Python102

Python for Data Science Bootcamp
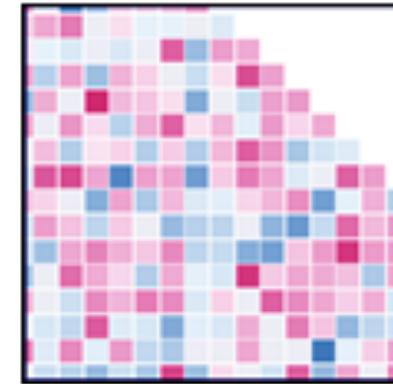
(5.2) Python for Data Visualization

Seaborn

AIAT Academy

Python for Data Visualization Outline

- Matplotlib

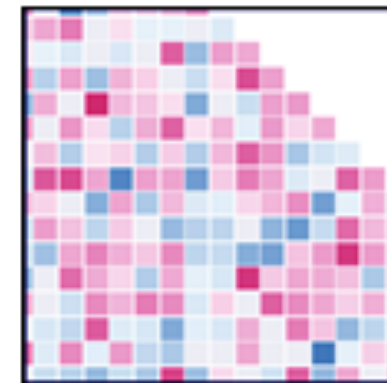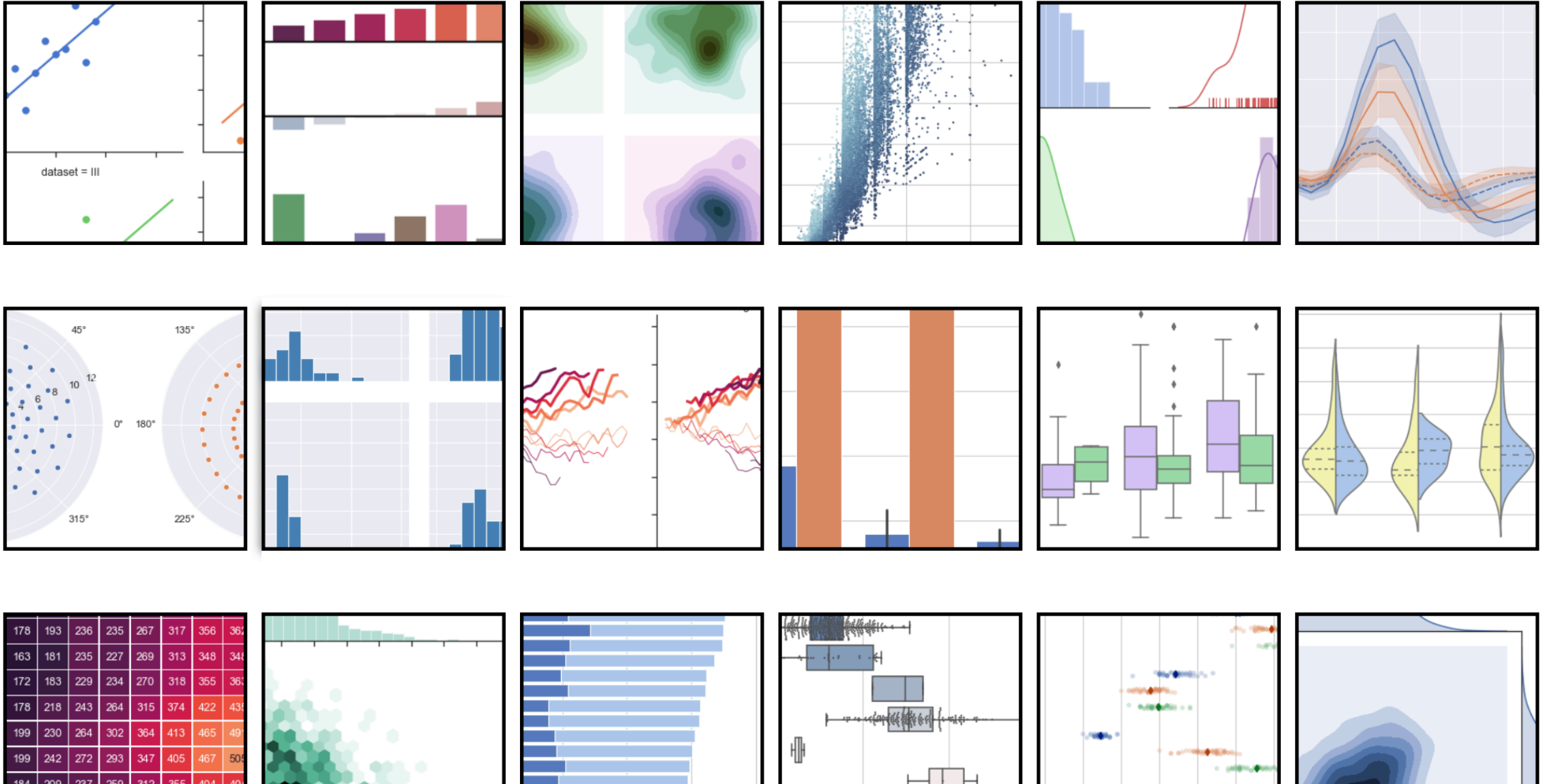- **Seaborn**

Seaborn

Introduction to Seaborn

- A Statistical plotting library

- Beautiful default styles

- Designed to work well with Pandas DataFrame object

- https://seaborn.pydata.org

Seaborn

Introduction to Seaborn (Example Gallery)

Seaborn Installation

- To install Seaborn, just going to your terminal or command prompt and typing


```
conda install seaborn
```

or

```
pip install seaborn
```

Seaborn Usages (Distribution Plots)

```
import seaborn as sns

import matplotlib.pyplot as plt

%matplotlib inline

tips = sns.load_dataset('tips')

tips.head()
```
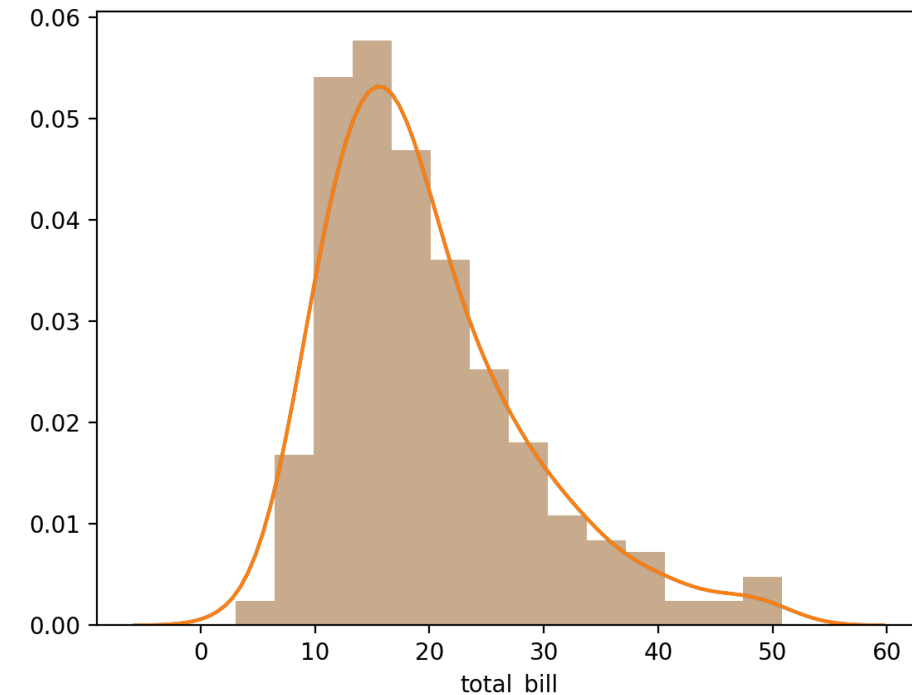
Seaborn comes with
build-in datasets

```
#     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
```

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

plt.show()
```
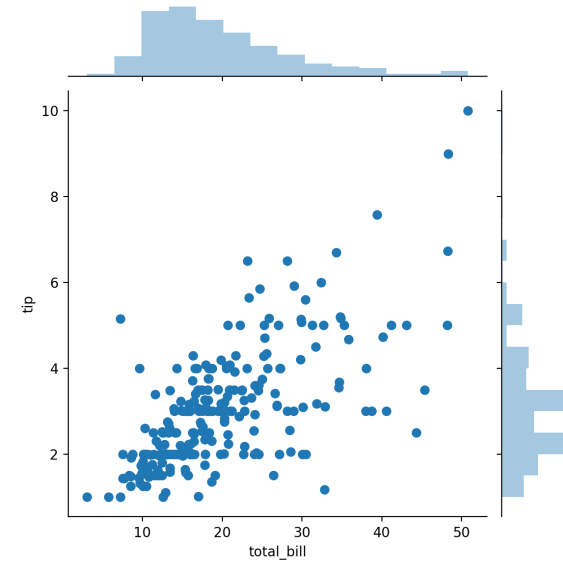
Seaborn Usages (Joint Plots)

```
tips.head()
```

```
#      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
```

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

plt.show()
```

Seaborn Usages (Pair Plots)

```
tips.head()

#     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
```

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

plt.show()
```

```
# plot every pairs from data
# by "hue"
```

Seaborn Usages (Bar Plots)
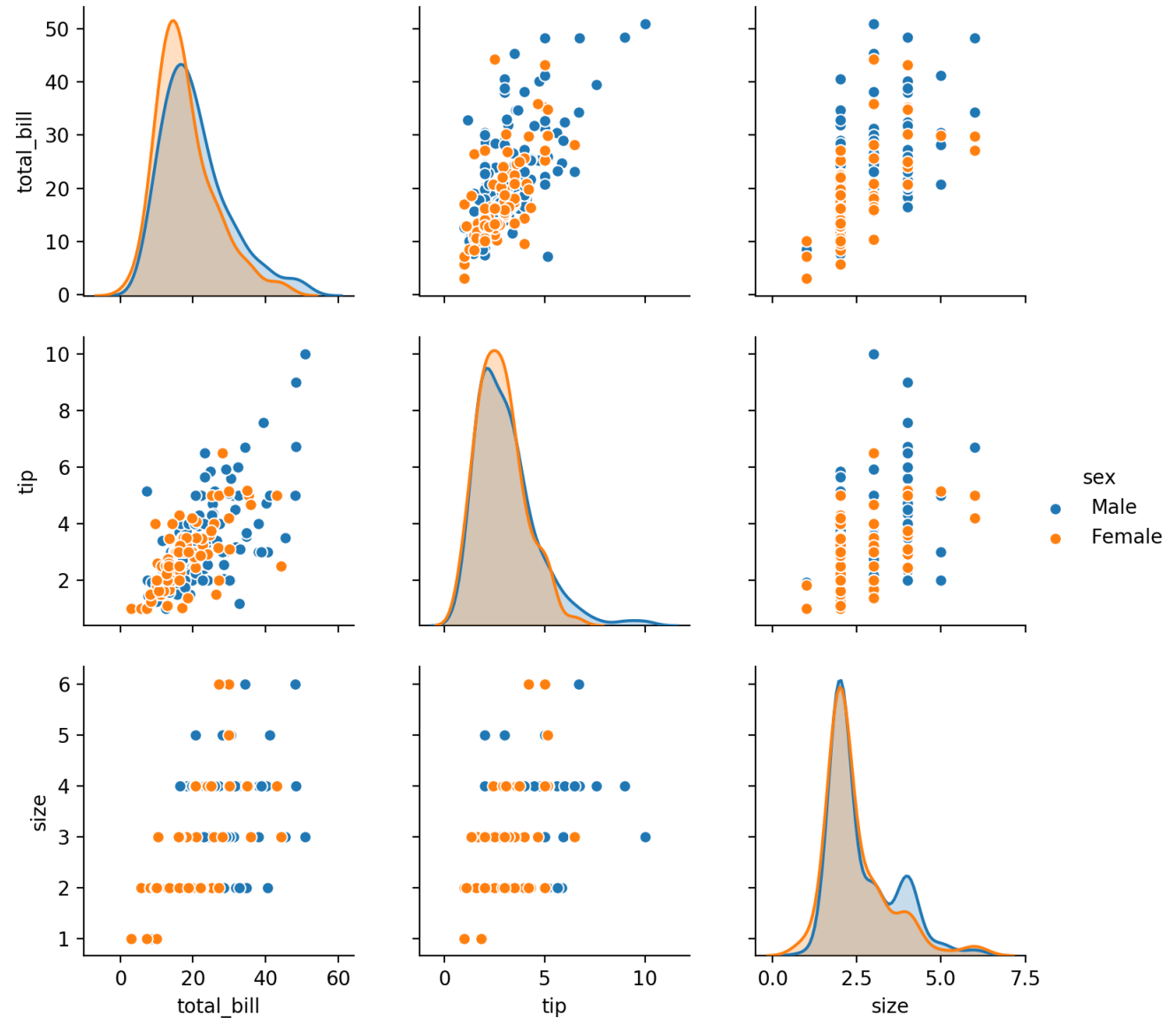
```
tips.head()
```

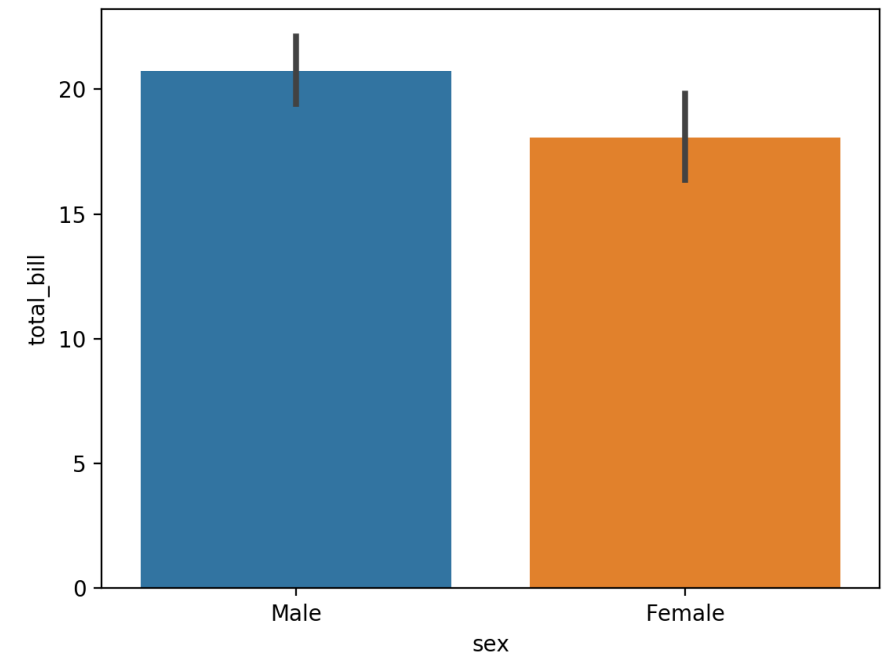```
total_bill   tip   sex smoker   day   time   size
0   16.99   1.01   Female   No   Sun   Dinner   2
1   10.34   1.66   Male   No   Sun   Dinner   3
2   21.01   3.50   Male   No   Sun   Dinner   3
3   23.68   3.31   Male   No   Sun   Dinner   2
4   24.59   3.61   Female   No   Sun   Dinner   4
```

```
sns.barplot(x="sex",y="total_bill",data=tips)
# barplot default in average plotting
plt.show()
```
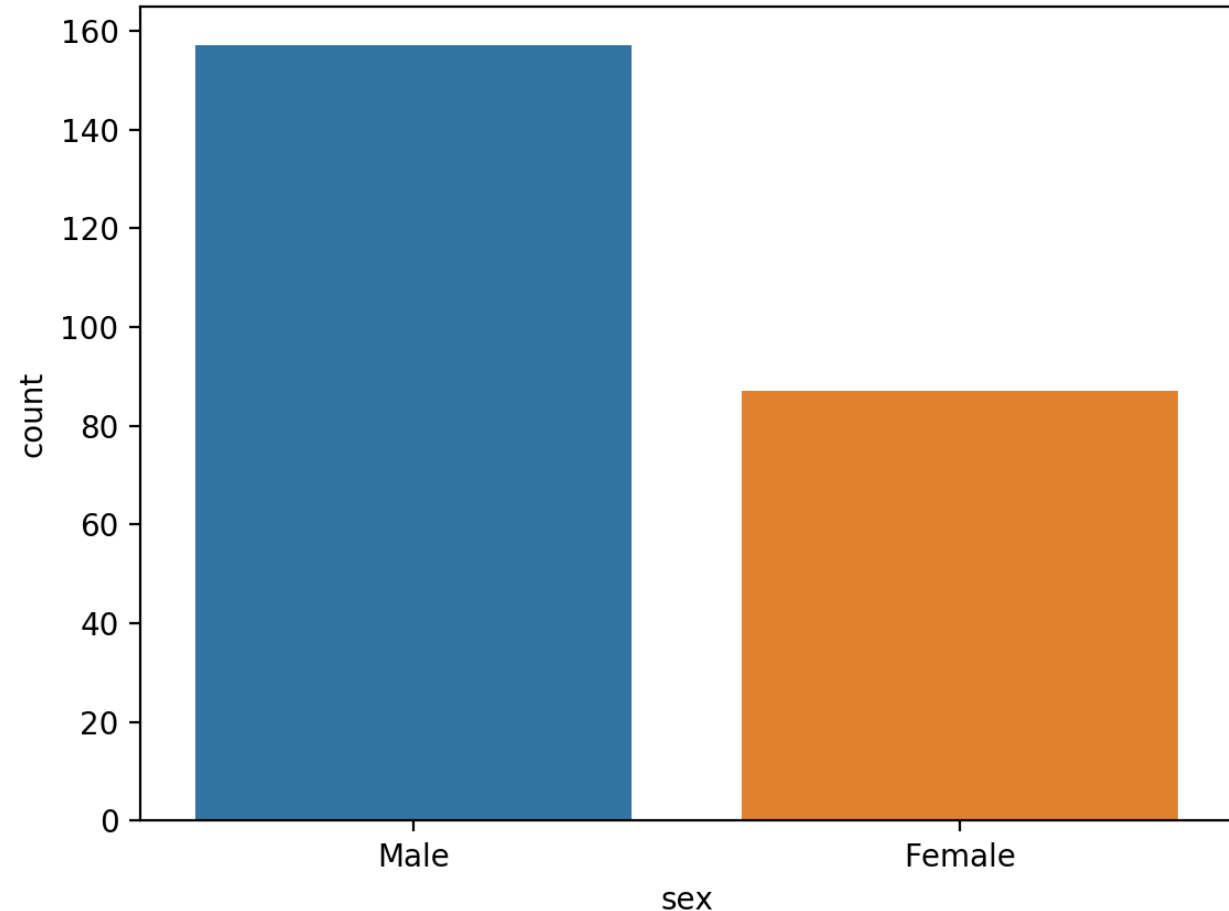
# Seaborn Usages (Count Plots)

`tips.head()`

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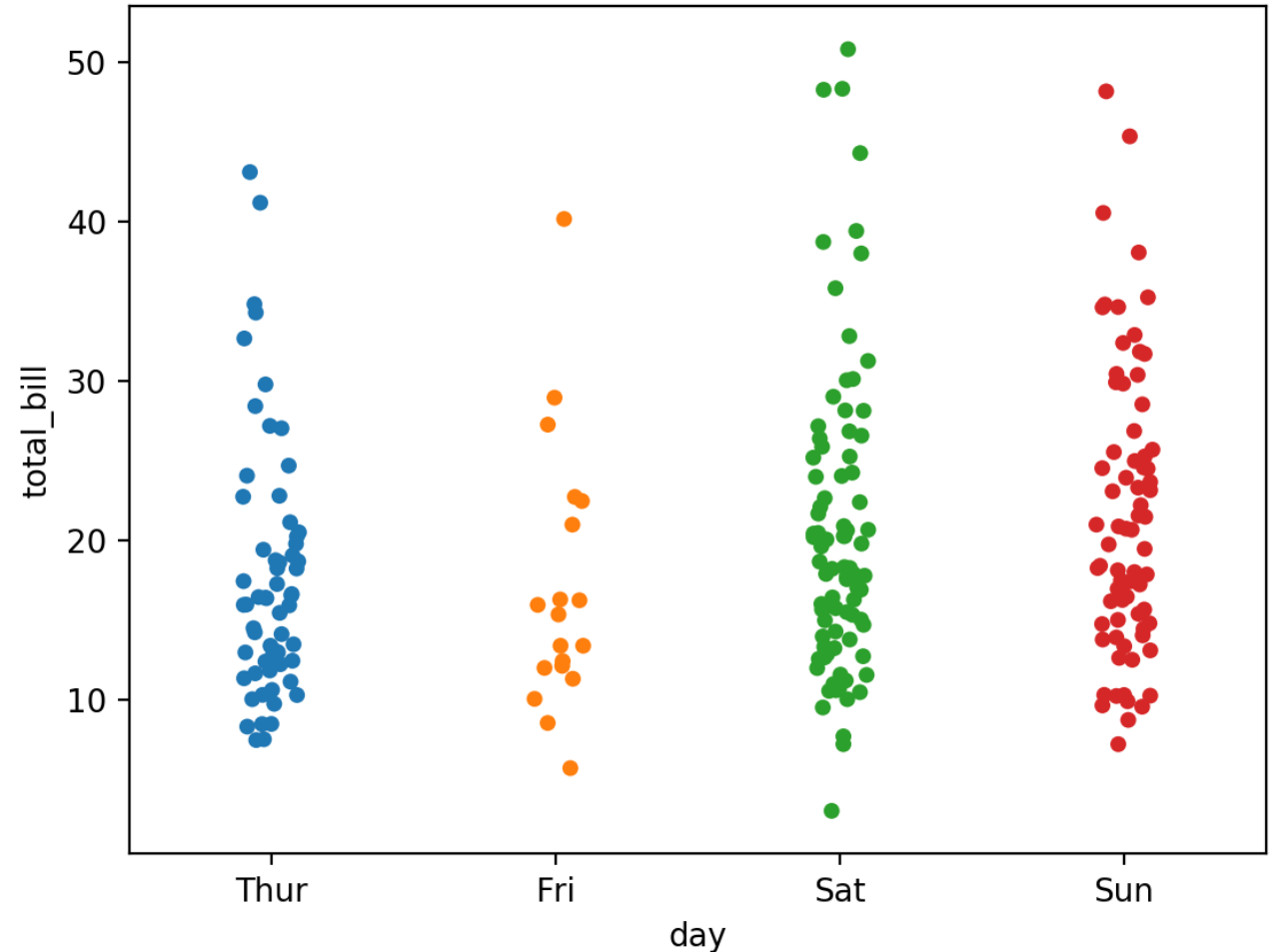
`sns.countplot(x="sex",data=tips)`

`plt.show()`

# Seaborn Usages (Strip Plots)

```
tips.head()

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

plt.show()
```
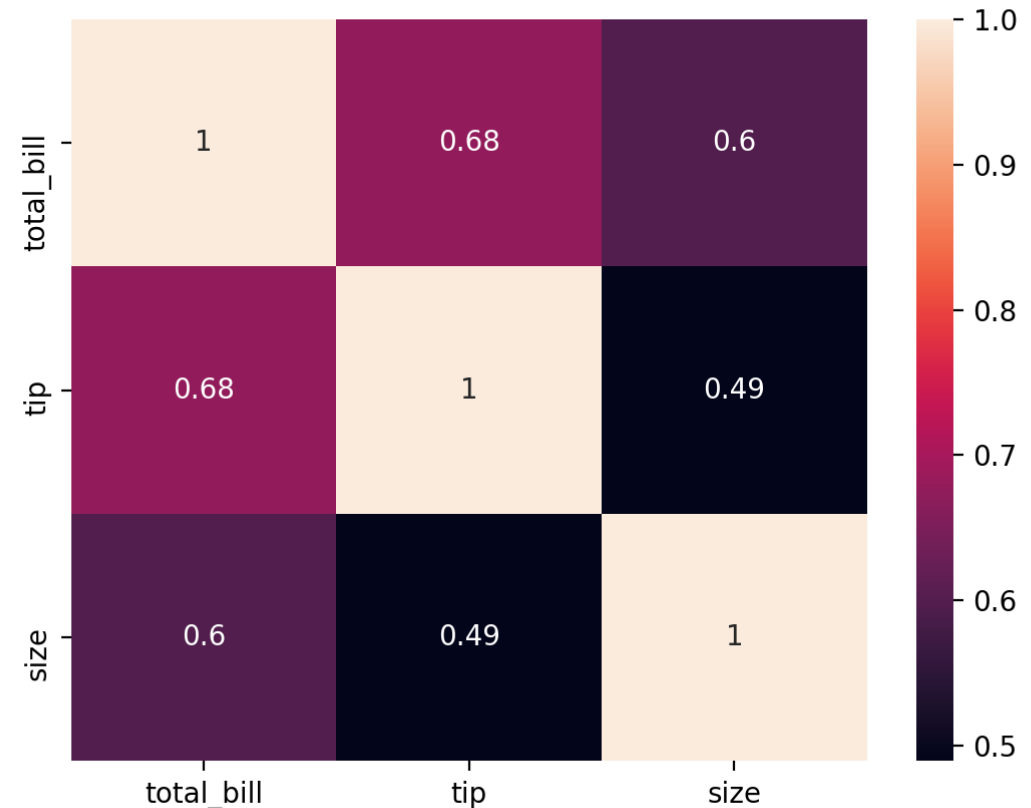
# Seaborn Usages (Matrix Plots)

```python
tc = tips.corr()

#                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

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

Seaborn Usages (Matrix Plots)

```python
flights = sns.load_dataset("flights")

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

year	1949	1950	1951	1952	...	1957	1958	1959	1960
month					...				
January	112	115	145	171	...	315	340	360	417
February	118	126	150	180	...	301	318	342	391
March	132	141	178	193	...	356	362	406	419
April	129	135	163	181	...	348	348	396	461
May	121	125	172	183	...	355	363	420	472
June	135	149	178	218	...	422	435	472	535
July	148	170	199	230	...	465	491	548	622
August	148	170	199	242	...	467	505	559	606
September	136	158	184	209	...	404	404	463	508
October	119	133	162	191	...	347	359	407	461
November	104	114	146	172	...	305	310	362	390
December	118	140	166	194	...	336	337	405	432

```
flights.head()
#     year     month   passengers
# 0   1949   January          112
# 1   1949   February         118
# 2   1949      March          132
# 3   1949      April          129
# 4   1949        May          121
```
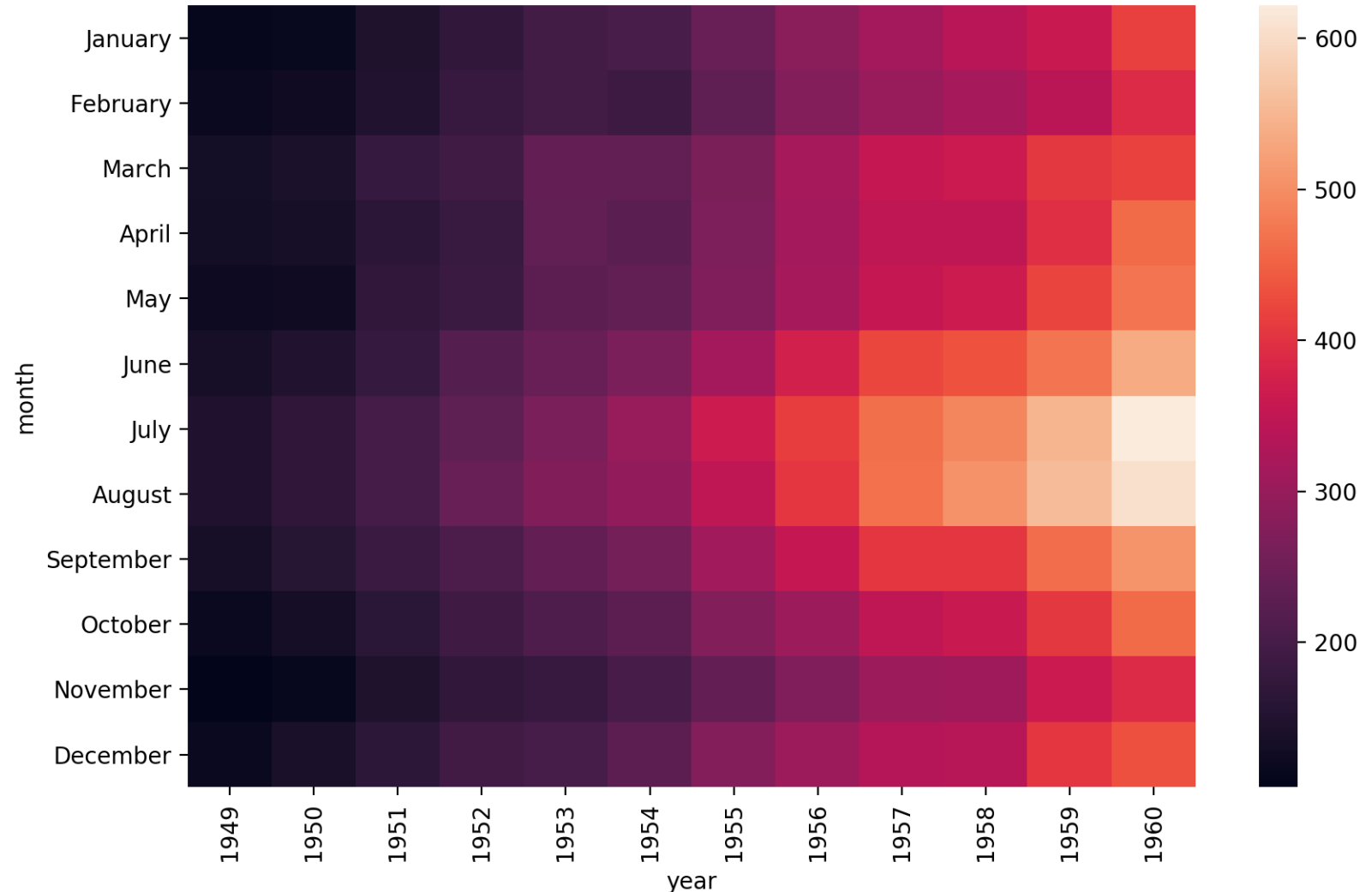
Seaborn Usages (Matrix Plots)

```python
flights = sns.load_dataset("flights")

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

sns.heatmap(fp)

plt.show()
```
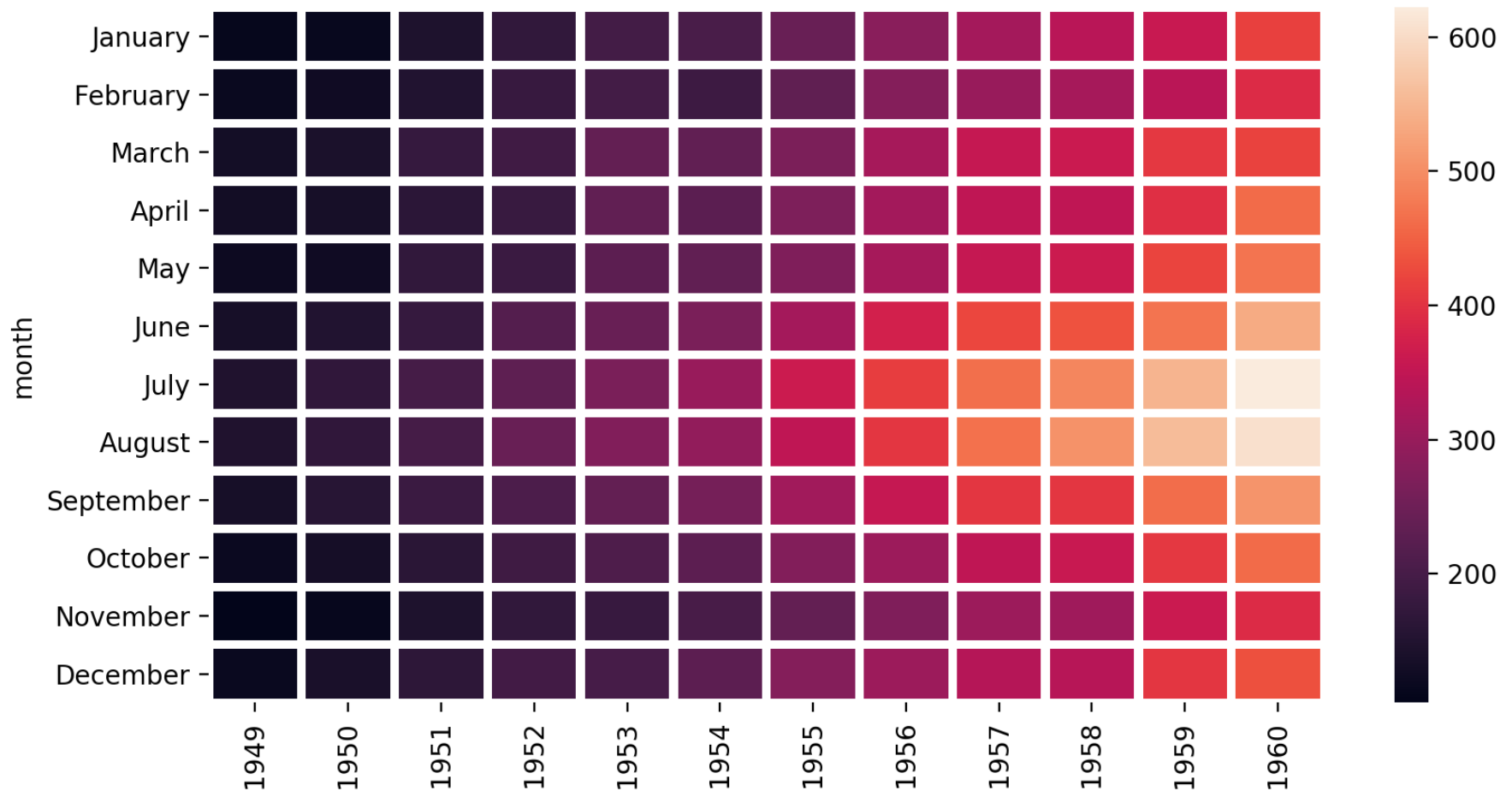
Seaborn Usages (Matrix Plots)

```python
flights = sns.load_dataset("flights")

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

sns.heatmap(fp,linecolor='white',linewidths=3)

plt.show()
```
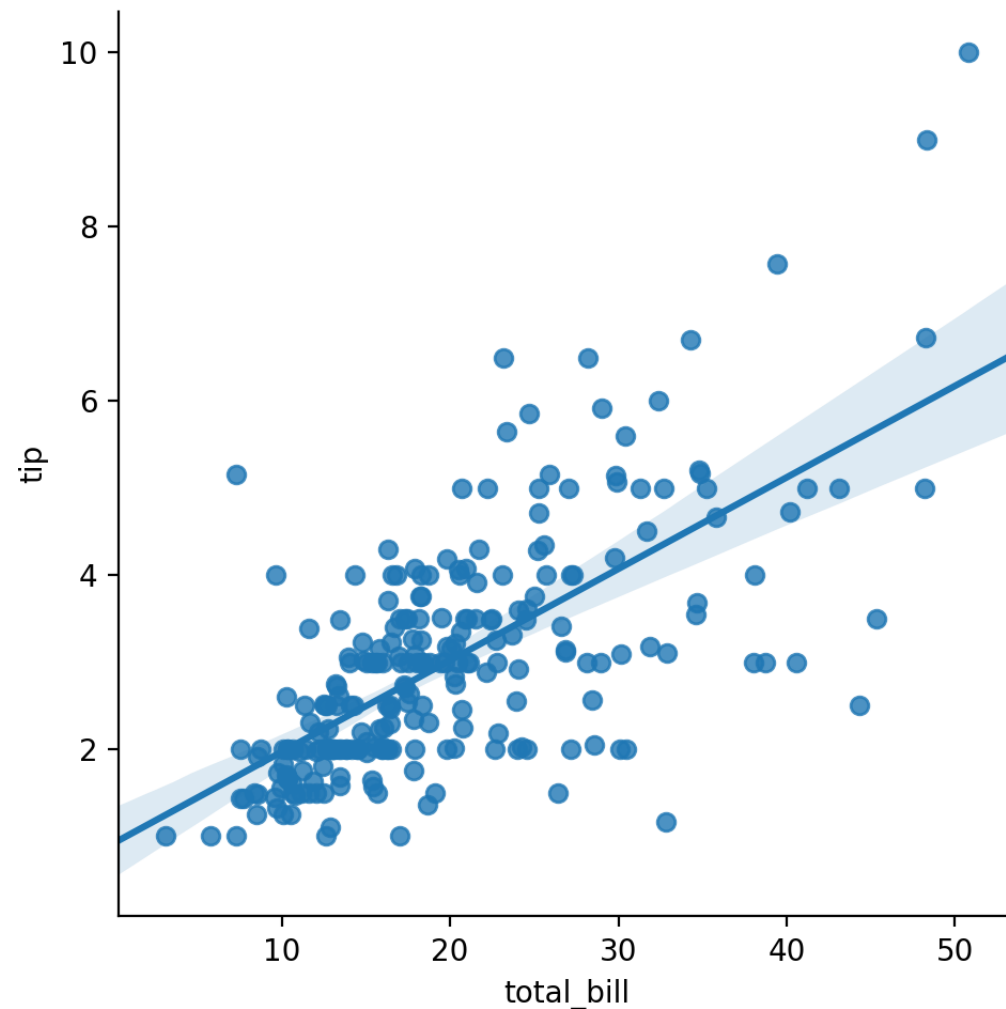
Seaborn Usages (Regression Plots)

```
tips.head()
```

```
#    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
```

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

plt.show()
```
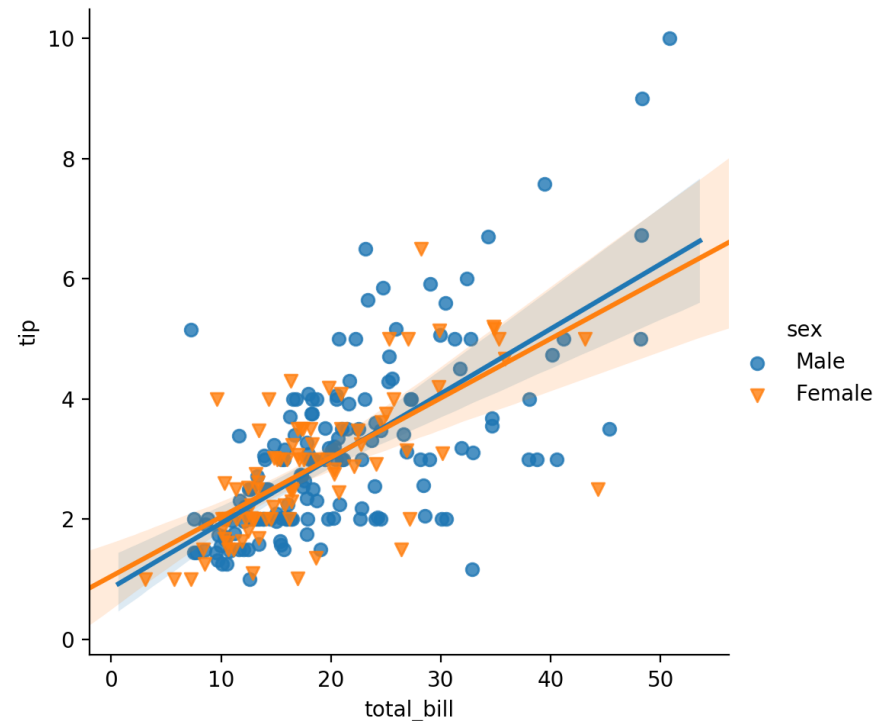
Seaborn Usages (Regression Plots)

```
tips.head()
```

```
#     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
```

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

```
plt.show()
```

Seaborn Usages (Regression Plots)

```
tips.head()
```

```
#     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
```

```python
sns.lmplot(x="total_bill",y="tip",hue="sex",data=tips, col="day")

plt.show()
```