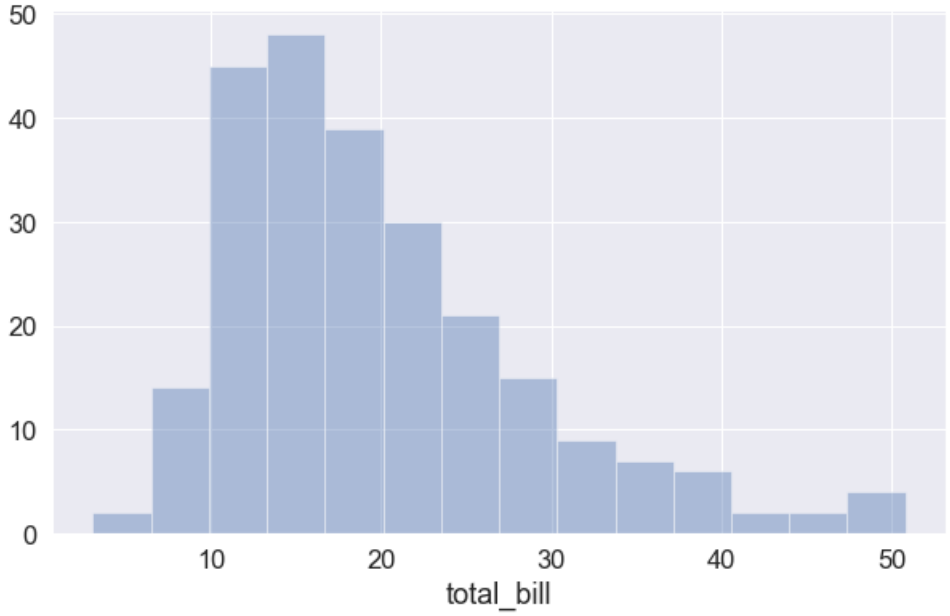


In [20]:

```
sns.distplot(tips['total_bill'], kde = False)
```

Out[20]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754c7c55c8>
```

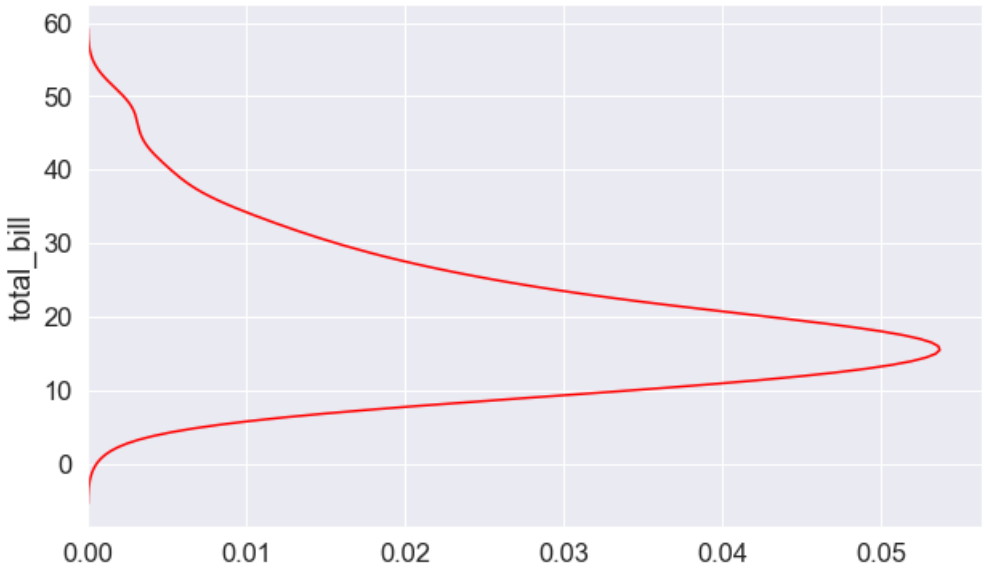


In [23]:

```
sns.distplot(tips['total_bill'], hist = False, vertical= True, color = 'red')
```

Out[23]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754c8bf648>
```



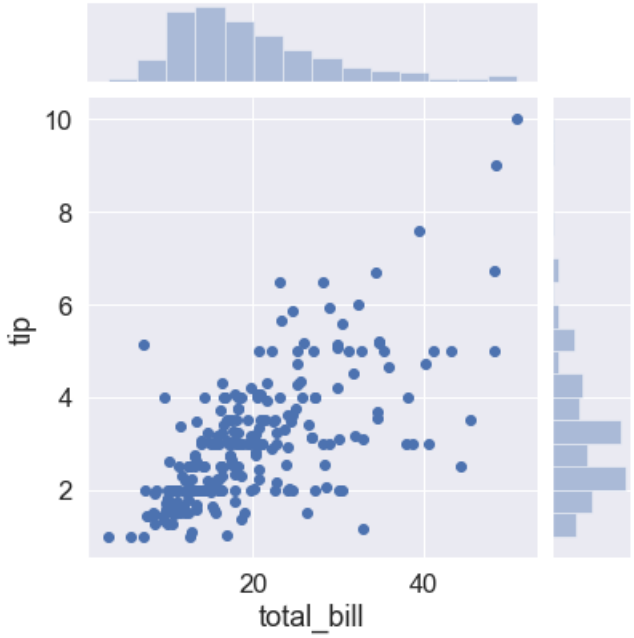
Join Plot

In [25]:

```
sns.jointplot(x = 'total_bill', y = 'tip', data = tips)
```

Out[25]:

```
<seaborn.axisgrid.JointGrid at 0x1754cacb248>
```

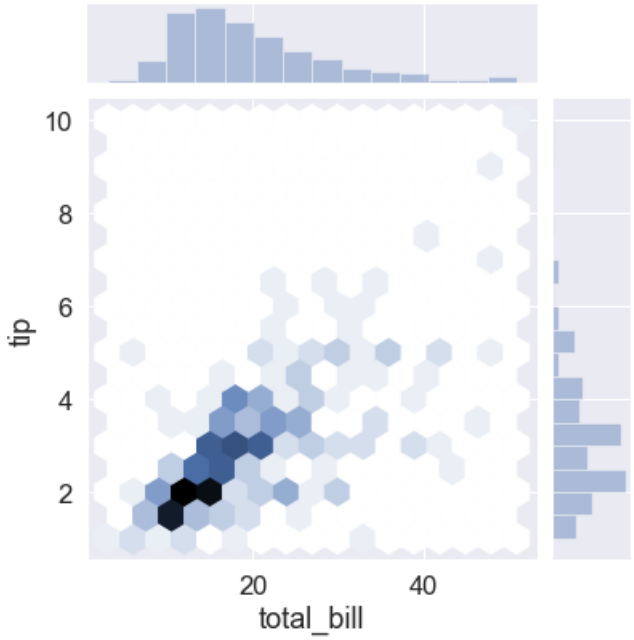


In [26]:

```
sns.jointplot(x = 'total_bill', y = 'tip', data = tips, kind = 'hex')
```

Out[26]:

```
<seaborn.axisgrid.JointGrid at 0x1754ca8ebc8>
```

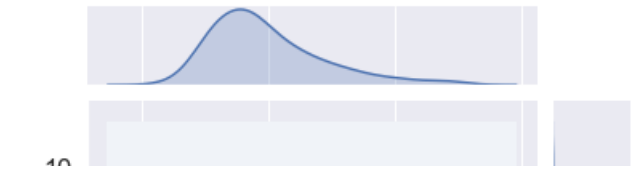


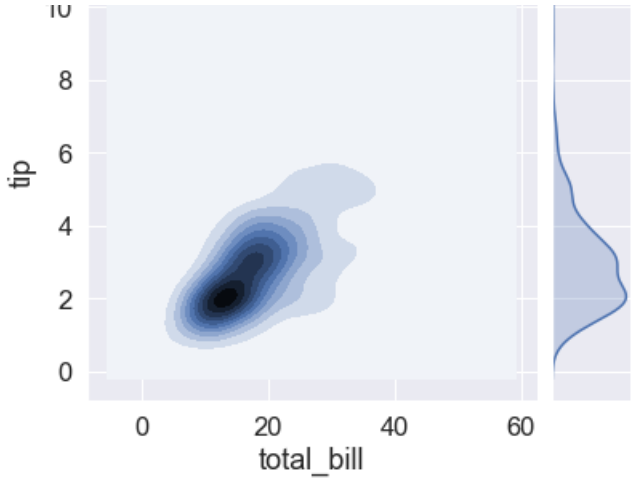
In [27]:

```
sns.jointplot(x = 'total_bill', y = 'tip', data = tips, kind = 'kde')
```

Out[27]:

```
<seaborn.axisgrid.JointGrid at 0x1754dbd7708>
```



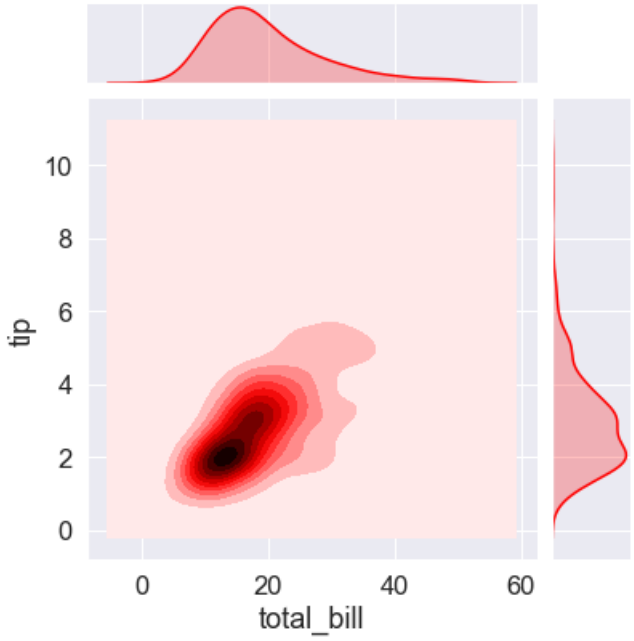


In [28]:

```
sns.jointplot(x = 'total_bill', y = 'tip', data = tips, kind = 'kde', color = 'red')
```

Out[28]:

```
<seaborn.axisgrid.JointGrid at 0x1754dcdbb88>
```



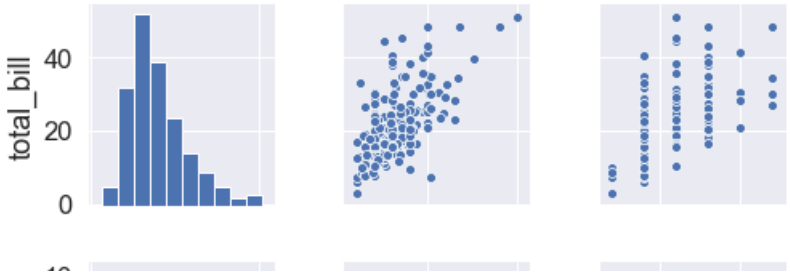
Pair Plot

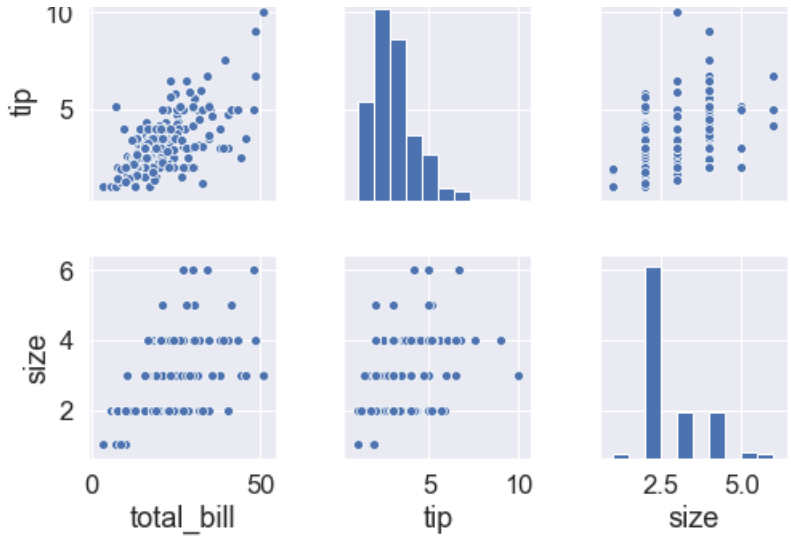
In [29]:

```
sns.pairplot(tips)
```

Out[29]:

```
<seaborn.axisgrid.PairGrid at 0x1754dbaacc8>
```





In [30]:

```
tips.head()
```

Out[30]:

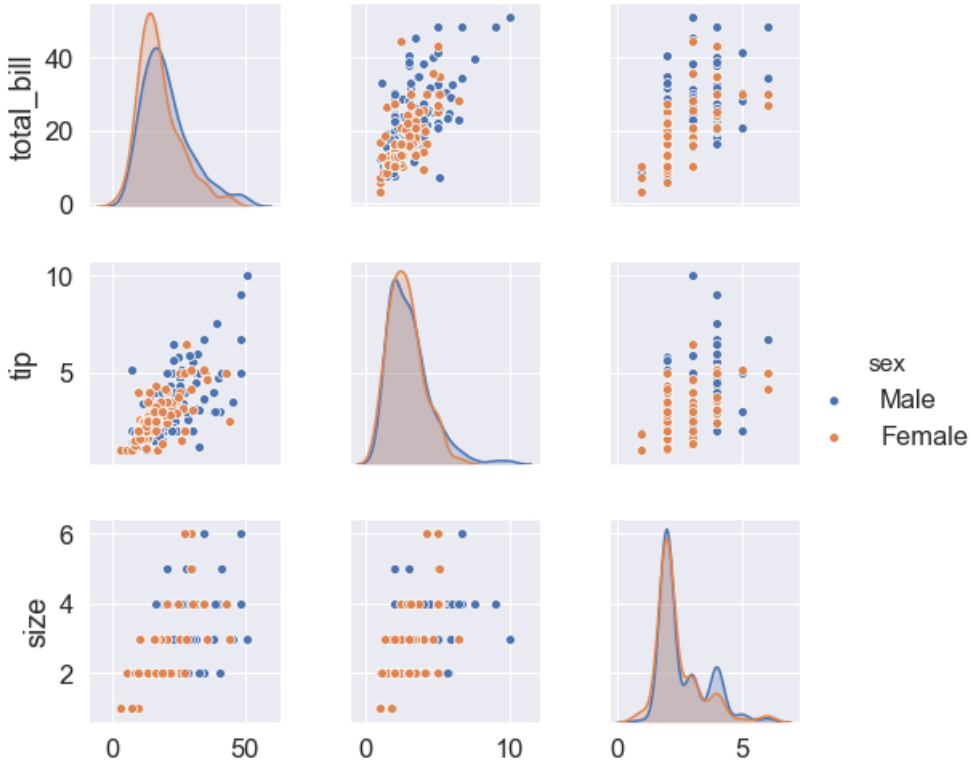
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

In [31]:

```
sns.pairplot(tips, hue = 'sex')
```

Out[31]:

```
<seaborn.axisgrid.PairGrid at 0x1754e2eb208>
```



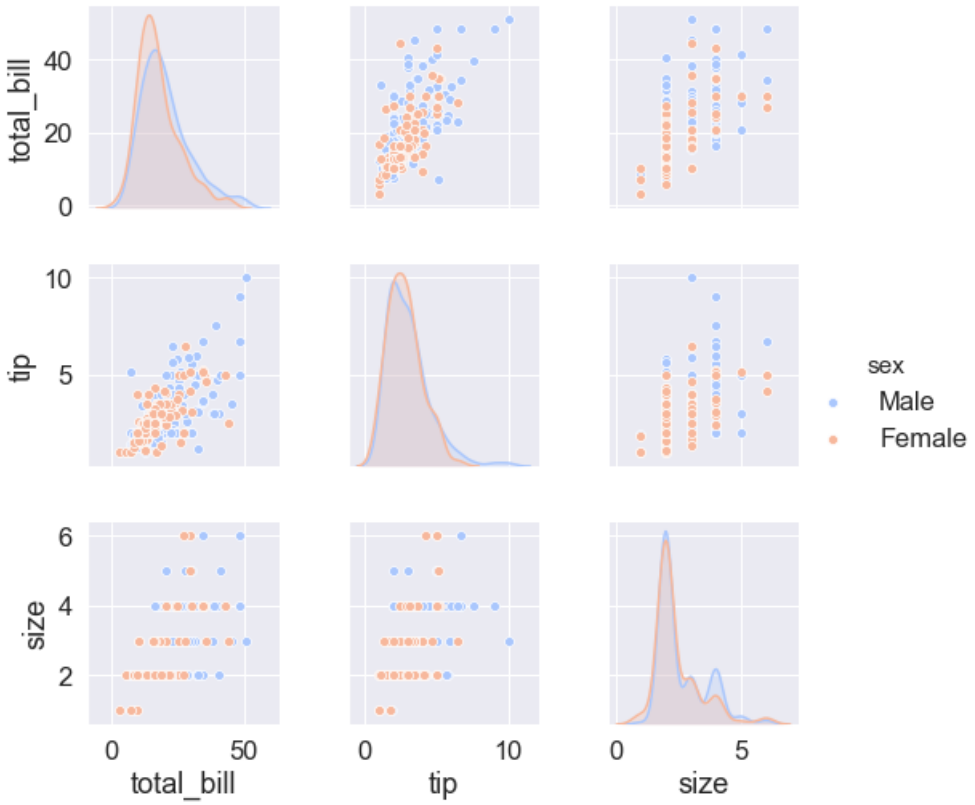
total_bill tip size

In [34]:

```
sns.pairplot(tips, hue = 'sex', palette = 'coolwarm')
```

Out[34]:

```
<seaborn.axisgrid.PairGrid at 0x1754e959a08>
```



Categorical Ploting

In [35]:

```
tips.head()
```

Out[35]:

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

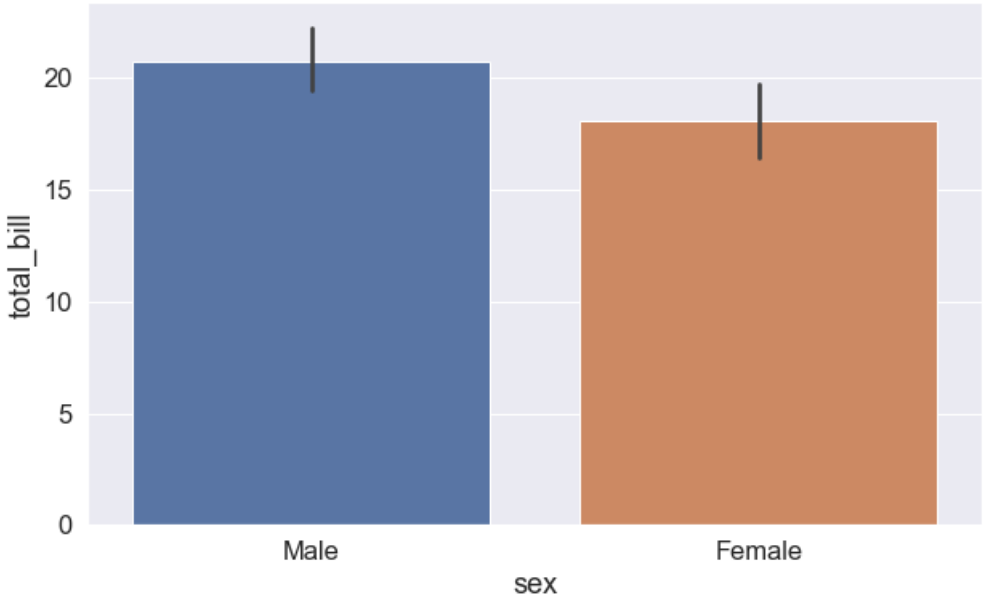
Barplot

In [36]:

```
sns.barplot(x ='sex' , y = 'total_bill' , data = tips)
```

Out[36]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754ed73248>
```

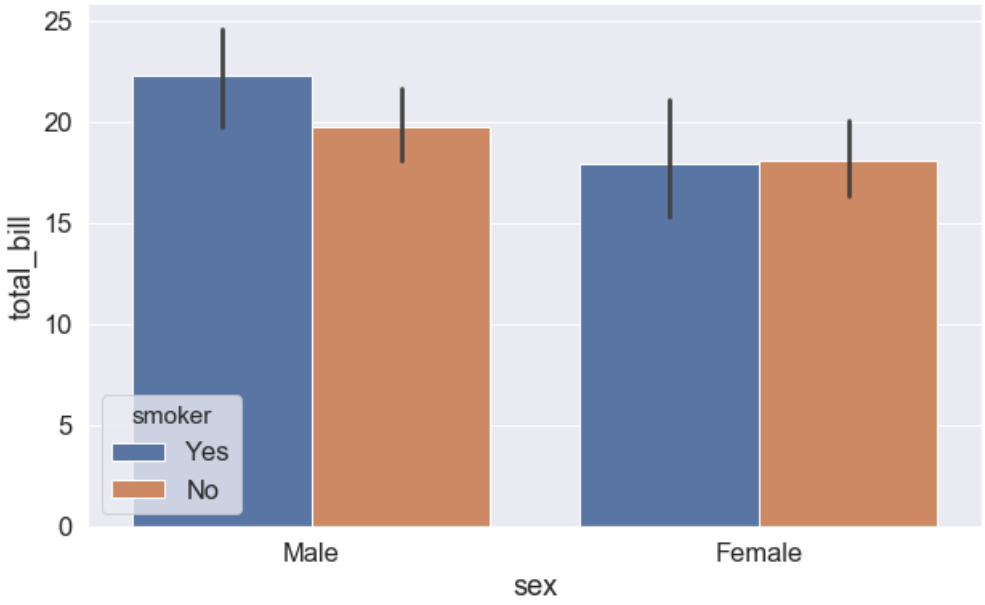


In [37]:

```
sns.barplot(x ='sex' , y = 'total_bill' , data = tips, hue = 'smoker')
```

Out[37]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754f0ded48>
```

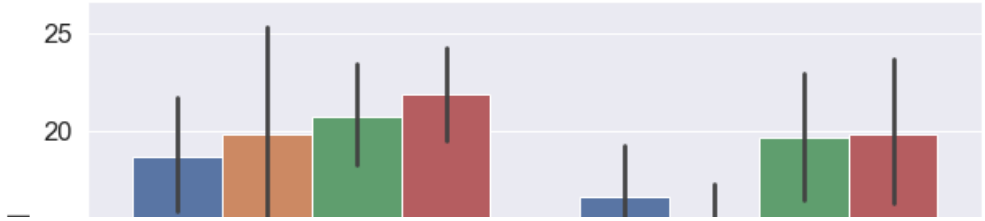


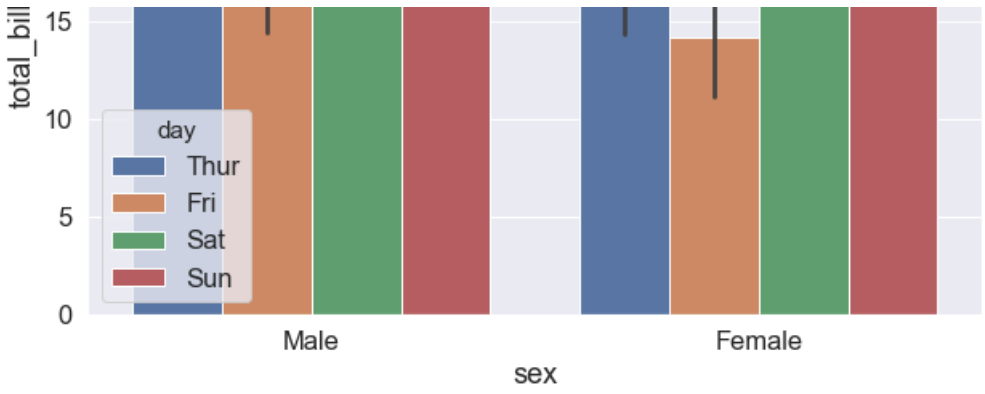
In [38]:

```
sns.barplot(x ='sex' , y = 'total_bill' , data = tips, hue = 'day')
```

Out[38]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754ee6e6c8>
```



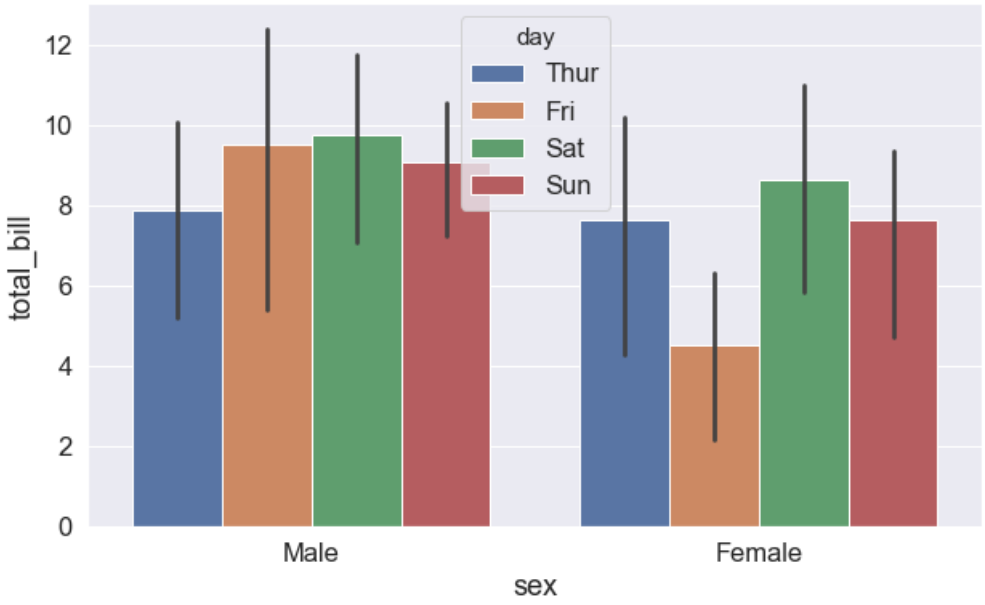


In [39]:

```
sns.barplot(x ='sex' , y = 'total_bill' , data = tips, hue = 'day', estimator = np.std)
```

Out[39]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754ee54548>
```



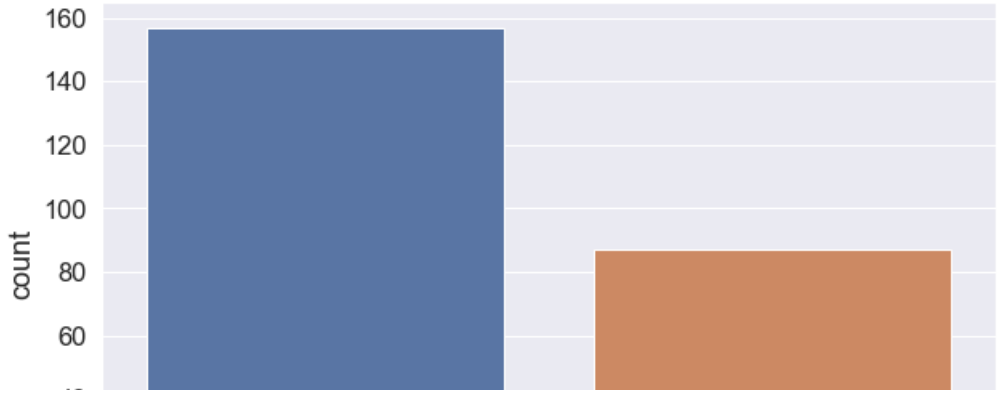
Count Plot

In [41]:

```
sns.countplot(x = 'sex', data = tips)
```

Out[41]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754f1a2508>
```

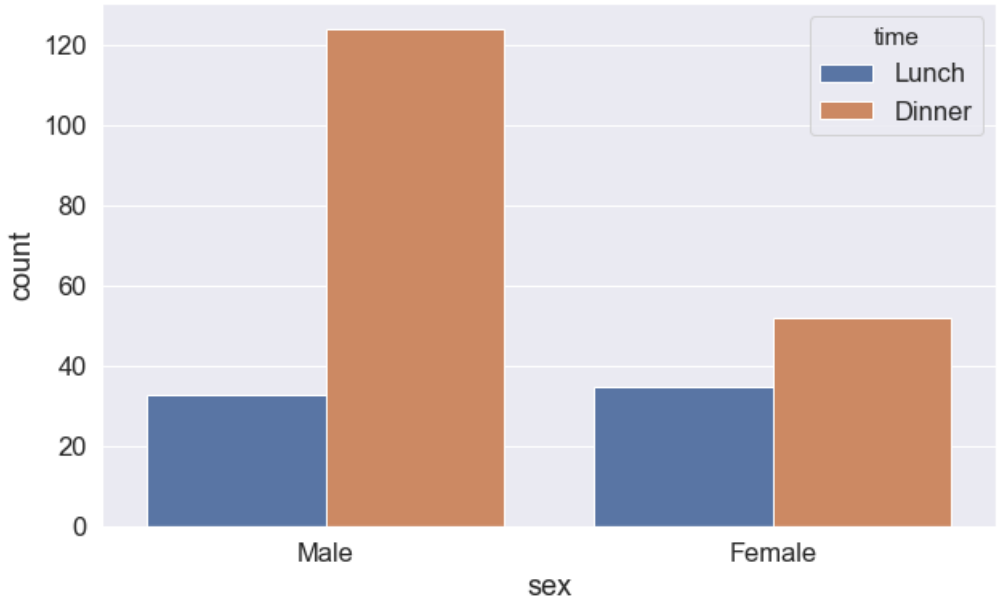


In [42]:

```
sns.countplot(x = 'sex', data = tips, hue = 'time')
```

Out[42]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754f1f6048>
```

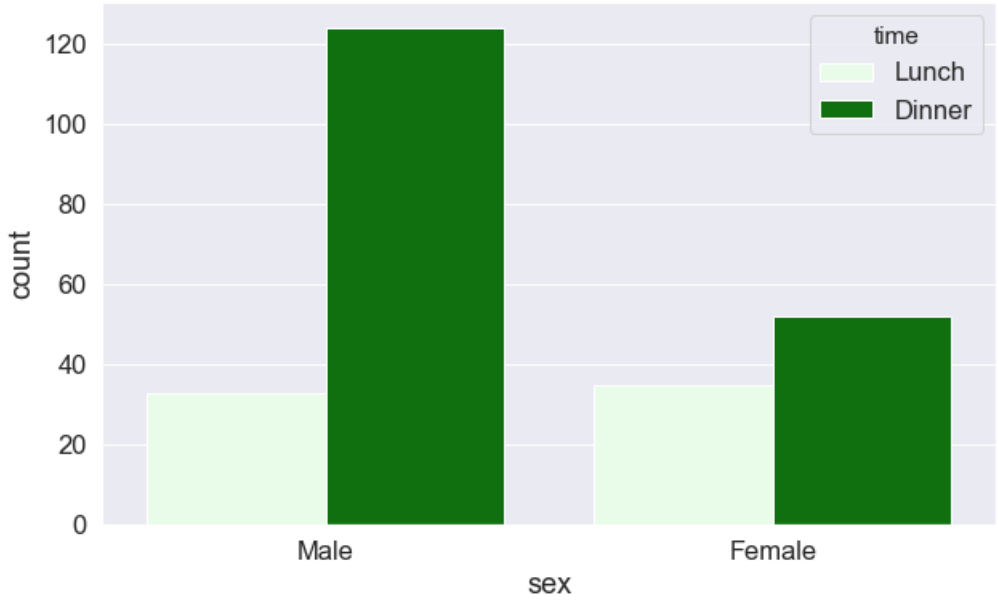


In [44]:

```
sns.countplot(x = 'sex', data = tips, hue = 'time', color = 'green')
```

Out[44]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17550236348>
```

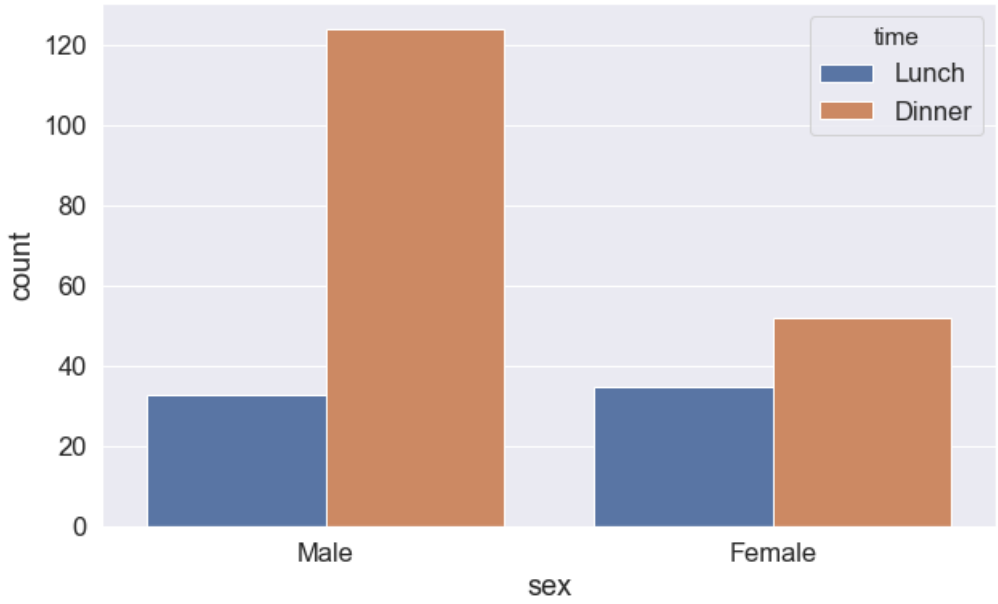


In [45]:

```
sns.countplot(x = 'sex', data = tips, hue = 'time')
```

Out[45]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x175502a79c8>
```



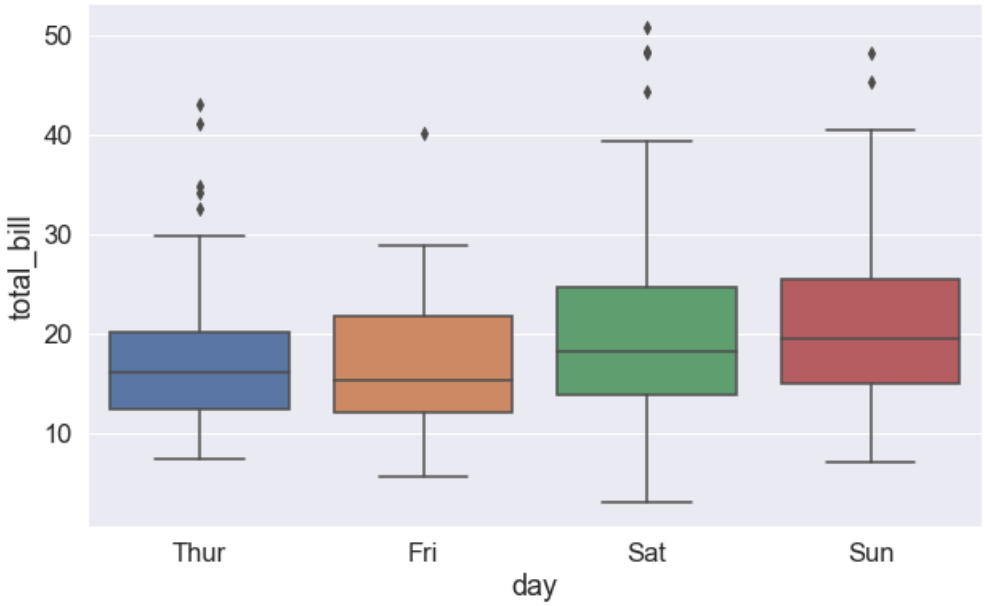
Box Plot

In [47]:

```
sns.boxplot(x = 'day', y = 'total_bill', data = tips)
```

Out[47]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x1754a7d0808>
```

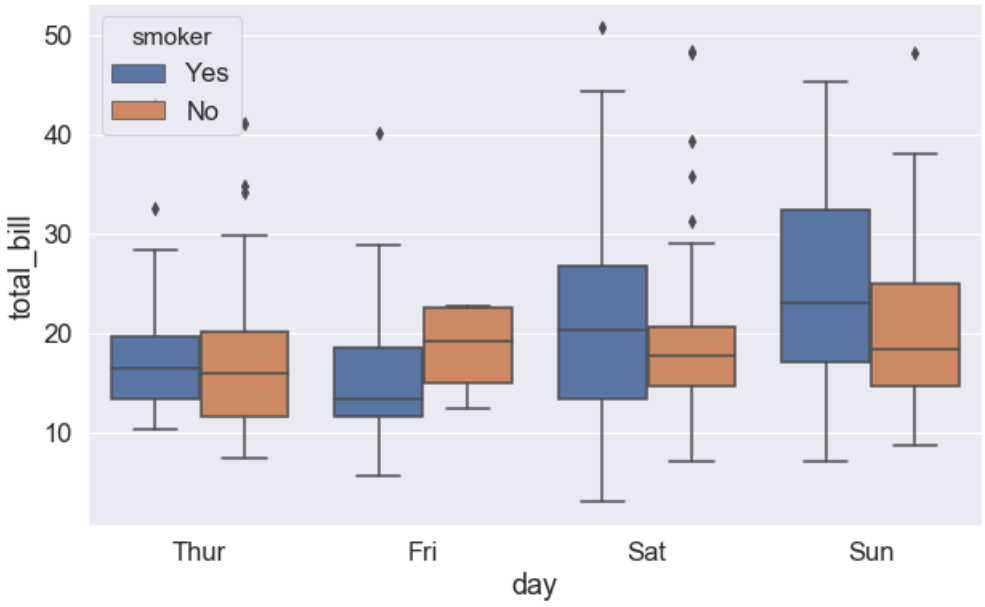


In [50]:

```
sns.boxplot(x = 'day', y = 'total_bill', data = tips, hue = 'smoker')
```

Out[50]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x175505e1908>
```



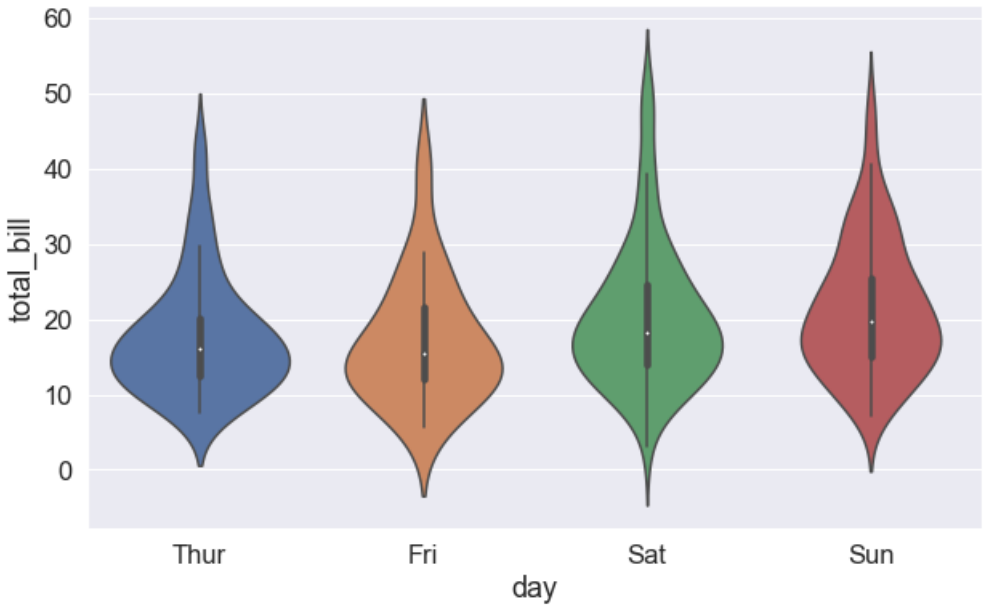
Violin Plot

In [51]:

```
sns.violinplot(x = 'day', y = 'total_bill', data = tips)
```

Out[51]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17550870308>
```

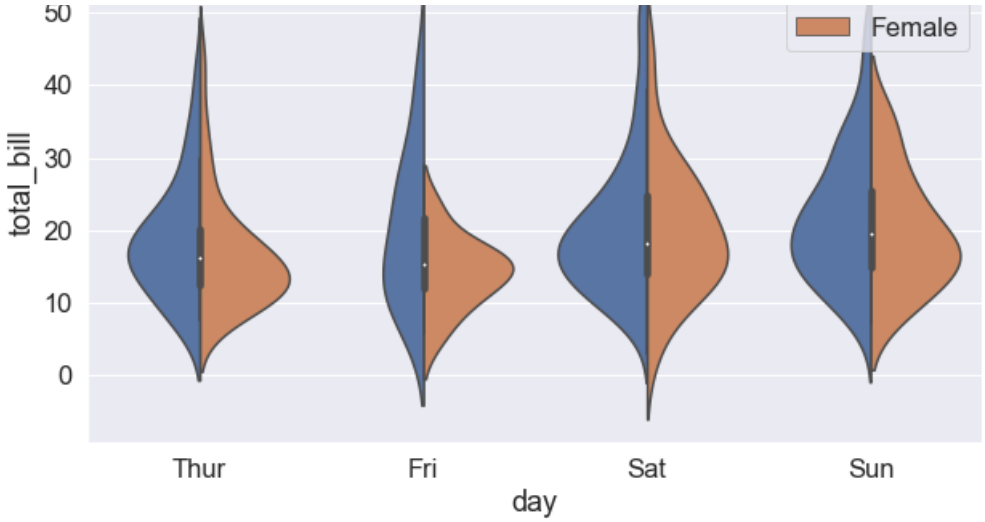


In [54]:

```
sns.violinplot(x = 'day', y = 'total_bill', data = tips, hue ='sex', split = True)
```

Out[54]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x175519ffc08>
```



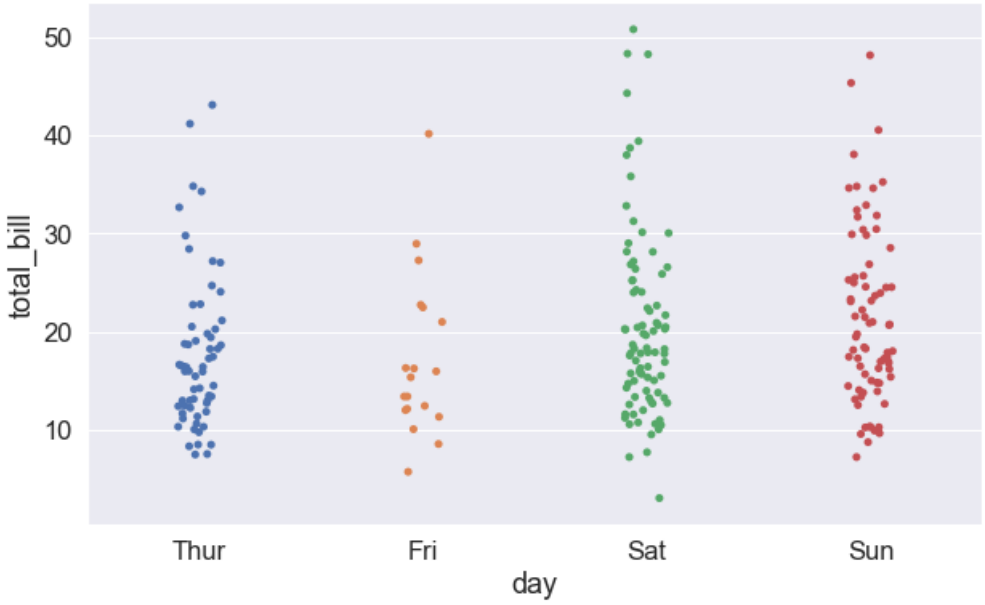
Strip Plot

In [56]:

```
sns.stripplot(x = 'day', y = 'total_bill', data = tips)
```

Out[56]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17551914d88>
```



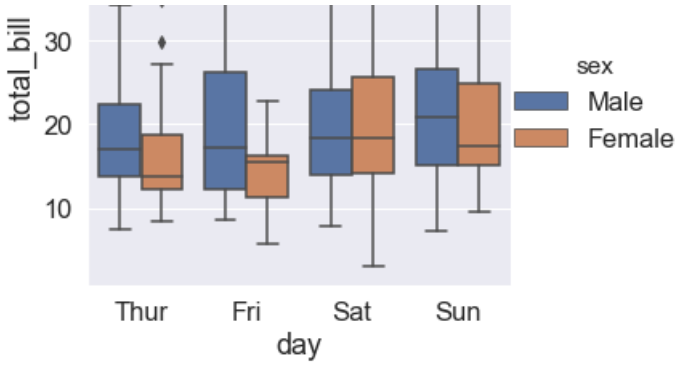
Factor Plot/ General Plot

In [64]:

```
sns.factorplot(x = 'day', y = 'total_bill', data = tips, kind = 'box', hue = 'sex')
```

Out[64]:

```
<seaborn.axisgrid.FacetGrid at 0x17551c1b548>
```



Matrix Plot

Heat Map

In [65]:

```
tips = sns.load_dataset('tips')
```

In [66]:

```
tips.head()
```

Out[66]:

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

In [67]:

```
tips.corr()
```

Out[67]:

	total_bill	tip	size
total_bill	1.000000	0.675734	0.598315
tip	0.675734	1.000000	0.489299
size	0.598315	0.489299	1.000000

In [68]:

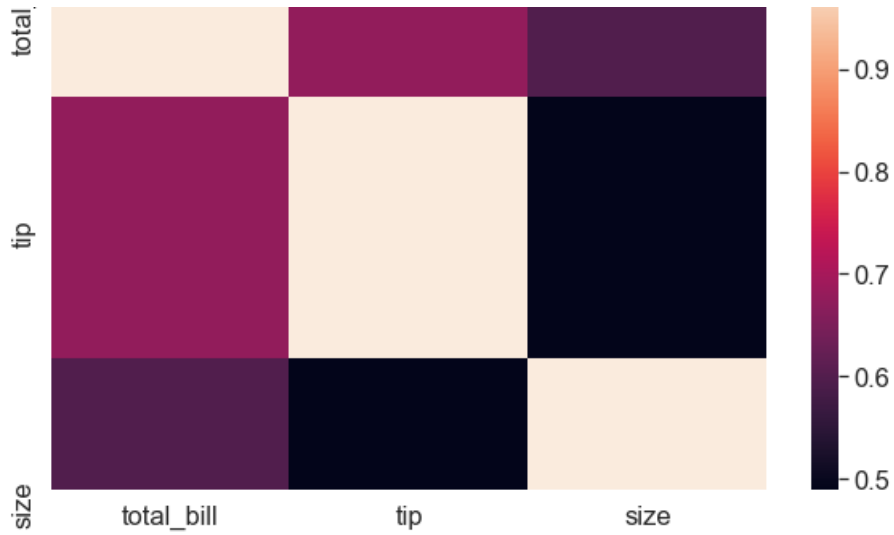
```
tc = tips.corr()
```

In [69]:

```
sns.heatmap(tc)
```

Out[69]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17552de8748>
```

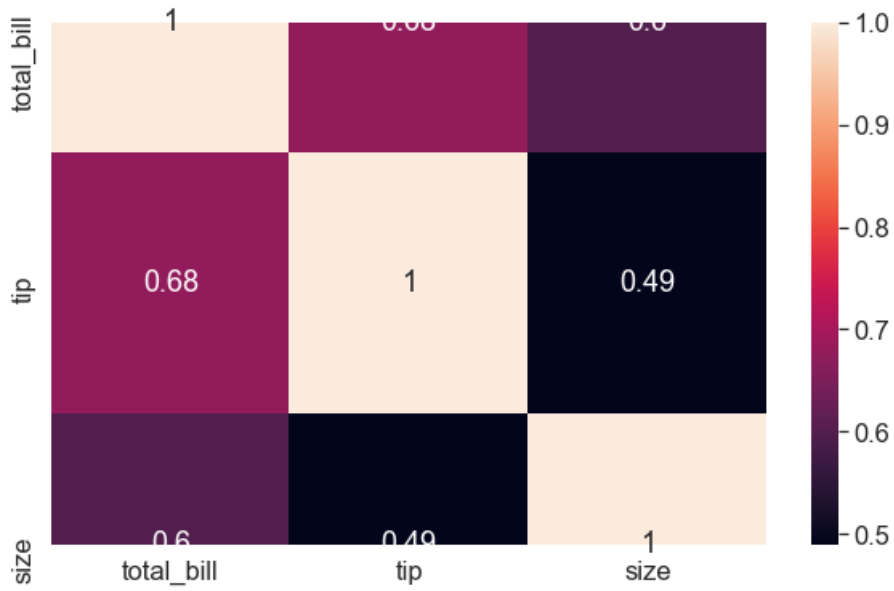


In [70]:

```
sns.heatmap(tc, annot = True)
```

Out[70]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17552fcdac8>
```

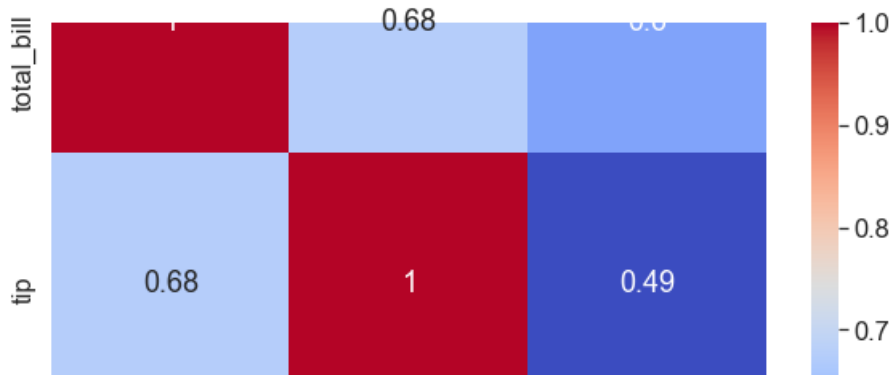


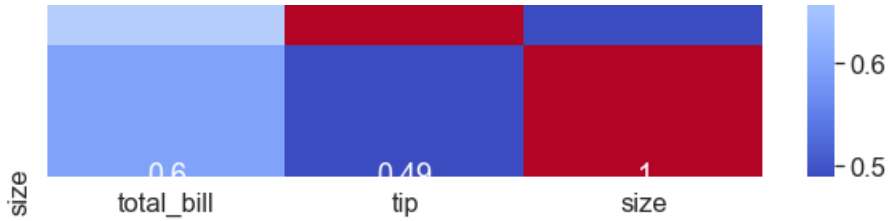
In [71]:

```
sns.heatmap(tc, annot = True, cmap = 'coolwarm')
```

Out[71]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17552efd9c8>
```





In [72]:

```
flights = sns.load_dataset('flights')
```

In [73]:

```
flights.head()
```

Out[73]:

	year	month	passengers
0	1949	January	112
1	1949	February	118
2	1949	March	132
3	1949	April	129
4	1949	May	121

In [74]:

```
fp = flights.pivot_table(index = 'month', columns = 'year', values = 'passengers')
```

In [75]:

```
fp.head()
```

Out[75]:

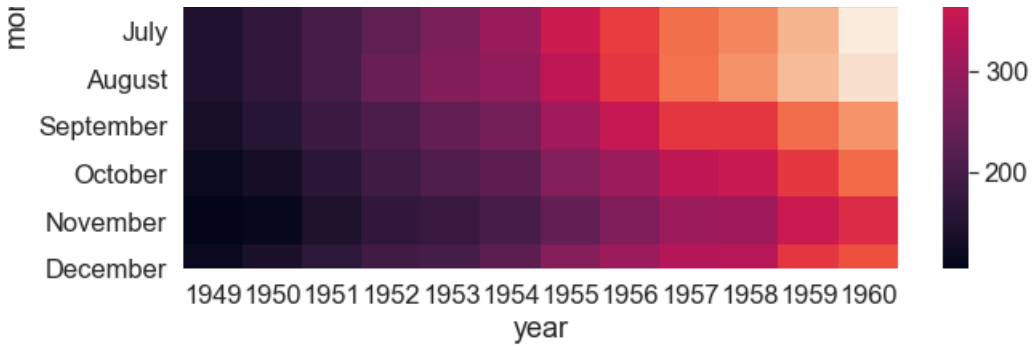
year	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960
month												
January	112	115	145	171	196	204	242	284	315	340	360	417
February	118	126	150	180	196	188	233	277	301	318	342	391
March	132	141	178	193	236	235	267	317	356	362	406	419
April	129	135	163	181	235	227	269	313	348	348	396	461
May	121	125	172	183	229	234	270	318	355	363	420	472

In [76]:

```
sns.heatmap(fp)
```

Out[76]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x17552fa4f08>
```

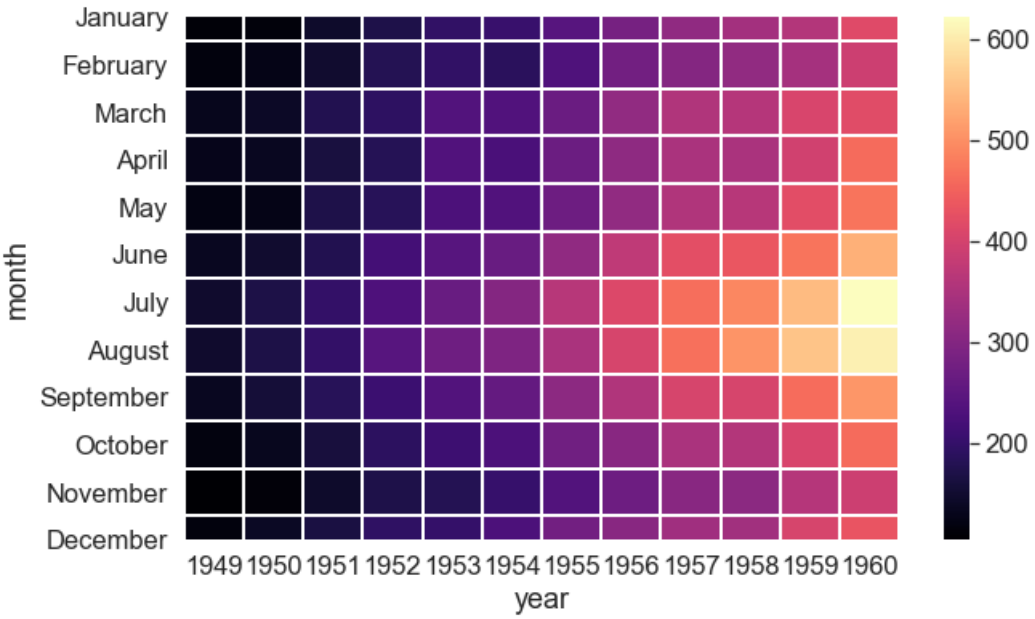


In [77]:

```
sns.heatmap(fp, cmap = 'magma', linecolor = 'white', linewidth = 1)
```

Out[77]:

```
<matplotlib.axes._subplots.AxesSubplot at 0x175530b8348>
```



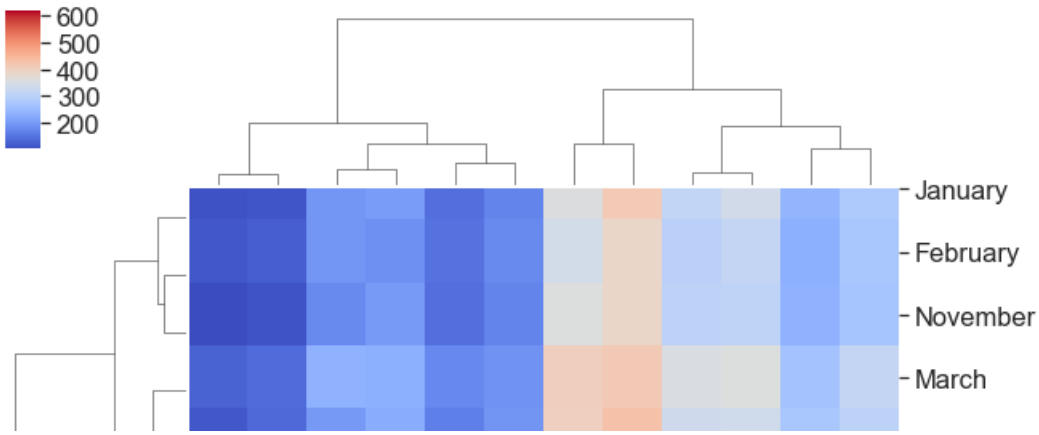
Clustermap

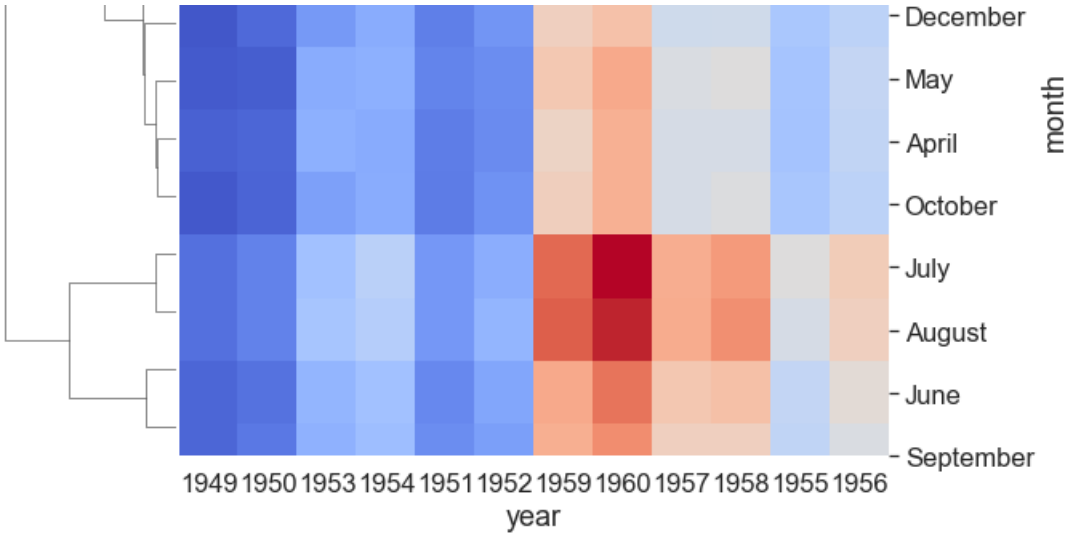
In [78]:

```
sns.clustermap(fp, cmap = 'coolwarm')
```

Out[78]:

```
<seaborn.matrix.ClusterGrid at 0x17552f7c688>
```





Grids

In [79]:

```
iris = sns.load_dataset('iris')
```

In [80]:

```
iris.head()
```

Out[80]:

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

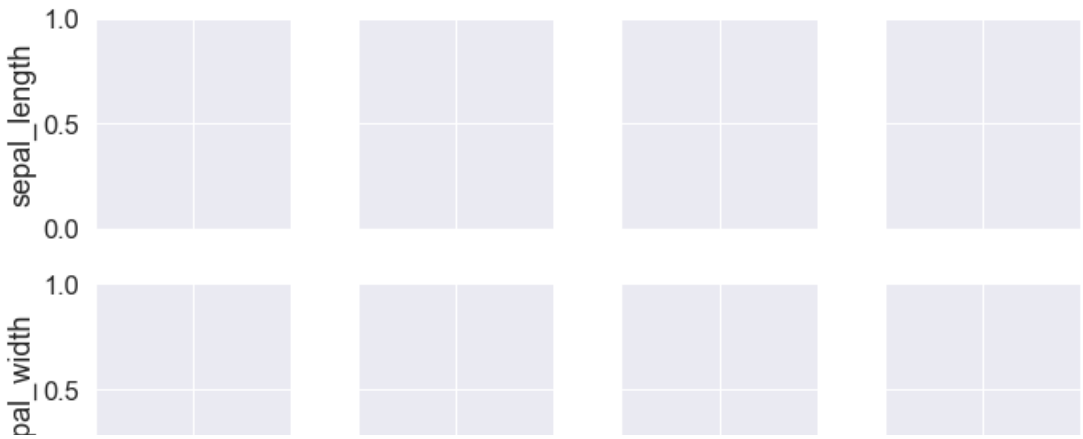
PairGrid

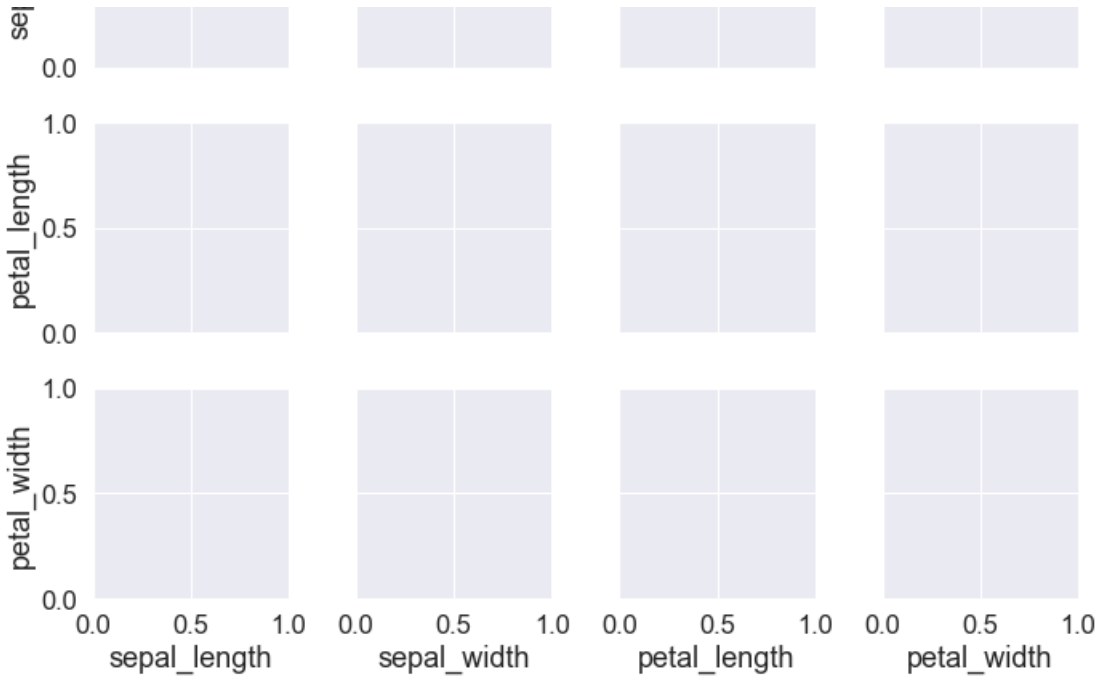
In [81]:

```
sns.PairGrid(iris)
```

Out[81]:

```
<seaborn.axisgrid.PairGrid at 0x175534d98c8>
```





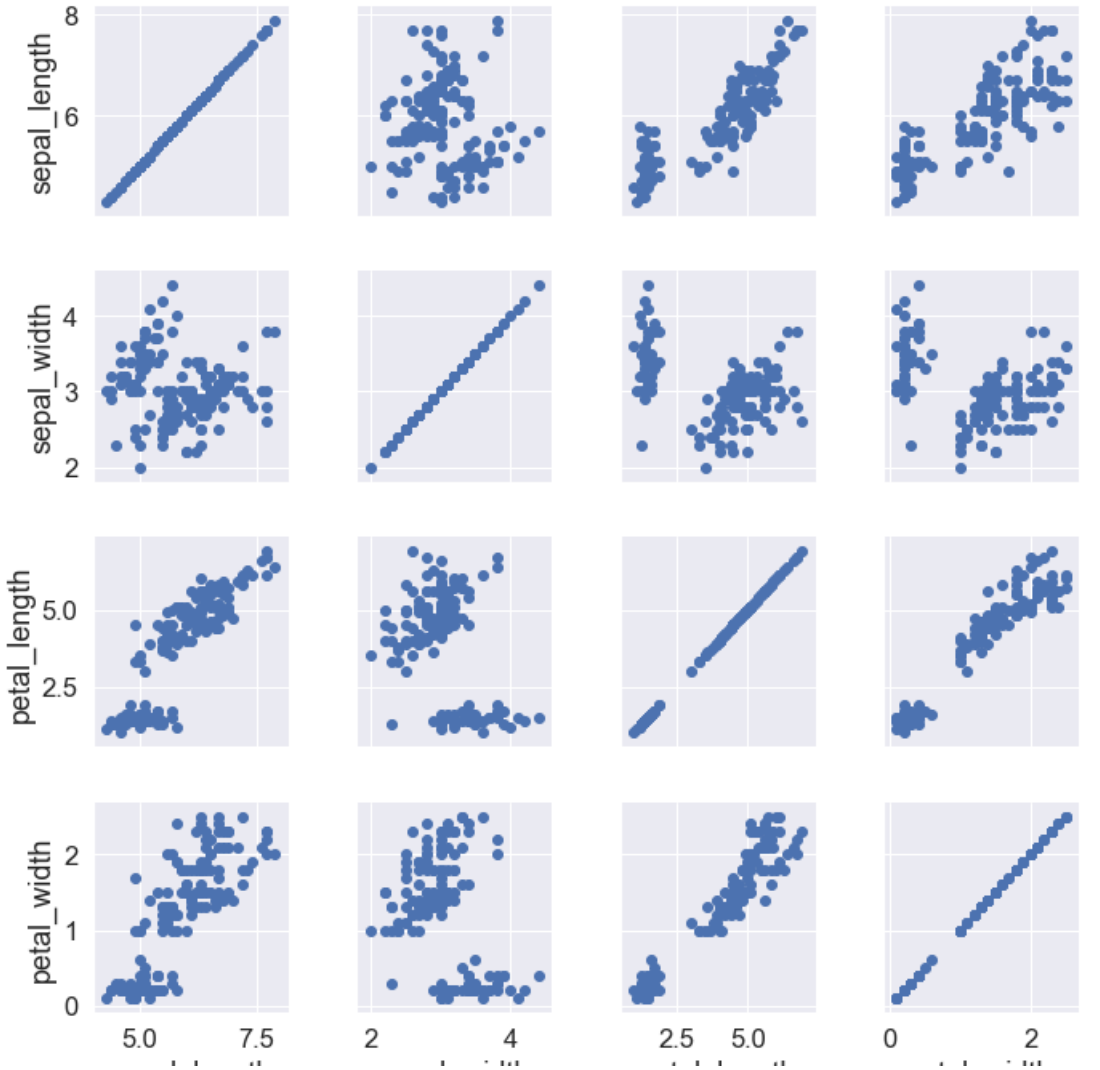
In [82]:

```
g = sns.PairGrid(iris)

g.map(plt.scatter)
```

Out[82]:

```
<seaborn.axisgrid.PairGrid at 0x17553c8e288>
```



sepal_length sepal_width petal_length petal_width

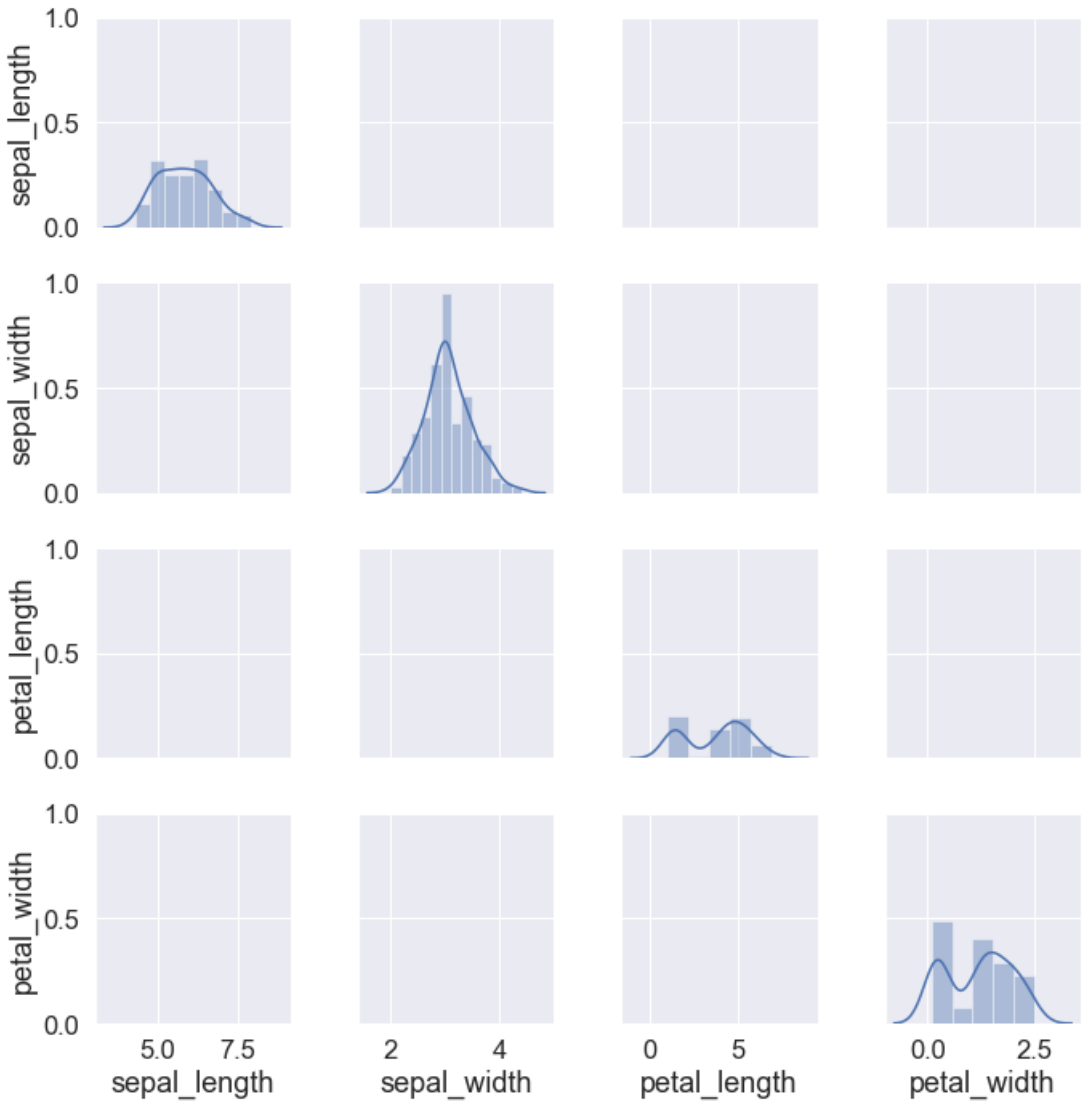
In [84]:

```
g = sns.PairGrid(iris)

g.map_diag(sns.distplot)
```

Out[84]:

```
<seaborn.axisgrid.PairGrid at 0x17554b3d908>
```



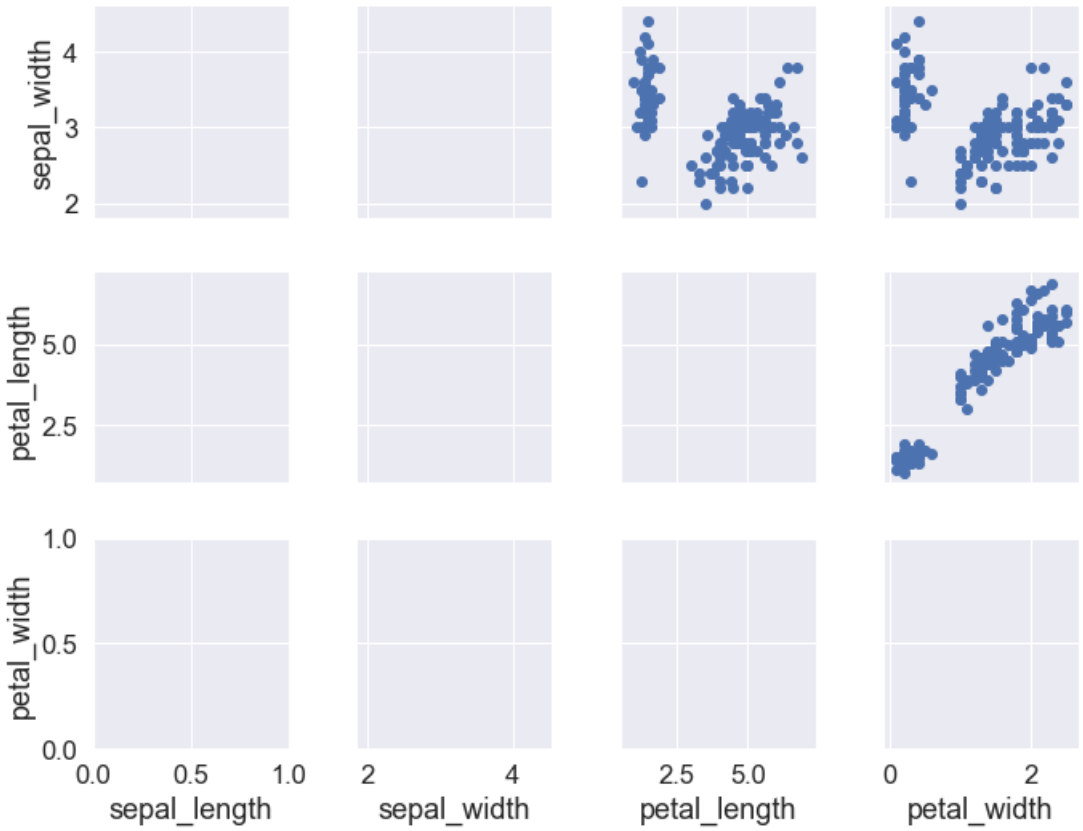
In [85]:

```
g = sns.PairGrid(iris)

g.map_upper(plt.scatter)
```

Out[85]:

```
<seaborn.axisgrid.PairGrid at 0x175561a6fc8>
```



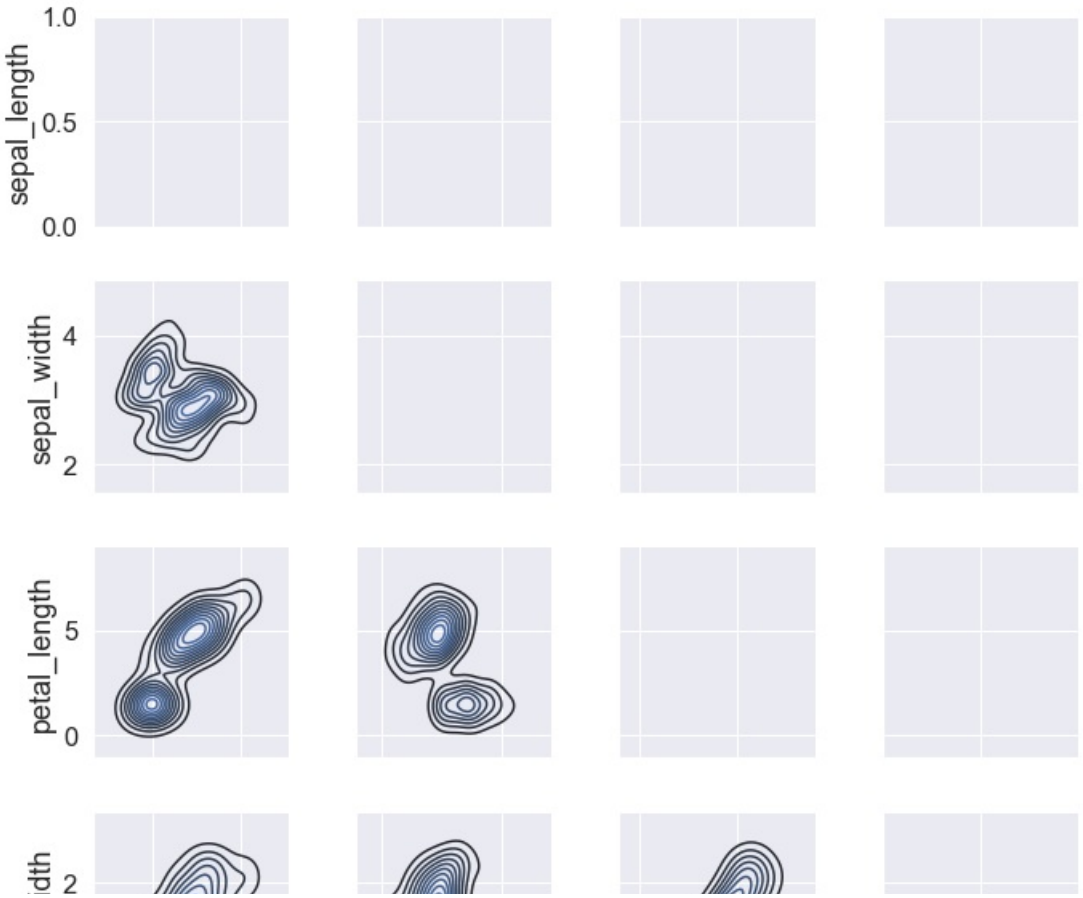
In [86]:

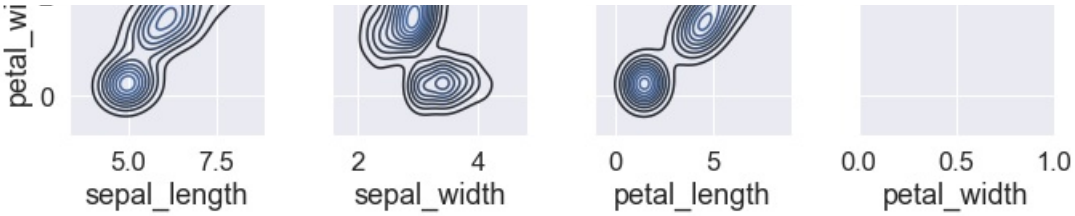
```
g = sns.PairGrid(iris)

g.map_lower(sns.kdeplot)
```

Out[86]:

```
<seaborn.axisgrid.PairGrid at 0x17554a90508>
```





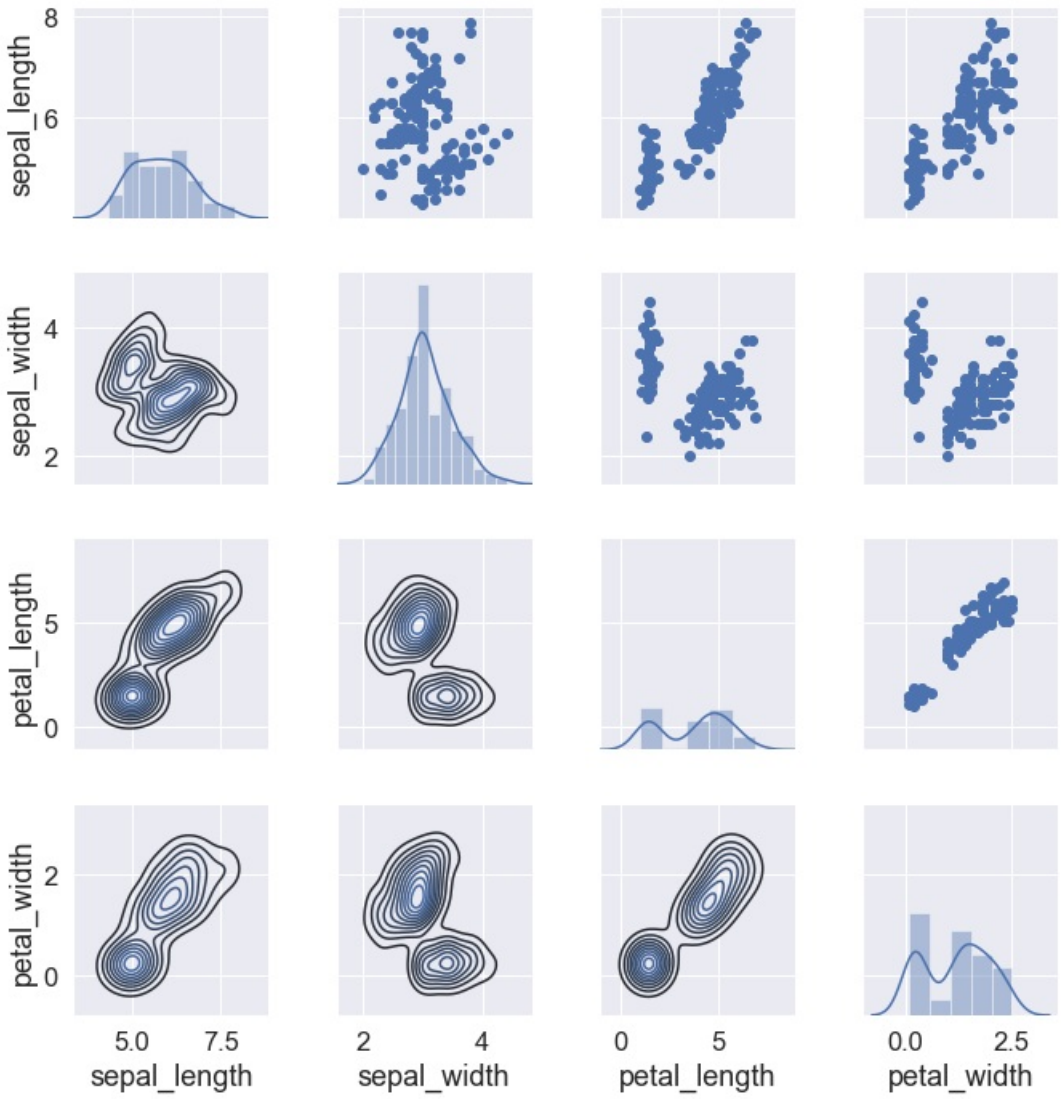
In [87]:

```
g = sns.PairGrid(iris)

g.map_diag(sns.distplot)
g.map_upper(plt.scatter)
g.map_lower(sns.kdeplot)
```

Out[87]:

```
<seaborn.axisgrid.PairGrid at 0x17556dbc248>
```



Faset Grid

In [88]:

```
tips.head()
```

Out[88]:

total_bill	tip	sex	smoker	day	time	size

	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

In [91]:

```
sns.set(font_scale = 1)
```

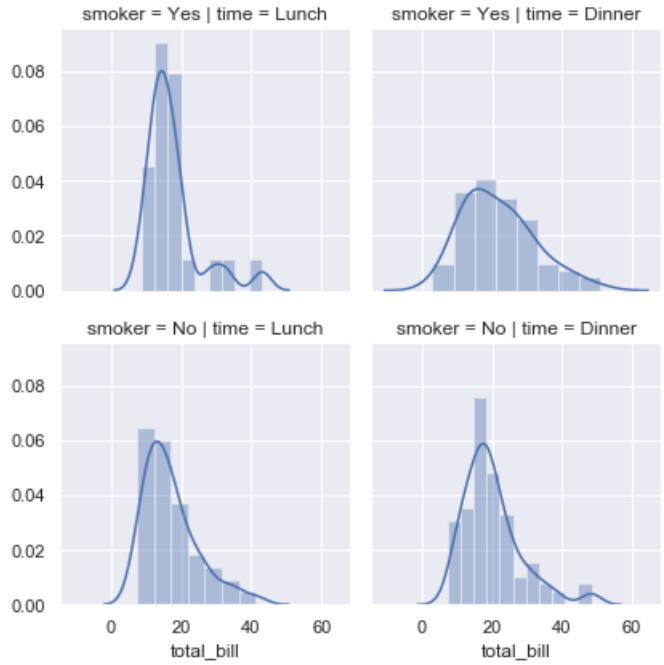
In [92]:

```
g = sns.FacetGrid(data = tips, col = 'time', row = 'smoker')

g.map(sns.distplot, 'total_bill')
```

Out[92]:

```
<seaborn.axisgrid.FacetGrid at 0x1755784bf88>
```



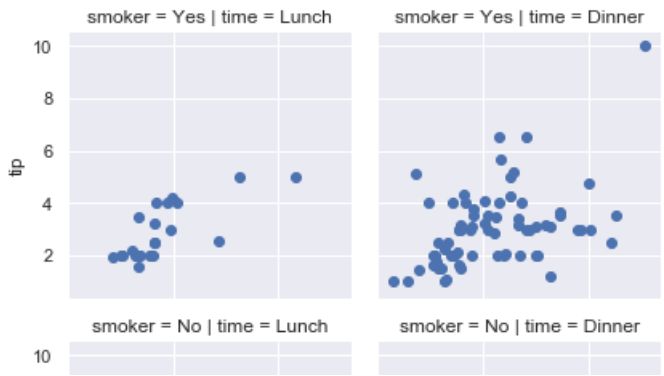
In [94]:

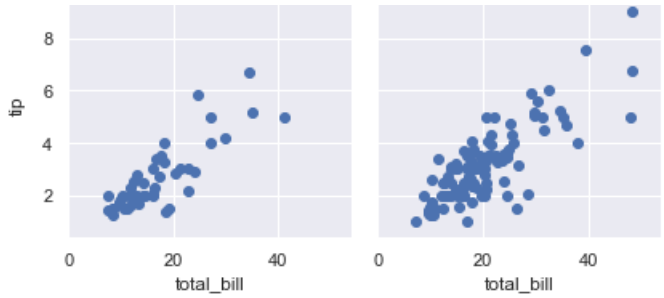
```
g = sns.FacetGrid(data = tips, col = 'time', row = 'smoker')

g.map(plt.scatter, 'total_bill', 'tip')
```

Out[94]:

```
<seaborn.axisgrid.FacetGrid at 0x17557eaed08>
```





Regression Plot

In [95]:

```
tips.head()
```

Out[95]:

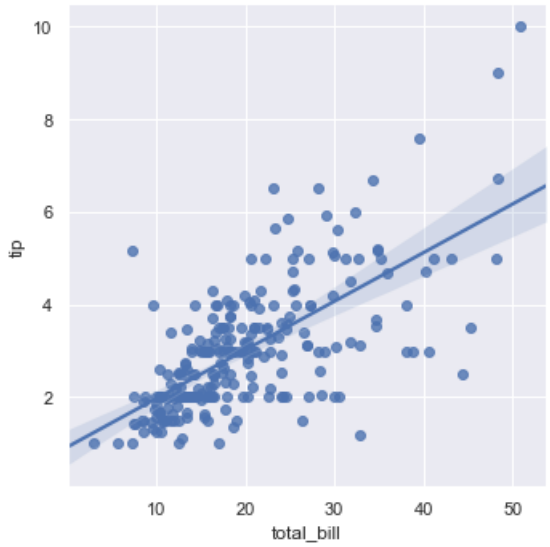
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

In [96]:

```
sns.lmplot(x = 'total_bill', y = 'tip', data = tips)
```

Out[96]:

```
<seaborn.axisgrid.FacetGrid at 0x17558003348>
```

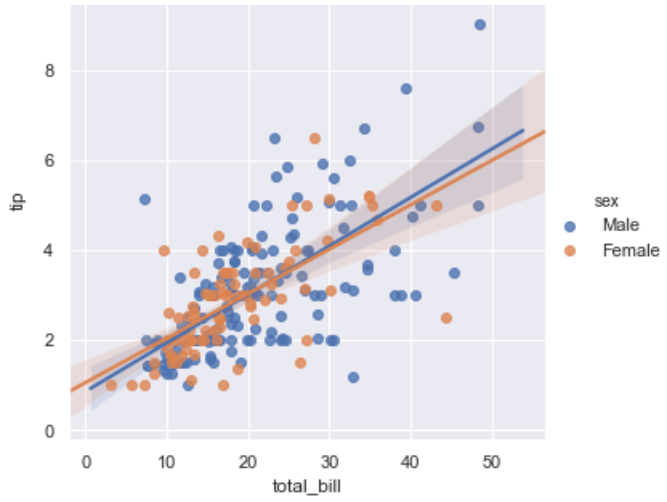


In [97]:

```
sns.lmplot(x = 'total_bill', y = 'tip', data = tips, hue = 'sex')
```

Out[97]:

```
<seaborn.axisgrid.FacetGrid at 0x17559037c08>
```

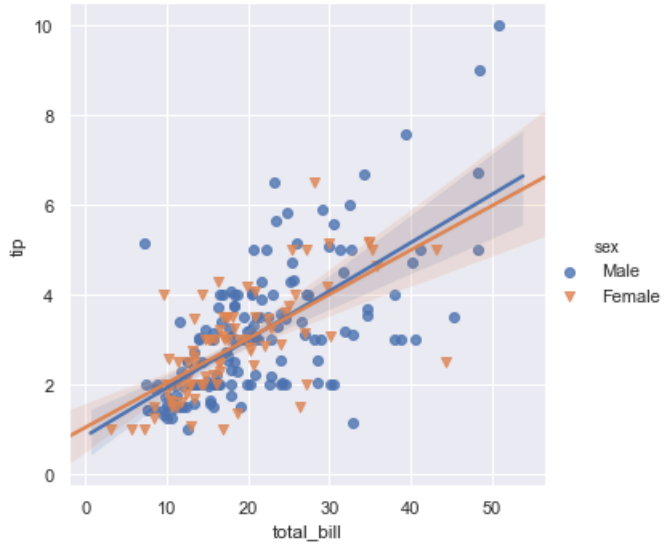


In [98]:

```
sns.lmplot(x = 'total_bill', y = 'tip', data = tips, hue = 'sex', markers = ['o','v'])
```

Out[98]:

```
<seaborn.axisgrid.FacetGrid at 0x1755905bb88>
```

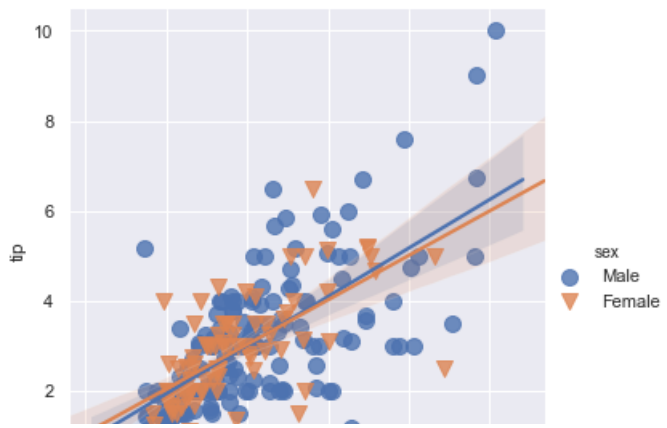


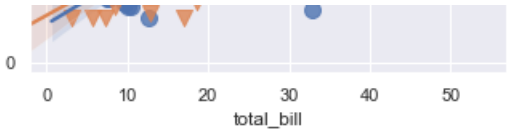
In [99]:

```
sns.lmplot(x = 'total_bill', y = 'tip', data = tips, hue = 'sex', markers = ['o','v'], scatter_kws
= {'s':100})
```

Out[99]:

```
<seaborn.axisgrid.FacetGrid at 0x175590fcc88>
```





In [101]:

```
sns.lmplot(x = 'total_bill', y = 'tip', data = tips, col = 'sex')
```

Out[101]:

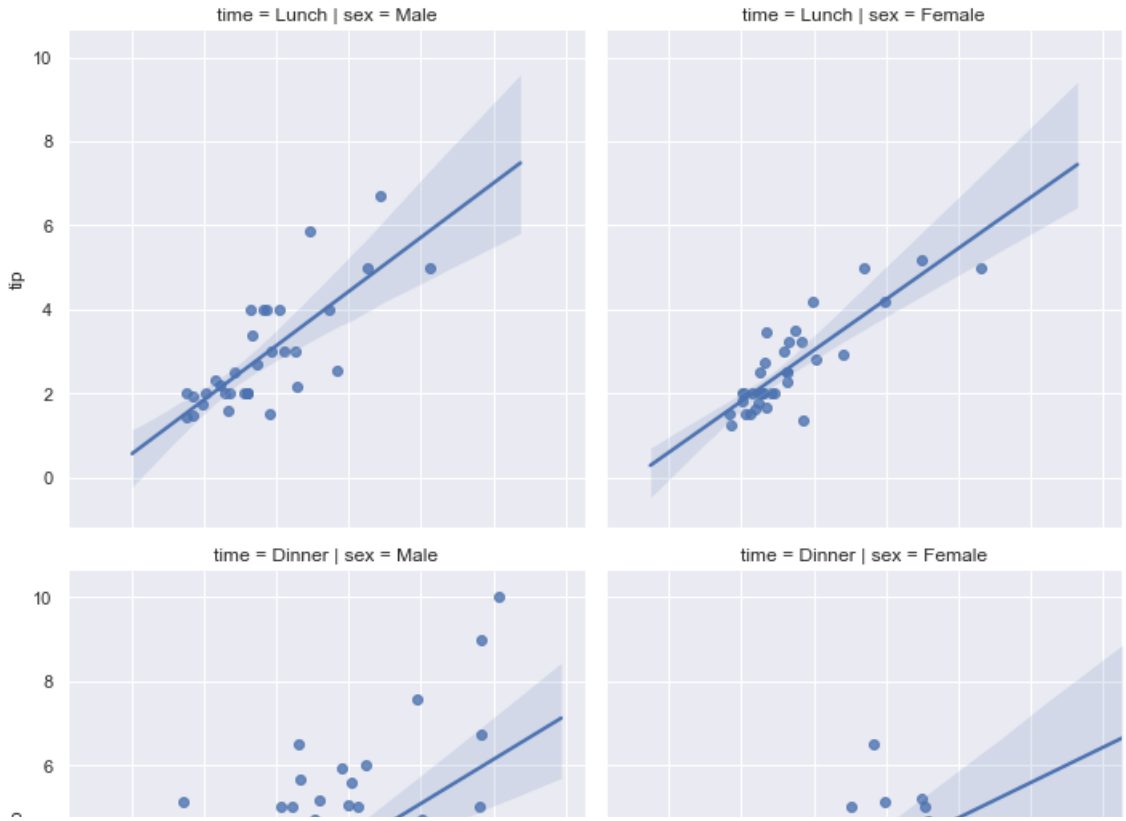
```
<seaborn.axisgrid.FacetGrid at 0x175592c48c8>
```

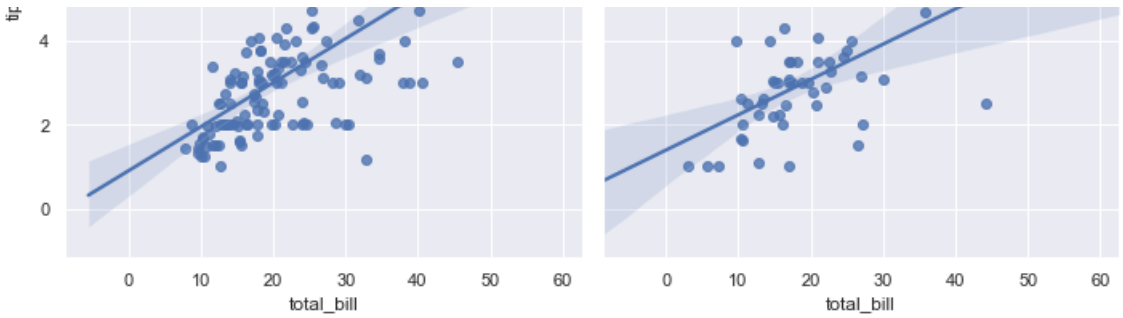
In [102]:

```
sns.lmplot(x = 'total_bill', y = 'tip', data = tips, col = 'sex', row = 'time')
```

Out[102]:

```
<seaborn.axisgrid.FacetGrid at 0x17559390808>
```




tip
4
2
0
0
10
20
30
40
50
60
total_bill
0
10
20
30
40
50
60
total_bill