BATCH : Data Science Turbo Course

LESSON : **Data Visualization**

DATE : 05.09.2023

SUBJECT : **Matplotlib & Seaborn**

ZOOM GİRİŞLERİNİZİ LÜTFEN **LMS** SİSTEMİ ÜZERİNDEN YAPINIZ

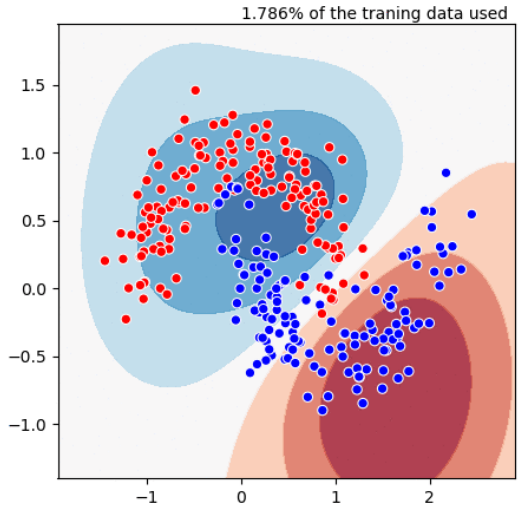
Data Visualization With Python

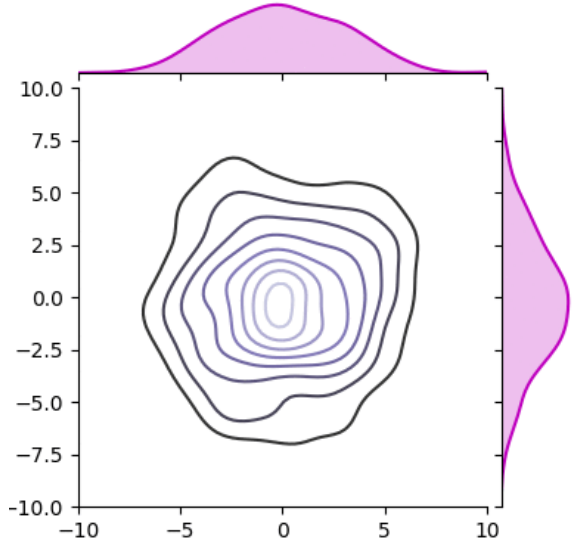
Eğitim Programı

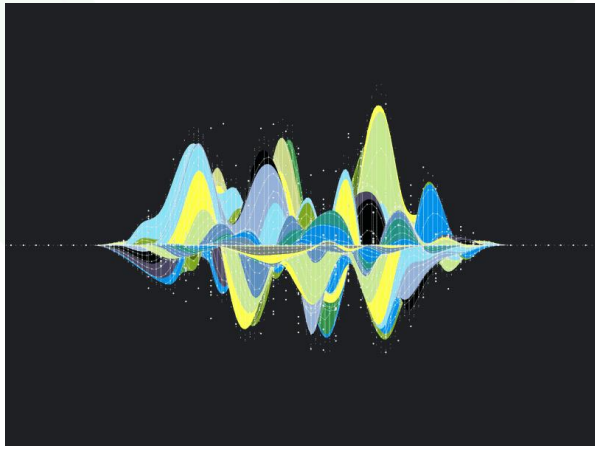
- Data Visualization Intro
- Matplotlib & Seaborn
- EDA & Visualization Project with Real Dataset

Kursun Kapsamı

matplotlib seaborn Project







Dersin Kapsamı

matplotlib

seaborn

```
import matplotlib.pyplot as plt
```

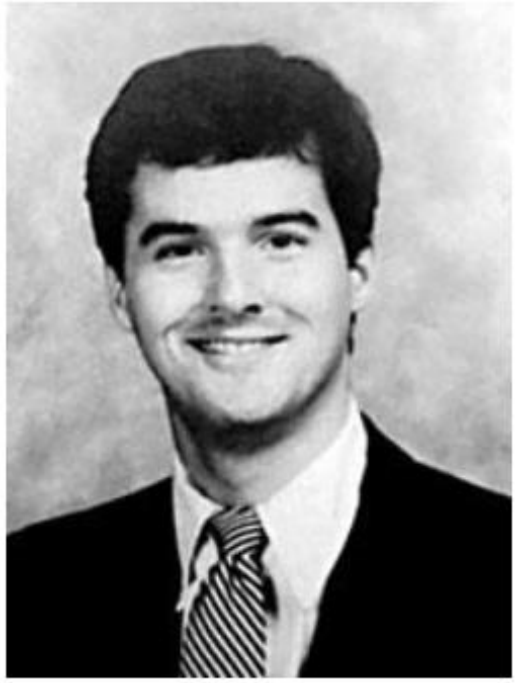
```
import seaborn as sns
```

Create figure, axes, subplots

Built on matplotlib and
can be used together with it

Matplotlib & Seaborn

- Python'da veriyi görselleştirmek için kullanılır.
- 2002 yılında John Hunter tarafından matlap tarzında bir arayüz oluşturmak amacıyla bir proje olarak başlatıldı.
- İlk sürümü 2003 yılında yayınlandı.
- Geliştirilerek seaborn kütüphanesi ortaya çıkmıştır.



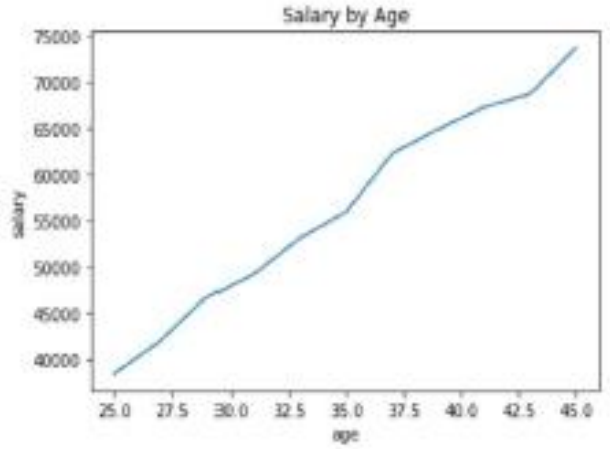
Matplotlib & Seaborn

FEATURES	MATPLOTLIB	SEABORN
Functionality	It is utilized for making basic graphs. Datasets are visualised with graphs styles. • Bar graphs, • Histograms, • Pie charts, • Scatter plots, • Lines and so on.	Seaborn contains a number of patterns and plots for data visualization. It uses fascinating themes. It helps in compiling whole data into a single plot.
Syntax	It uses comparatively complex and lengthy syntax.	It uses comparatively simple syntax which is easier to learn and understand.

Two Methods

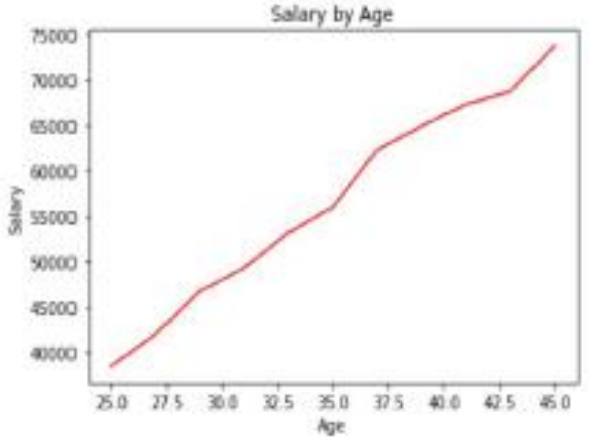
Functional Method

```
plt.plot(age, salary)
plt.xlabel("age")
plt.ylabel("salary")
plt.title("Salary by Age")

plt.show()
```



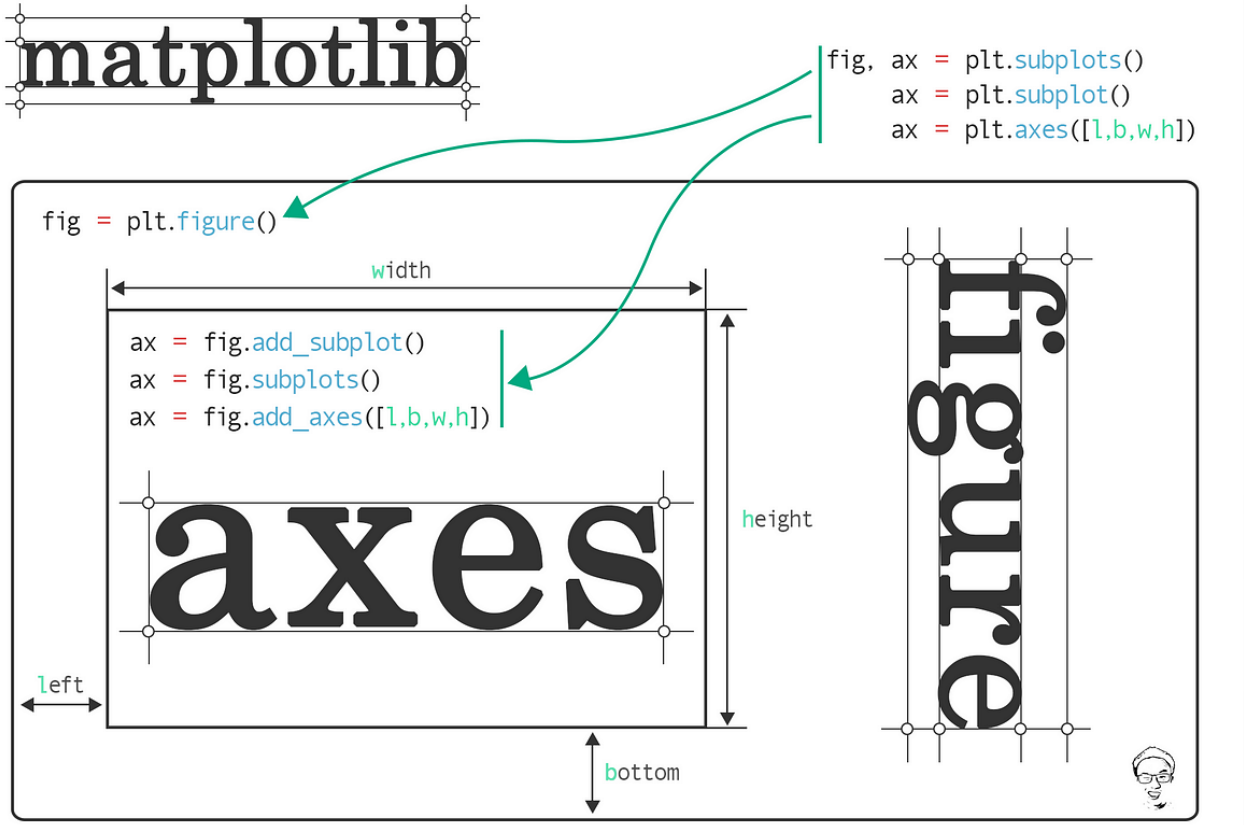
Object Oriented

```
fig, ax = plt.subplots()

ax.plot(age, salary, "r")
ax.set_xlabel("Age")
ax.set_ylabel("Salary")
ax.set_title("Salary by Age")
```



Axis – Axes – Figure ?

Figure, Axes, Axis nasıl anlaşılmalıdır?



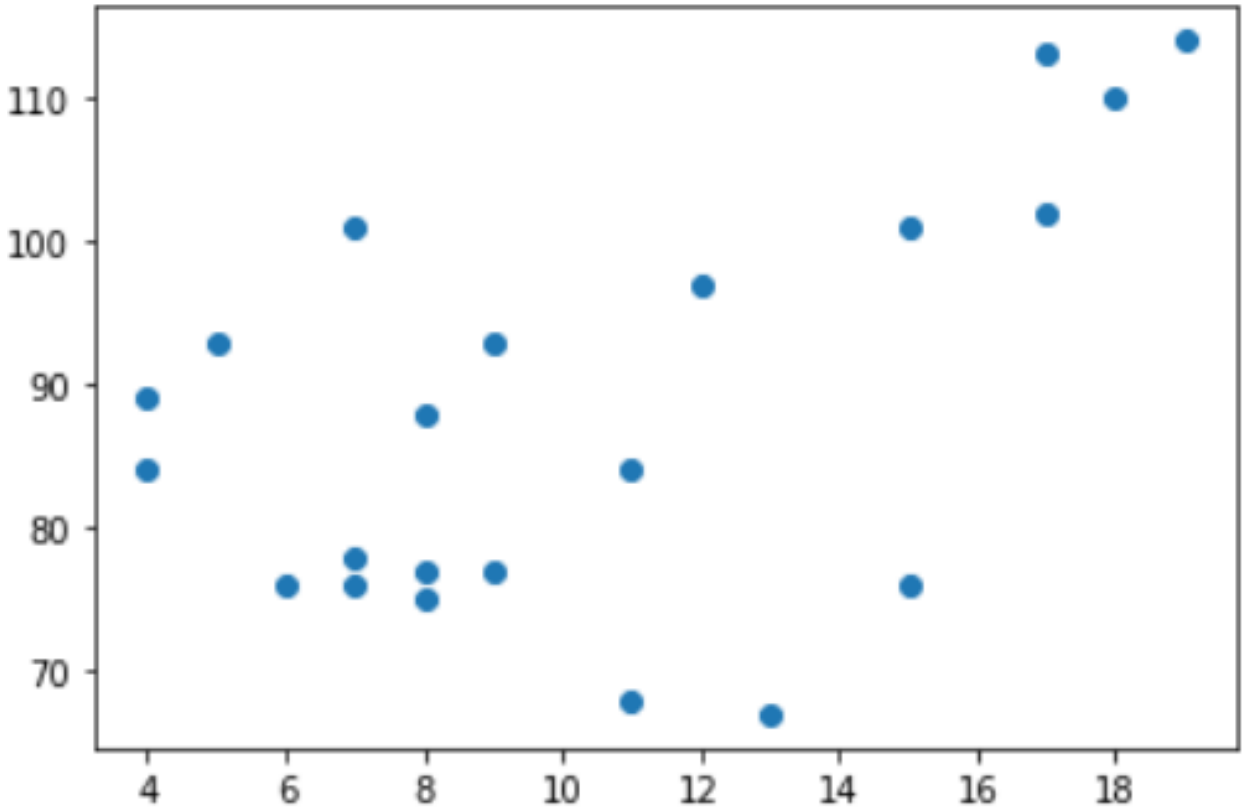
Scatter Plot

- Data seti oluşturma

```
x = np.array([4, 6, 8, 7, 7, 15, 5 ,9 ,4,
              11, 13, 8, 7, 8, 11, 9, 12,
              15 ,17 ,17 ,19, 18 ])

y = np.array([89, 76, 77, 78, 101, 76, 93,
              77 ,84 ,68, 67, 75, 76, 88,
              84, 93, 97, 101 ,102 ,113 ,114, 110 ])
```

- Grafik oluşturma

```
plt.scatter(x, y)
plt.show()
```



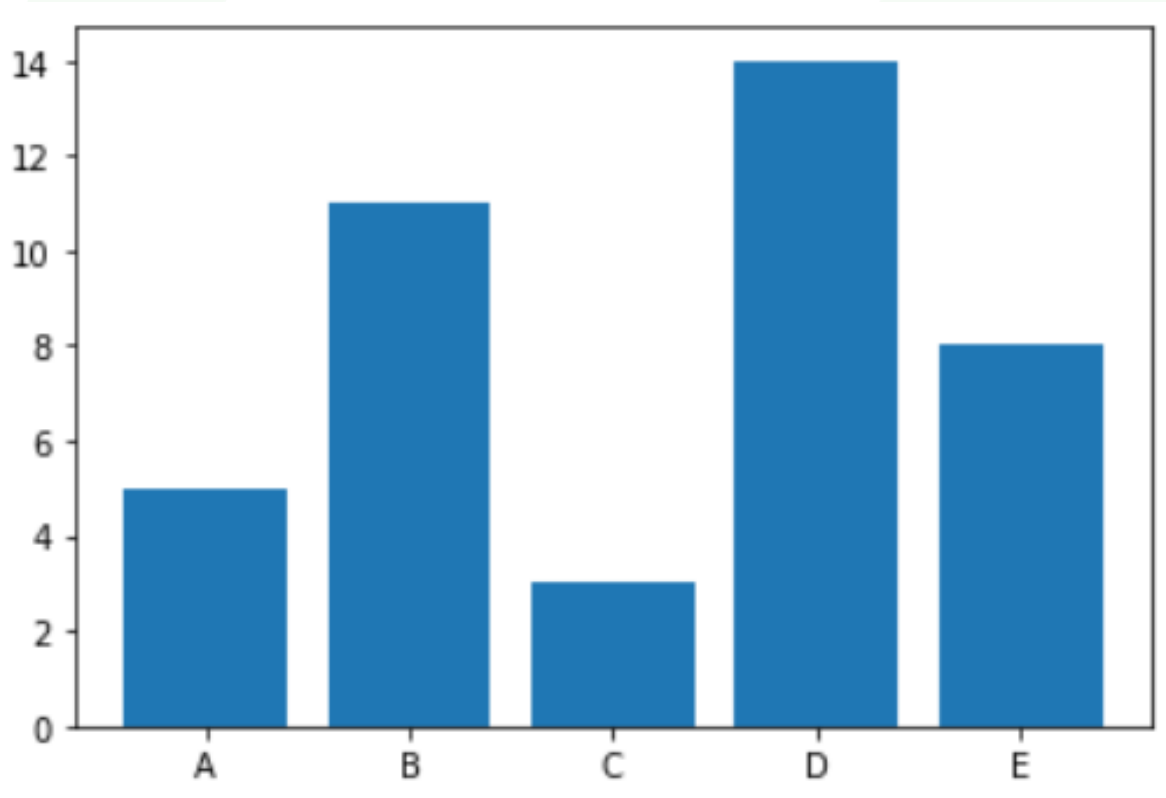
Bar Chart

- Data seti oluşturma

```
x = np.array(["A", "B", "C", "D", "E"])
y = np.array([5, 11, 3, 14, 8])
```

- Grafik oluşturma

```
plt.bar(x, y)
plt.show()
```



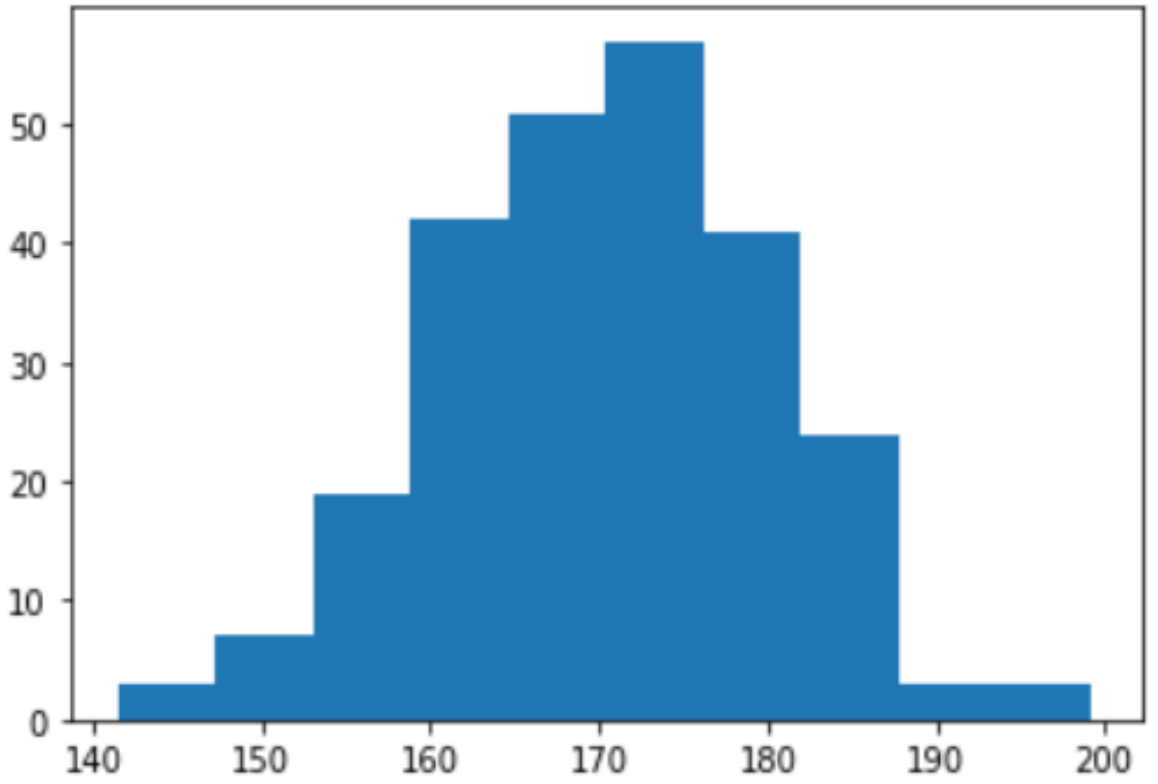
Histogram

- Data seti oluşturma

```
1  x = np.random.normal(170, 10, 250)
```

- Grafik oluşturma

```
1  plt.hist(x)
2  plt.show()
```



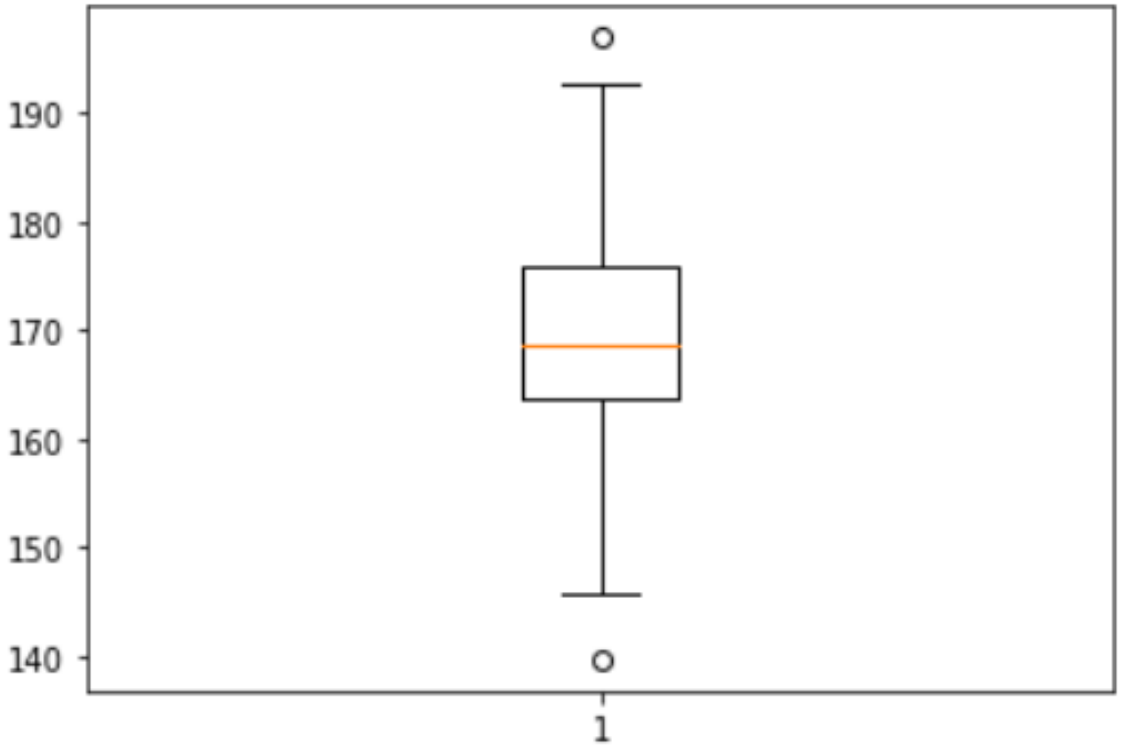
Box Plot

- Data seti oluşturma

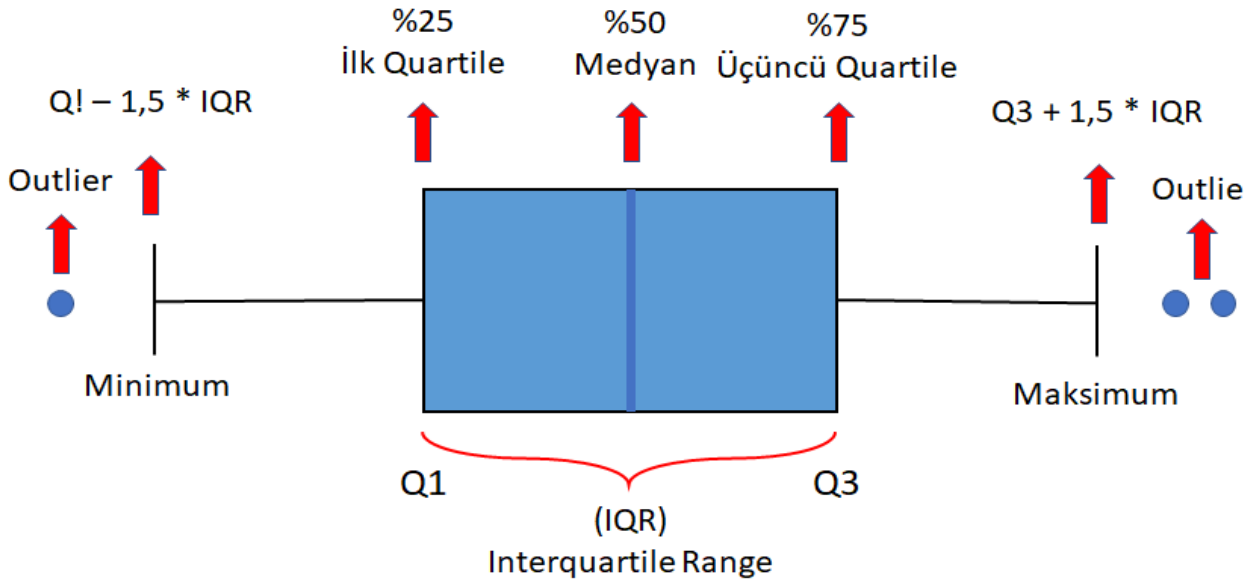
```
1  x = np.random.normal(170, 10, 250)
```

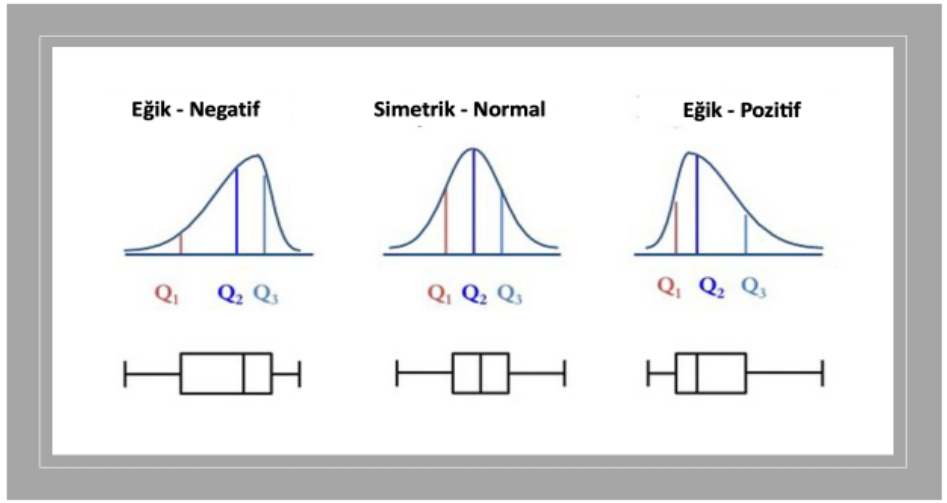
- Grafik oluşturma

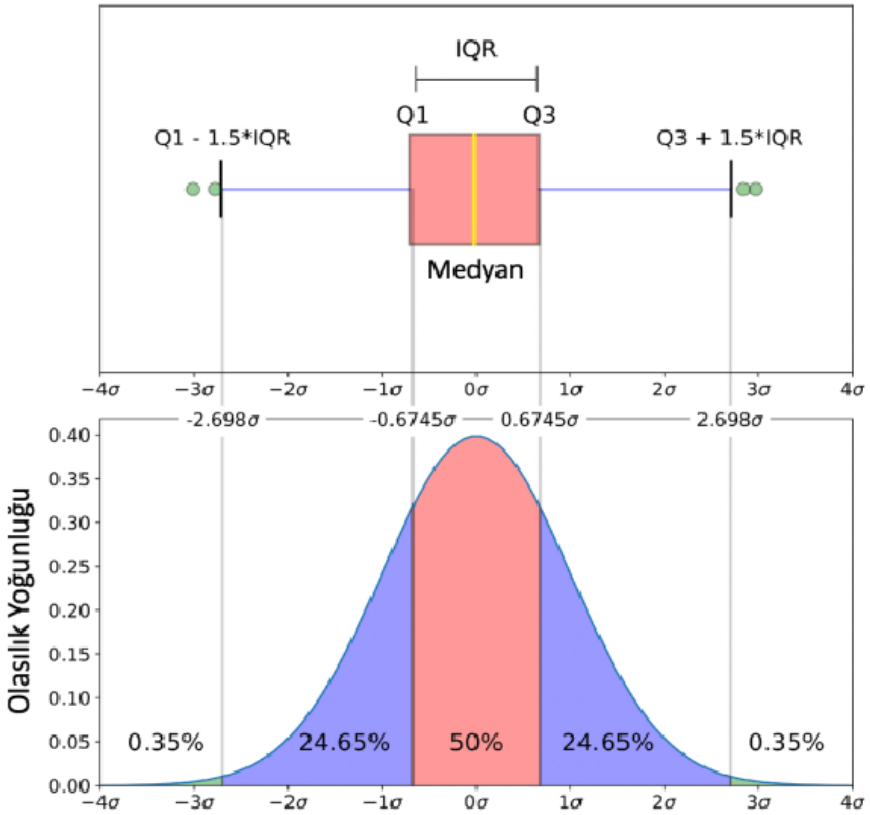
```
1  plt.boxplot(x)
2  plt.show()
```



Box Plot


%25
İlk Quartile
%50
Medyan
%75
Üçüncü Quartile
Q! – 1,5 * IQR
Q3 + 1,5 * IQR
Outlier
Outlie
Minimum
Maksimum
Q1
Q3
(IQR)
Interquartile Range


Eğik - Negatif
Simetrik - Normal
Eğik - Pozitif
Q1 Q2 Q3
Q1 Q2 Q3
Q1 Q2 Q3


IQR
Q1
Q3
Q1 - 1.5*IQR
Q3 + 1.5*IQR
Medyan
−4σ −3σ −2σ −1σ 0σ 1σ 2σ 3σ 4σ
-2.698σ
-0.6745σ
0.6745σ
2.698σ
Olasılık Yoğunluğu
0.40 0.35 0.30 0.25 0.20 0.15 0.10 0.05 0.00
0.35%
24.65%
50%
24.65%
0.35%

Pie Chart

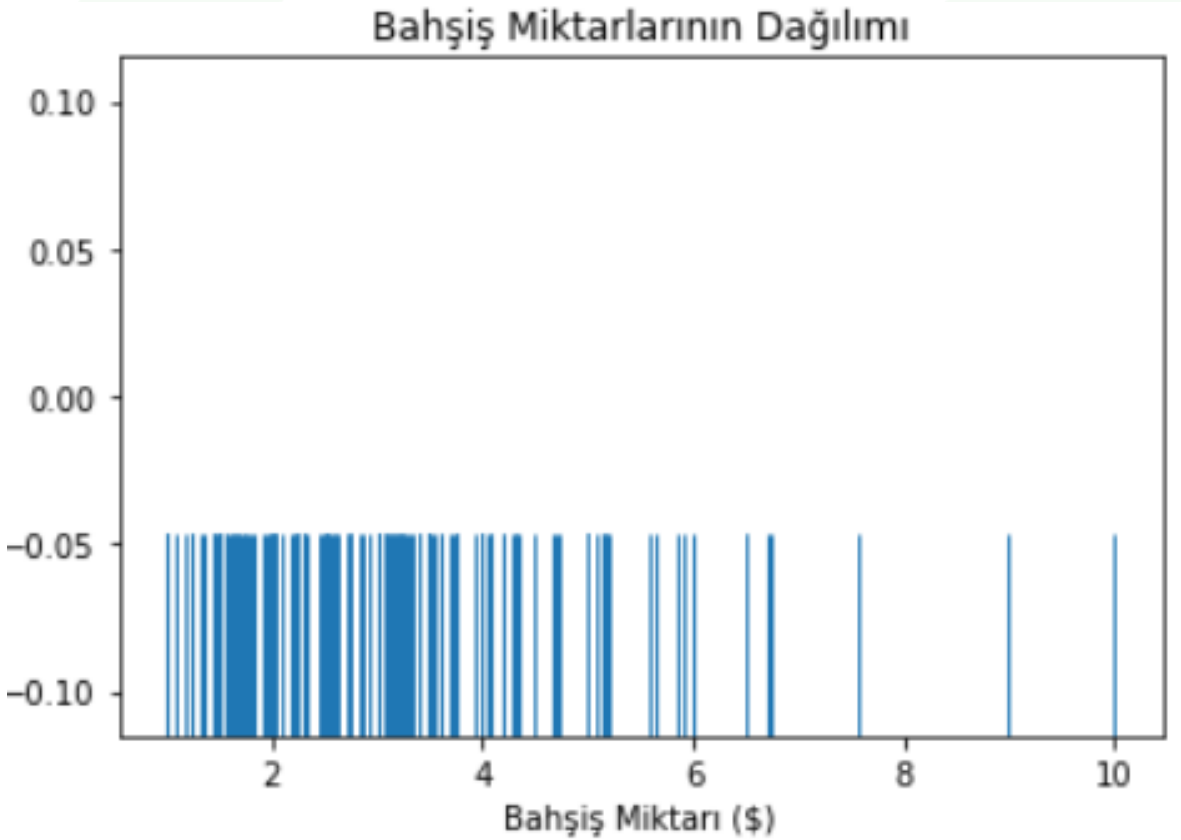
- Genellikle nominal ve ordinal değişkenlerle kullanılır.
- Oransal olarak dilimlerin toplamı %100'e eşittir.
- Her dilim değişken içindeki bir niteliği temsil eder

Seaborn Plot Types

Distributions Plots Dağılım Grafikleri	Categorical Plots (Kategorik Grafikler)	Comparison Plots Karşılaştırma Grafikleri
• displot • histplot • kdeplot • rugplot	• barplot • countplot • boxplot • swarmplot • violinplot	• jointplot • pairplot • catplot • matrixplot • gridplot

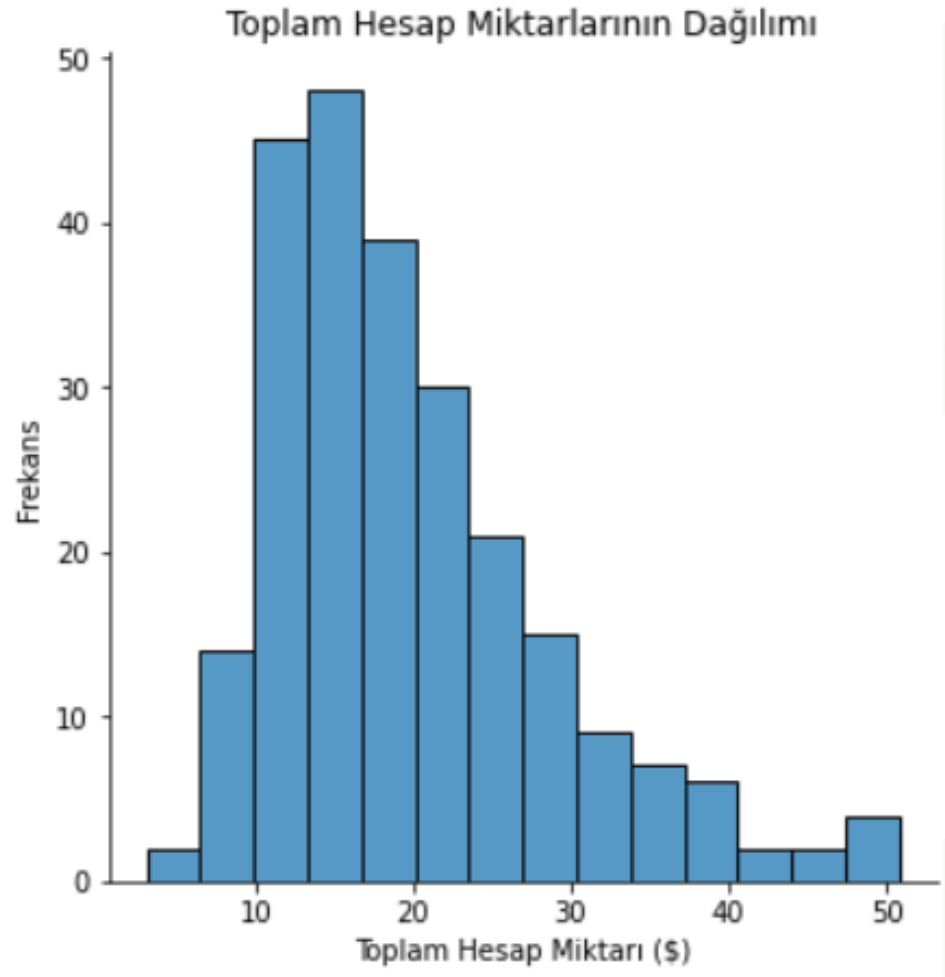
Distribution Plots - rugplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Rugplot oluşturma
sns.rugplot(x="tip", data = tips, height = 0.3)

# Eksen etiketleri ve başlık
plt.xlabel('Bahşiş Miktarı ($)')
plt.title('Bahşiş Miktarlarının Dağılımı')

# Grafiki göster
plt.show()
```



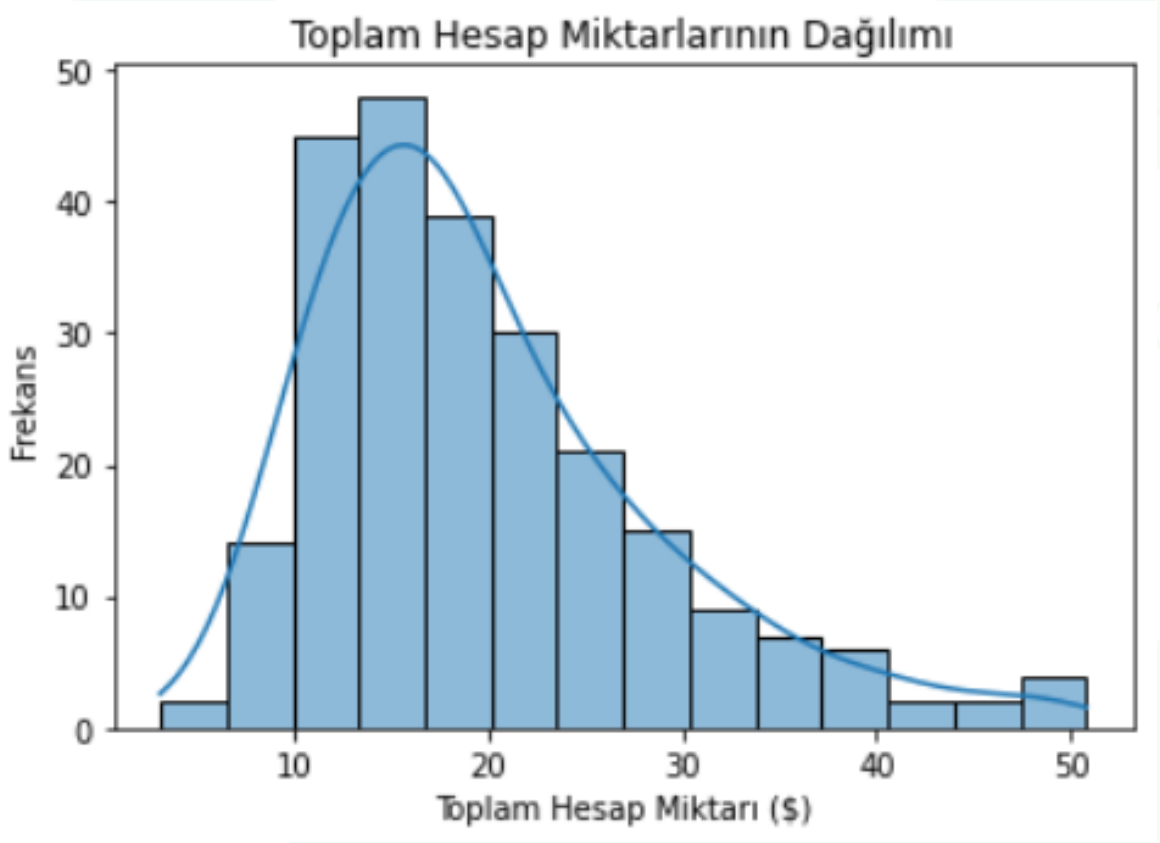
Distribution Plots - displot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Displot oluşturma
sns.displot(tips['total_bill'])

# Eksen etiketleri ve başlık
plt.xlabel('Toplam Hesap Miktarı ($)')
plt.ylabel('Frekans')
plt.title('Toplam Hesap Miktarlarının Dağılımı')

# Grafiki göster
plt.show()
```



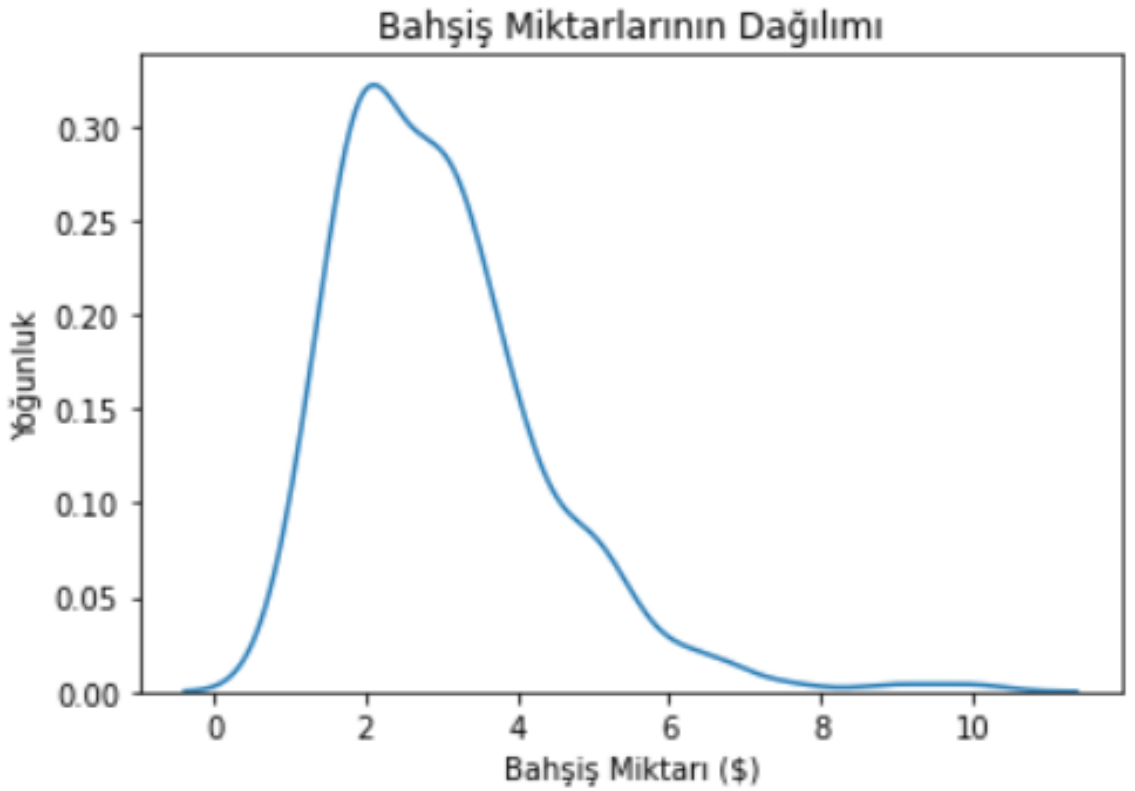
Distribution Plots - histplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Histplot oluşturma
sns.histplot(tips['total_bill'], kde=True)

# Eksen etiketleri ve başlık
plt.xlabel('Toplam Hesap Miktarı ($)')
plt.ylabel('Frekans')
plt.title('Toplam Hesap Miktarlarının Dağılımı')

# Grafiki göster
plt.show()
```



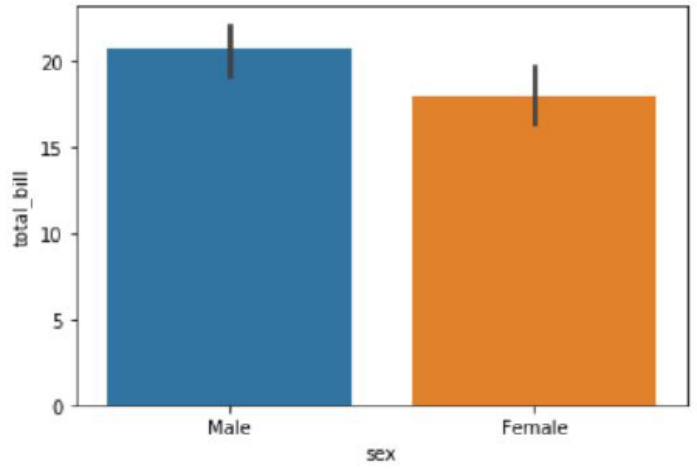
Distribution Plots - kde

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# KDE plot oluşturma
sns.kdeplot(tips['tip'])

# Eksen etiketleri ve başlık
plt.xlabel('Bahşiş Miktarı ($)')
plt.ylabel('Yoğunluk')
plt.title('Bahşiş Miktarlarının Dağılımı')

# Grafiki göster
plt.show()
```

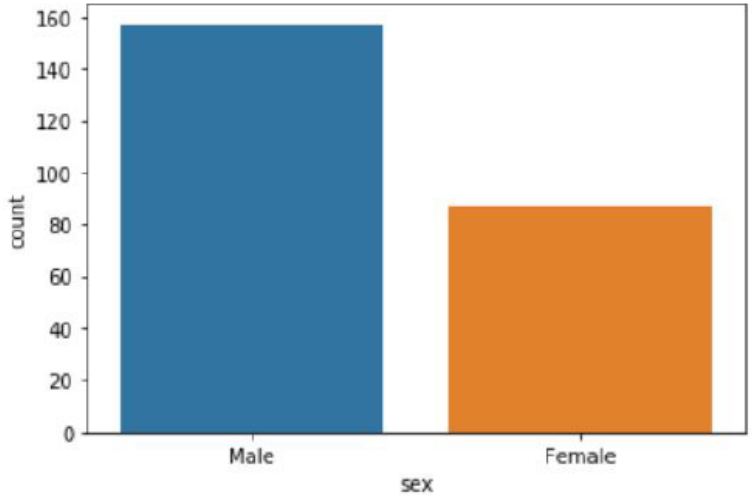


Categorical Plots

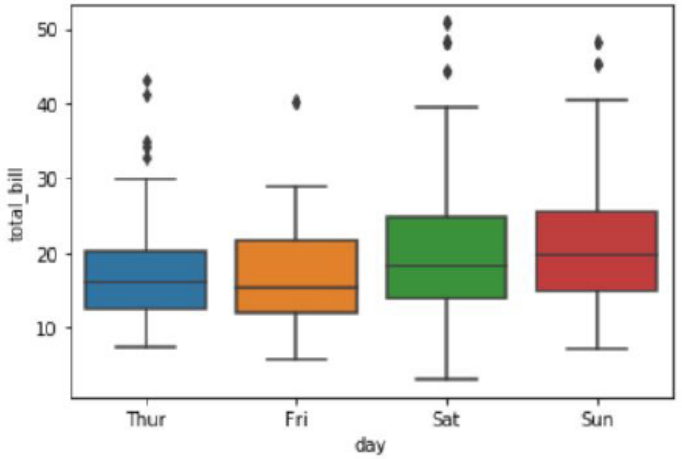
barplot



countplot

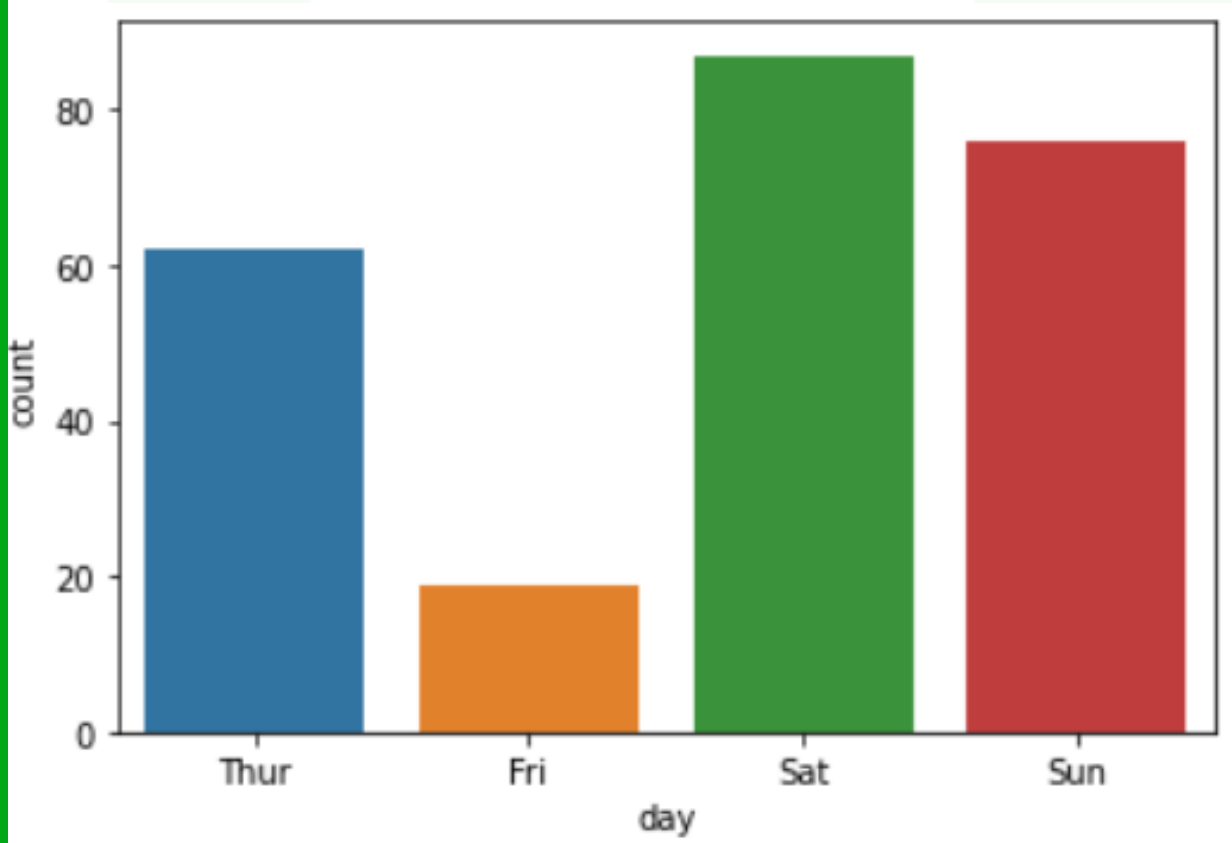


boxplot



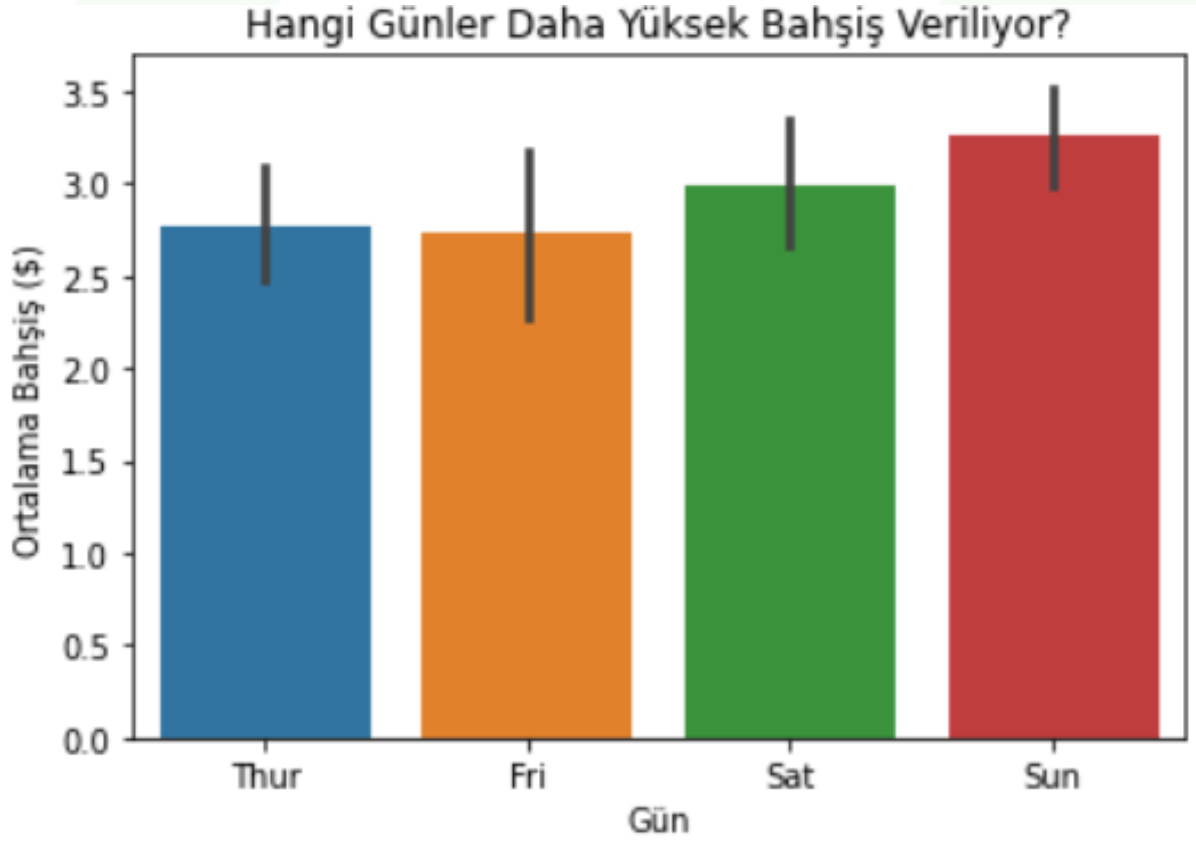
Categorical Plots - countplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Countplot oluşturma
sns.countplot(x="day", data=tips)

# Grafiki göster
plt.show()
```

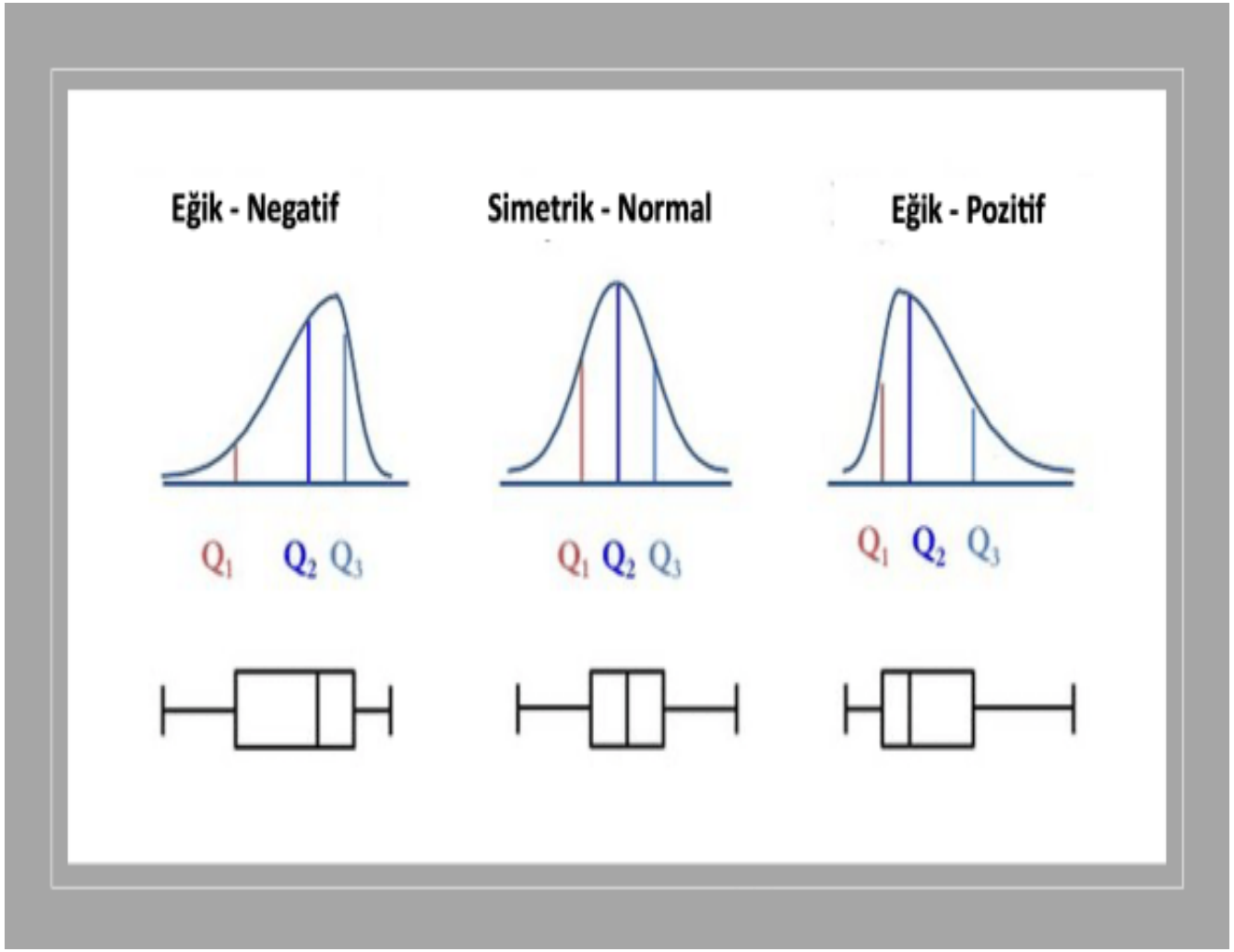


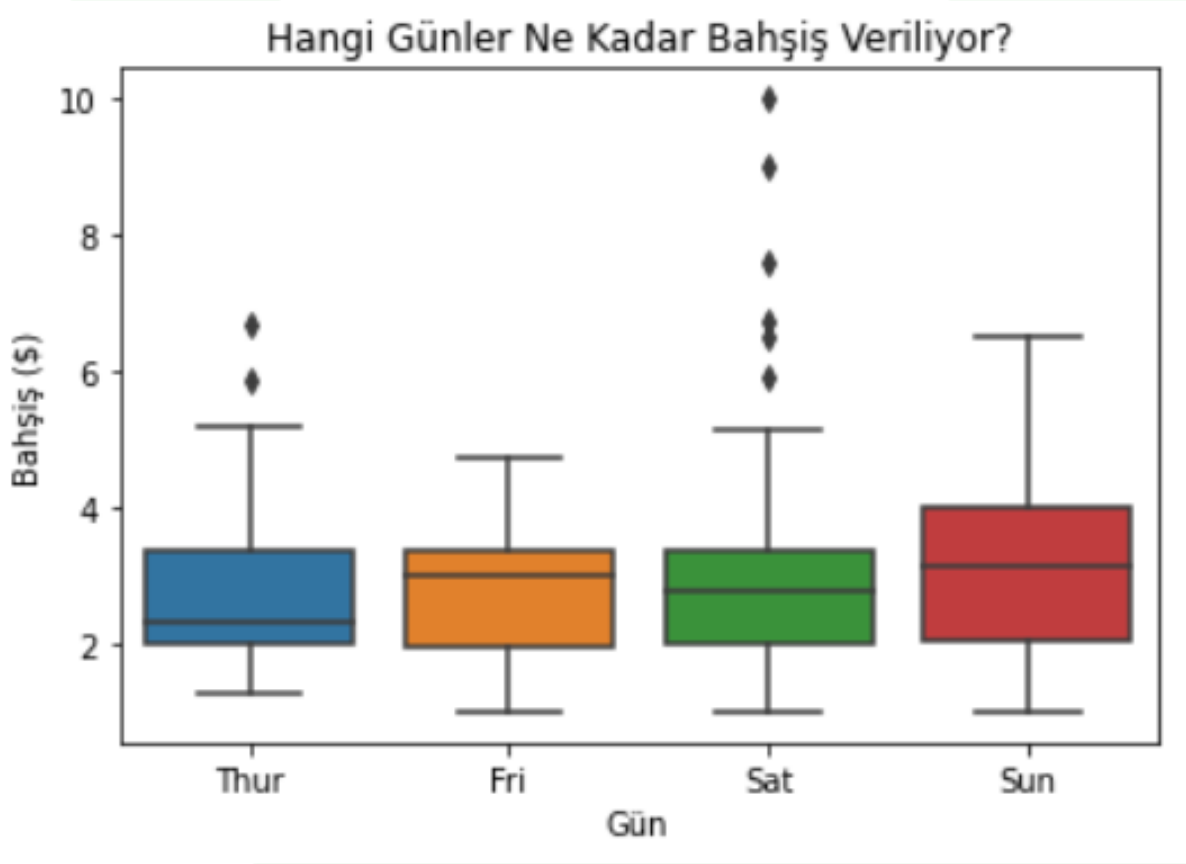
Categorical Plots - barplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Barplot oluşturma
sns.barplot(x='day', y='tip', data=tips)

# Eksen etiketleri ve başlık
plt.xlabel('Gün')
plt.ylabel('Ortalama Bahşiş ($)')
plt.title('Hangi Günler Daha Yüksek Bahşiş Veriliyor?')

# Grafiki göster
plt.show()
```



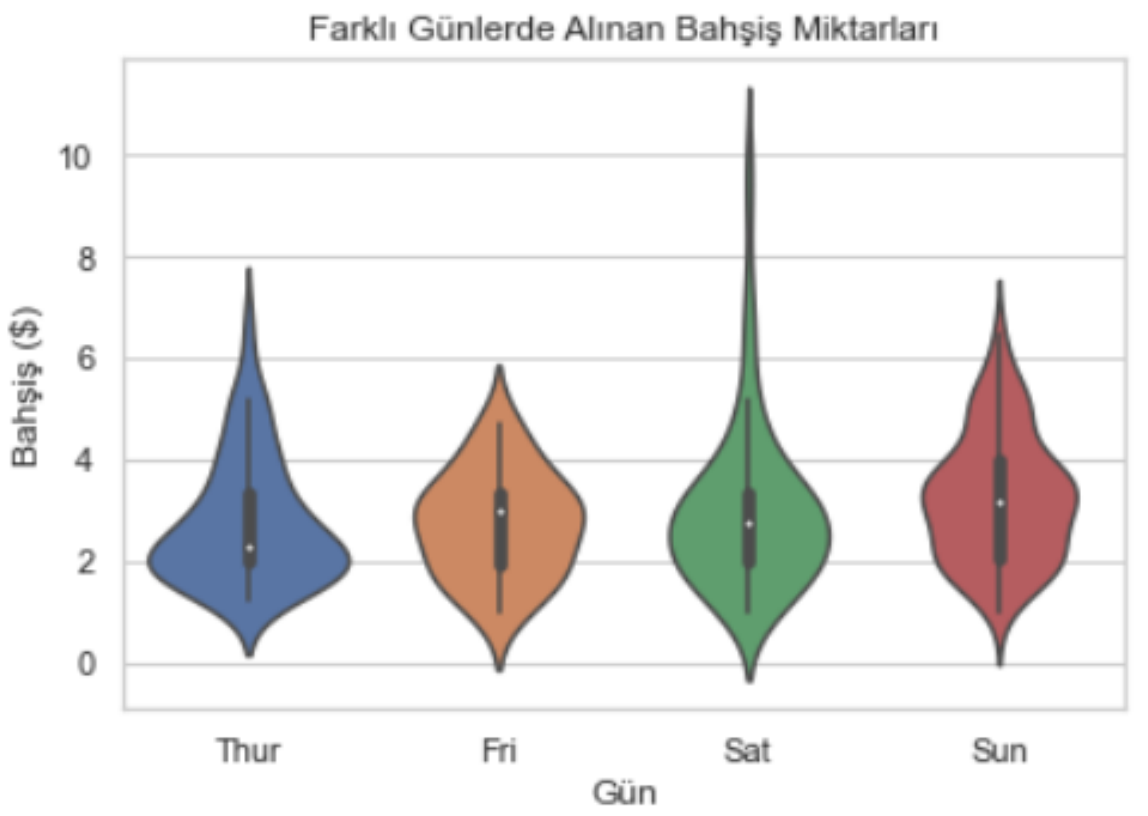
Categorical Plots - boxplot





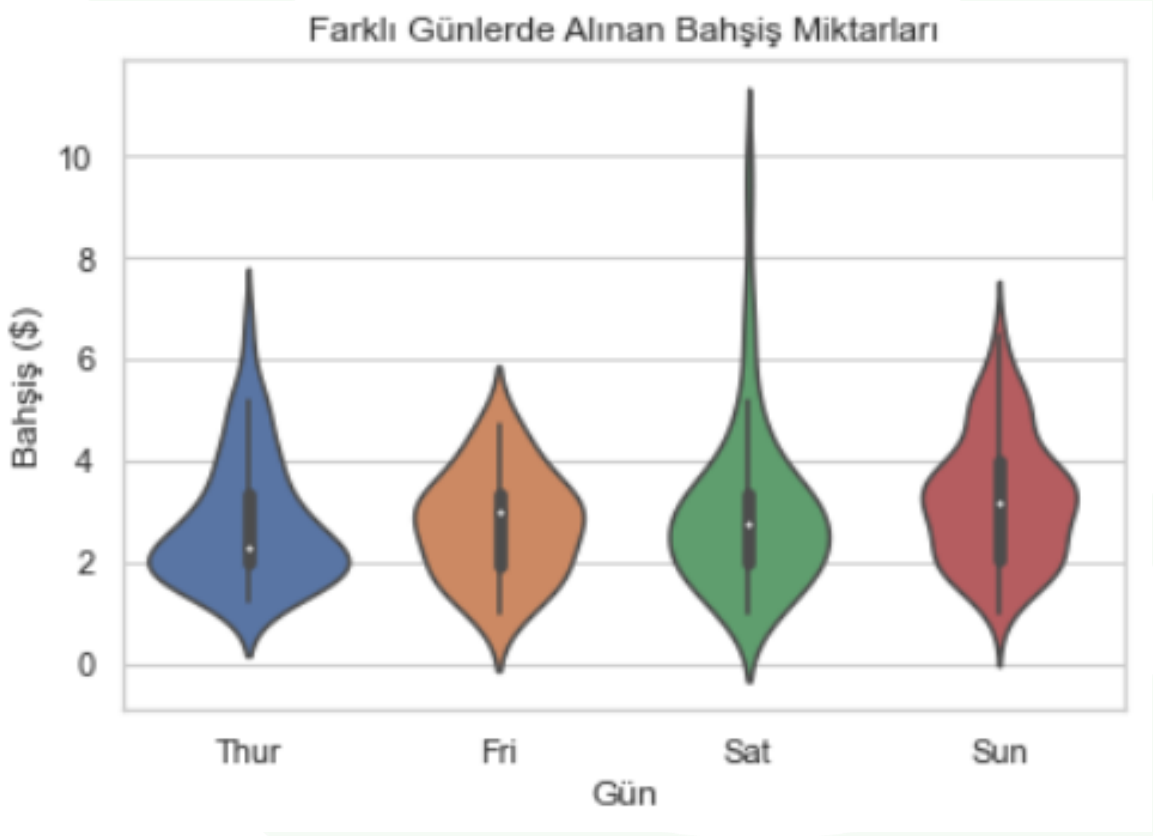
Categorical Plots – swarmplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Violin plot oluşturma
sns.violinplot(x="day", y="tip", data=tips)

# Başlık ve etiketler
plt.title('Farklı Günlerde Alınan Bahşiş Miktarları')
plt.xlabel('Gün')
plt.ylabel('Bahşiş ($)')

# Grafiki göster
plt.show()
```



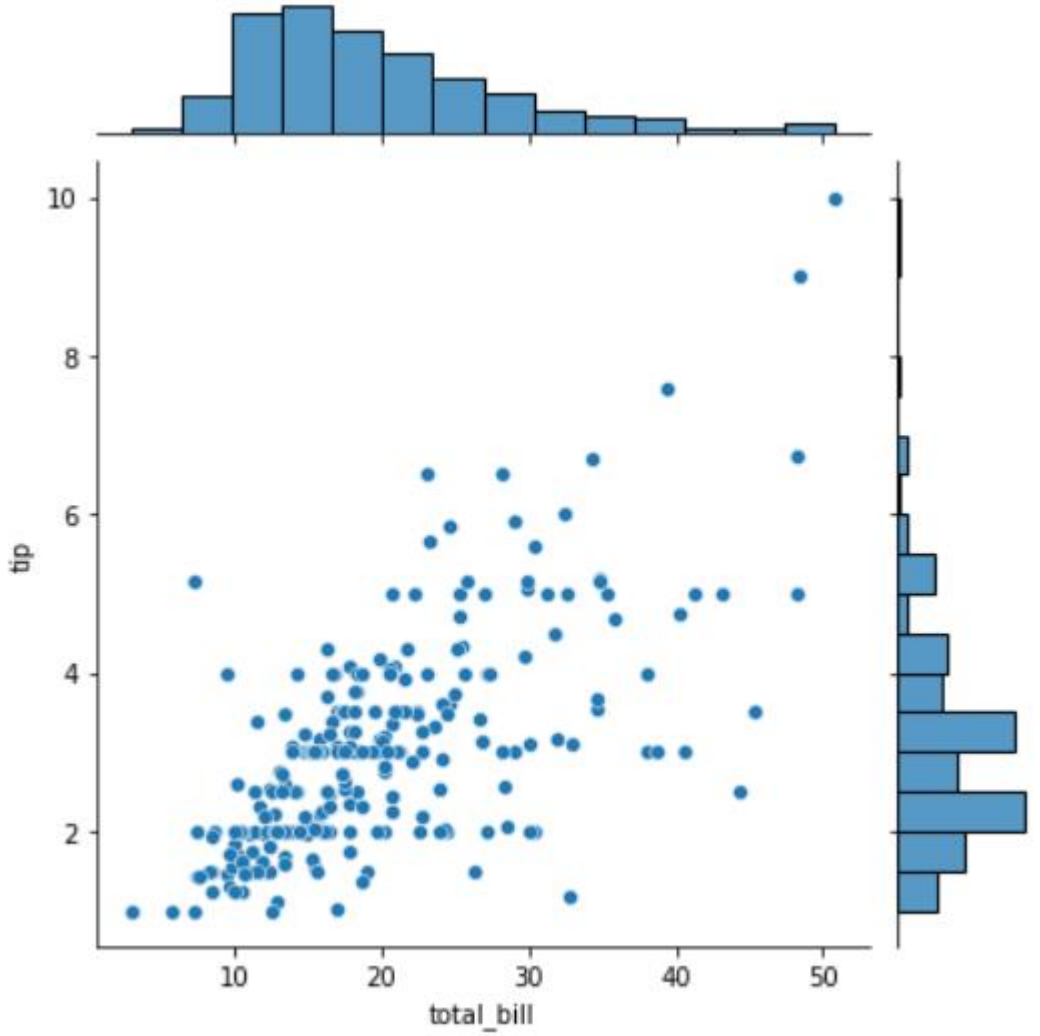
Categorical Plots – violinplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Violin plot oluşturma
sns.violinplot(x="day", y="tip", data=tips)

# Başlık ve etiketler
plt.title('Farklı Günlerde Alınan Bahşiş Miktarları')
plt.xlabel('Gün')
plt.ylabel('Bahşiş ($)')

# Grafiki göster
plt.show()
```



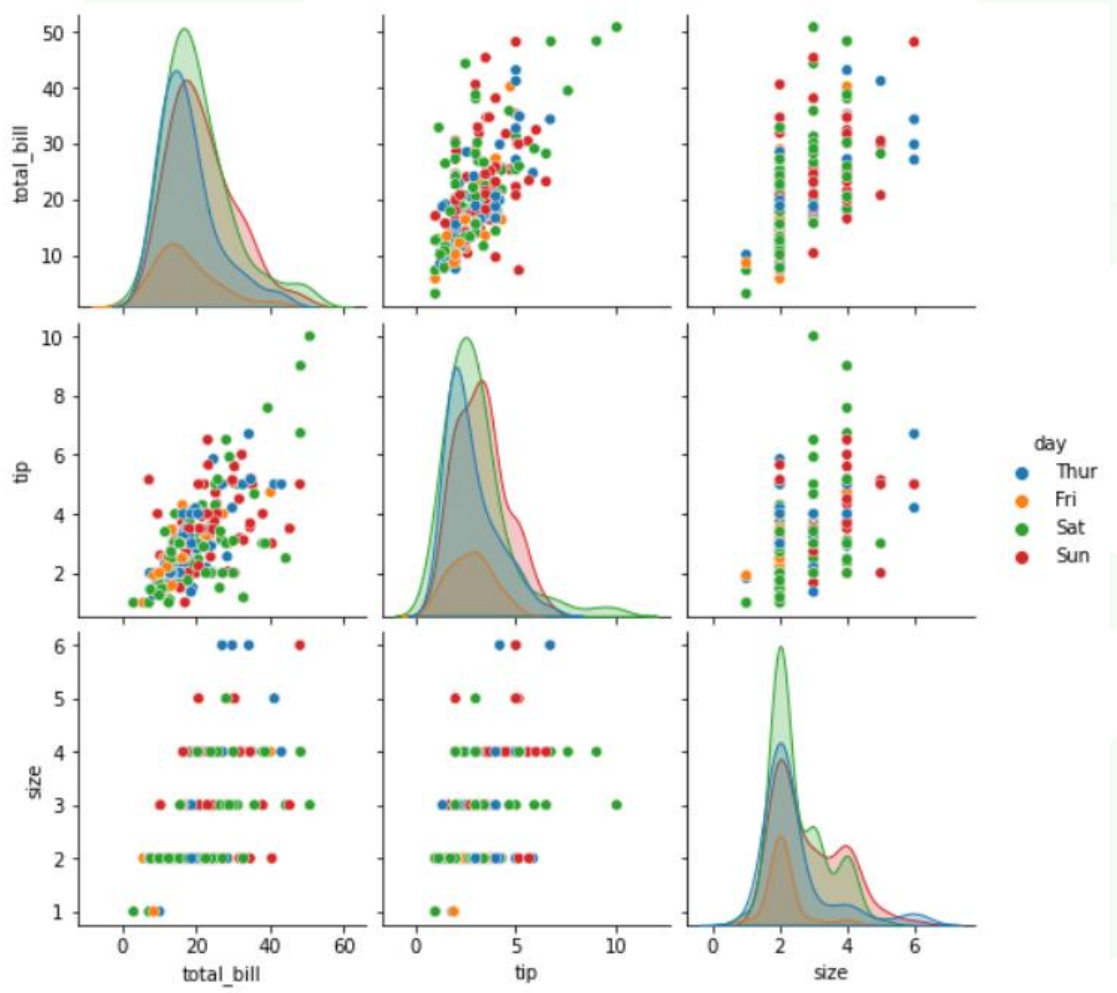
Comparison Plots - jointplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Jointplot oluşturma
sns.jointplot(x='total_bill', y='tip', data=tips)

# Grafiki göster
plt.show()
```



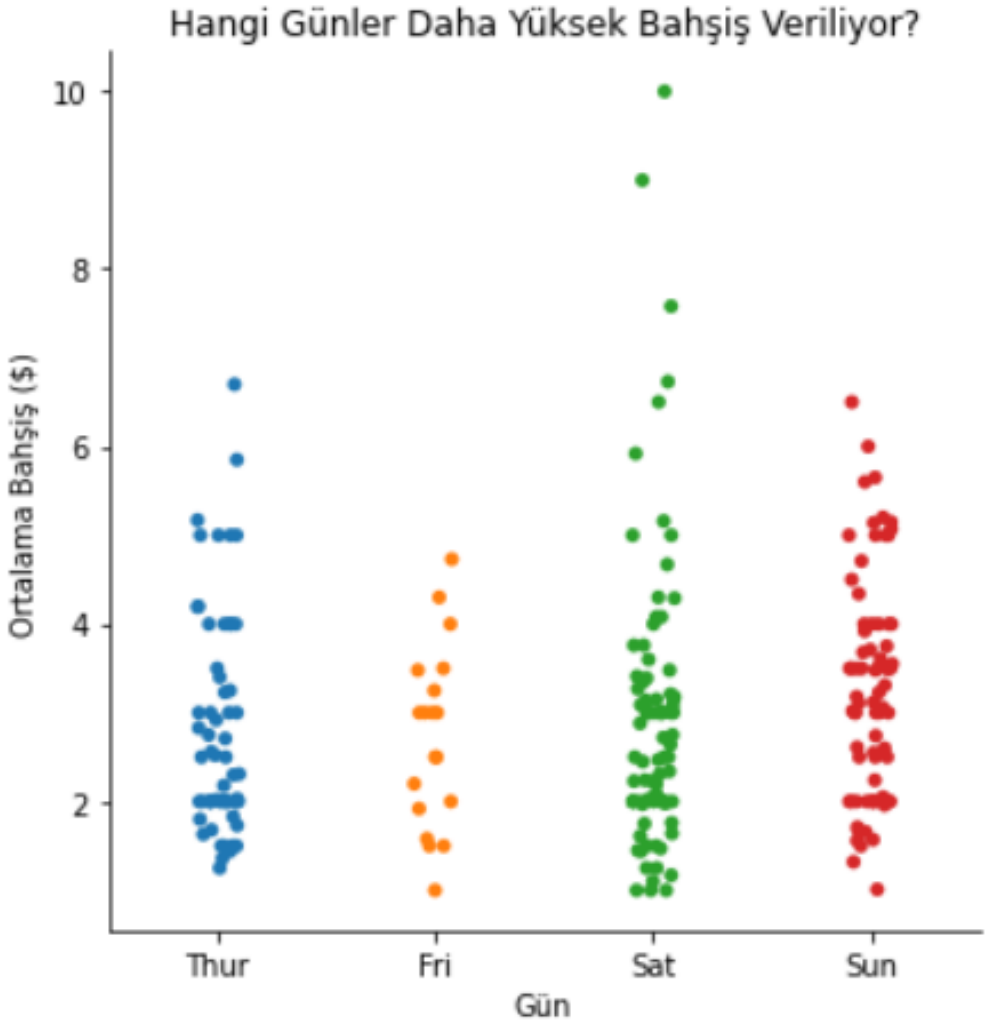
Comparison Plots - pairplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Pairplot oluşturma
sns.pairplot(tips, hue = "day")

# Grafiki göster
plt.show()
```



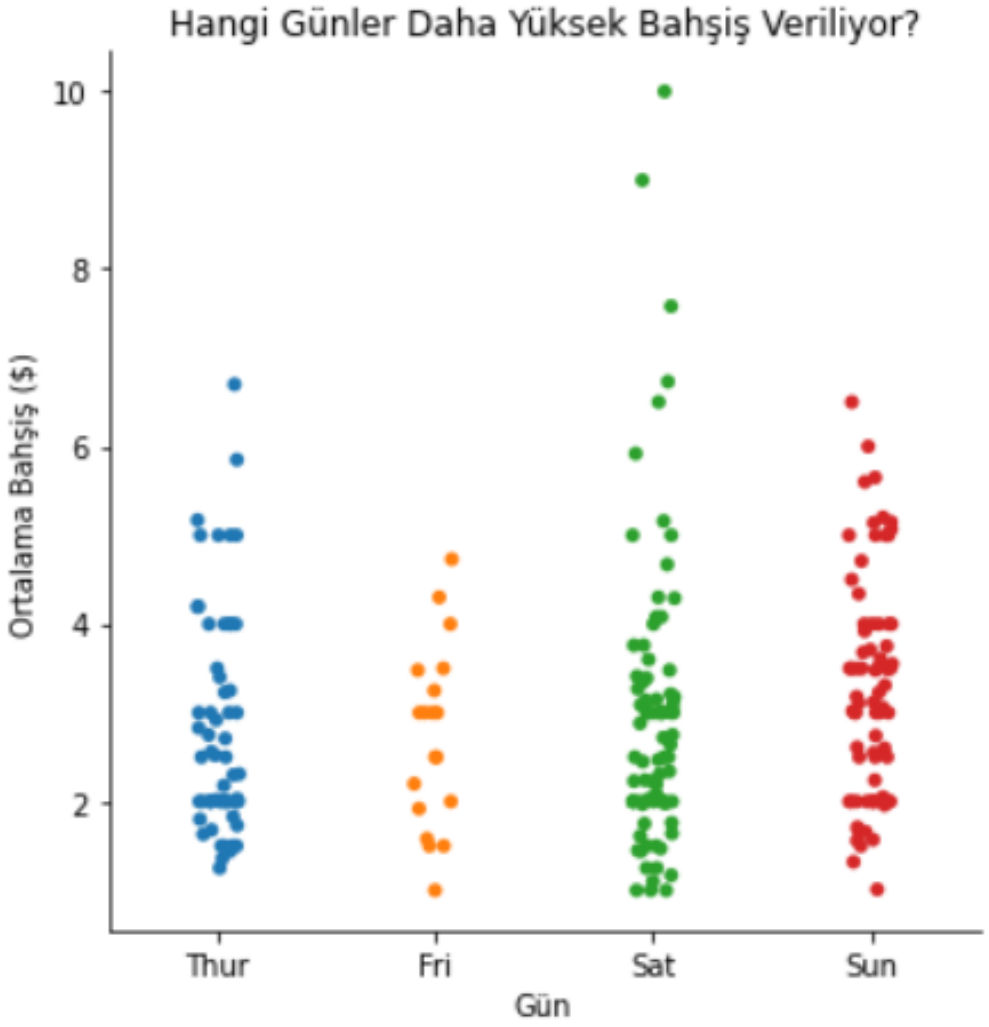
Comparison Plots - catplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Catplot oluşturma (bar plot türünde)
sns.catplot(x='day', y='tip', data=tips)

# Eksen etiketleri ve başlık
plt.xlabel('Gün')
plt.ylabel('Ortalama Bahşiş ($)')
plt.title('Hangi Günler Daha Yüksek Bahşiş Veriliyor?')

# Grafiki göster
plt.show()
```



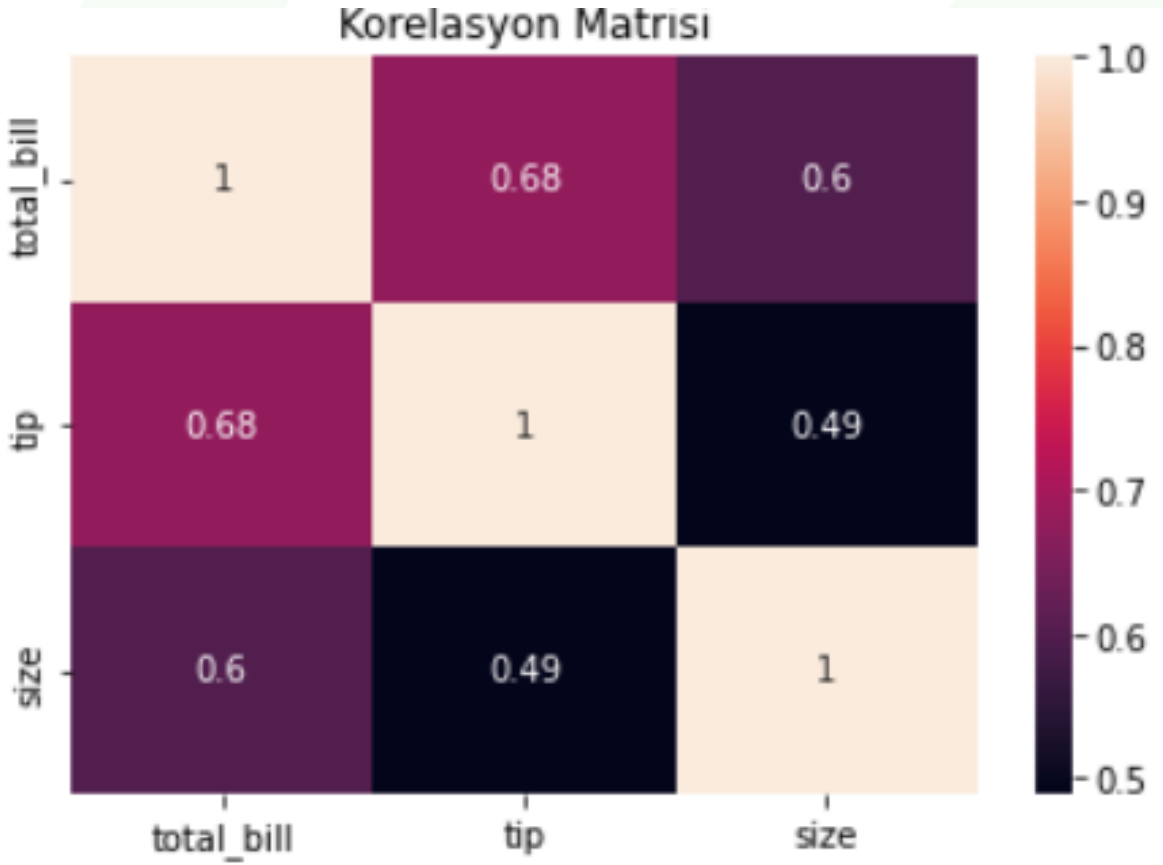
Comparison Plots - matrixplot

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Catplot oluşturma (bar plot türünde)
sns.catplot(x='day', y='tip', data=tips)

# Eksen etiketleri ve başlık
plt.xlabel('Gün')
plt.ylabel('Ortalama Bahşiş ($)')
plt.title('Hangi Günler Daha Yüksek Bahşiş Veriliyor?')

# Grafiki göster
plt.show()
```



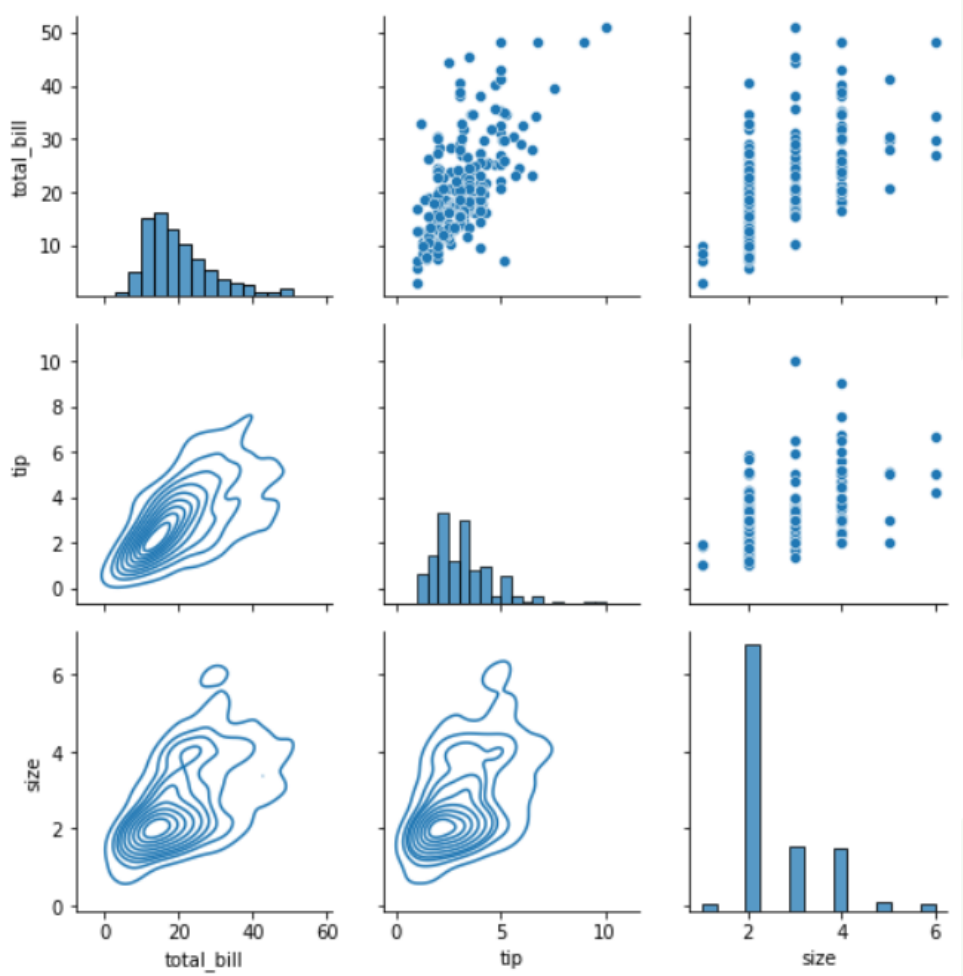
Matrix Plots

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# Korelasyon matrisini hesapla
correlation_matrix = tips.corr()

# Isı haritası oluşturma
sns.heatmap(correlation_matrix, annot=True)

# Eksen etiketleri ve başlık
plt.title('Korelasyon Matrisi')

# Grafiki göster
plt.show()
```



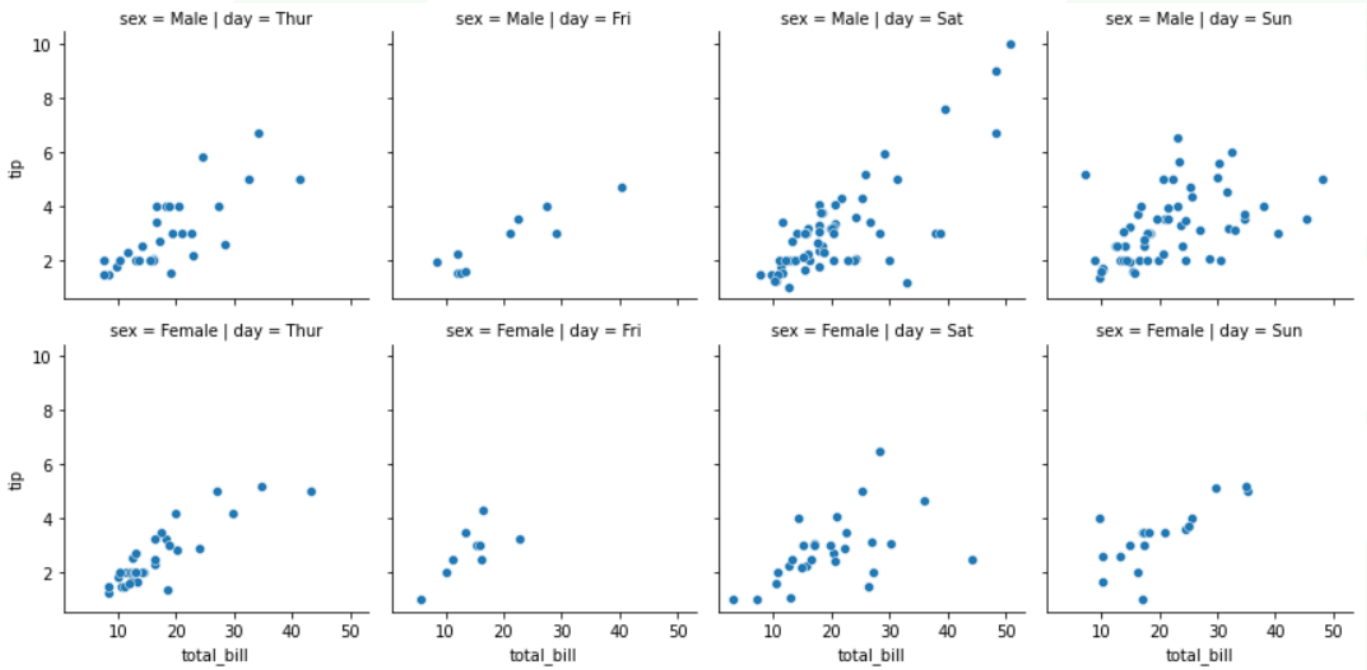
Grids

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# PairGrid oluşturma
g = sns.PairGrid(tips)

# Üst üçgen matrise scatter plot ekleyelim
g.map_upper(sns.scatterplot)

# Köşegen matrise histogram ekleyelim
g.map_diag(sns.histplot)

# Alt üçgen matrise korelasyon katsayısı ekleyelim
g.map_lower(sns.kdeplot)

# Grafiki göster
plt.show()
```



Grids

```python
import seaborn as sns
import matplotlib.pyplot as plt

# Tips veri setini yükle
tips = sns.load_dataset("tips")

# FacetGrid oluşturma
g = sns.FacetGrid(tips, col="day", row="sex")

# Her bir alt grafiğe bir scatter plot ekleyelim
g.map(sns.scatterplot, "total_bill", "tip")

# Grafiki göster
plt.show()
```




Coffee break...
06:50