CS23334-FUNDAMENTALS OF DATA SCIENCE

DEVA DHARSHINI P(240701107)

6.) EDA FOR QUALITATIVE AND QUANTITATIVE DATA

Aim:

To perform Exploratory Data Analysis (EDA) on qualitative (categorical) and quantitative (numerical) data to understand data patterns, relationships, and distributions.

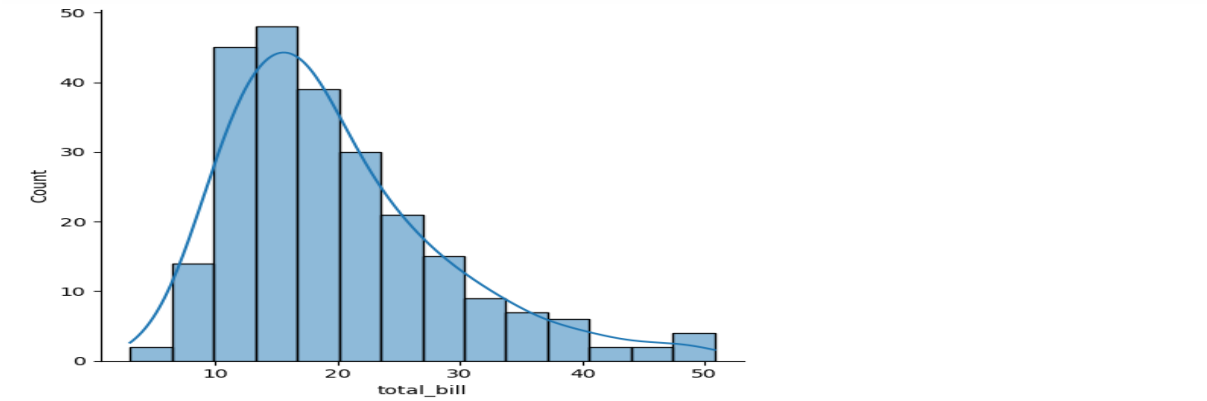
Code:

```
import seaborn as sns
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
tips=sns.load_dataset('tips')
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

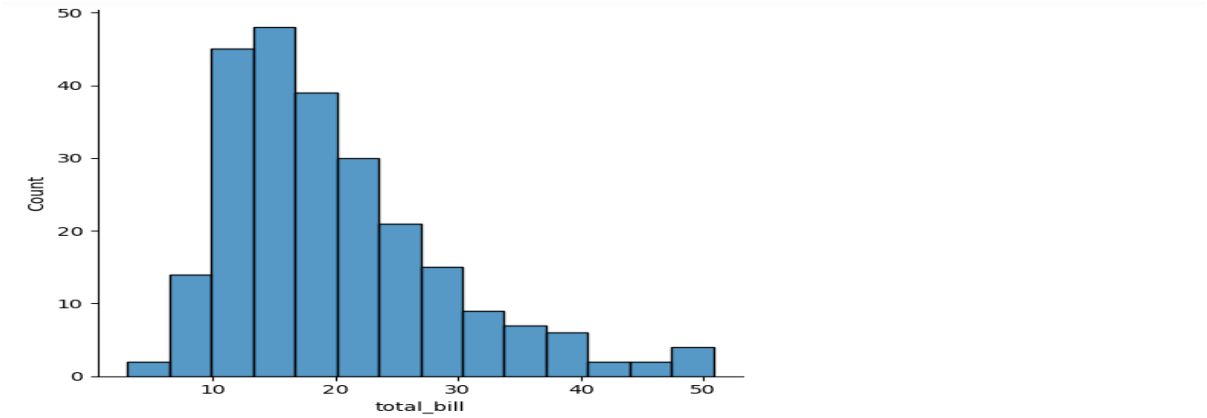
```
sns.displot(tips.total_bill,kde=True)
```

```
<seaborn.axisgrid.FacetGrid at 0x279712f6d10>
```



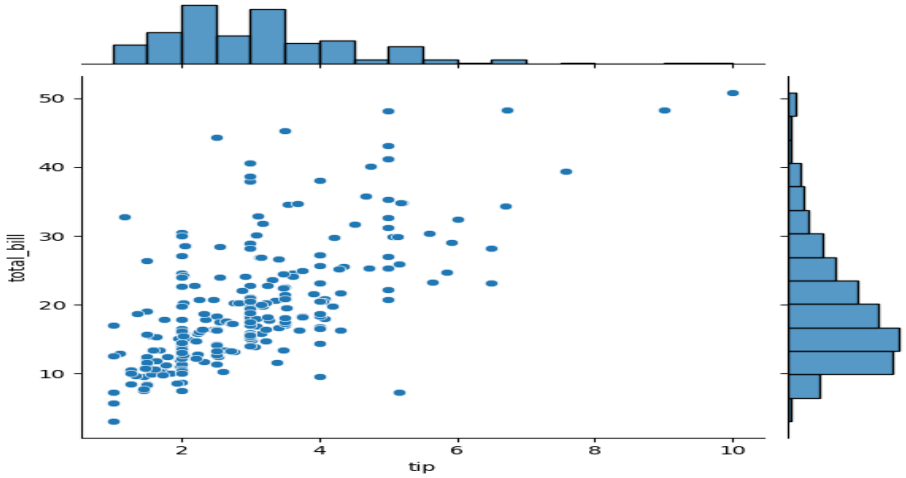

```
sns.displot(tips.total_bill,kde=False)
```

```
<seaborn.axisgrid.FacetGrid at 0x27976b46470>
```



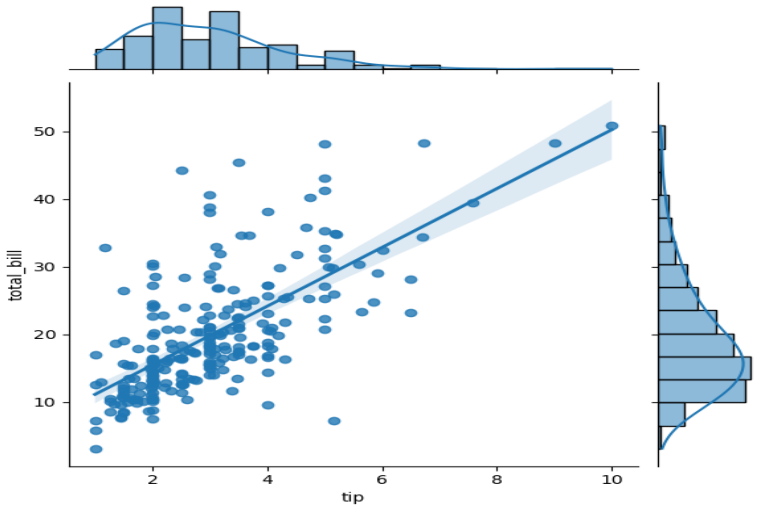

```
sns.jointplot(x=tips.tip,y=tips.total_bill)
```

```
<seaborn.axisgrid.JointGrid at 0x27979cd5cf0>
```



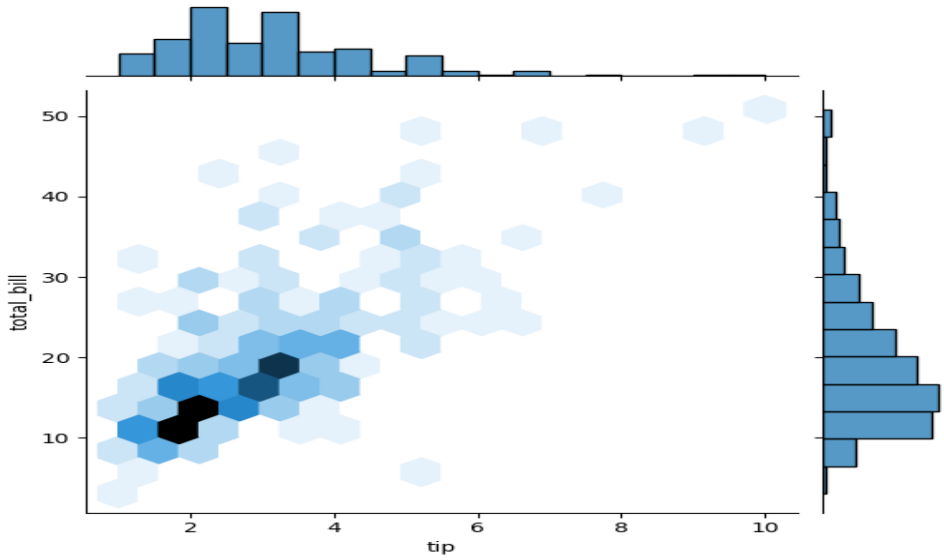

```
sns.jointplot(x=tips.tip,y=tips.total_bill,kind="reg")
```

```
<seaborn.axisgrid.JointGrid at 0x27971391e10>
```



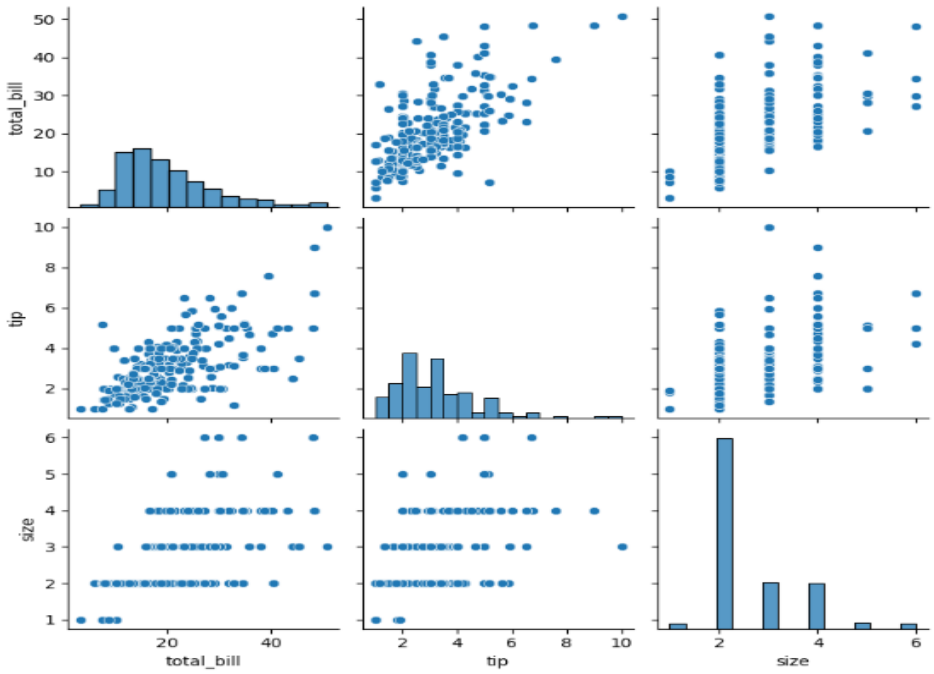

```
sns.jointplot(x=tips.tip,y=tips.total_bill,kind="hex")
```

```
<seaborn.axisgrid.JointGrid at 0x2797a31d690>
```



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

```
<seaborn.axisgrid.PairGrid at 0x2797a31d510>
```

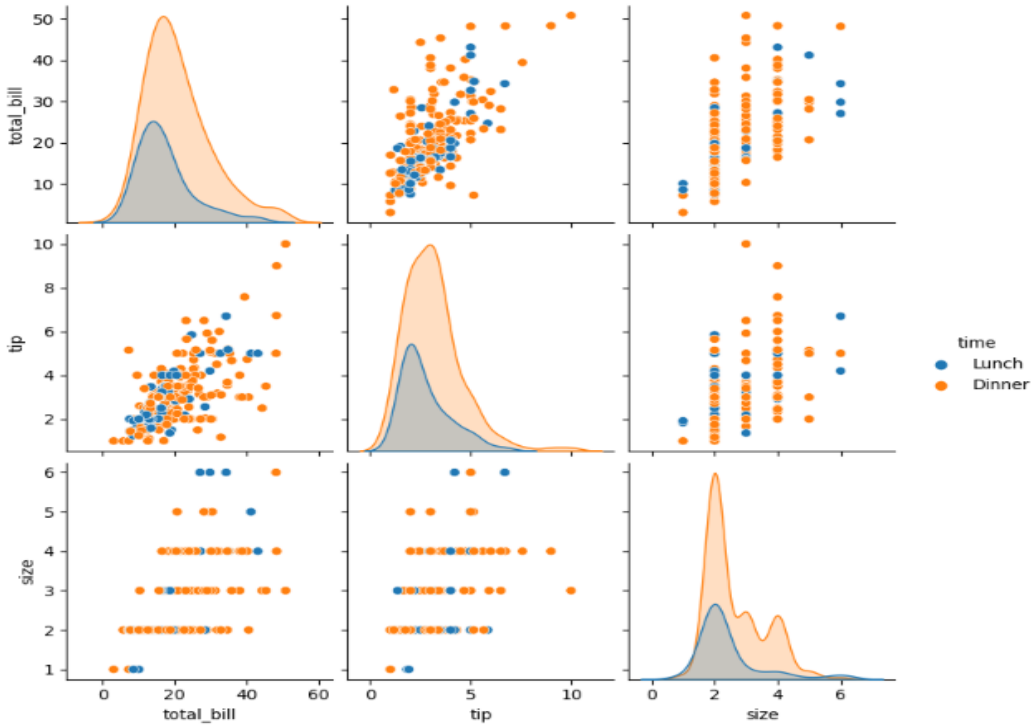



```
tips.time.value_counts()
```

```
Dinner    176
Lunch      68
Name: time, dtype: int64
```

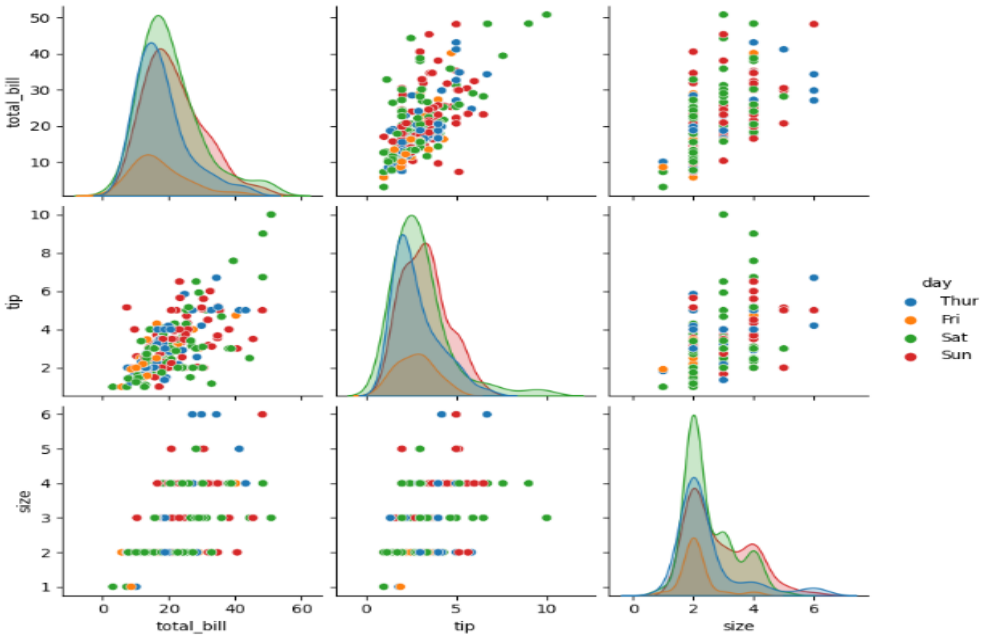
```
sns.pairplot(tips,hue='time')
```

```
<seaborn.axisgrid.PairGrid at 0x27976b455a0>
```



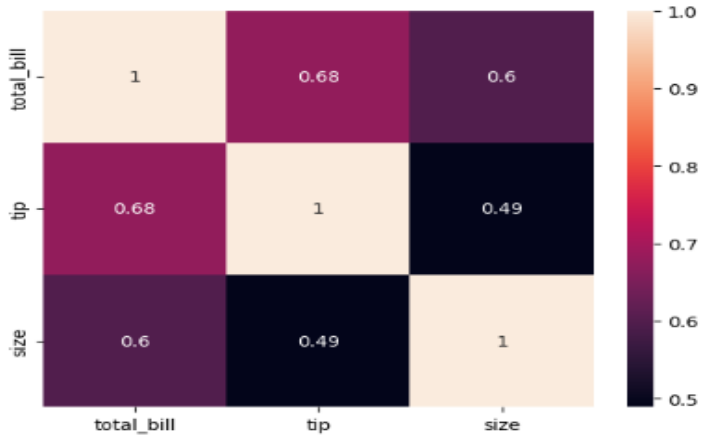
```
sns.pairplot(tips,hue='day')
```

```
<seaborn.axisgrid.PairGrid at 0x2797c0b4250>
```



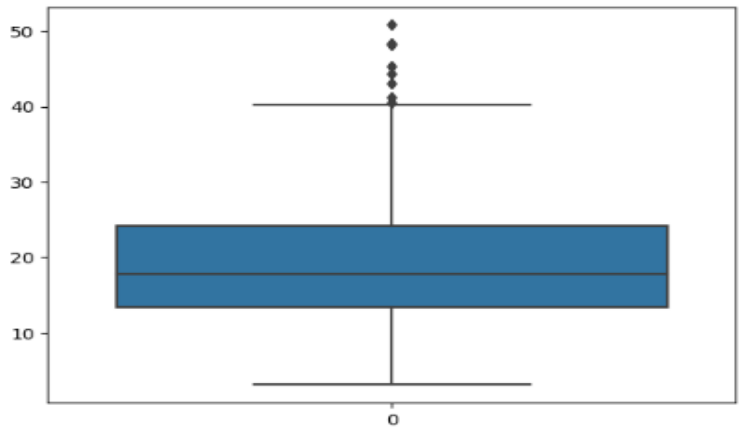

```
sns.heatmap(tips.corr(numeric_only=True),annot=True)
```

```
<Axes: >
```



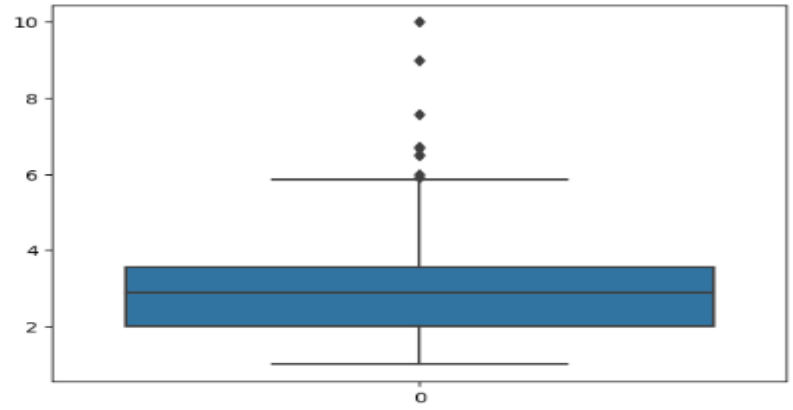

```
sns.boxplot(tips.total_bill)
```

```
<Axes: >
```



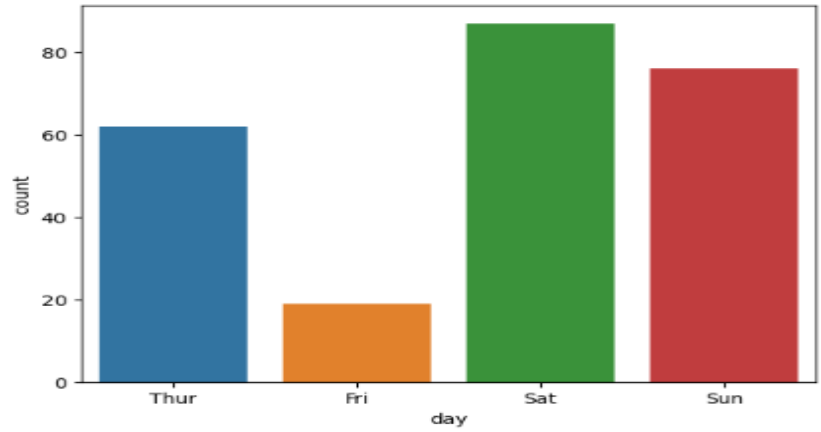
```
sns.boxplot(tips.tip)
```

```
<Axes: >
```



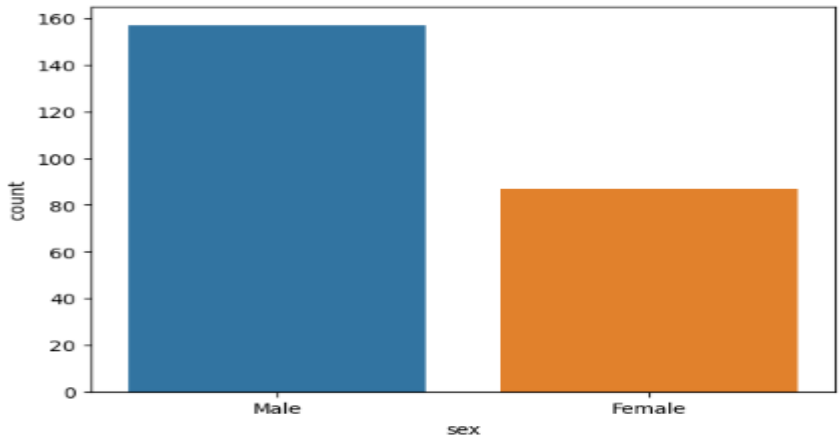

```
sns.countplot(data=tips,x='day')
```

```
<Axes: xlabel='day', ylabel='count'>
```




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

```
<Axes: xlabel='sex', ylabel='count'>
```



Result:

The EDA provided clear insights into both qualitative and quantitative variables, revealing trends, distributions, and correlations useful for further data analysis and model building.