In [1]:
```
# import all modules

import pandas as pd
import datetime as dt
from pandas_datareader import data
from matplotlib import pyplot as plt
import numpy as np
import seaborn as sns
%matplotlib inline
```

In [2]:
```
tips = sns.load_dataset('tips')   #use the sns module to load default tips dat
aframe within sns, store it as tips. data_source generating
```

In [3]:
```
tips.head() #show the top 5 rows of tips
```

Out[3]:

	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 [4]:
```
tips.describe()
```

Out[4]:

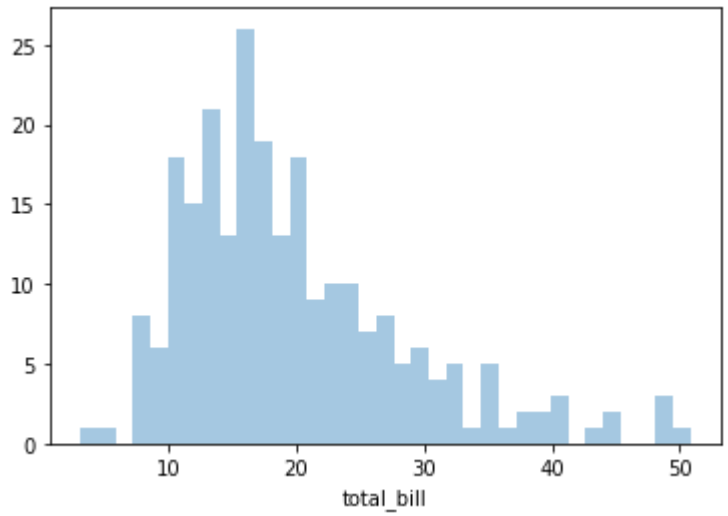
	total_bill	tip	size
count	244.000000	244.000000	244.000000
mean	19.785943	2.998279	2.569672
std	8.902412	1.383638	0.951100
min	3.070000	1.000000	1.000000
25%	13.347500	2.000000	2.000000
50%	17.795000	2.900000	2.000000
75%	24.127500	3.562500	3.000000
max	50.810000	10.000000	6.000000

In [5]:
```
tips.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 244 entries, 0 to 243
Data columns (total 7 columns):
 #   Column      Non-Null Count  Dtype   
---  ------      --------------  -----   
 0   total_bill  244 non-null    float64 
 1   tip         244 non-null    float64 
 2   sex         244 non-null    category
 3   smoker      244 non-null    category
 4   day         244 non-null    category
 5   time        244 non-null    category
 6   size        244 non-null    int64   
dtypes: category(4), float64(2), int64(1)
memory usage: 7.3 KB
```

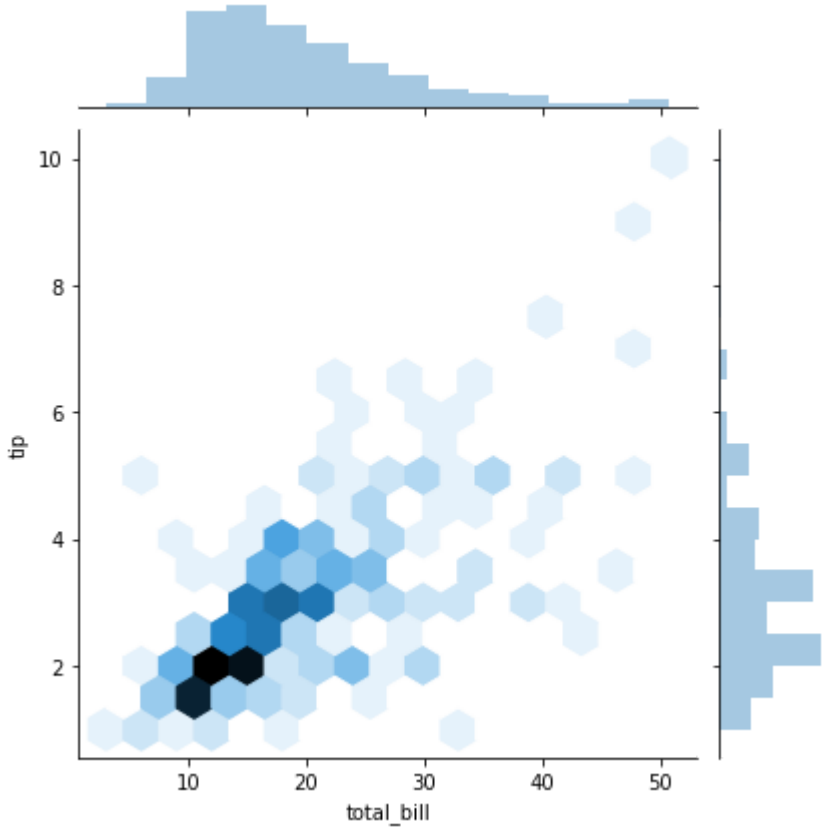
In [6]:
```
sns.distplot(tips['total_bill'],kde=False,bins=35)  #shows the total bill as x axis and counts of different bills as y axis
```

Out[6]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x1546a197ee0>
```



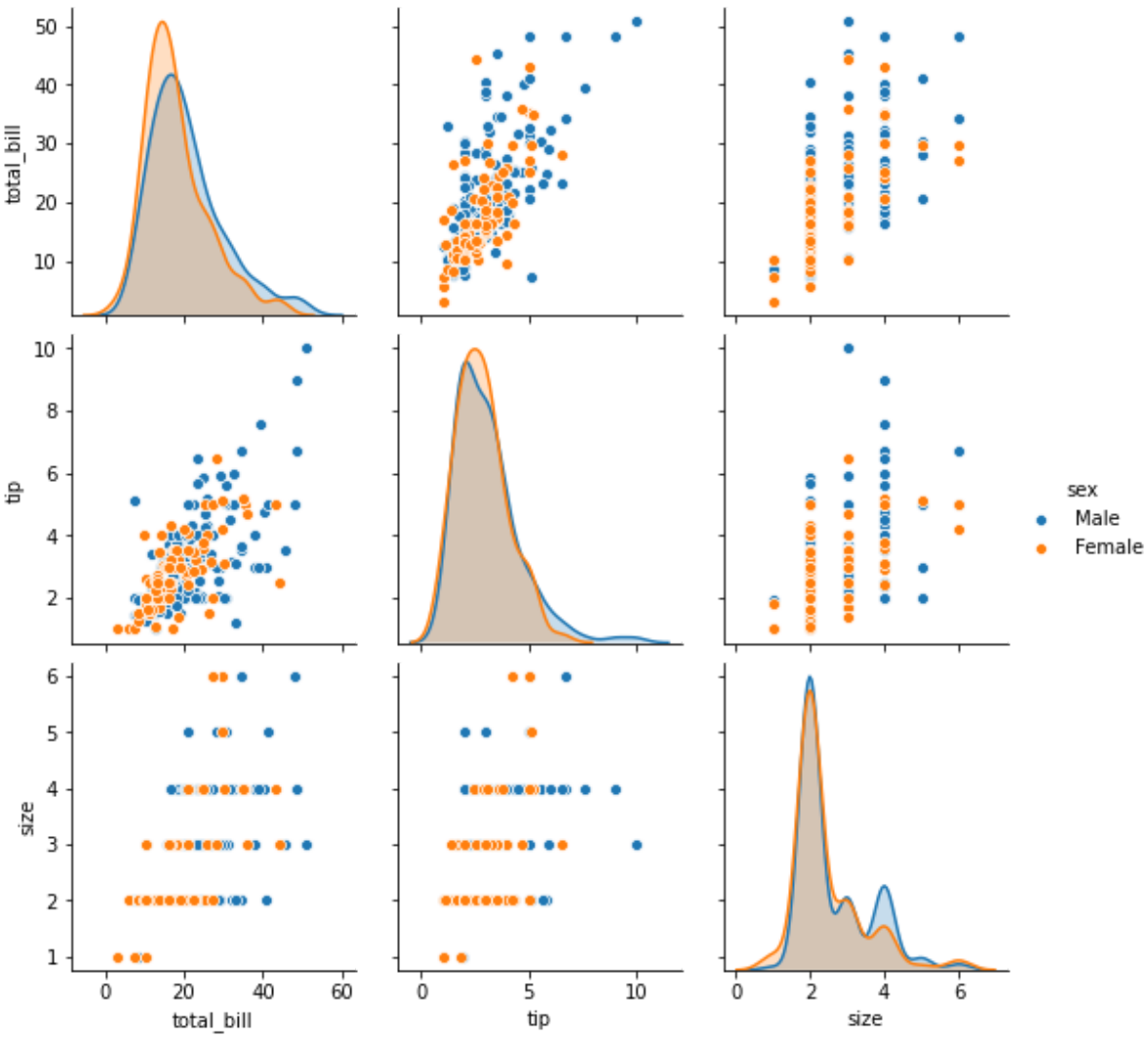
In [7]:
```
sns.jointplot(x= "total_bill", y="tip", data = tips, kind="hex") #shows a join
tplot with total bill as x axis, tip as y axis, and hex as the kind.
```

Out[7]: <seaborn.axisgrid.JointGrid at 0x1546a920250>



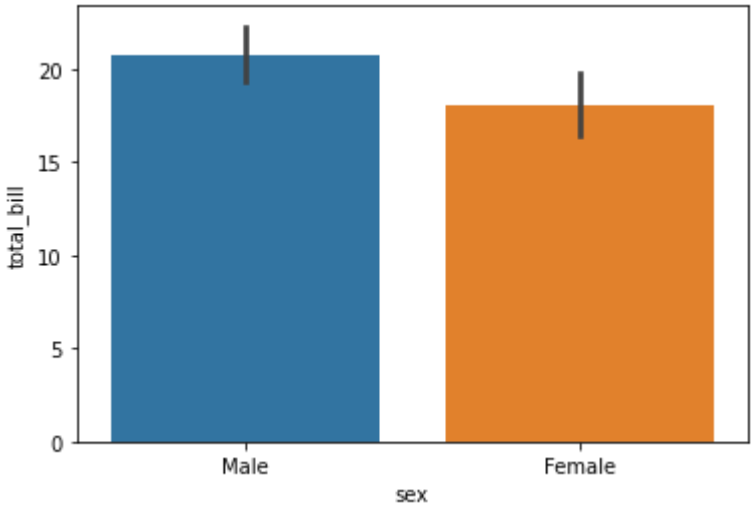
In [8]:
```
sns.pairplot(tips, hue="sex") #do jointplots for every single combination of n
umeric columns in the tips dataframe, different sex column will have different
color
```

Out[8]: <seaborn.axisgrid.PairGrid at 0x1546aa9f400>



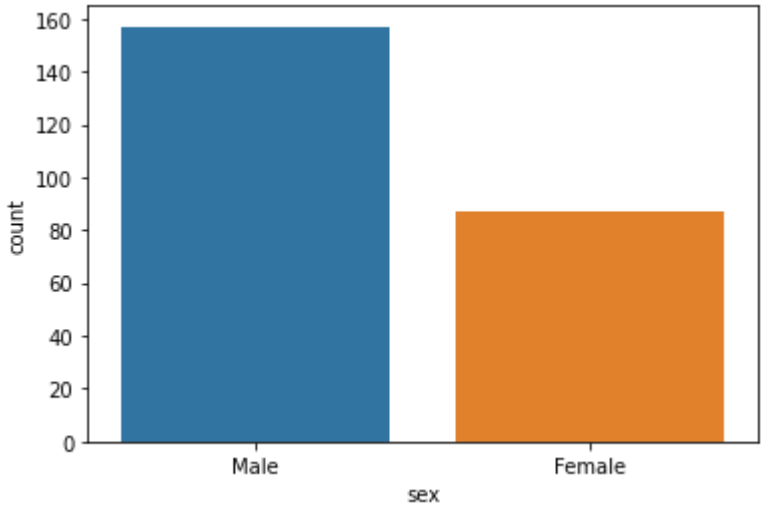
In [11]:
```
sns.barplot(x = "sex", y ="total_bill", data=tips)
```

Out[11]: <matplotlib.axes._subplots.AxesSubplot at 0x1546b2ca8e0>



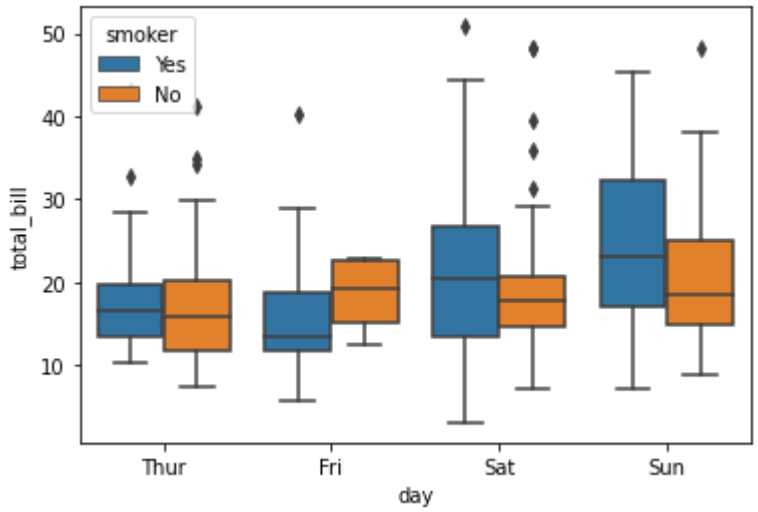
In [12]:
```
sns.countplot(x="sex", data=tips)
```

Out[12]: <matplotlib.axes._subplots.AxesSubplot at 0x1546b3185e0>



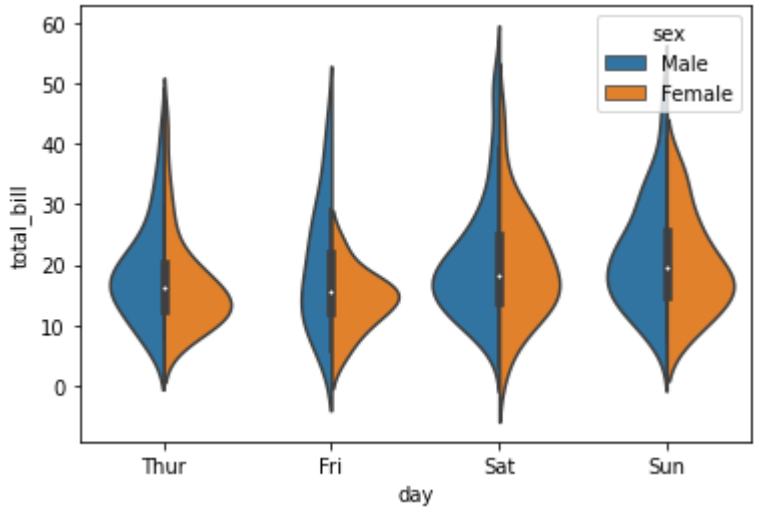
In [13]:
```
sns.boxplot(x="day", y="total_bill", data=tips, hue="smoker") #categolize by s
mokers or not, doesnt include outliners
```

Out[13]: <matplotlib.axes._subplots.AxesSubplot at 0x1546b368820>



In [14]:
```
sns.violinplot(x="day", y= "total_bill", data=tips, hue = "sex", split=True) #
include outliners, has more info than boxplot.
```

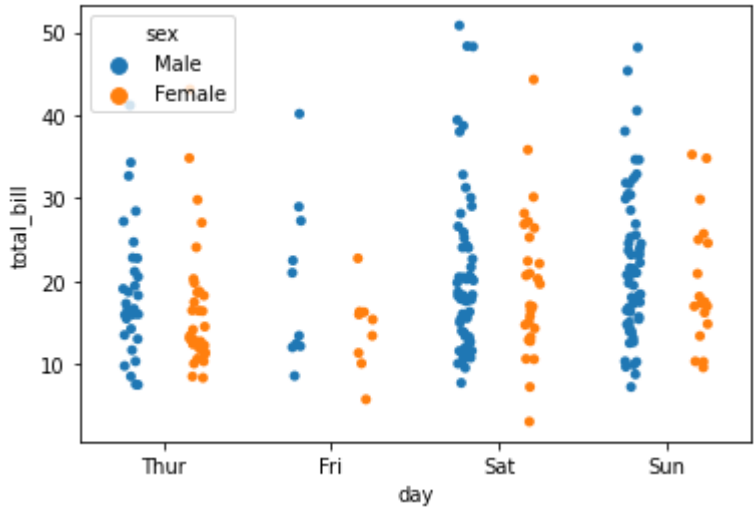
Out[14]: <matplotlib.axes._subplots.AxesSubplot at 0x1546b446760>



In [15]:
```
sns.stripplot(x="day", y="total_bill", data=tips, jitter=True, hue="sex", split=True) #jitter = True to seperate stack points
```

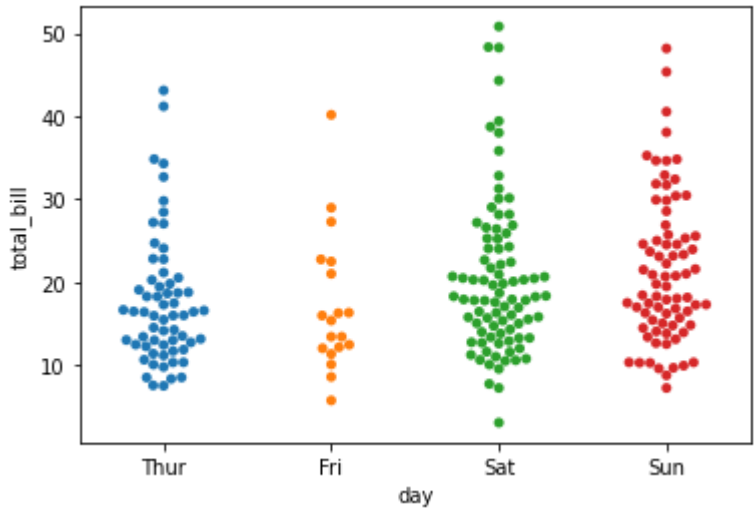
```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\categorical.py:2781: UserWarning: The `split` parameter has been renamed to `dodge`.
  warnings.warn(msg, UserWarning)
```

Out[15]: <matplotlib.axes._subplots.AxesSubplot at 0x1546b4f9b20>



In [16]:
```
sns.swarmplot(x='day', y='total_bill', data=tips) #swarmplot is a combo of violinplot and stripplot together
```

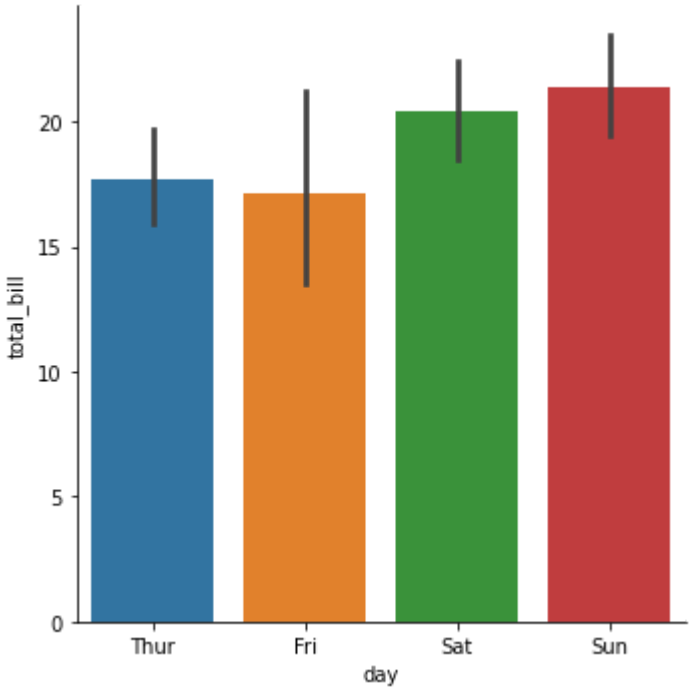
Out[16]: <matplotlib.axes._subplots.AxesSubplot at 0x1546b5aebb0>



In [17]:
```
sns.factorplot(x='day', y='total_bill', data=tips, kind='bar') #equals to a barplot, all plots can be called in this way.
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\categorical.py:3666: UserWarning: The `factorplot` function has been renamed to `catplot`. The original name will be removed in a future release. Please update your code. Note that the default `kind` in `factorplot` (`'point'`) has changed `'strip'` in `catplot`.
  warnings.warn(msg)
```

Out[17]:
```
<seaborn.axisgrid.FacetGrid at 0x1546b5faa30>
```



In [18]:
```
tips1 = tips.groupby("time")
```

In [23]:
```
tips1.head()
```

Out[23]:

	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
77	27.20	4.00	Male	No	Thur	Lunch	4
78	22.76	3.00	Male	No	Thur	Lunch	2
79	17.29	2.71	Male	No	Thur	Lunch	2
80	19.44	3.00	Male	Yes	Thur	Lunch	2
81	16.66	3.40	Male	No	Thur	Lunch	2

In [24]:
```
tips1.size()
```

Out[24]:
```
time
Lunch      68
Dinner    176
dtype: int64
```

In [25]:
```
tips1.last()
```

Out[25]:

time	total_bill	tip	sex	smoker	day	size
Lunch	10.09	2.0	Female	Yes	Fri	2
Dinner	18.78	3.0	Female	No	Thur	2

In [26]:
```
tips1.first()
```

Out[26]:

time	total_bill	tip	sex	smoker	day	size
Lunch	27.20	4.00	Male	No	Thur	4
Dinner	16.99	1.01	Female	No	Sun	2

In [28]:
```
tips1.groups
```

Out[28]:
```
{'Lunch': Int64Index([ 77,  78,  79,  80,  81,  82,  83,  84,  85,  86,  87,
88,  89,
             117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129,
             130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142,
             143, 144, 145, 146, 147, 148, 149, 191, 192, 193, 194, 195, 196,
             197, 198, 199, 200, 201, 202, 203, 204, 205, 220, 221, 222, 223,
             224, 225, 226],
            dtype='int64'),
 'Dinner': Int64Index([  0,   1,   2,   3,   4,   5,   6,   7,   8,   9,
             ...
             234, 235, 236, 237, 238, 239, 240, 241, 242, 243],
            dtype='int64', length=176)}
```

In [31]:
```
tips1.get_group("Lunch")
```

Out[31]:

	total_bill	tip	sex	smoker	day	size
77	27.20	4.00	Male	No	Thur	4
78	22.76	3.00	Male	No	Thur	2
79	17.29	2.71	Male	No	Thur	2
80	19.44	3.00	Male	Yes	Thur	2
81	16.66	3.40	Male	No	Thur	2
...	...	...	...	...	...	...
222	8.58	1.92	Male	Yes	Fri	1
223	15.98	3.00	Female	No	Fri	3
224	13.42	1.58	Male	Yes	Fri	2
225	16.27	2.50	Female	Yes	Fri	2
226	10.09	2.00	Female	Yes	Fri	2

68 rows × 6 columns

In [34]:
```
tips1.get_group("Dinner")
```

Out[34]:

	total_bill	tip	sex	smoker	day	size
0	16.99	1.01	Female	No	Sun	2
1	10.34	1.66	Male	No	Sun	3
2	21.01	3.50	Male	No	Sun	3
3	23.68	3.31	Male	No	Sun	2
4	24.59	3.61	Female	No	Sun	4
...	...	...	...	...	...	...
239	29.03	5.92	Male	No	Sat	3
240	27.18	2.00	Female	Yes	Sat	2
241	22.67	2.00	Male	Yes	Sat	2
242	17.82	1.75	Male	No	Sat	2
243	18.78	3.00	Female	No	Thur	2

176 rows × 6 columns

In [35]:
```
tips1.max()
```

Out[35]:

time	total_bill	tip	size
Lunch	43.11	6.7	6
Dinner	50.81	10.0	6

In [37]:
```
tips1.min()
```

Out[37]:

time	total_bill	tip	size
Lunch	7.51	1.25	1
Dinner	3.07	1.00	1

In [38]:
```
tips1.sum()
```

Out[38]:

time	total_bill	tip	size
Lunch	1167.47	185.51	164
Dinner	3660.30	546.07	463

In [39]:
```
tips1.mean()
```

Out[39]:

time	total_bill	tip	size
Lunch	17.168676	2.728088	2.411765
Dinner	20.797159	3.102670	2.630682

In [41]:
```
tips1["tip"].sum()
```

Out[41]:
```
time
Lunch     185.51
Dinner    546.07
Name: tip, dtype: float64
```

In [42]:
```
tips1["total_bill"].sum()
```

Out[42]:
```
time
Lunch     1167.47
Dinner    3660.30
Name: total_bill, dtype: float64
```

In [45]:
```
tips2 = tips.groupby("day")
```

In [46]:
```
tips2.head()
```

Out[46]:

	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
19	20.65	3.35	Male	No	Sat	Dinner	3
20	17.92	4.08	Male	No	Sat	Dinner	2
21	20.29	2.75	Female	No	Sat	Dinner	2
22	15.77	2.23	Female	No	Sat	Dinner	2
23	39.42	7.58	Male	No	Sat	Dinner	4
77	27.20	4.00	Male	No	Thur	Lunch	4
78	22.76	3.00	Male	No	Thur	Lunch	2
79	17.29	2.71	Male	No	Thur	Lunch	2
80	19.44	3.00	Male	Yes	Thur	Lunch	2
81	16.66	3.40	Male	No	Thur	Lunch	2
90	28.97	3.00	Male	Yes	Fri	Dinner	2
91	22.49	3.50	Male	No	Fri	Dinner	2
92	5.75	1.00	Female	Yes	Fri	Dinner	2
93	16.32	4.30	Female	Yes	Fri	Dinner	2
94	22.75	3.25	Female	No	Fri	Dinner	2

In [ ]:
```

```