Data Visualization With Seaborn

Seaborn is a Python visualization library based on Matplotlib that provides a high-level interface for drawing attractive and informative statistical graphics. Seaborn helps in creating complex visualizations with just a few lines of code. In this lesson, we will cover the basics of Seaborn, including creating various types of plots and customizing them.

In [2]:
```
!pip install seaborn
```

```
Requirement already satisfied: seaborn in e:\software\ide\anaconda\lib\site-packages (0.13.2)
Requirement already satisfied: numpy!=1.24.0,>=1.20 in e:\software\ide\anaconda\lib\site-packages (from seaborn) (1.26.4)
Requirement already satisfied: pandas>=1.2 in e:\software\ide\anaconda\lib\site-packages (from seaborn) (2.2.2)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in e:\software\ide\anaconda\lib\site-packages (from seaborn) (3.8.4)
Requirement already satisfied: contourpy>=1.0.1 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.2.0)
Requirement already satisfied: cycler>=0.10 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.51.0)
Requirement already satisfied: kiwisolver>=1.3.1 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.4.4)
Requirement already satisfied: packaging>=20.0 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (23.2)
Requirement already satisfied: pillow>=8 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (10.3.0)
Requirement already satisfied: pyparsing>=2.3.1 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (3.0.9)
Requirement already satisfied: python-dateutil>=2.7 in e:\software\ide\anaconda\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in e:\software\ide\anaconda\lib\site-packages (from pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in e:\software\ide\anaconda\lib\site-packages (from pandas>=1.2->seaborn) (2023.3)
Requirement already satisfied: six>=1.5 in e:\software\ide\anaconda\lib\site-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.16.0)
```

In [4]:
```
import seaborn as sns
```

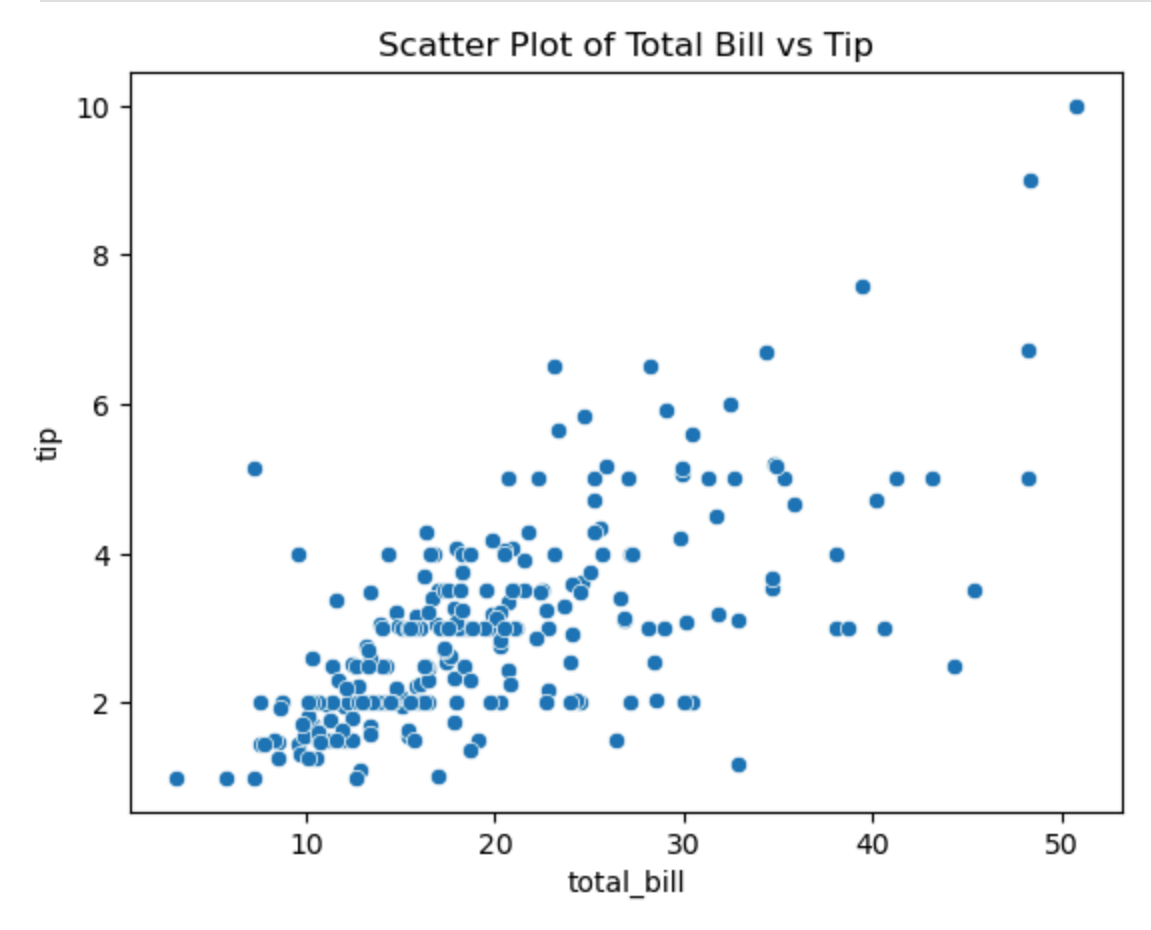
In [6]:
```
### Basic Plotting With Seaborn
tips=sns.load_dataset('tips')
tips
```

Out[6]:

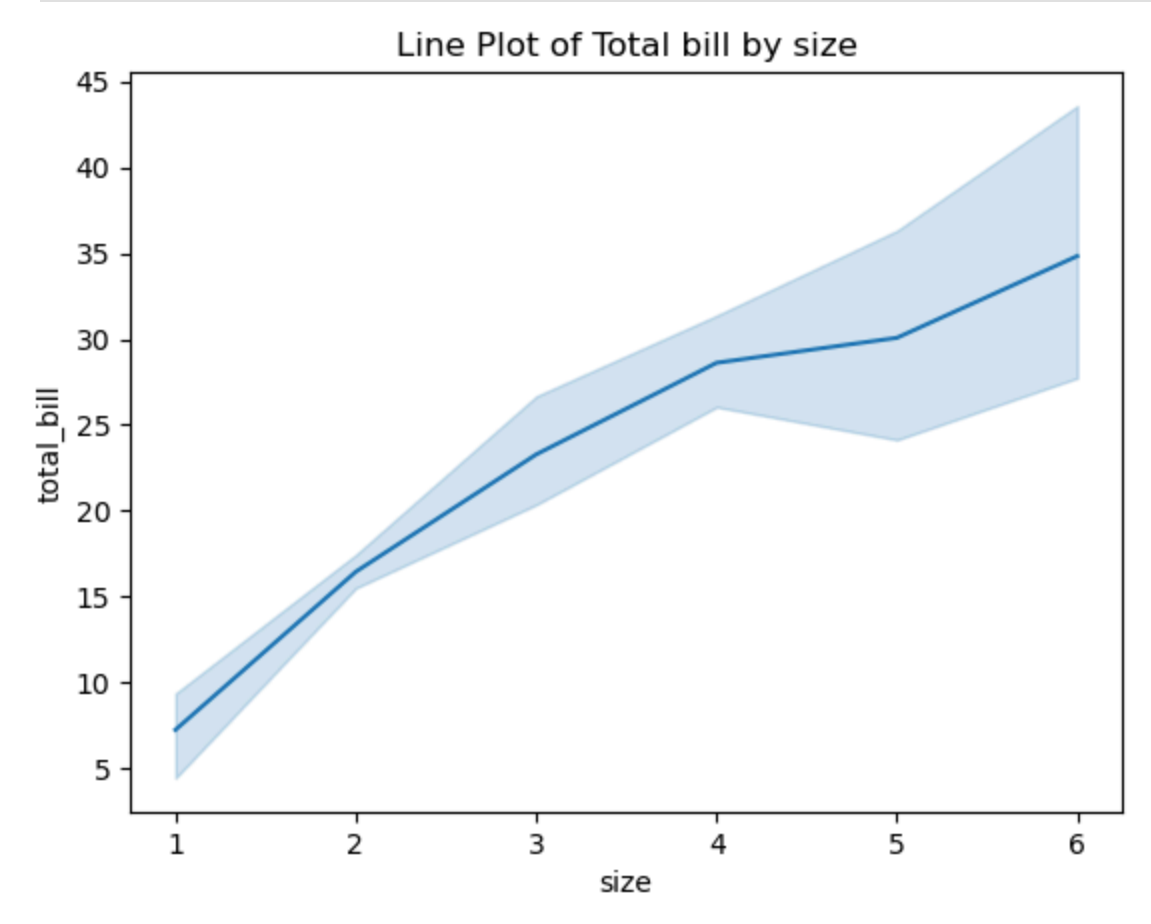
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4
...	...	...	...	...	...	...	...
239	29.03	5.92	Male	No	Sat	Dinner	3
240	27.18	2.00	Female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

244 rows × 7 columns

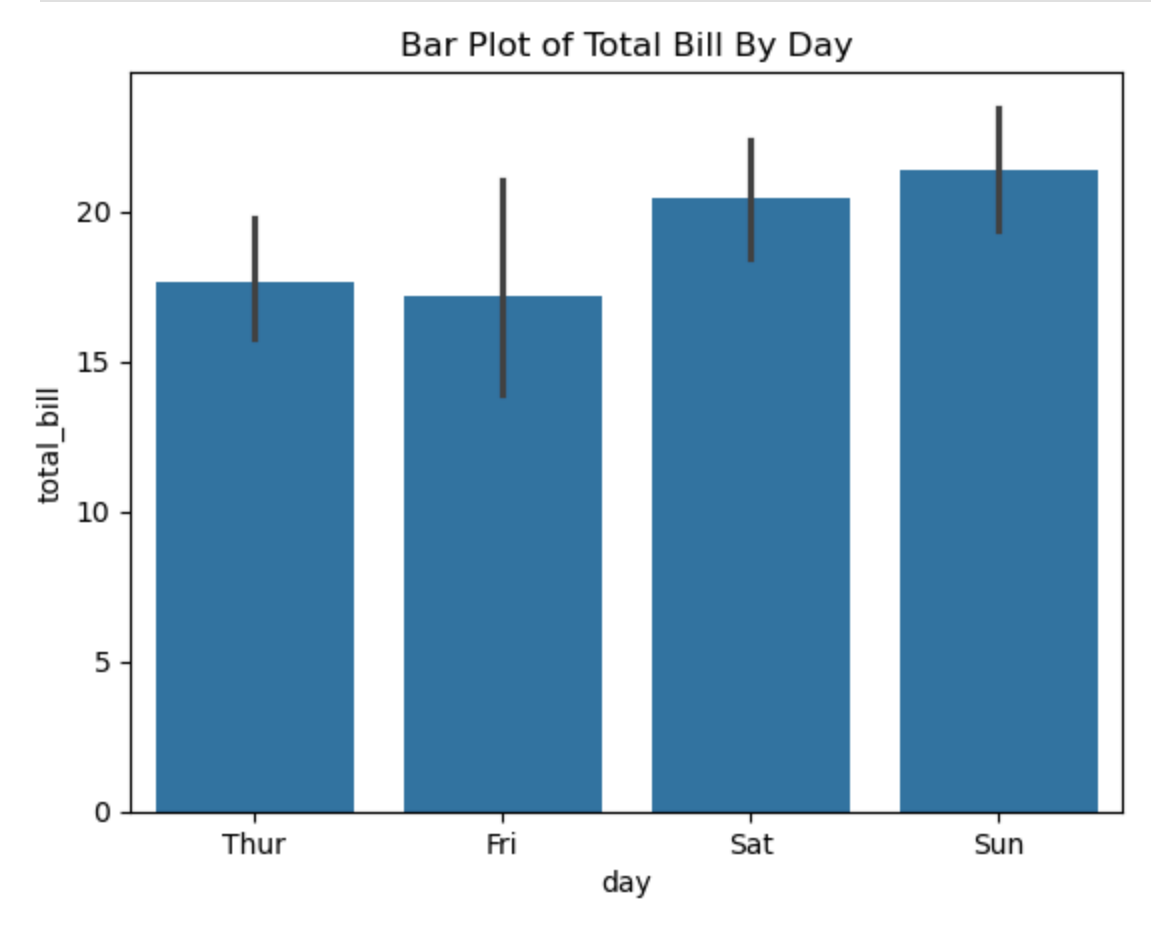
In [8]:
```
##create a scatter plot
import matplotlib.pyplot as plt
sns.scatterplot(x='total_bill', y='tip', data=tips)
plt.title("Scatter Plot of Total Bill vs Tip")
plt.show()
```



In [10]:
```
## Line Plot
sns.lineplot(x='size',y='total_bill', data=tips)
plt.title("Line Plot of Total bill by size")
plt.show()
```

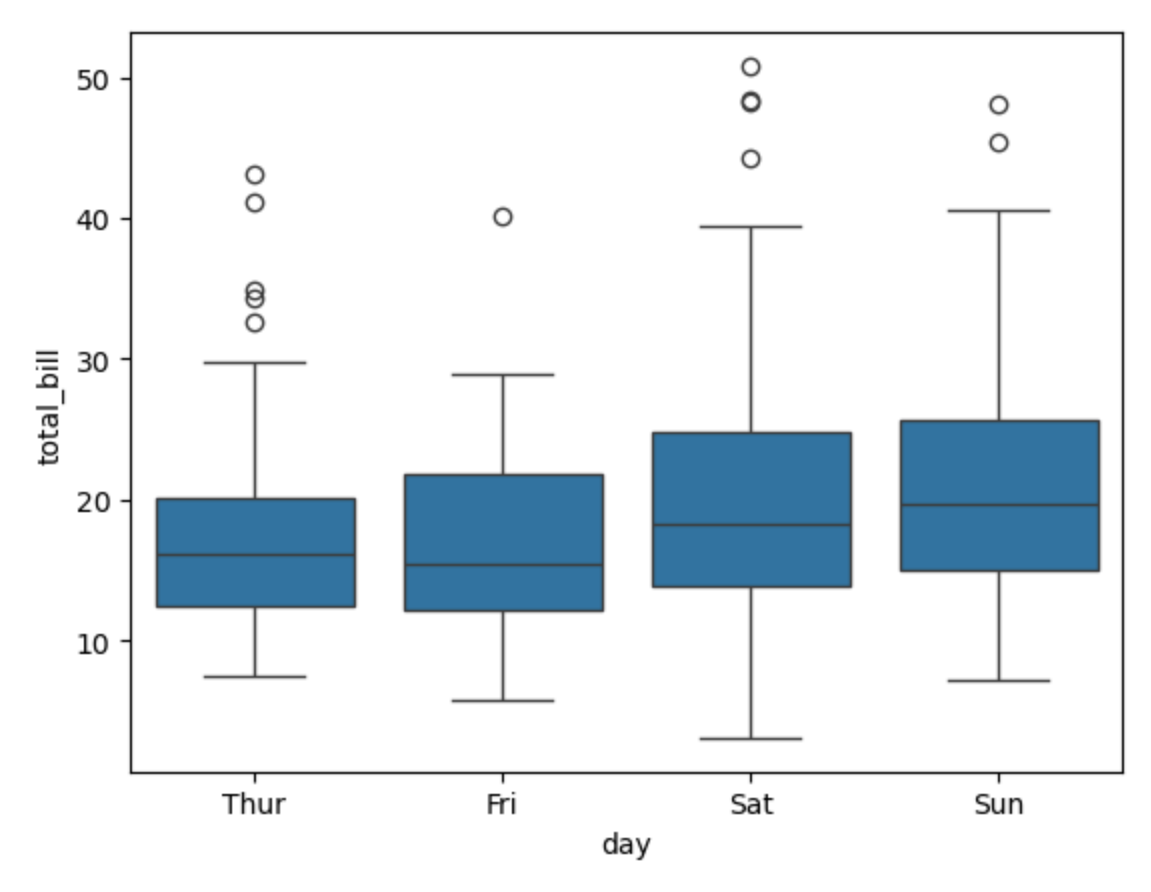


In [11]:
```
## Categorical Plots
## BAr Plot
import seaborn as sns
import matplotlib.pyplot as plt
tips=sns.load_dataset('tips')
sns.barplot(x='day', y='total_bill', data=tips)
plt.title('Bar Plot of Total Bill By Day')
plt.show()
```



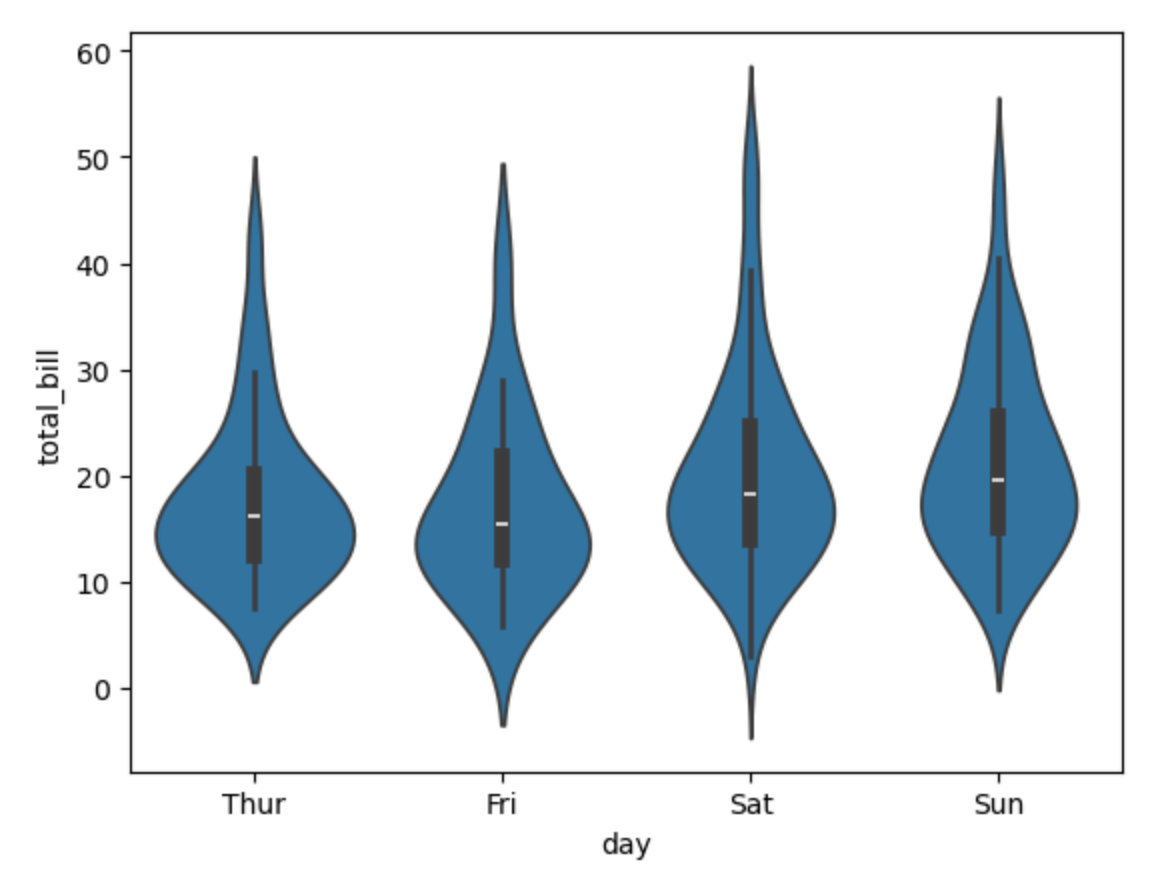
In [15]:
```
#Boxplot
sns.boxplot(x="day",y="total_bill",data=tips)
```

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



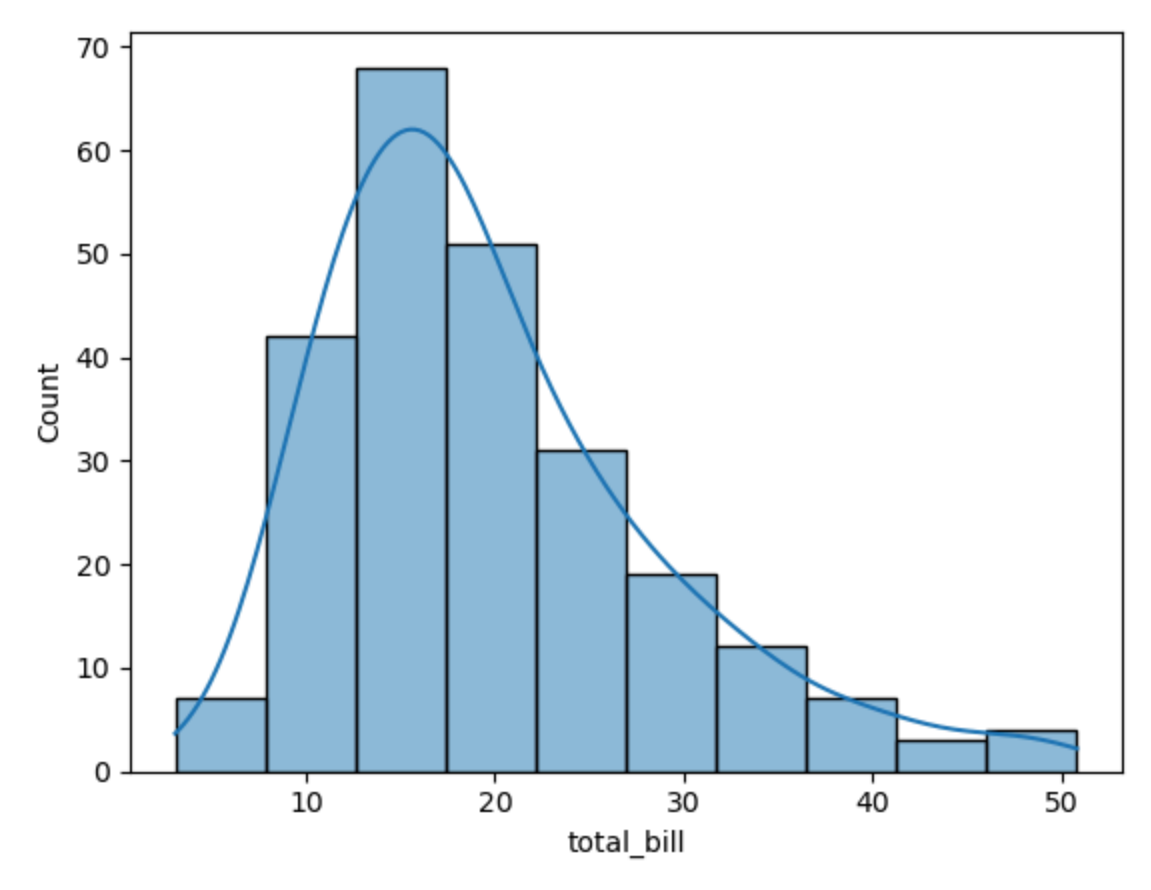
In [17]:
```
#violinplot
sns.violinplot(x="day",y="total_bill",data=tips)
```

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



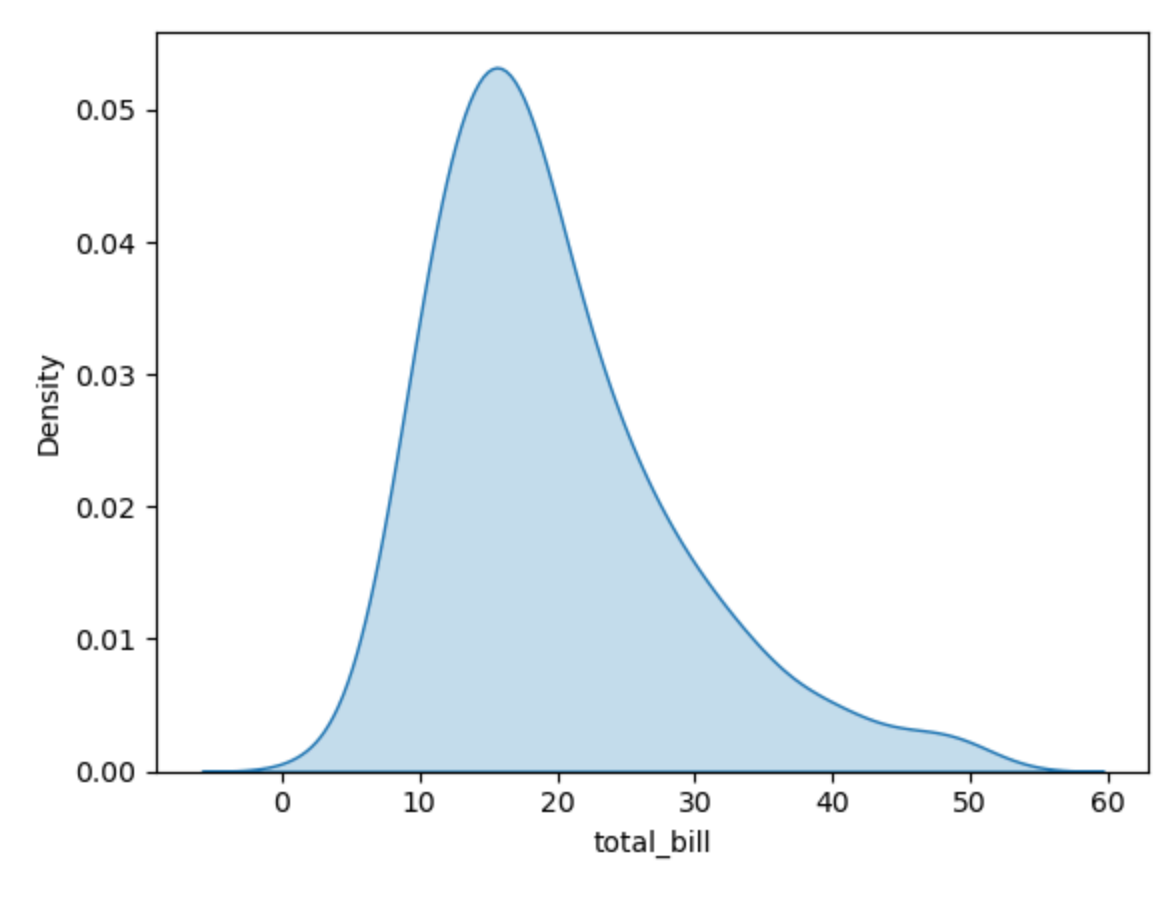
In [19]:
```
### Histograms
sns.histplot(tips['total_bill'], bins=10, kde=True)
```

Out[19]:
```
<Axes: xlabel='total_bill', ylabel='Count'>
```



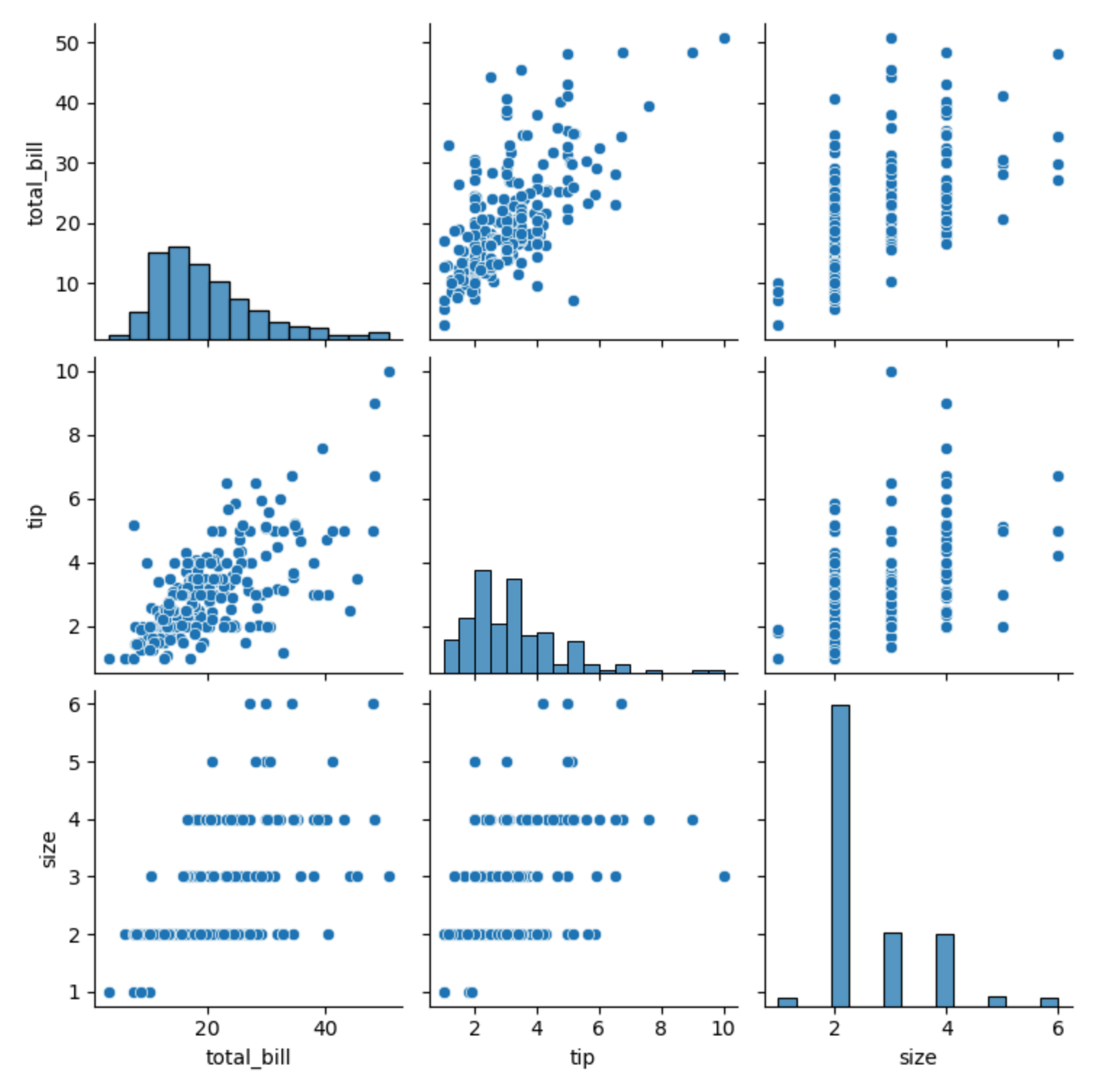
In [23]:
```
##kde plot
sns.kdeplot(tips['total_bill'], fill=True)
```

Out[23]:
```
<Axes: xlabel='total_bill', ylabel='Density'>
```



In [25]:
```
#pairplot
sns.pairplot(tips)
```

Out[25]:
```
<seaborn.axisgrid.PairGrid at 0x1b07c541280>
```



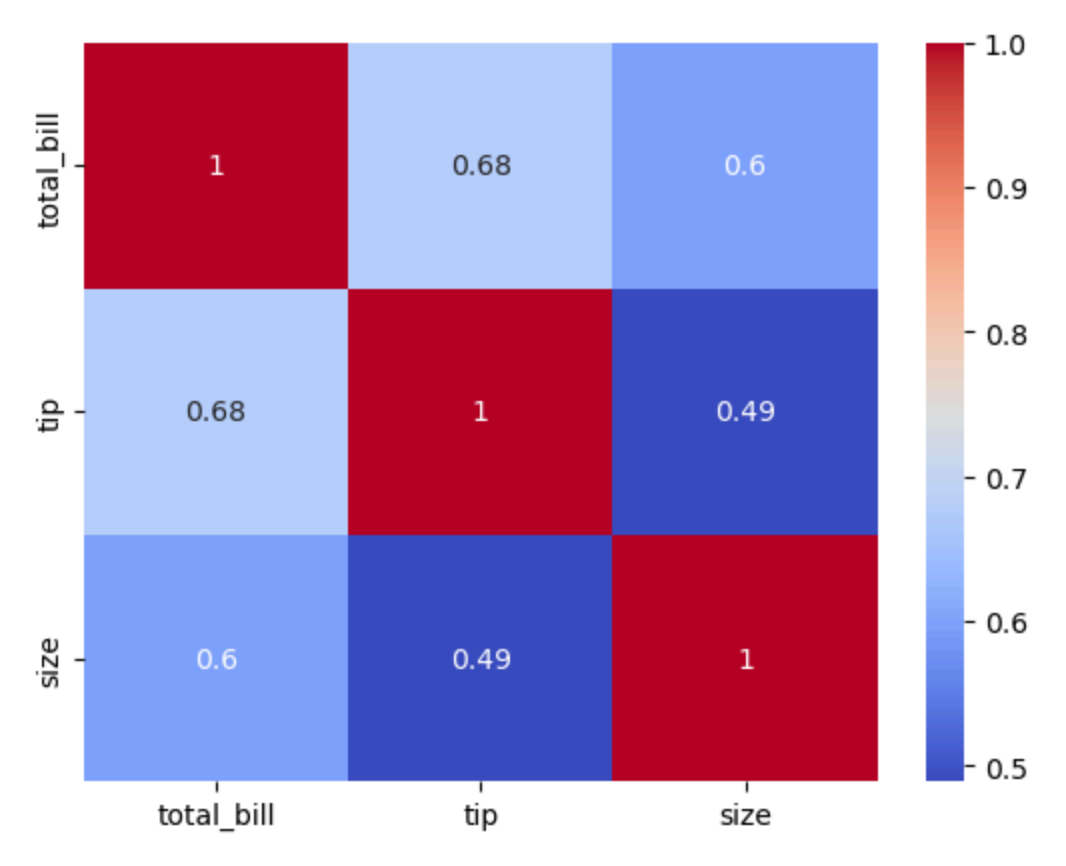
In [27]:
```
## HEatmap
corr=tips [['total_bill', 'tip', 'size']].corr()
corr
```

Out[27]:

	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 [29]:
```
sns.heatmap(corr,annot=True,cmap='coolwarm')
```

Out[29]:
```
<Axes: >
```

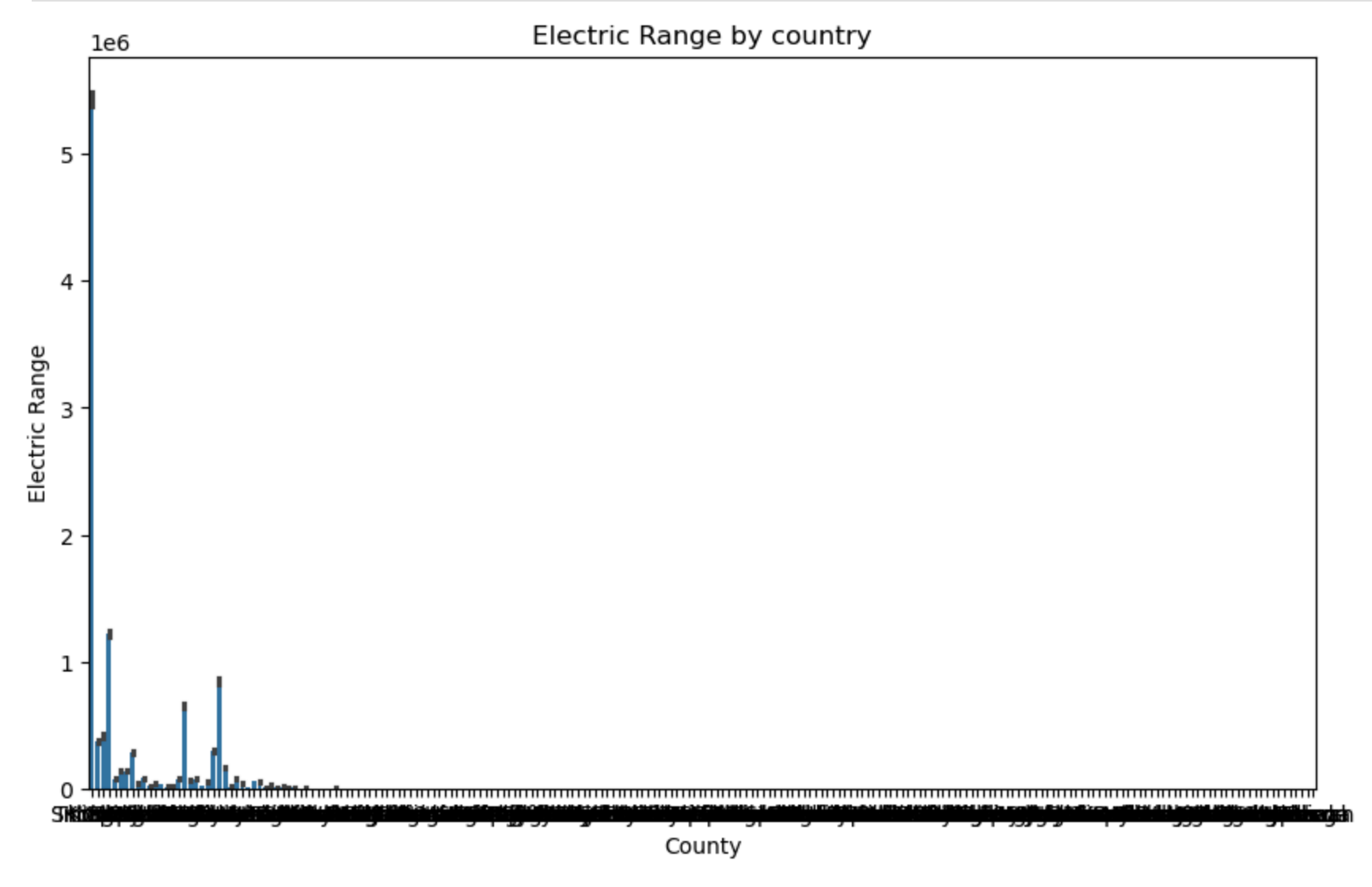


In [35]:
```
import pandas as pd
sales_df=pd.read_csv('data.csv')
sales_df.head()
```

Out[35]:

	VIN (1-10)	County	City	State	Postal Code	Model Year	Make	Model	Electric Vehicle Type	Clean Alternative Fuel Vehicle (CAFV) Eligibility	Electric Range	Base MSRP	Legislative District	DOL Vehicle ID	Vehicle Location	Electric Utility	2020 Census Tract
0	5YJ3E1EBXK	King	Seattle	WA	98178.0	2019	TESLA	MODEL 3	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	220.0	0.0	37.0	477309682	POINT (-122.23825 47.49461)	CITY OF SEATTLE - (WA)|CITY OF TACOMA - (WA)	5.303301e+10
1	5YJYGDEE3L	Kitsap	Poulsbo	WA	98370.0	2020	TESLA	MODEL Y	Battery Electric Vehicle (BEV)	Clean Alternative Fuel Vehicle Eligible	291.0	0.0	23.0	109705683	POINT (-122.64681 47.73689)	PUGET SOUND ENERGY INC	5.303509e+10
2	KM8KRDAF5P	Kitsap	Olalla	WA	98359.0	2023	HYUNDAI	IONIQ 5	Battery Electric Vehicle (BEV)	Eligibility unknown as battery range has not b...	0.0	0.0	26.0	230390492	POINT (-122.54729 47.42602)	PUGET SOUND ENERGY INC	5.303509e+10
3	5UXTA6C0XM	Kitsap	Seabeck	WA	98380.0	2021	BMW	X5	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	30.0	0.0	35.0	267929112	POINT (-122.81585 47.64509)	PUGET SOUND ENERGY INC	5.303509e+10
4	JTMAB3FV7P	Thurston	Rainier	WA	98576.0	2023	TOYOTA	RAV4 PRIME	Plug-in Hybrid Electric Vehicle (PHEV)	Clean Alternative Fuel Vehicle Eligible	42.0	0.0	2.0	236505139	POINT (-122.68993 46.88897)	PUGET SOUND ENERGY INC	5.306701e+10

In [42]:
```
## Plot total sales by product
plt.figure(figsize=(10,6))
sns.barplot(x='County', y="Electric Range", data=sales_df, estimator=sum)
plt.title('Electric Range by country')
plt.xlabel('County')
plt.ylabel('Electric Range')
plt.show()
```



In [ ]: