

Experiment to understand EDA-Quantitative and Qualitative analysis.

Description: Understand the importance of EDA-Quantitative and Qualitative analysis.

Aim:

To understand the importance of Exploratory Data Analysis (EDA) by performing both quantitative and qualitative analysis on a dataset.

Procedure:

1. Load a sample dataset (e.g., the Titanic or Tips dataset) using pandas.
2. Perform **quantitative analysis** by summarizing numerical features — use functions like `describe()`, compute mean, median, mode, and visualize distributions with histograms or boxplots.
3. Perform **qualitative analysis** on categorical features — use `value_counts()`, bar charts, and count plots to explore category distributions.
4. Identify patterns, missing values, and relationships between numerical and categorical variables using correlation and cross-tabulations.
5. Summarize insights from both analyses to understand data trends and prepare the dataset for modeling.

In [43]:

```
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()
```

Out[43]:

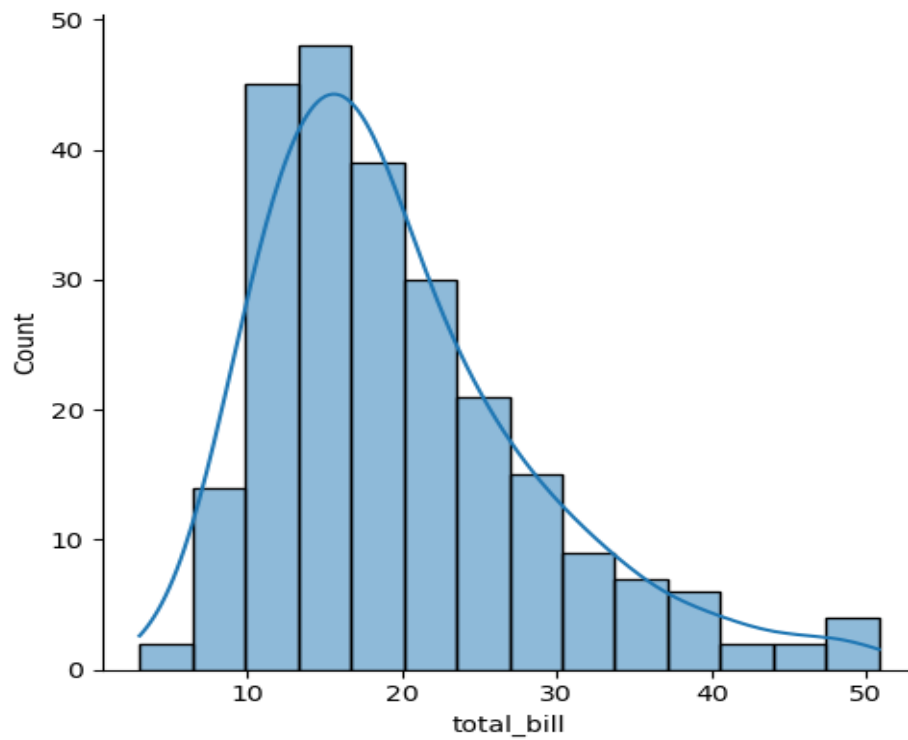
	total_bill	tip	sex	smoker	day	time	size
0	16.99	1.01	Female	No	Sun	Dinner	2
1	10.34	1.66	Male	No	Sun	Dinner	3
2	21.01	3.50	Male	No	Sun	Dinner	3
3	23.68	3.31	Male	No	Sun	Dinner	2
4	24.59	3.61	Female	No	Sun	Dinner	4

In [44]:

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

Out[44]:

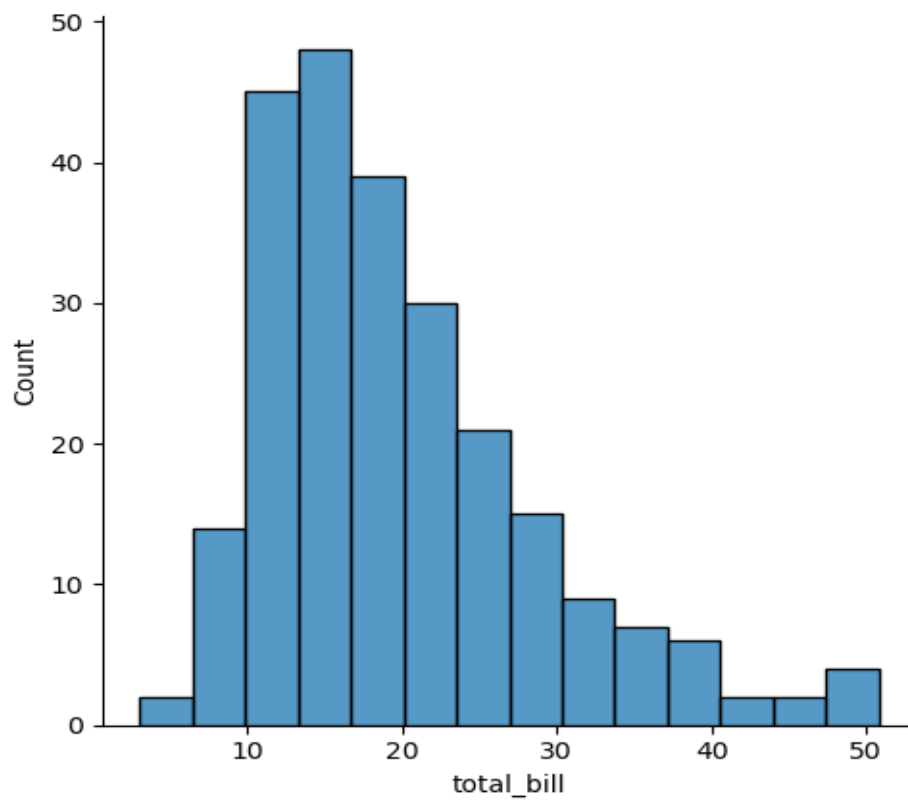
<seaborn.axisgrid.FacetGrid at 0x20a0373ffd0>



In [45]:

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

Out[45]:<seaborn.axisgrid.FacetGrid at 0x20a0380b070>

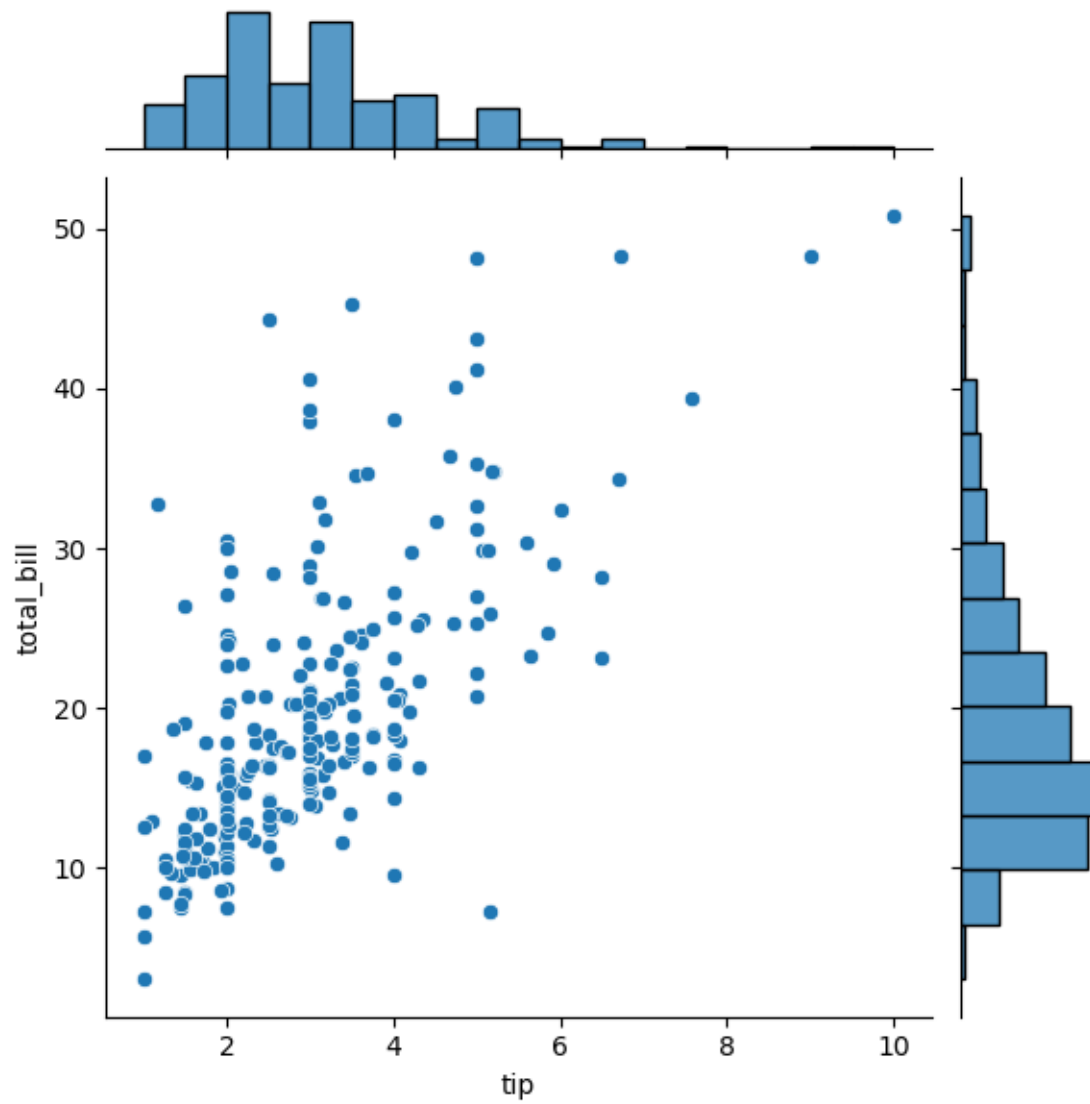


In [46]:

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

Out[46]:

```
<seaborn.axisgrid.JointGrid at 0x20a038a49a0>
```

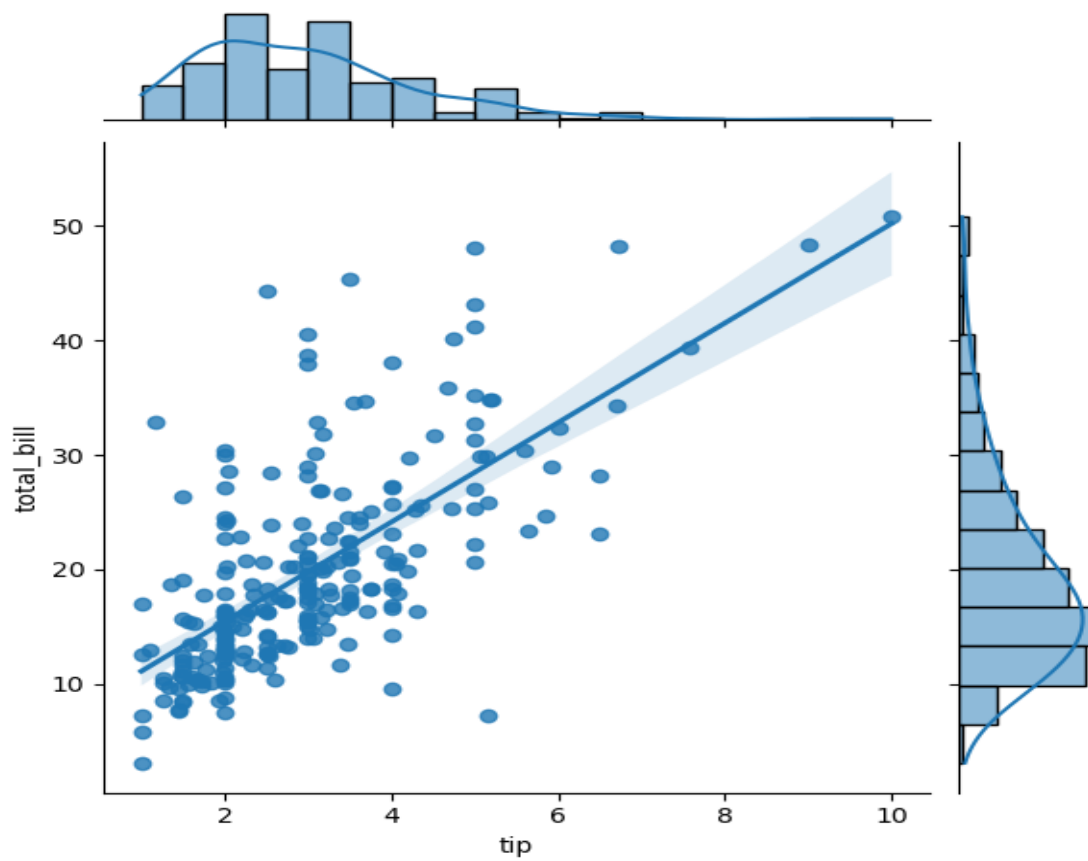


In [47]:

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

Out[47]:

```
<seaborn.axisgrid.JointGrid at 0x20a038295d0>
```

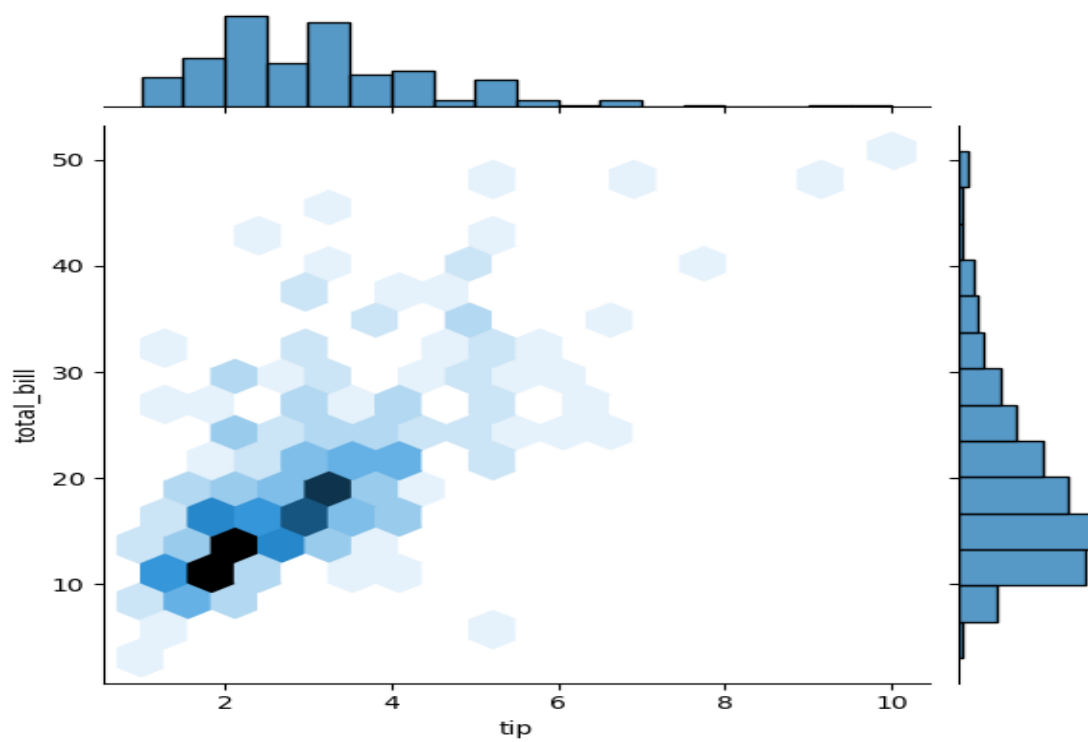


In [48]:

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

Out[48]:

```
<seaborn.axisgrid.JointGrid at 0x20a062f6980>
```

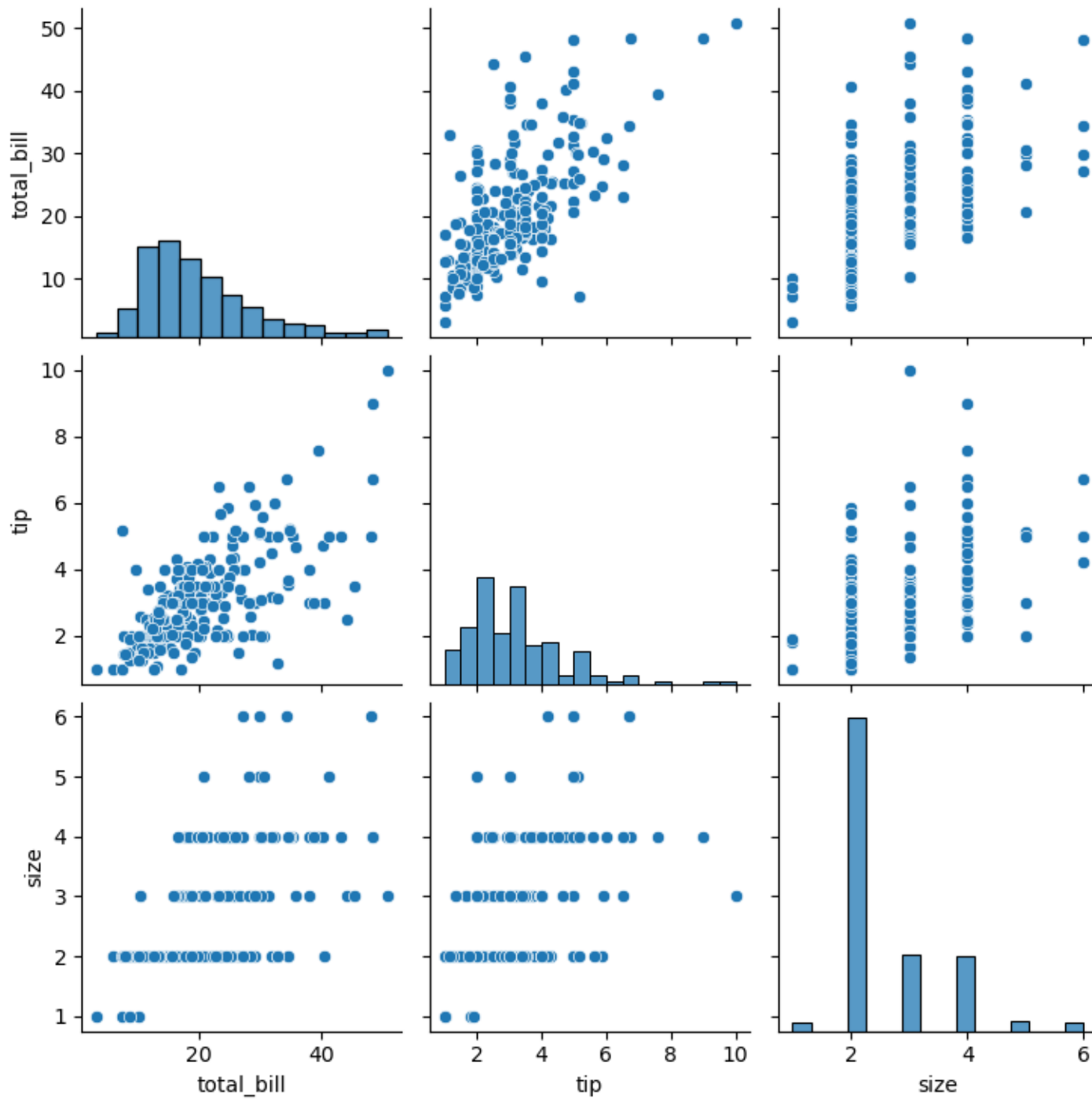


In [49]:

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

Out[49]:

```
<seaborn.axisgrid.PairGrid at 0x20a062f6530>
```



In [50]:

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

Out[50]:

```
Dinner    176
```

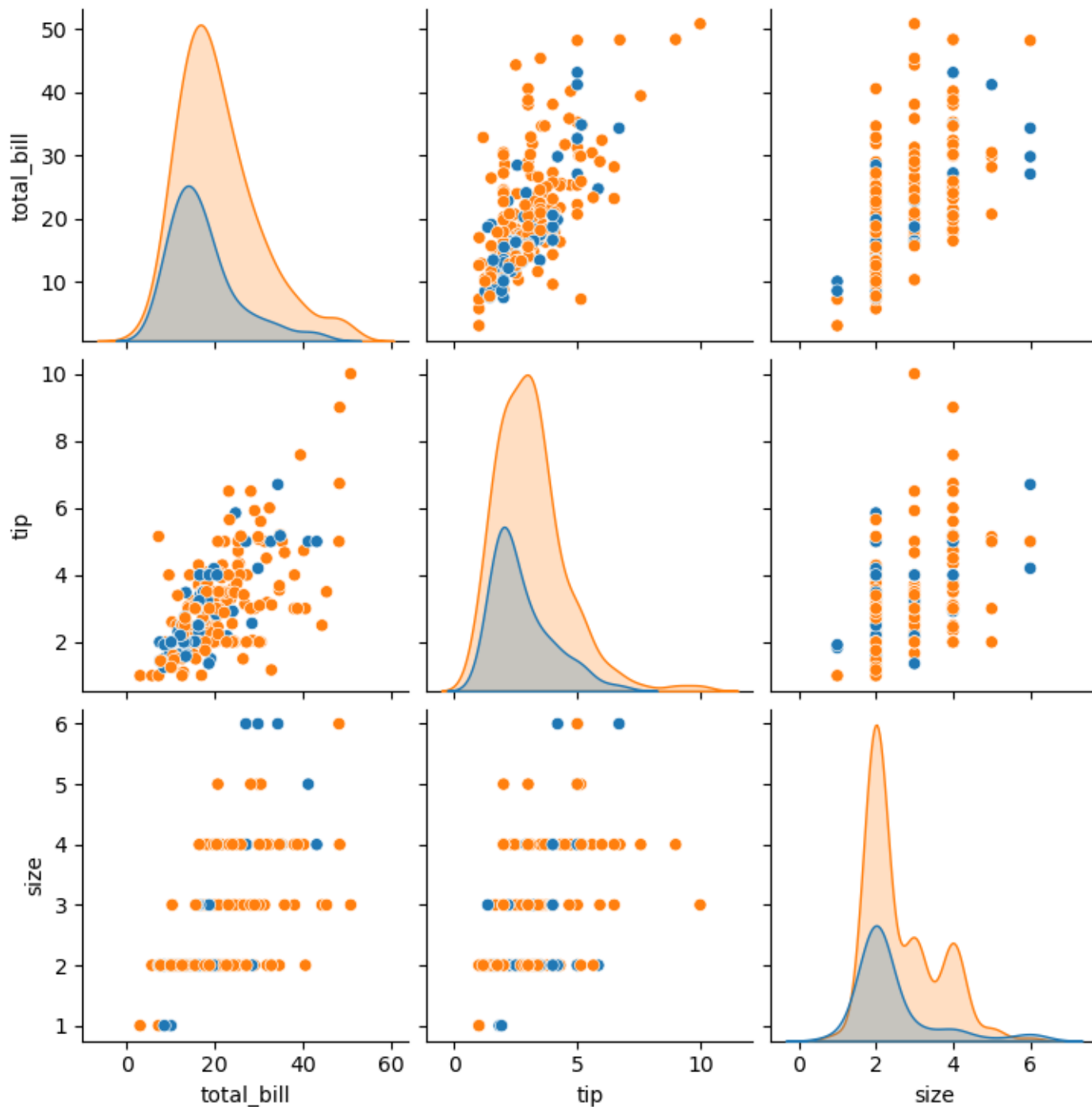
```
Lunch      68  
Name: time, dtype: int64
```

In [51]:

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

Out[51]:

```
<seaborn.axisgrid.PairGrid at 0x20a070d6530>
```

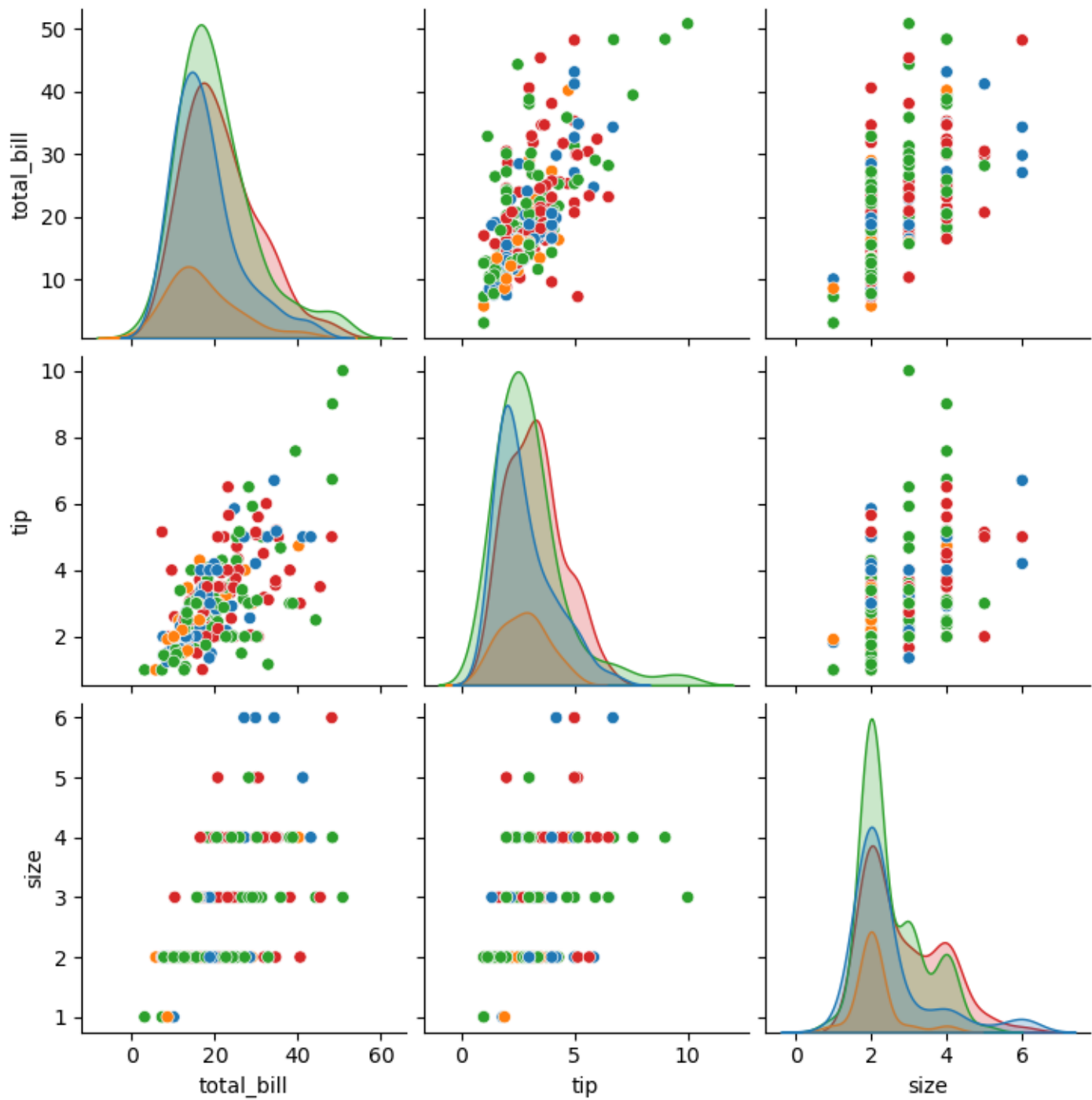


In [52]:

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

Out[52]:

```
<seaborn.axisgrid.PairGrid at 0x20a0793dcc0>
```

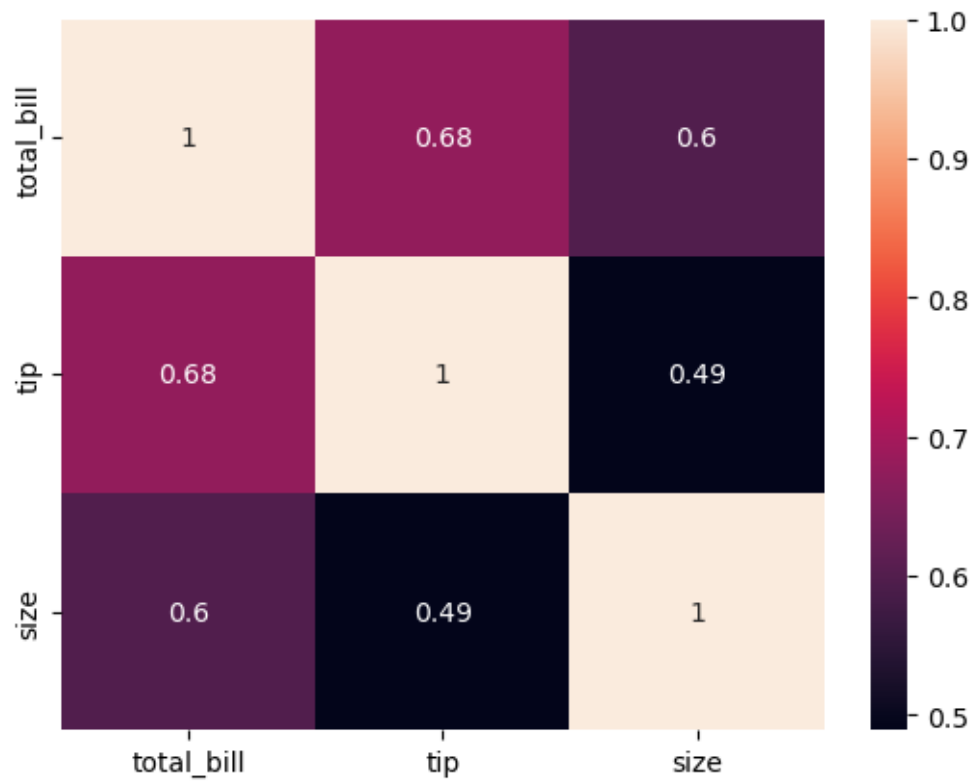


In [53]:

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

Out[53]:

```
<Axes: >
```

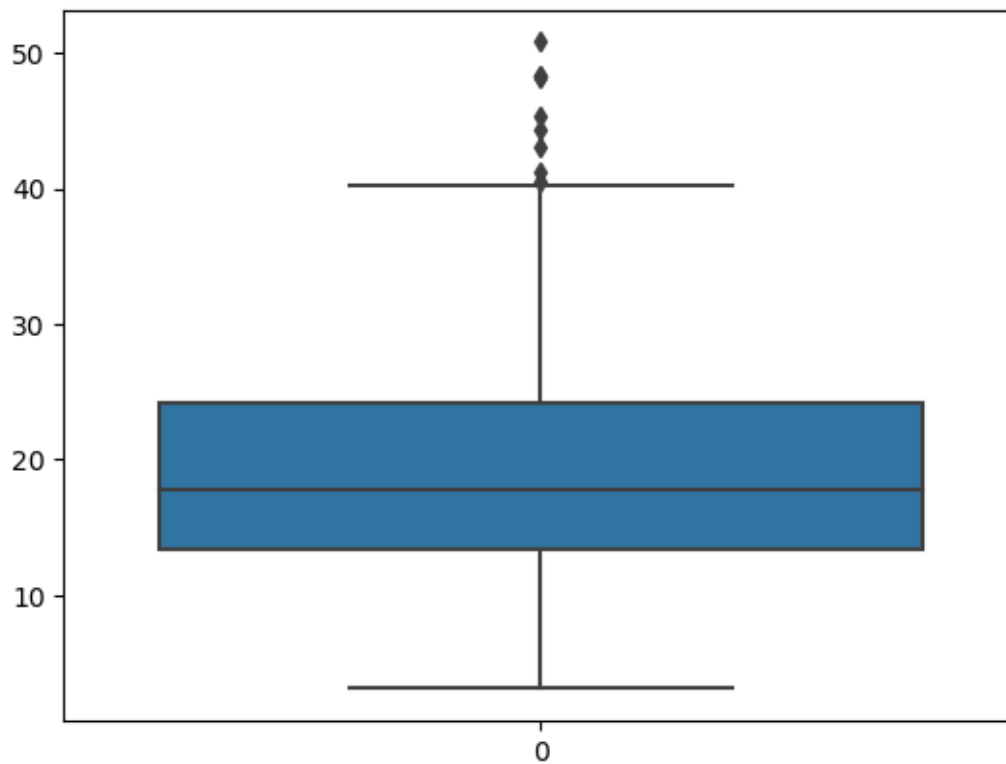


In [54]:

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

Out[54]:

<Axes: >

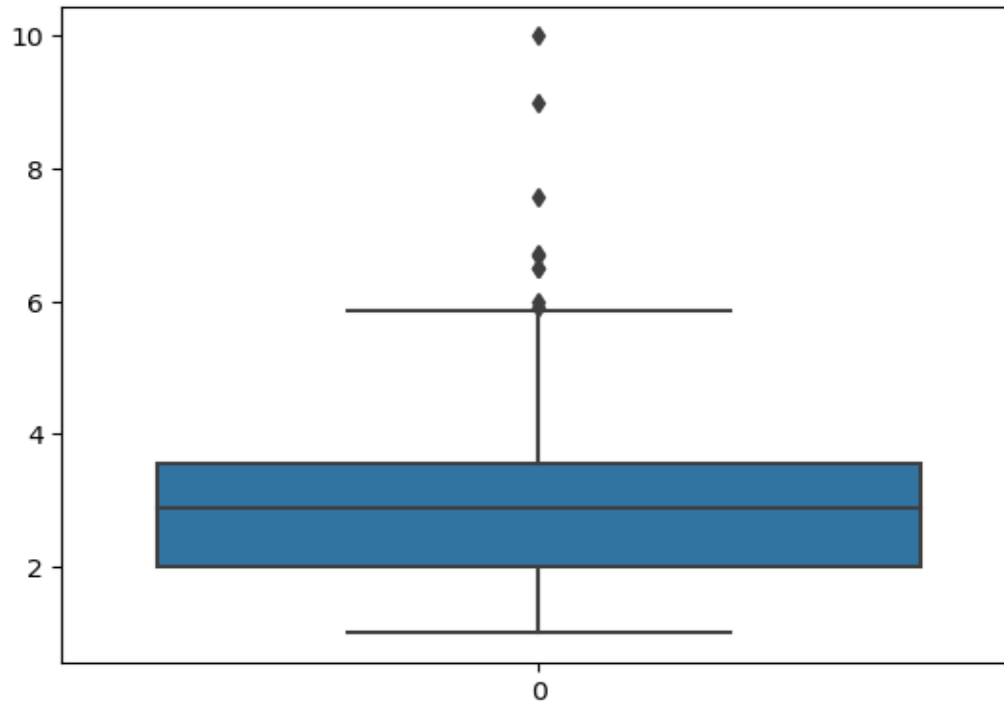


In [55]:

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

Out[55]:

<Axes: >

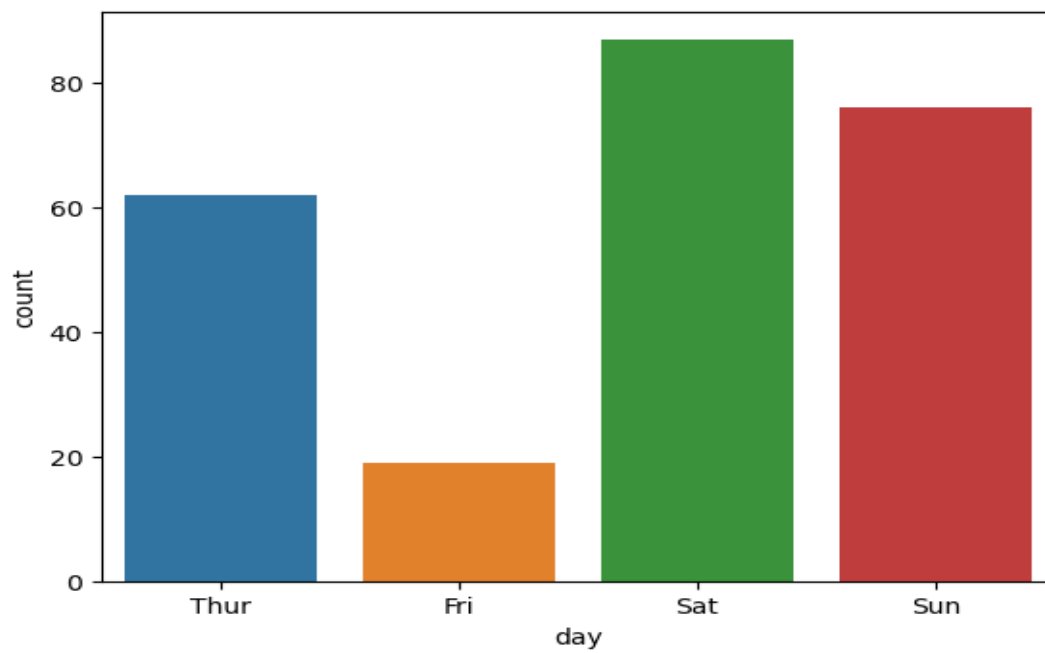


In [57]:

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

Out[57]:

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

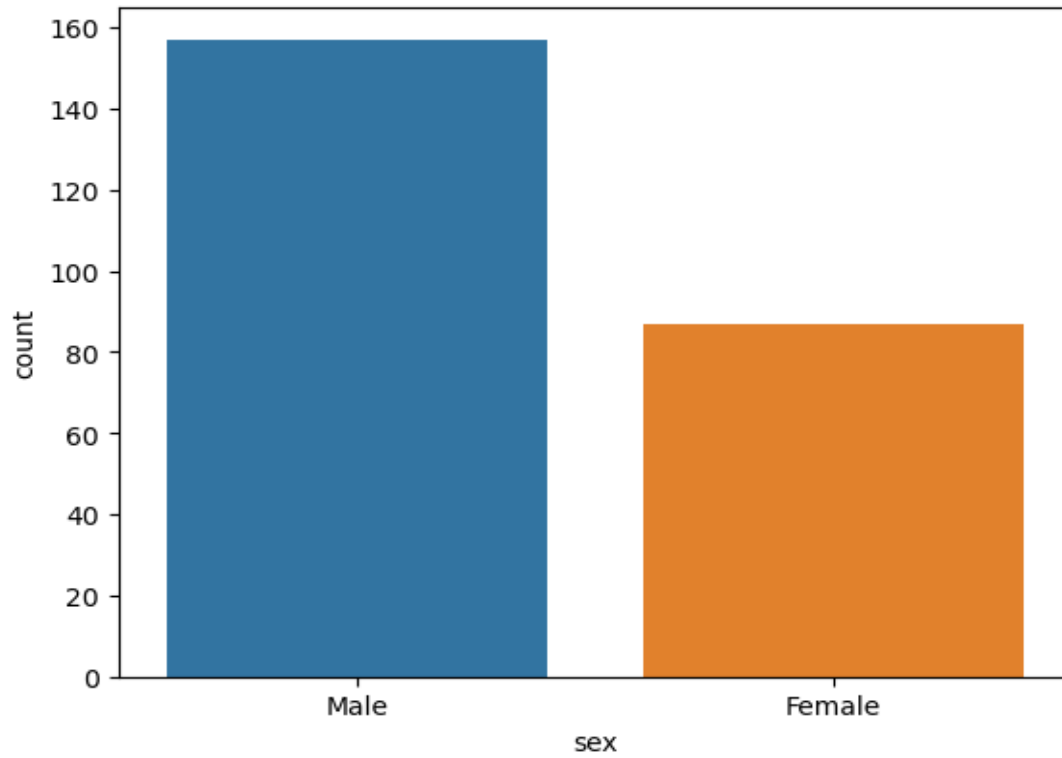


In [59]:

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

Out[59]:

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

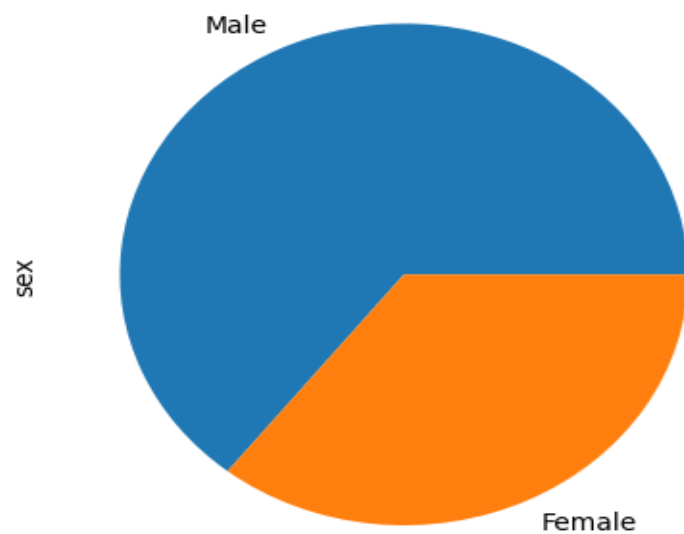


In [60]:

```
tips.sex.value_counts().plot(kind='pie')
```

Out[60]:

```
<Axes: ylabel='sex'>
```

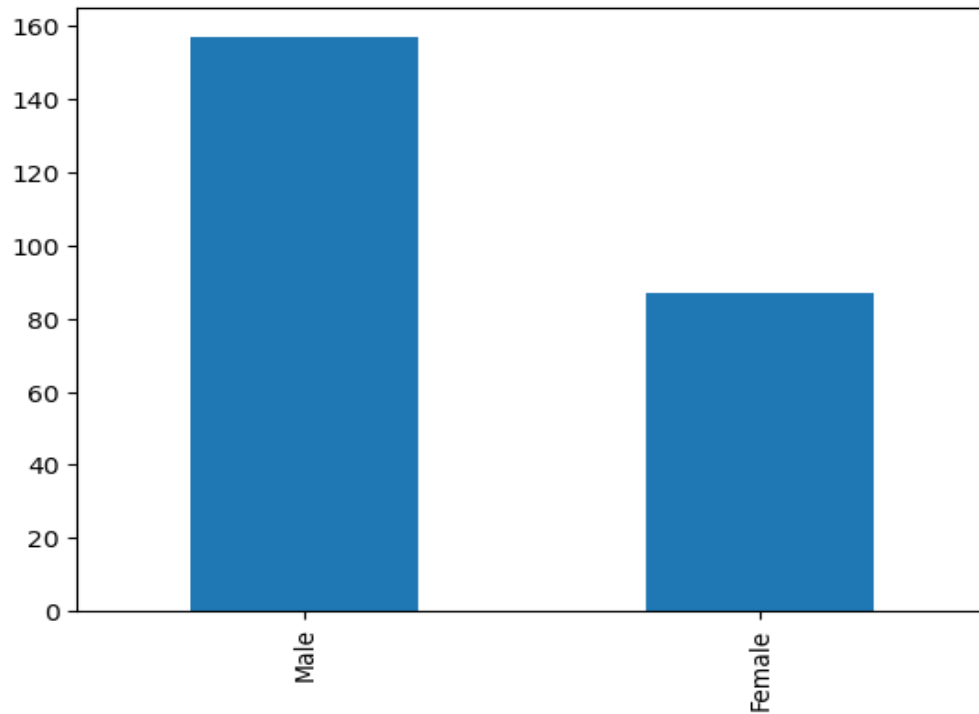


In [61]:

```
tips.sex.value_counts().plot(kind='bar')
```

Out[61]:

<Axes: >

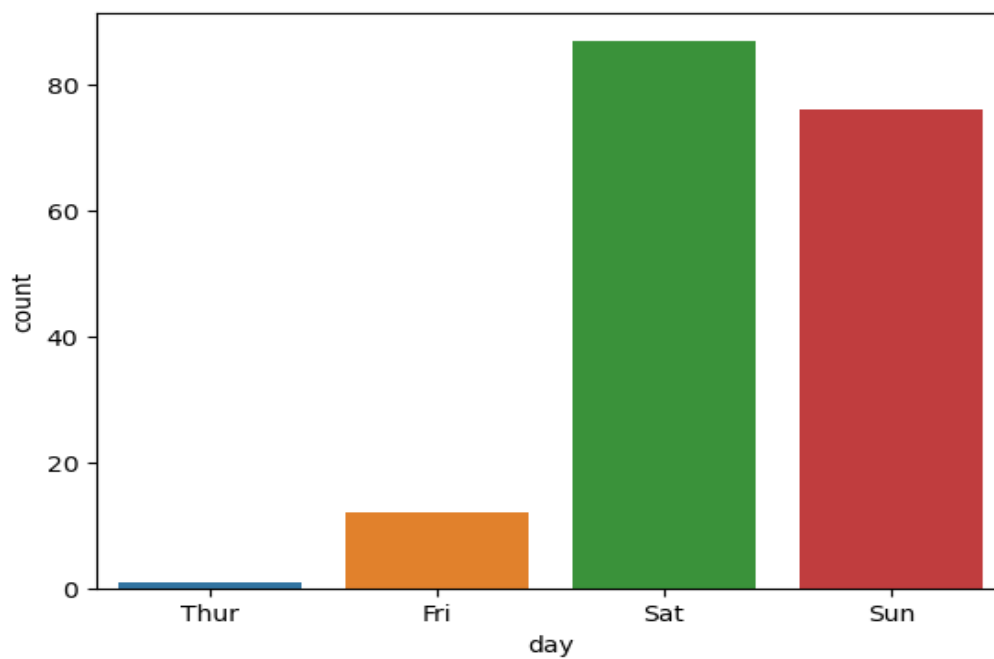


In [64]:

```
dinner_data = tips[tips['time'] == 'Dinner']  
sns.countplot(x='day', data=dinner_data)
```

Out[64]:

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



Result:

EDA helped identify data patterns, distributions, and potential anomalies.

Quantitative analysis revealed numeric trends and outliers, while qualitative analysis explained categorical patterns.

This process improved understanding of the dataset and guided data preprocessing and model selection.