

Seaborn Stuff

March 3, 2021

1 Importing Stuff

```
[1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
sns.set_theme(style="darkgrid")

%matplotlib inline
```

2 Example Datasets

```
[2]: print(sns.get_dataset_names())
```

```
['anagrams', 'anscombe', 'attention', 'brain_networks', 'car_crashes',
'diamonds', 'dots', 'exercise', 'flights', 'fmri', 'gammas', 'geyser', 'iris',
'mpg', 'penguins', 'planets', 'tips', 'titanic']
```

```
[3]: crash_data = sns.load_dataset('car_crashes')
iris_data = sns.load_dataset('iris')
tips_data = sns.load_dataset('tips')
```

```
[4]: crash_data
```

```
[4]:
```

	total	speeding	alcohol	not_distracted	no_previous	ins_premium	\
0	18.8	7.332	5.640	18.048	15.040	784.55	
1	18.1	7.421	4.525	16.290	17.014	1053.48	
2	18.6	6.510	5.208	15.624	17.856	899.47	
3	22.4	4.032	5.824	21.056	21.280	827.34	
4	12.0	4.200	3.360	10.920	10.680	878.41	
5	13.6	5.032	3.808	10.744	12.920	835.50	
6	10.8	4.968	3.888	9.396	8.856	1068.73	
7	16.2	6.156	4.860	14.094	16.038	1137.87	
8	5.9	2.006	1.593	5.900	5.900	1273.89	
9	17.9	3.759	5.191	16.468	16.826	1160.13	

10	15.6	2.964	3.900	14.820	14.508	913.15
11	17.5	9.450	7.175	14.350	15.225	861.18
12	15.3	5.508	4.437	13.005	14.994	641.96
13	12.8	4.608	4.352	12.032	12.288	803.11
14	14.5	3.625	4.205	13.775	13.775	710.46
15	15.7	2.669	3.925	15.229	13.659	649.06
16	17.8	4.806	4.272	13.706	15.130	780.45
17	21.4	4.066	4.922	16.692	16.264	872.51
18	20.5	7.175	6.765	14.965	20.090	1281.55
19	15.1	5.738	4.530	13.137	12.684	661.88
20	12.5	4.250	4.000	8.875	12.375	1048.78
21	8.2	1.886	2.870	7.134	6.560	1011.14
22	14.1	3.384	3.948	13.395	10.857	1110.61
23	9.6	2.208	2.784	8.448	8.448	777.18
24	17.6	2.640	5.456	1.760	17.600	896.07
25	16.1	6.923	5.474	14.812	13.524	790.32
26	21.4	8.346	9.416	17.976	18.190	816.21
27	14.9	1.937	5.215	13.857	13.410	732.28
28	14.7	5.439	4.704	13.965	14.553	1029.87
29	11.6	4.060	3.480	10.092	9.628	746.54
30	11.2	1.792	3.136	9.632	8.736	1301.52
31	18.4	3.496	4.968	12.328	18.032	869.85
32	12.3	3.936	3.567	10.824	9.840	1234.31
33	16.8	6.552	5.208	15.792	13.608	708.24
34	23.9	5.497	10.038	23.661	20.554	688.75
35	14.1	3.948	4.794	13.959	11.562	697.73
36	19.9	6.368	5.771	18.308	18.706	881.51
37	12.8	4.224	3.328	8.576	11.520	804.71
38	18.2	9.100	5.642	17.472	16.016	905.99
39	11.1	3.774	4.218	10.212	8.769	1148.99
40	23.9	9.082	9.799	22.944	19.359	858.97
41	19.4	6.014	6.402	19.012	16.684	669.31
42	19.5	4.095	5.655	15.990	15.795	767.91
43	19.4	7.760	7.372	17.654	16.878	1004.75
44	11.3	4.859	1.808	9.944	10.848	809.38
45	13.6	4.080	4.080	13.056	12.920	716.20
46	12.7	2.413	3.429	11.049	11.176	768.95
47	10.6	4.452	3.498	8.692	9.116	890.03
48	23.8	8.092	6.664	23.086	20.706	992.61
49	13.8	4.968	4.554	5.382	11.592	670.31
50	17.4	7.308	5.568	14.094	15.660	791.14

ins_losses abbrev

0	145.08	AL
1	133.93	AK
2	110.35	AZ
3	142.39	AR

4	165.63	CA
5	139.91	CO
6	167.02	CT
7	151.48	DE
8	136.05	DC
9	144.18	FL
10	142.80	GA
11	120.92	HI
12	82.75	ID
13	139.15	IL
14	108.92	IN
15	114.47	IA
16	133.80	KS
17	137.13	KY
18	194.78	LA
19	96.57	ME
20	192.70	MD
21	135.63	MA
22	152.26	MI
23	133.35	MN
24	155.77	MS
25	144.45	MO
26	85.15	MT
27	114.82	NE
28	138.71	NV
29	120.21	NH
30	159.85	NJ
31	120.75	NM
32	150.01	NY
33	127.82	NC
34	109.72	ND
35	133.52	OH
36	178.86	OK
37	104.61	OR
38	153.86	PA
39	148.58	RI
40	116.29	SC
41	96.87	SD
42	155.57	TN
43	156.83	TX
44	109.48	UT
45	109.61	VT
46	153.72	VA
47	111.62	WA
48	152.56	WV
49	106.62	WI
50	122.04	WY

3 Distribution Plots

```
[5]: sns.distplot(crash_data['total'])  
sns.distplot(crash_data['speeding'])
```

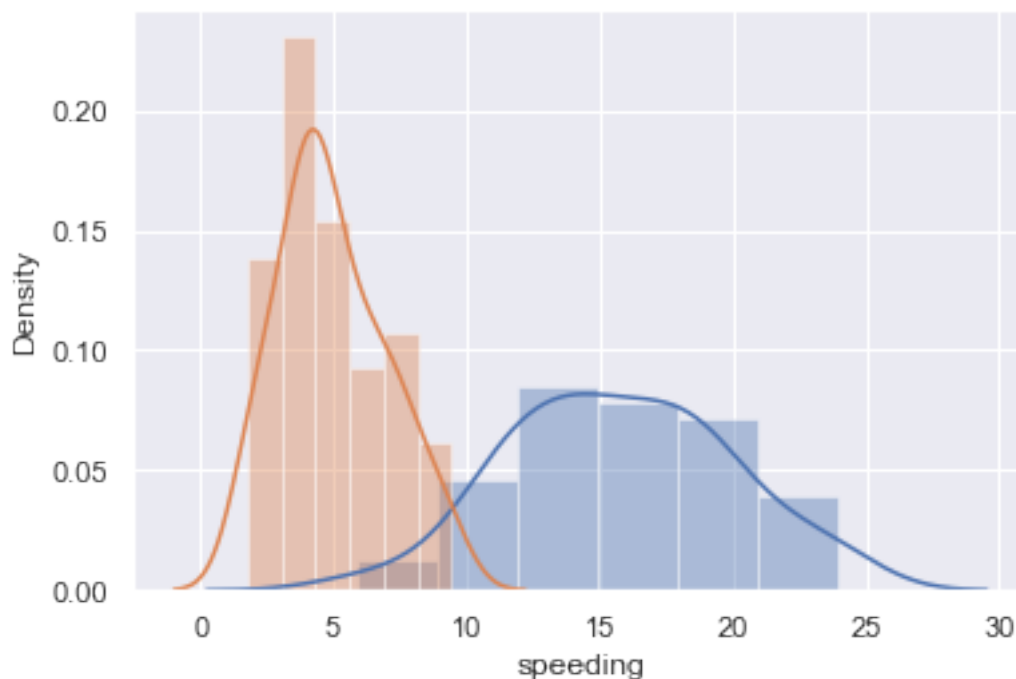
/Library/Frameworks/Python.framework/Versions/3.8/lib/python3.8/site-packages/seaborn/distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

```
warnings.warn(msg, FutureWarning)
```

/Library/Frameworks/Python.framework/Versions/3.8/lib/python3.8/site-packages/seaborn/distributions.py:2557: FutureWarning: `distplot` is a deprecated function and will be removed in a future version. Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

```
warnings.warn(msg, FutureWarning)
```

```
[5]: <AxesSubplot:xlabel='speeding', ylabel='Density'>
```

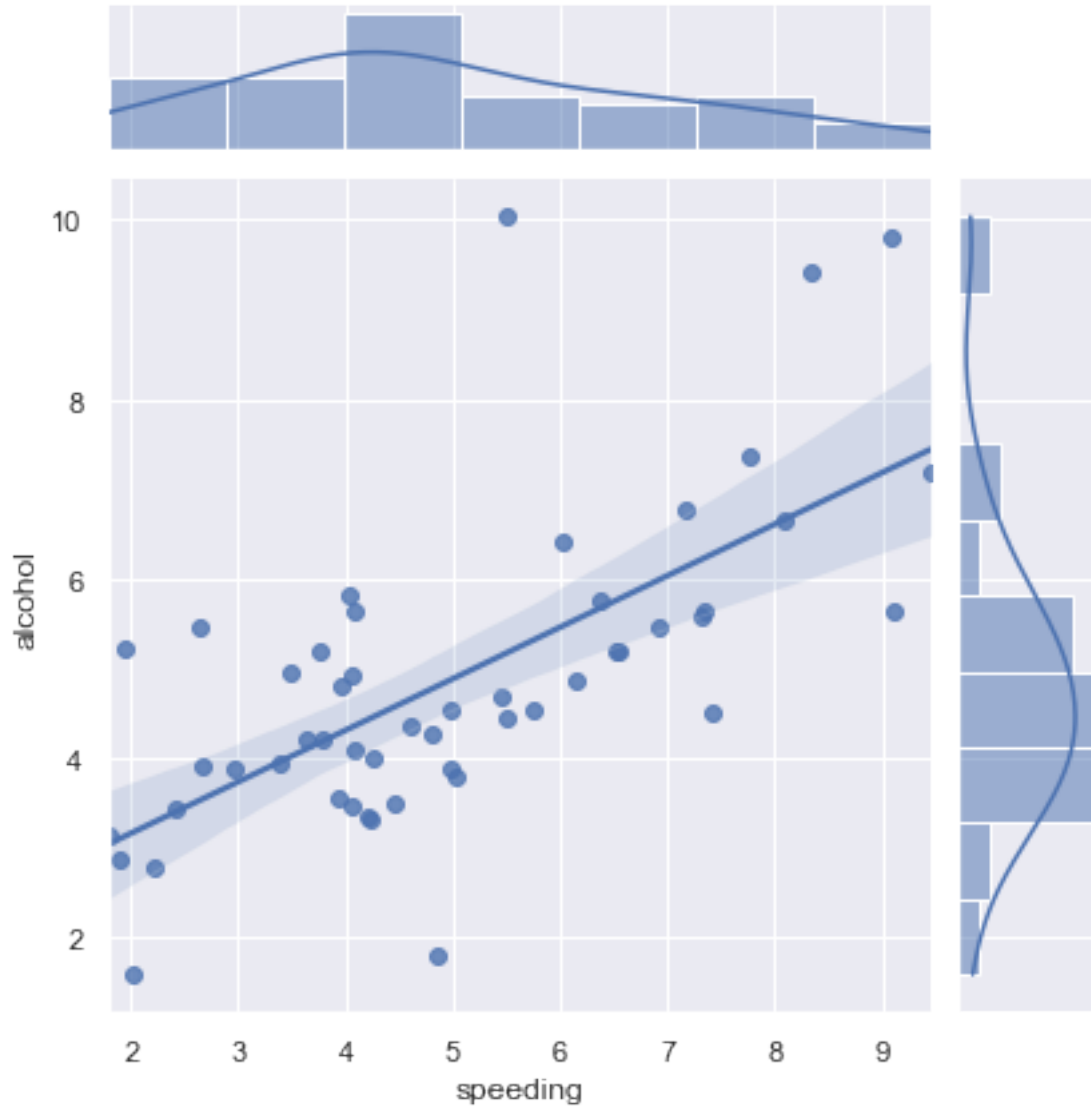


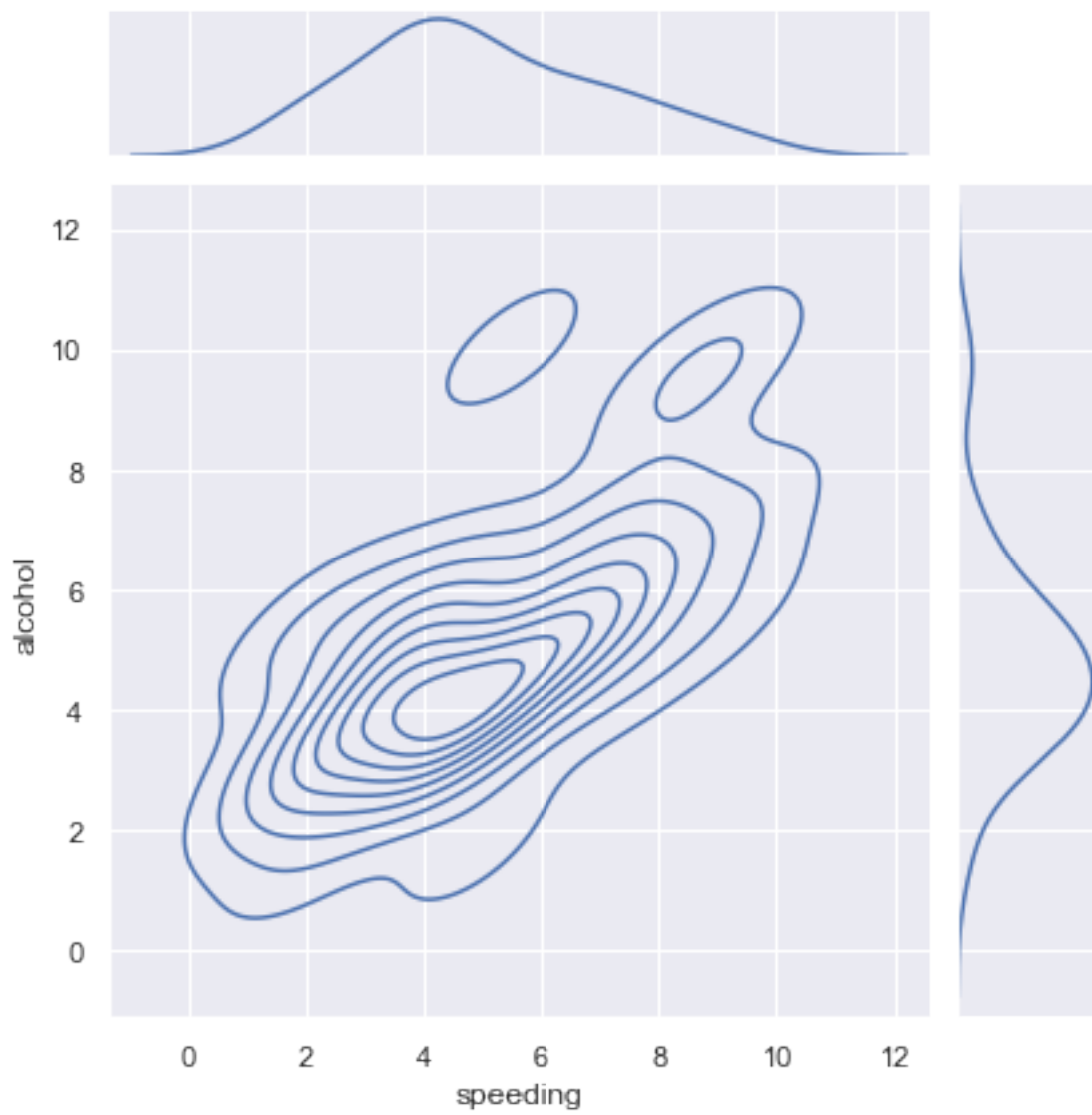
```
[6]: #Joint plot
```

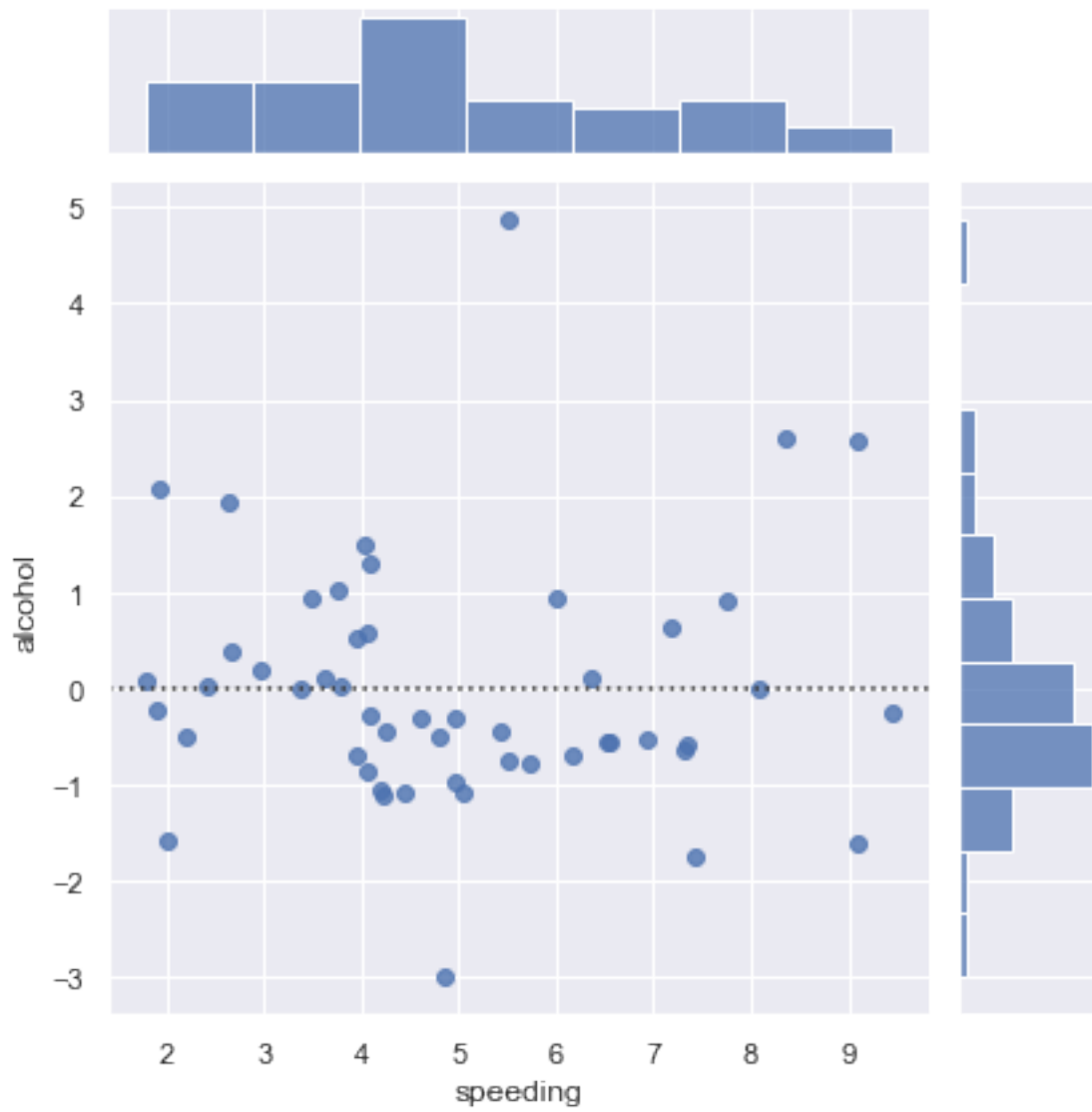
```
sns.jointplot(x='speeding', y='alcohol', data=crash_data, kind='reg')  
sns.jointplot(x='speeding', y='alcohol', data=crash_data, kind='kde')
```

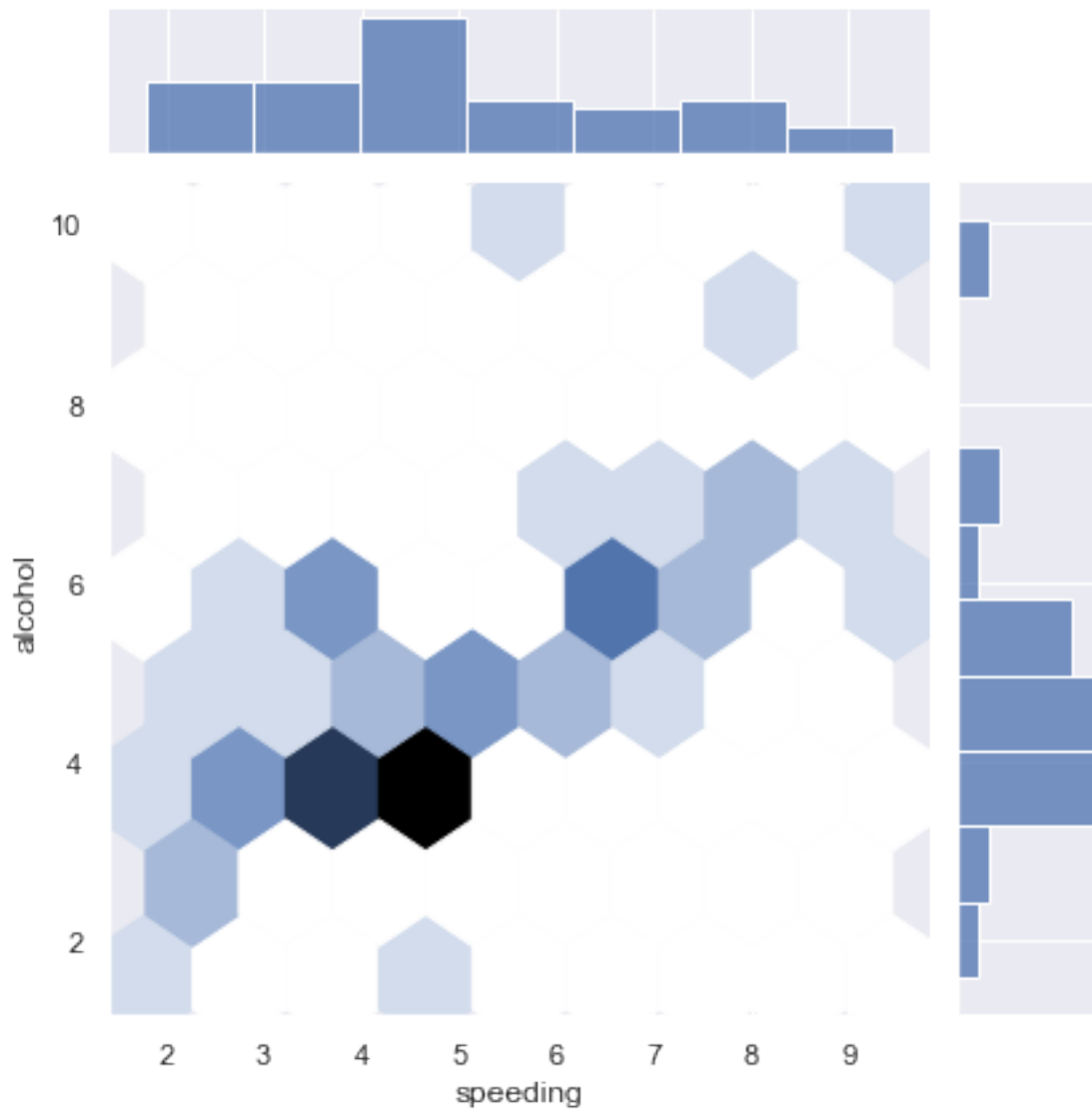
```
sns.jointplot(x='speeding', y='alcohol', data=crash_data, kind='resid')  
sns.jointplot(x='speeding', y='alcohol', data=crash_data, kind='hex')
```

[6]: <seaborn.axisgrid.JointGrid at 0x12c1563d0>





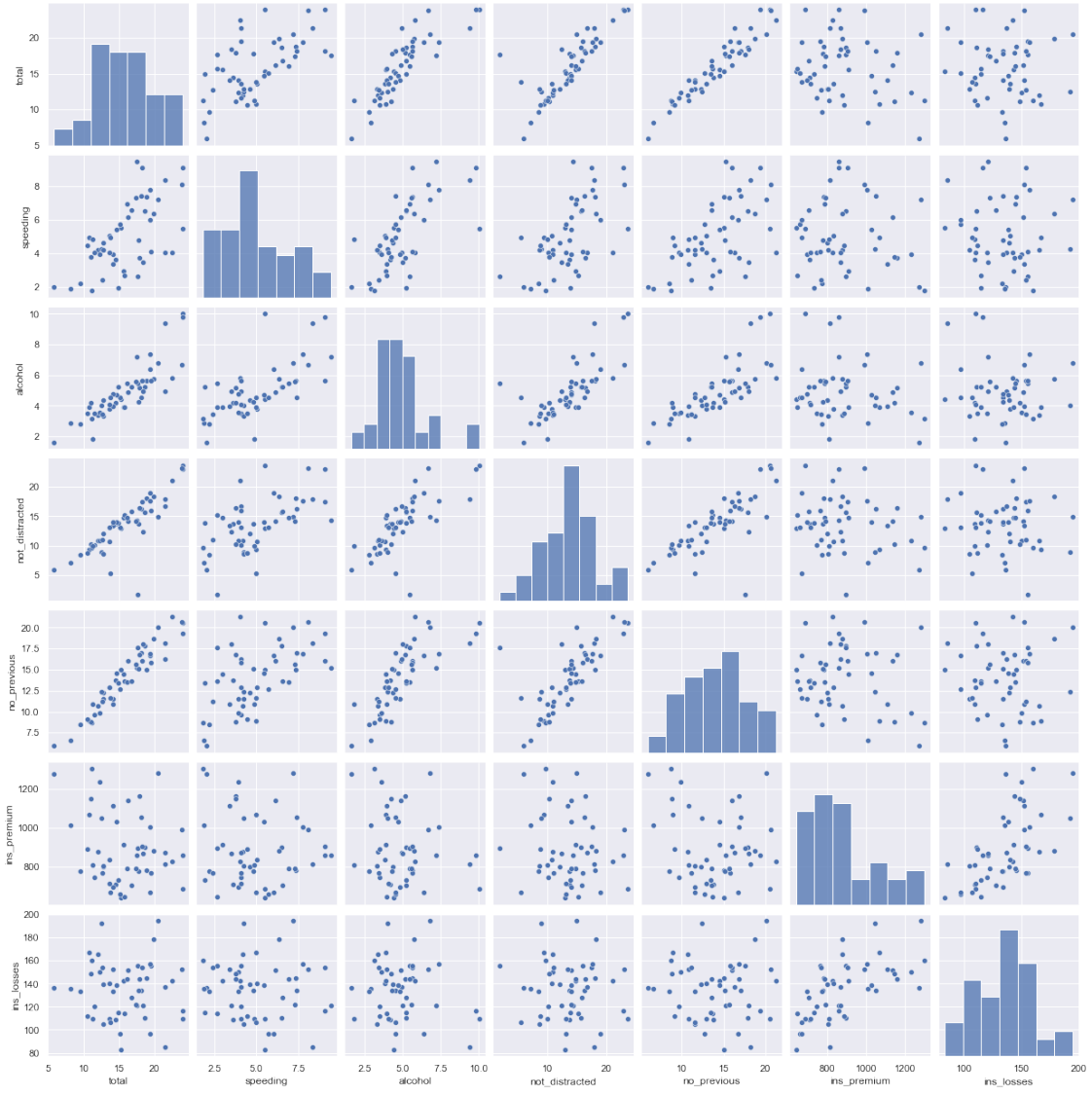


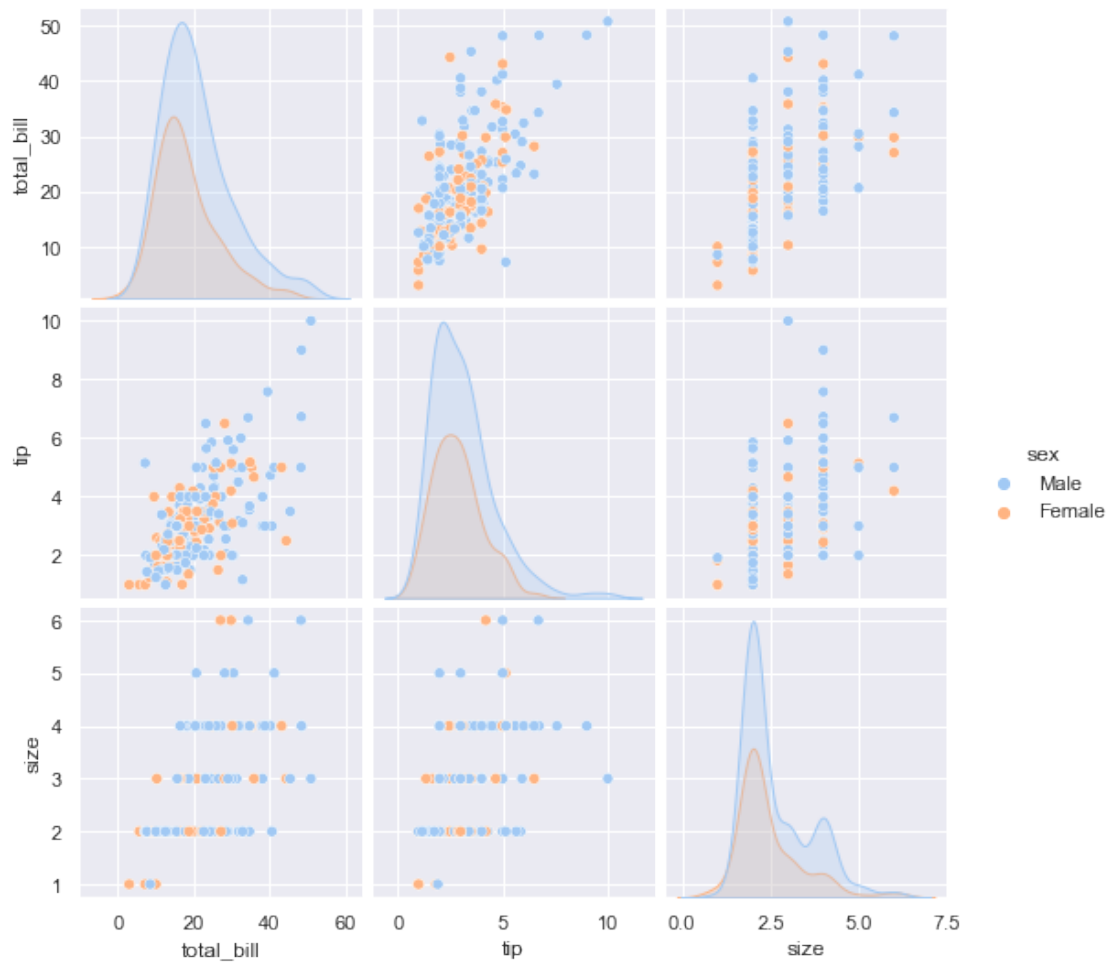


```
[7]: #Pair plots
sns.pairplot(crash_data)

sns.pairplot(tips_data, hue='sex', palette='pastel')
```

```
[7]: <seaborn.axisgrid.PairGrid at 0x12c8992e0>
```

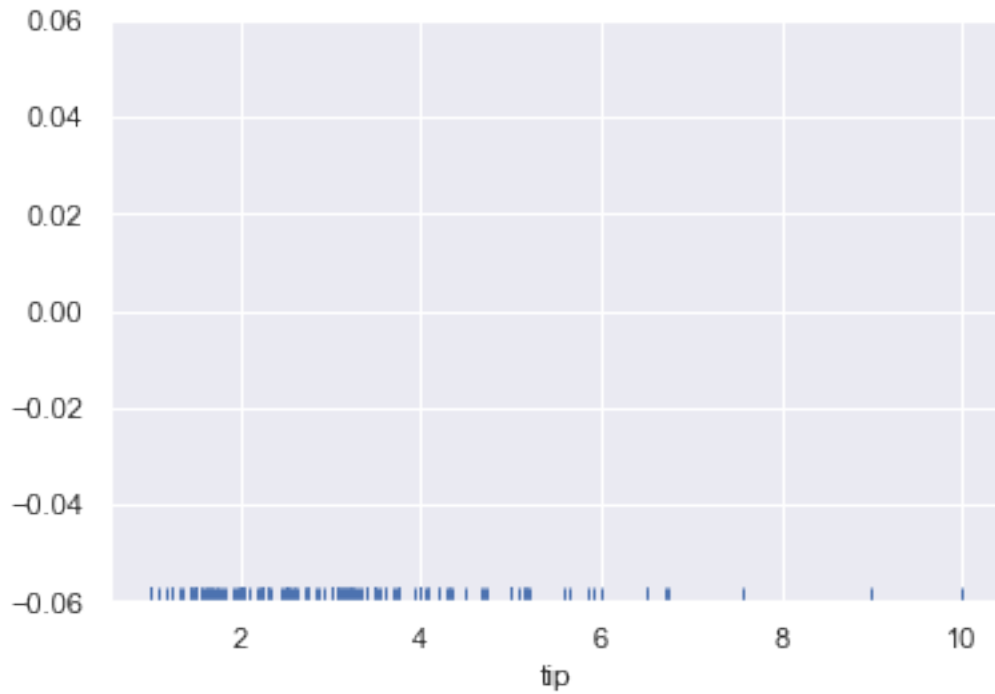





[8]: *#Rug Plot*

```
sns.rugplot(tips_data['tip'])
```

[8]: <AxesSubplot:xlabel='tip'>



4 Basic Styling

```
[9]: sns.set_style('darkgrid')

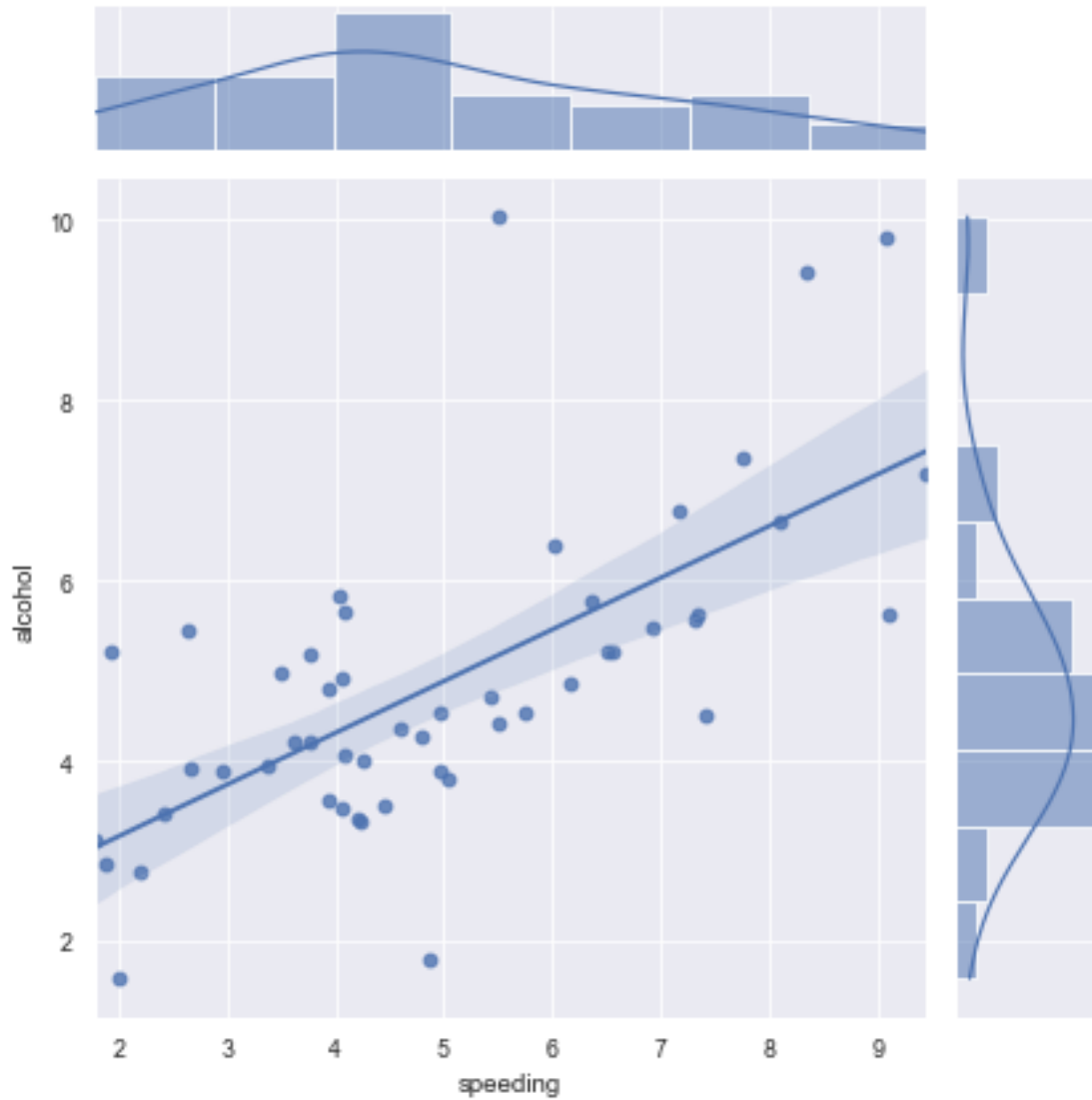
plt.figure(figsize=(8,4))

sns.set_context('paper')
sns.jointplot(x='speeding', y='alcohol', data=crash_data, kind='reg')

#https://matplotlib.org/stable/tutorials/colors/colormaps.html
```

```
[9]: <seaborn.axisgrid.JointGrid at 0x129e7c190>
```

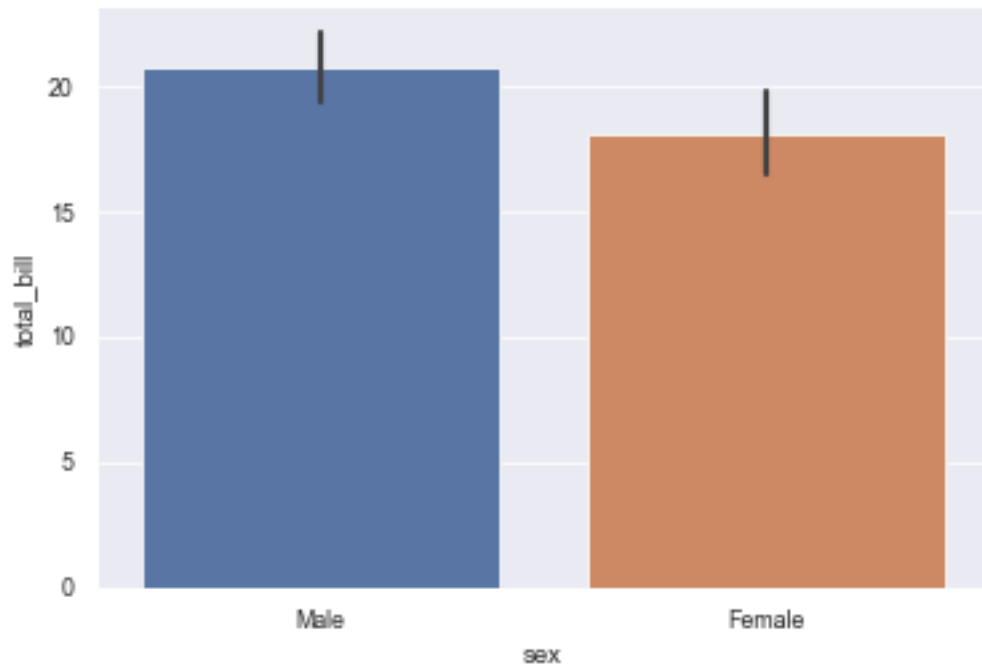
```
<Figure size 576x288 with 0 Axes>
```



5 Catagorical Plotting

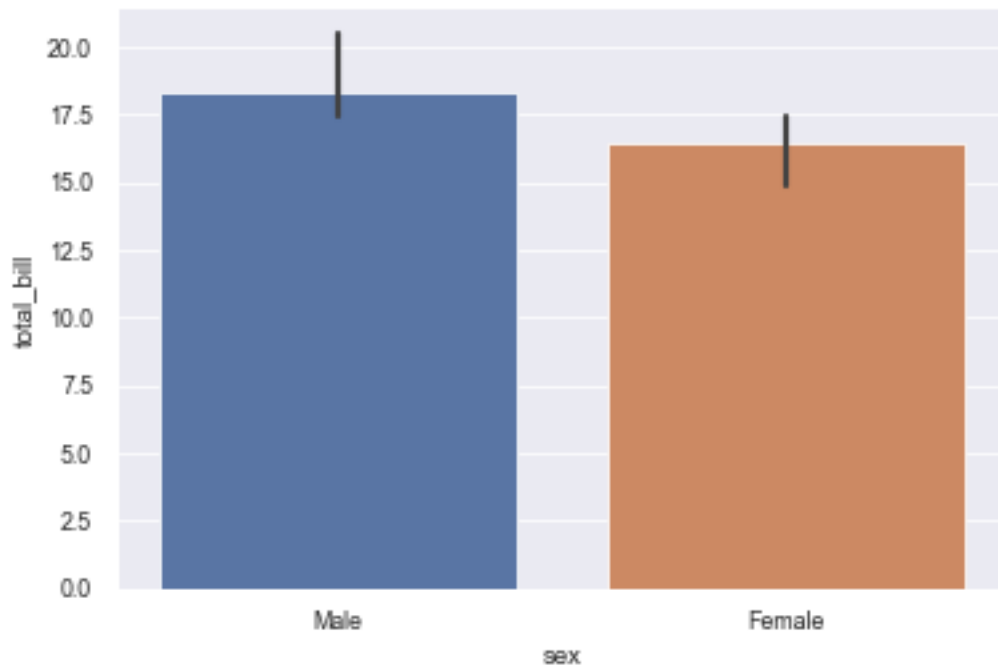
```
[10]: #Bar Plot
sns.barplot(x='sex', y='total_bill', data=tips_data, estimator=np.mean)
```

```
[10]: <AxesSubplot:xlabel='sex', ylabel='total_bill'>
```



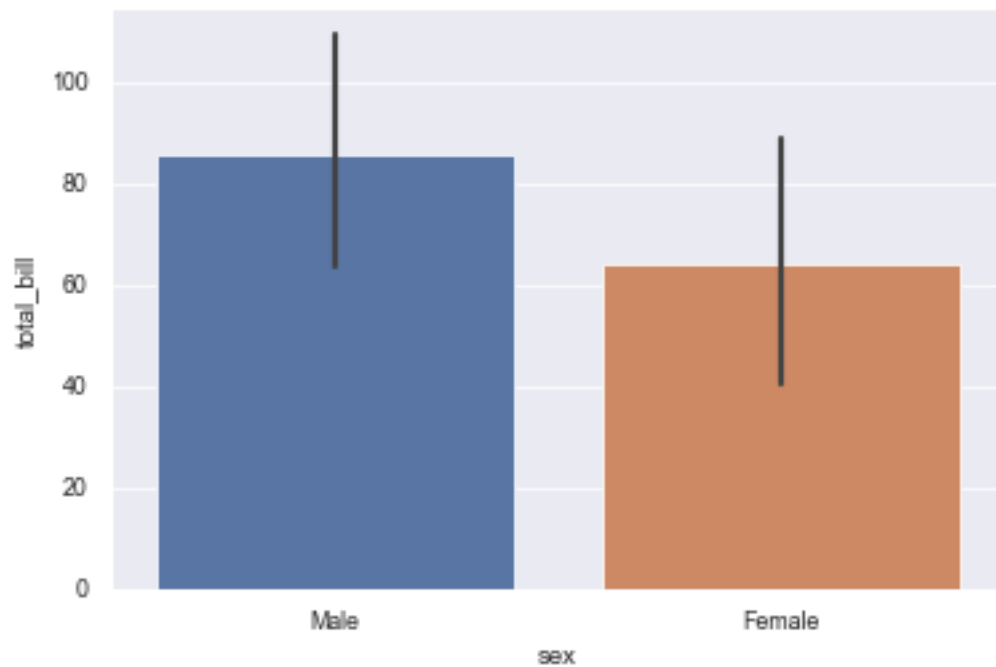
```
[11]: sns.barplot(x='sex', y='total_bill', data=tips_data, estimator=np.median)
```

```
[11]: <AxesSubplot:xlabel='sex', ylabel='total_bill'>
```



```
[12]: sns.barplot(x='sex', y='total_bill', data=tips_data, estimator=np.cov)
```

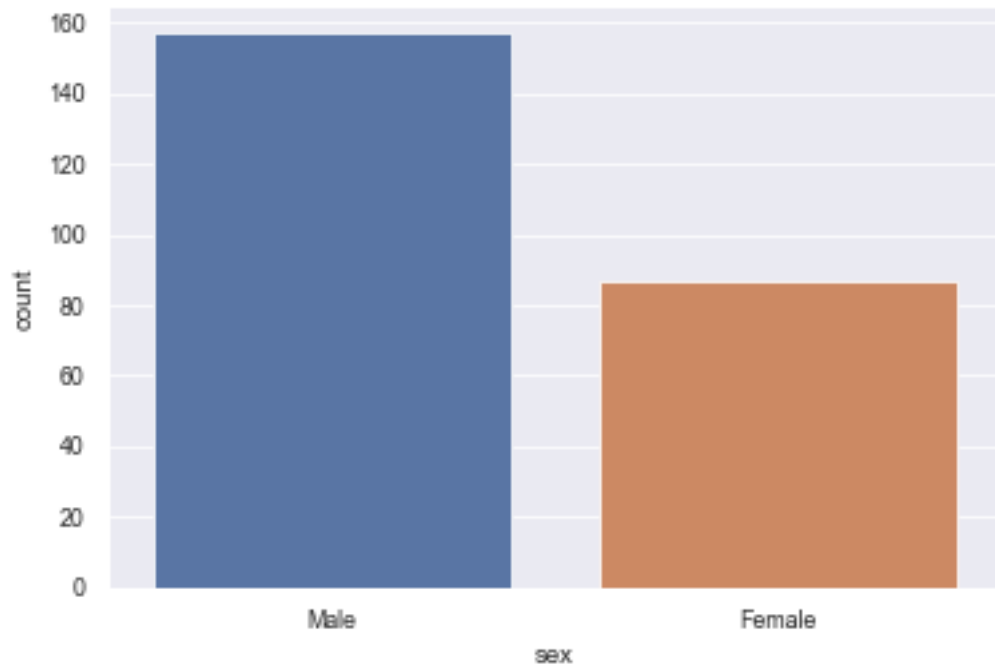
```
[12]: <AxesSubplot:xlabel='sex', ylabel='total_bill'>
```



```
[13]: #Count Plot
```

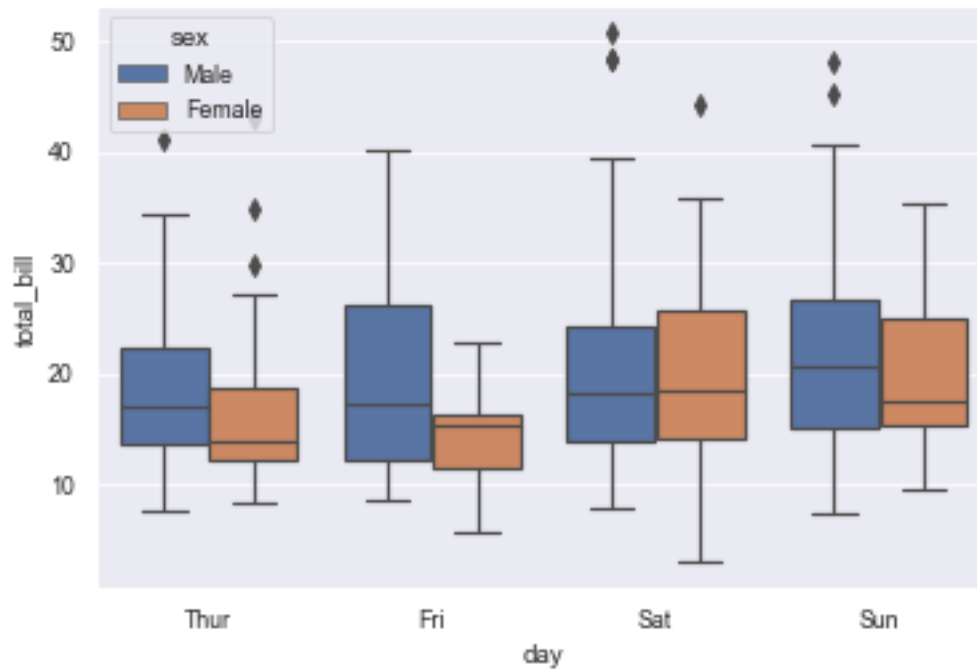
```
sns.countplot(x='sex', data=tips_data)
```

```
[13]: <AxesSubplot:xlabel='sex', ylabel='count'>
```



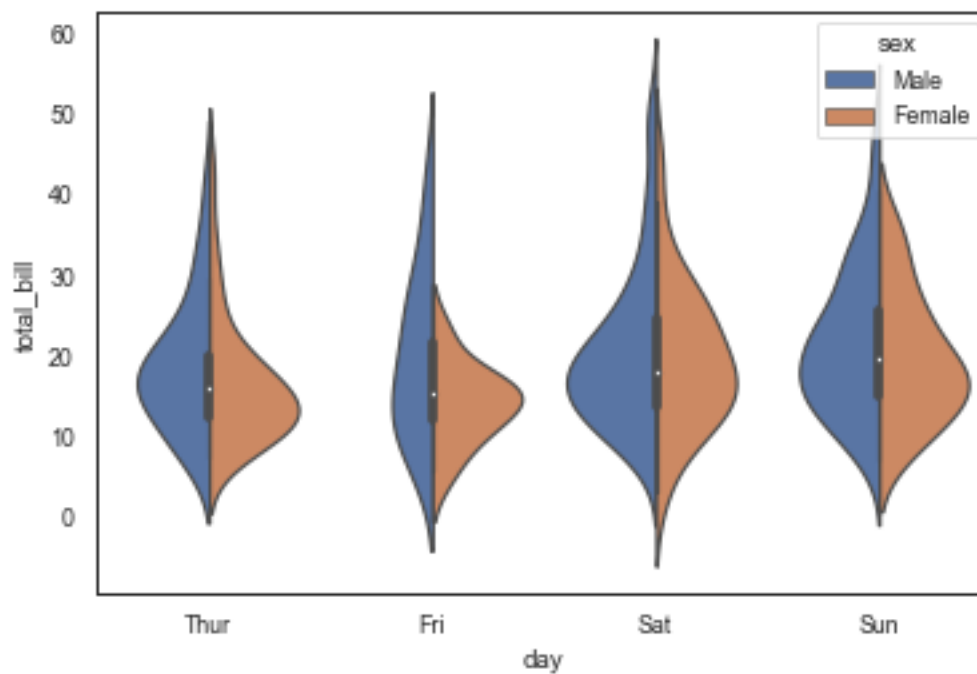
```
[14]: #Box Plot  
  
sns.boxplot(x='day', y='total_bill', data=tips_data, hue='sex')  
  
#plt.legend(loc=0)
```

```
[14]: <AxesSubplot:xlabel='day', ylabel='total_bill'>
```



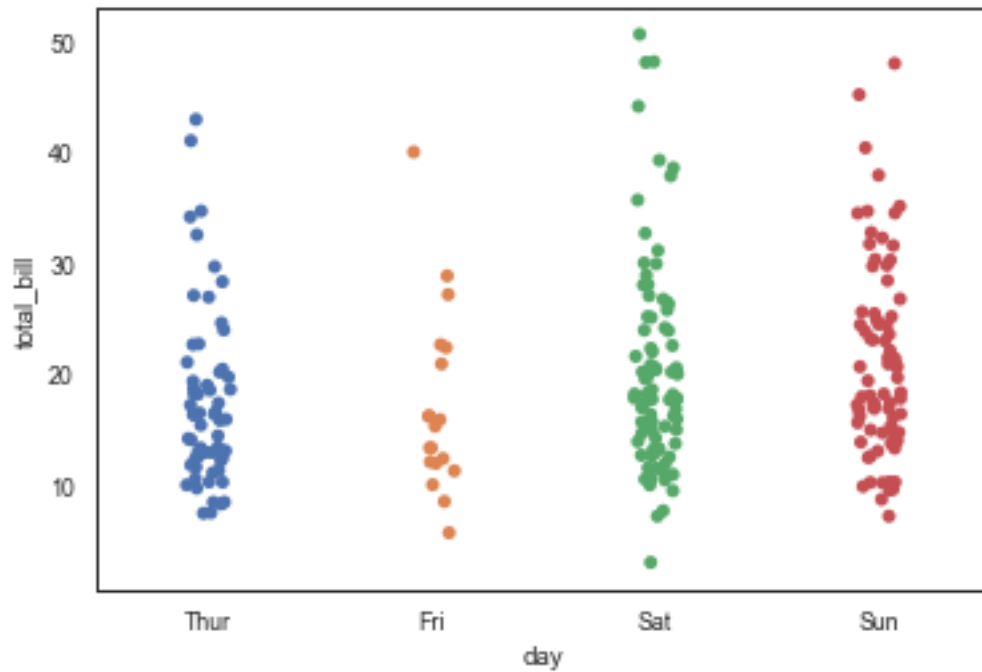
```
[15]: #Violin Plot
sns.set_style('white')
sns.violinplot(x='day', y='total_bill', data=tips_data, hue='sex',split=True)
```

```
[15]: <AxesSubplot:xlabel='day', ylabel='total_bill'>
```

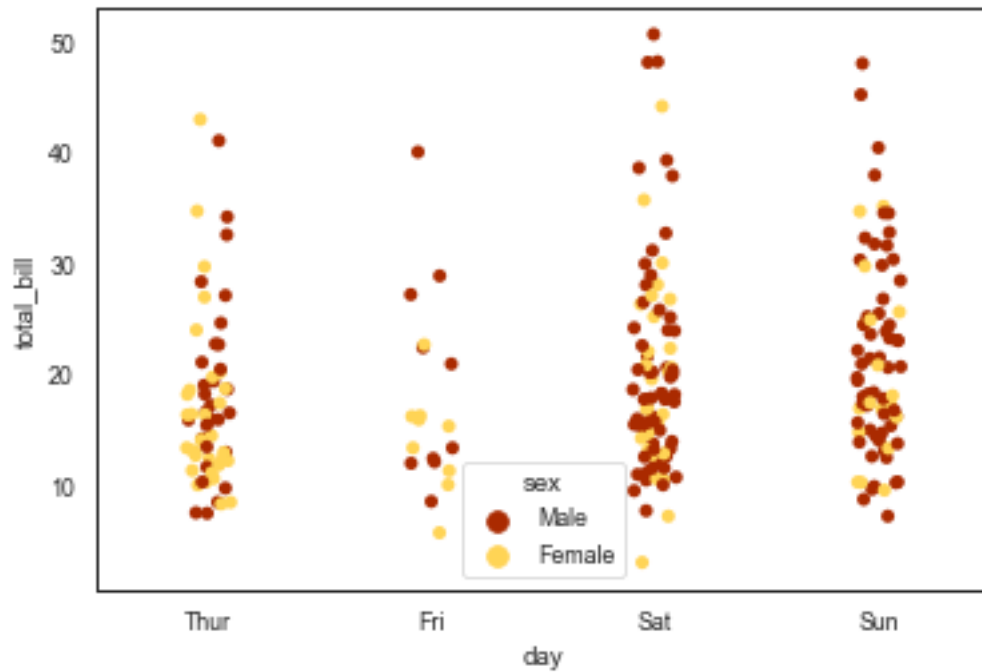



```
[16]: #Strip Plot

sns.stripplot(x='day', y='total_bill', data=tips_data)
plt.show()
sns.stripplot(x='day', y='total_bill', data=tips_data,
             hue='sex', palette='afmhot')
```

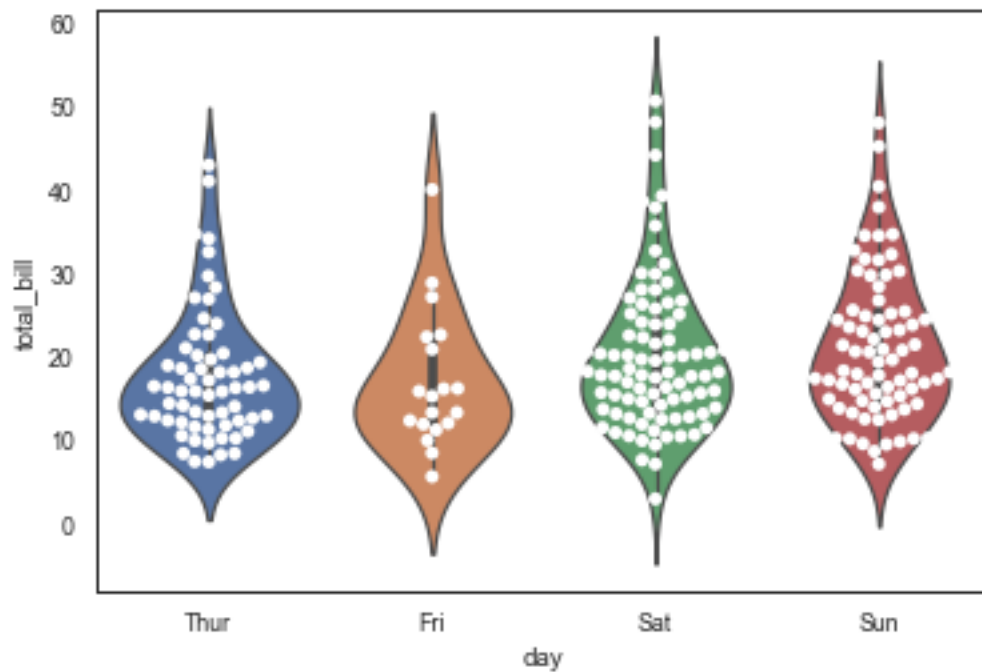


```
[16]: <AxesSubplot:xlabel='day', ylabel='total_bill'>
```



```
[17]: #Swarm Plot
sns.violinplot(x='day', y='total_bill', data=tips_data)
sns.swarmplot(x='day', y='total_bill', data=tips_data, color='white')
```

```
[17]: <AxesSubplot:xlabel='day', ylabel='total_bill'>
```



6 Matrix Plot

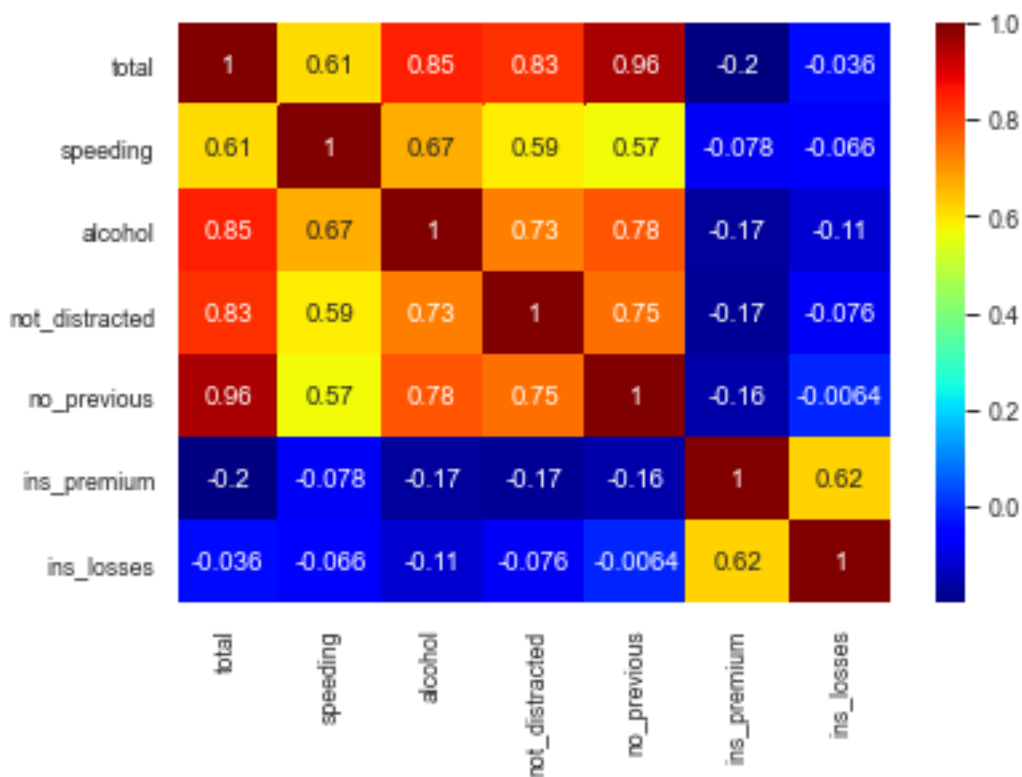
```
[18]: #Heatmap

#Get correlation matrix

corr_mtx = crash_data.corr()

sns.heatmap(corr_mtx, annot=True, cmap='jet')
```

[18]: <AxesSubplot:>

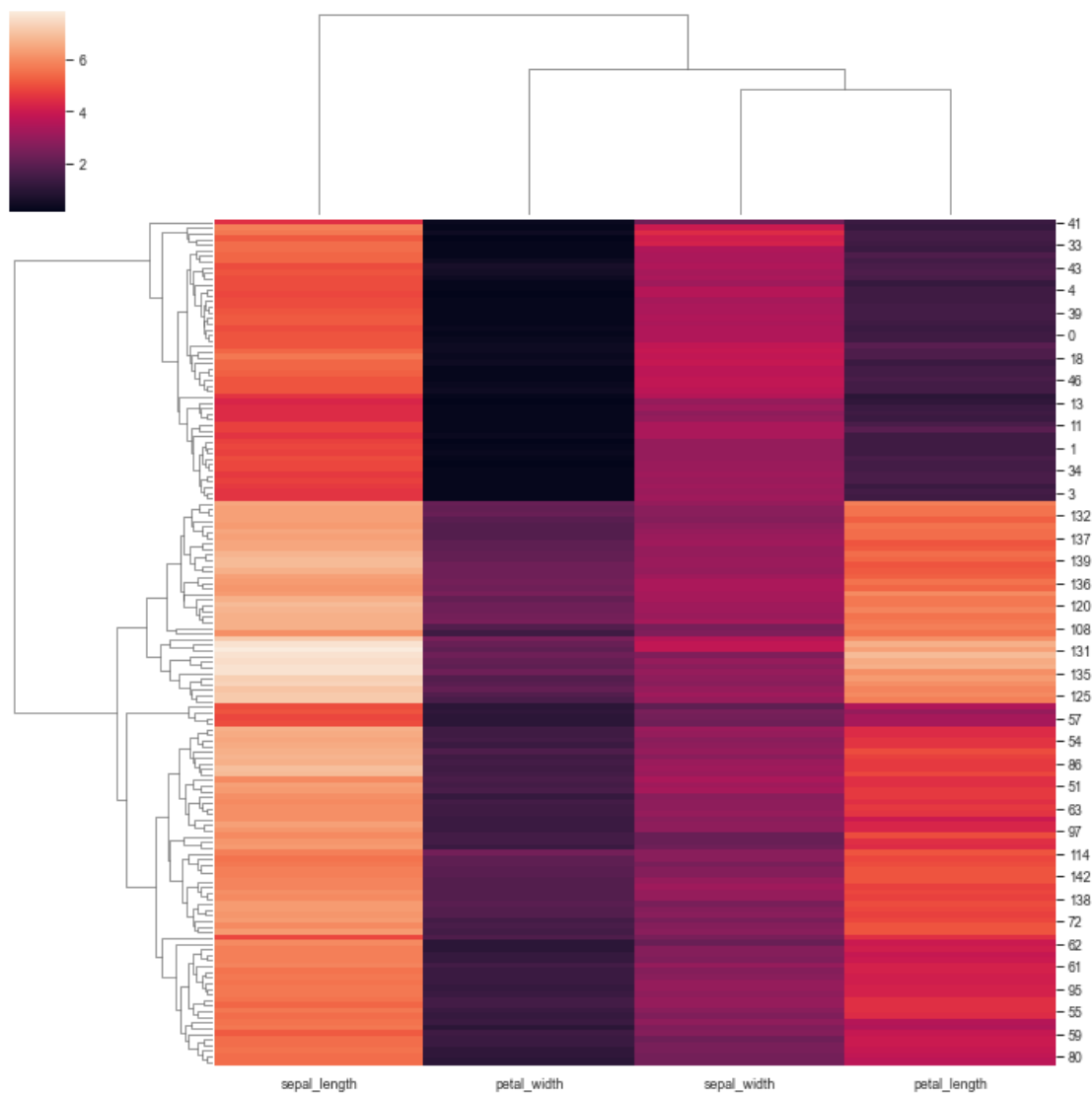


```
[21]: #Cluster Map

iris_data.pop('species')

sns.clustermap(iris_data)
```

[21]: <seaborn.matrix.ClusterGrid at 0x12f5f4820>



```
[22]: #Pair Grid
iris_pg = sns.PairGrid(tips_data, hue='sex')

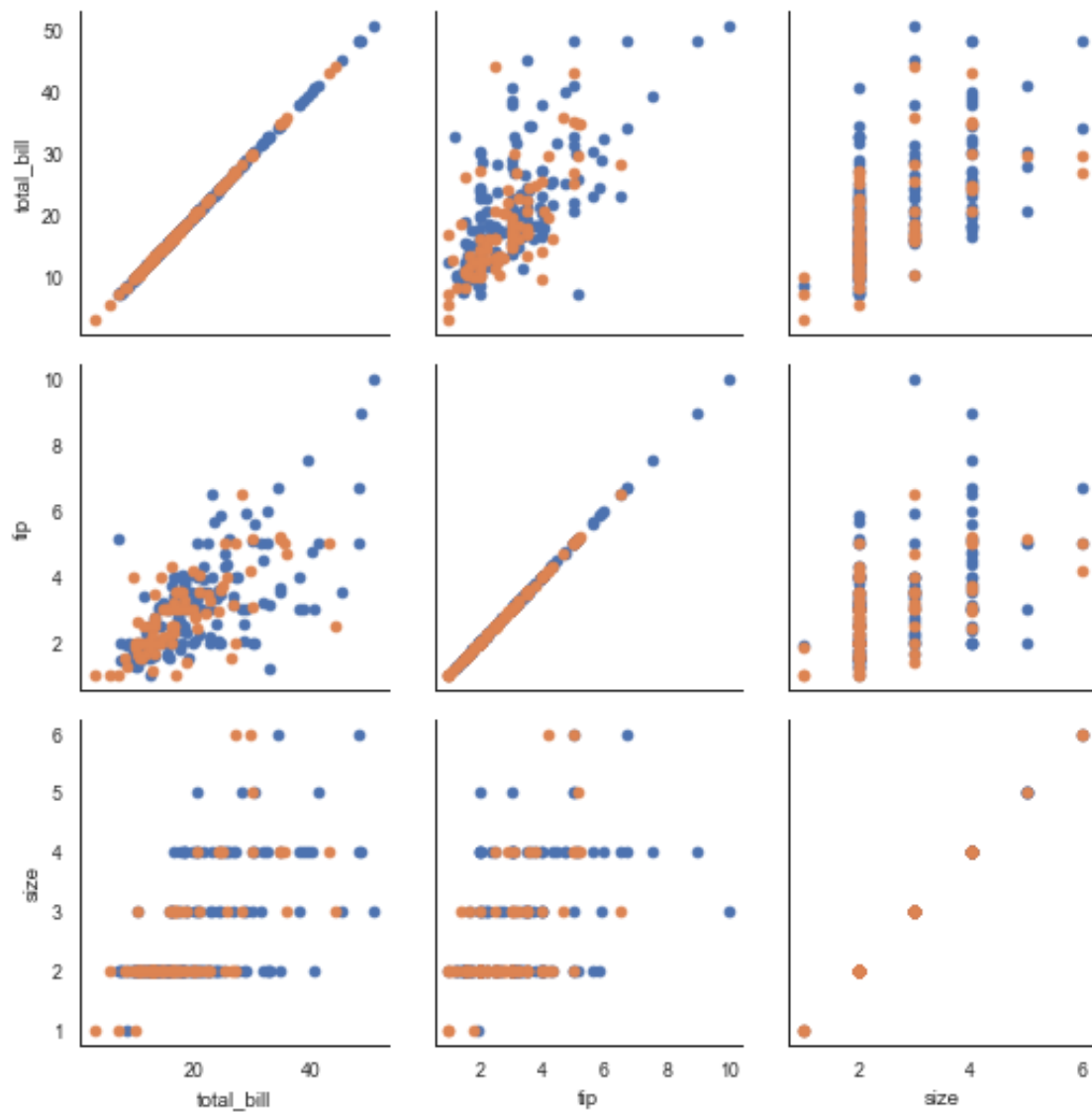
iris_pg.map(plt.scatter)

plt.show()

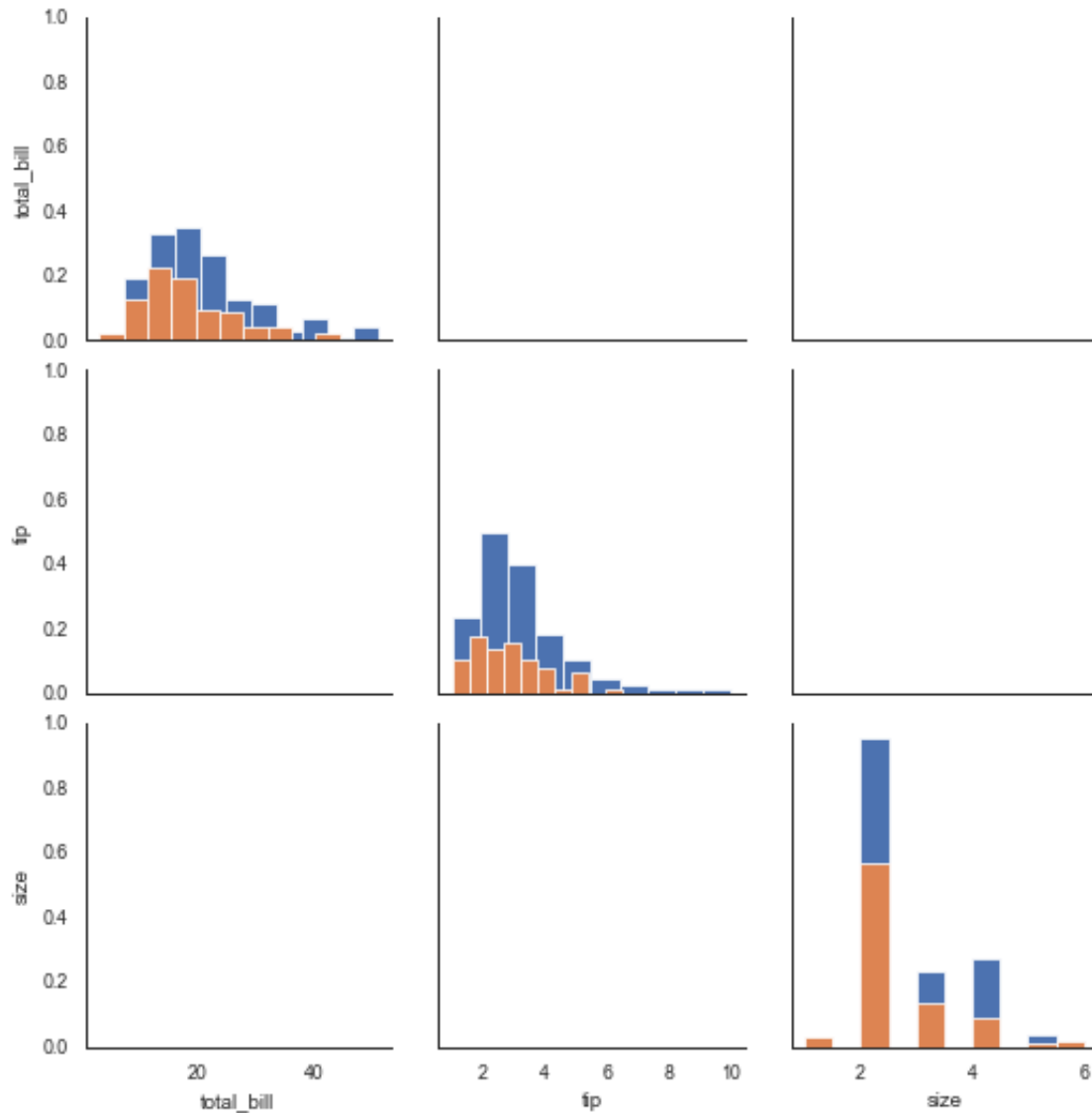
iris_pg2 = sns.PairGrid(tips_data, hue='sex')

iris_pg2.map_diag(plt.hist)
```

#Loads of customizability here, <https://seaborn.pydata.org/generated/seaborn.pairplot.html?highlight=pairplot#seaborn.pairplot>



[22]: <seaborn.axisgrid.PairGrid at 0x12f8e5610>



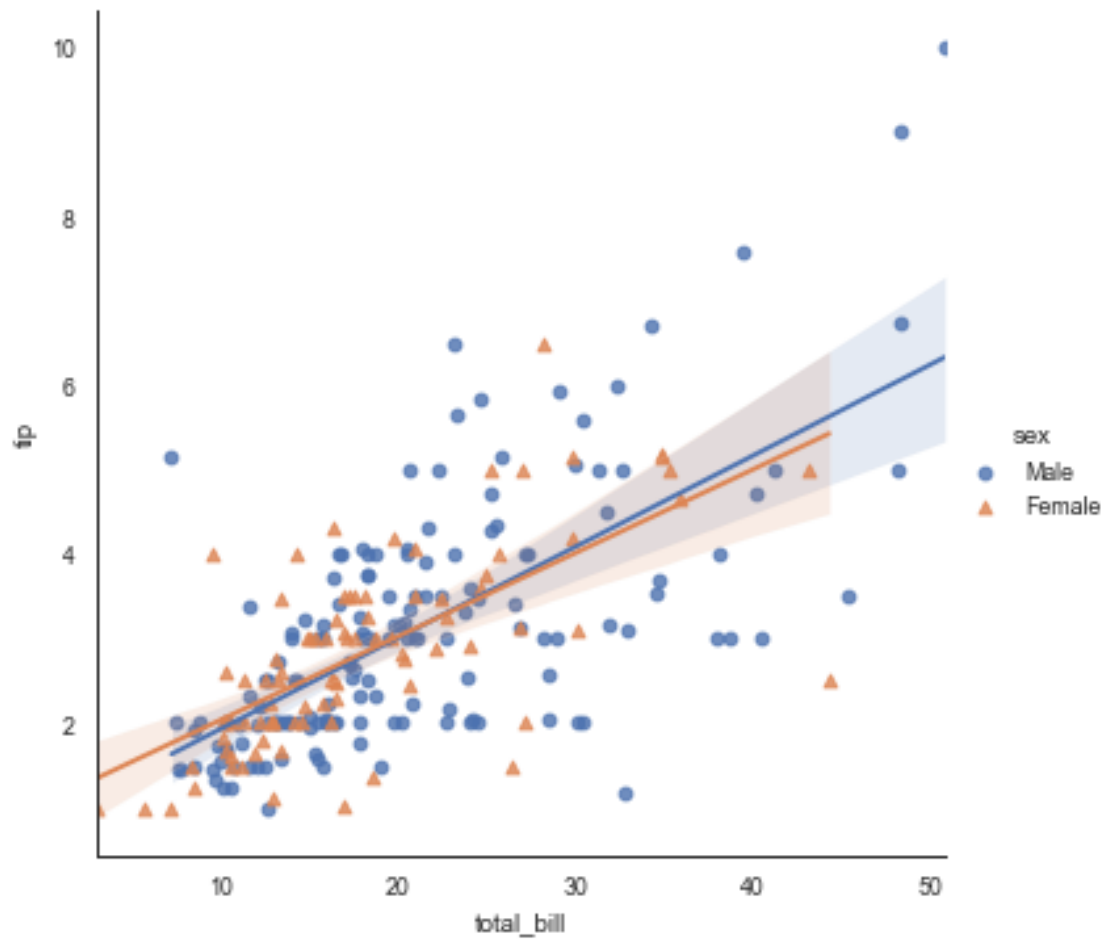
7 Regression Plots

```
[23]: tips_data.head()

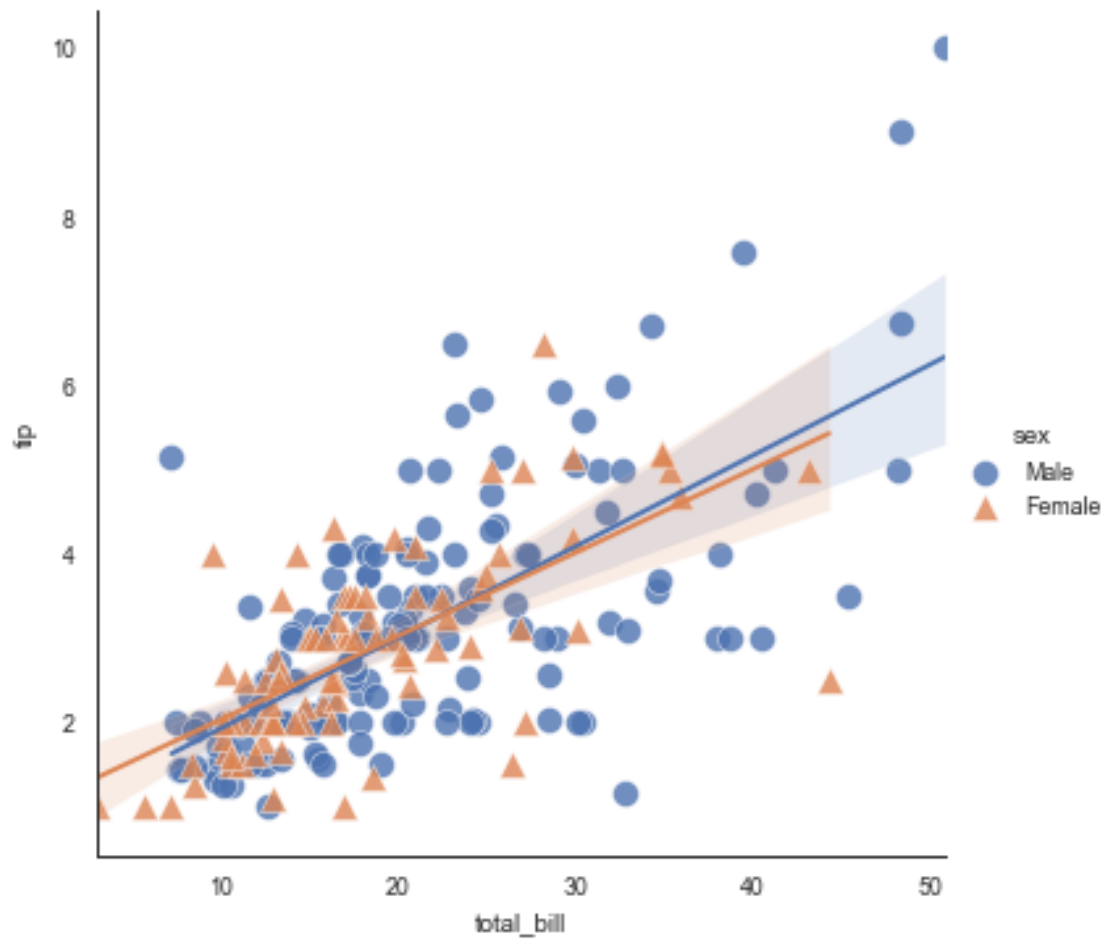
plt.figure(figsize=(8,6))
sns.set_context('paper')
sns.lmplot(x='total_bill', y='tip', data=tips_data, hue='sex',
           markers=['o', '^'])
plt.show()
```

```
sns.lmplot(x='total_bill', y='tip', data=tips_data, hue='sex',
↳ markers=['o', '^'], scatter_kws={'s':100, 'linewidth':0.5,
↳ 'edgecolor': 'white'})
```

<Figure size 576x432 with 0 Axes>

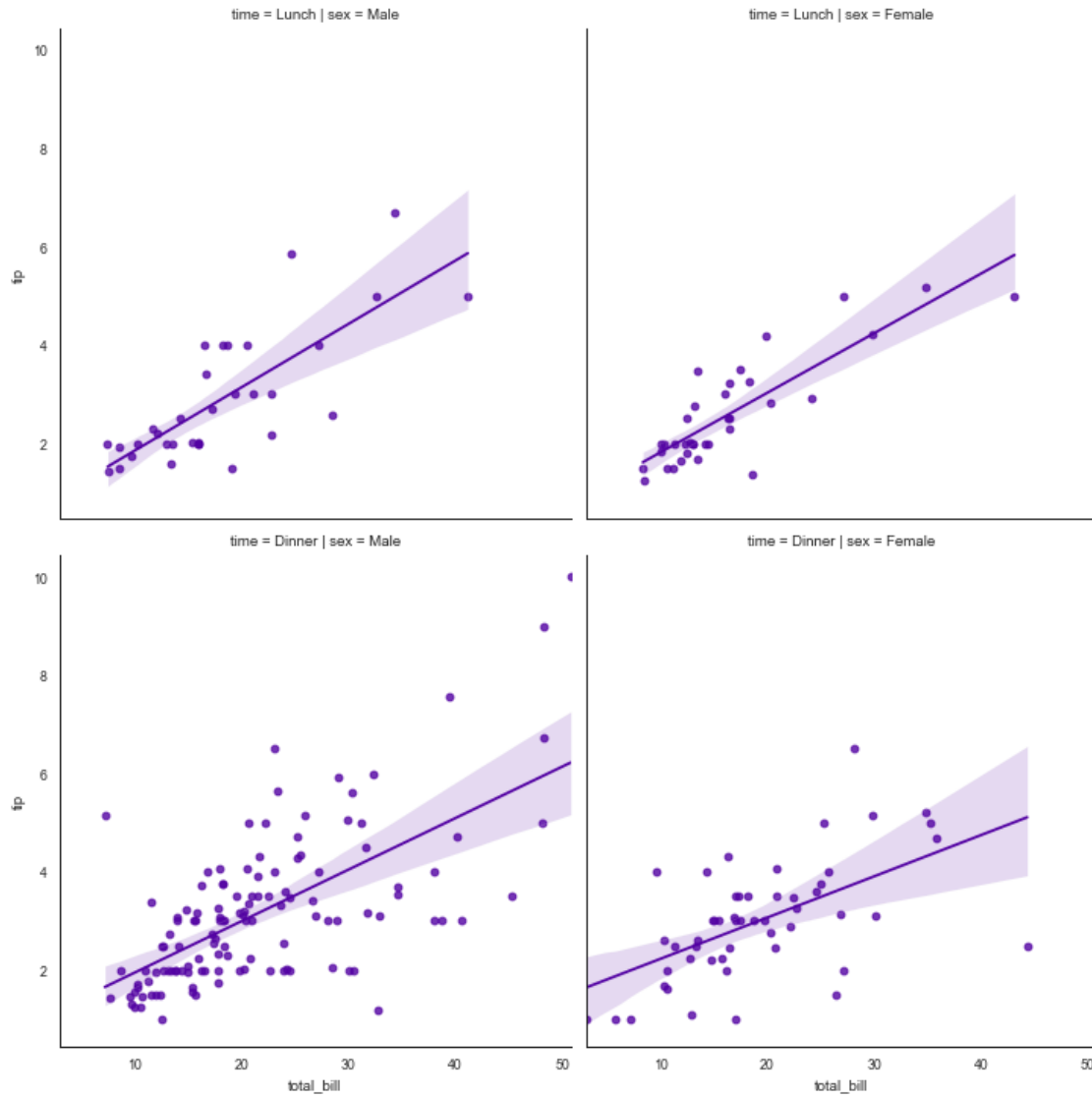


[23]: <seaborn.axisgrid.FacetGrid at 0x1307520a0>



```
[24]: sns.set_palette('plasma')
sns.set_context('paper')
sns.lmplot(x='total_bill', y='tip', data=tips_data, col='sex', row='time')
```

```
[24]: <seaborn.axisgrid.FacetGrid at 0x13056c4f0>
```

[]: