İçindekiler

[Machine Learning Days 2](#_Toc43076324)

[Data Visualization 2](#_Toc43076325)

[Veri Setinin Hikayesi 3](#_Toc43076326)

[Veri Görselleştirme 6](#_Toc43076327)

[Relational Plots with Matplotlib 6](#_Toc43076328)

[Scatter plot with Subplots 7](#_Toc43076329)

[Categorical Plots with Matplotlib 8](#_Toc43076330)

[Histogram 8](#_Toc43076331)

[Bar Plot 9](#_Toc43076332)

[Figure Kaydetme 10](#_Toc43076333)

[Seaborn 10](#_Toc43076334)

[Count Plot & Cat Plot 11](#_Toc43076335)

[Scatter Plot 12](#_Toc43076336)

[Line Plot 13](#_Toc43076337)

[Scatter Subplots 13](#_Toc43076338)

[Heatmap 16](#_Toc43076339)

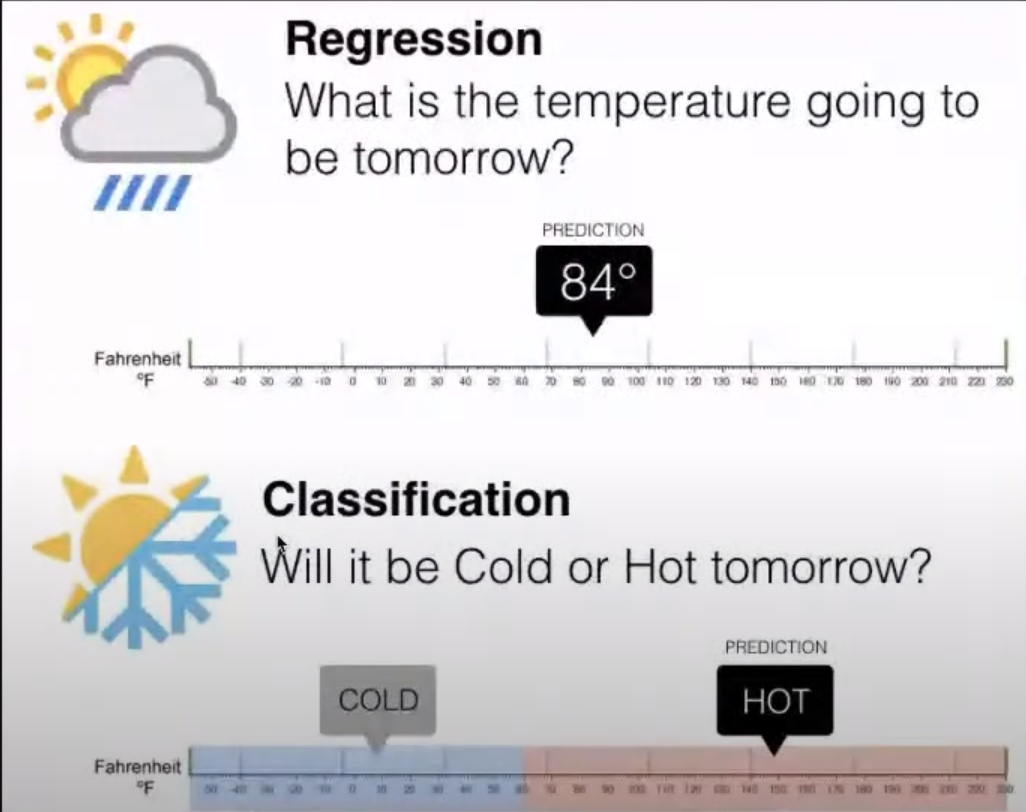
[Categoric Plot 16](#_Toc43076340)

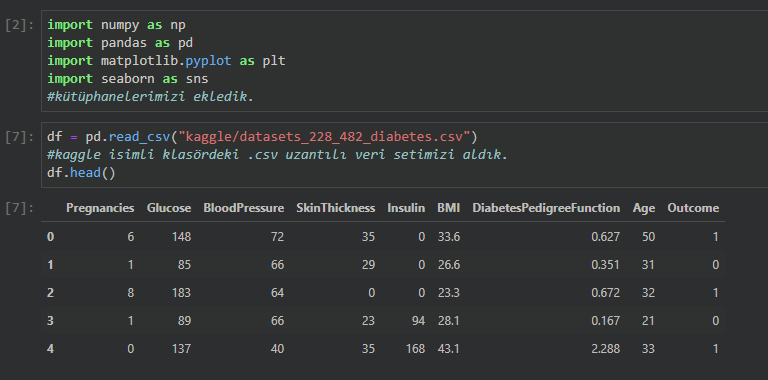
[Box Plot 17](#_Toc43076341)

# Machine Learning Days

# Data Visualization

**Numeric** ve **Categoric** veri tiplerimiz var.  
Kedi-köpek ya da sıcak-soğuk gibi nitel veriler **categoric** verilerdir.  
Eğer categoric verilerle tahminleme yapıyorsak **Classification** problemi çözüyoruz.  
İnsan yaşları gibi numeric verilerle tahminleme yapıyorsak **Regression** problemi çözüyoruz.



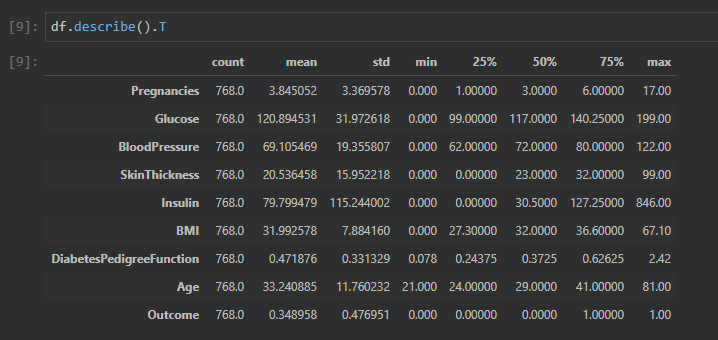


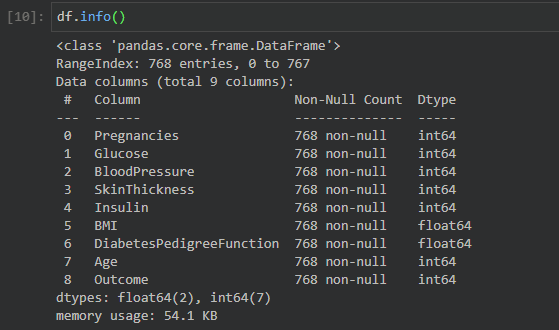
## Veri Setinin Hikayesi

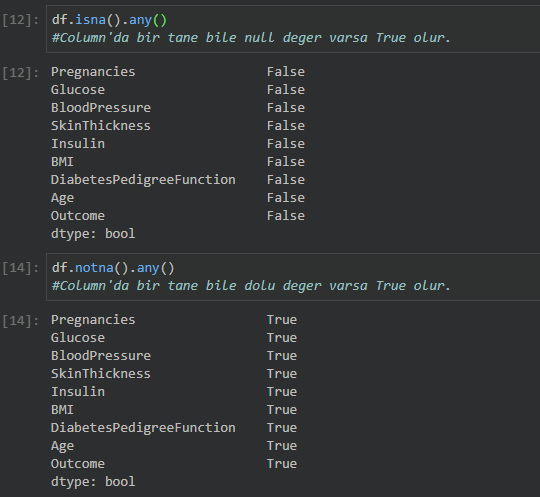
Veri kümesinin amacı, veri kümesine dahil edilen belirli tanı ölçümlerine dayanarak bir hastanın diyabet olup olmadığını teşhis amaçlı olarak tahmin etmektir.  
Veri kümeleri birkaç tıbbi öngörücü değişken ve bir hedef değişkenden oluşur, **Outcome**.  
Tahmin değişkenleri hastanın sahip olduğu gebelik sayısını, BMI'sını, insülin seviyesini, yaşını vb. İçerir.

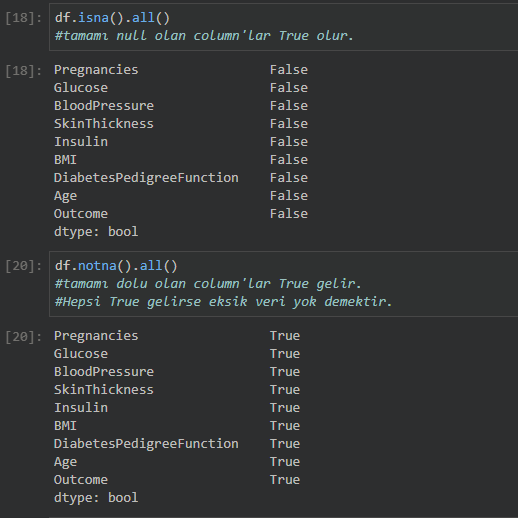
* **Pregnancies**: Hamile sayısı
* **Glucose**: Oral glukoz tolerans testinde 2 saatteki plazma glikoz konsantrasyonu
* **BloodPressure**: Diyastolik kan basıncı (mm Hg)
* **SkinThickness**: Triceps deri kat kalınlığı (mm)
* **Insulin**: 2 saatlik serum insülini (mu U / ml)
* **BMI**: Vücut kitle indeksi (kg olarak ağırlık / (m olarak yükseklik) ^ 2)
* **DiabetesPedigreeFunctio**: Diyabet soyağacı işlevi
* **Age**: Yaş
* **Outcome**: Sonuç (1 yada 0)

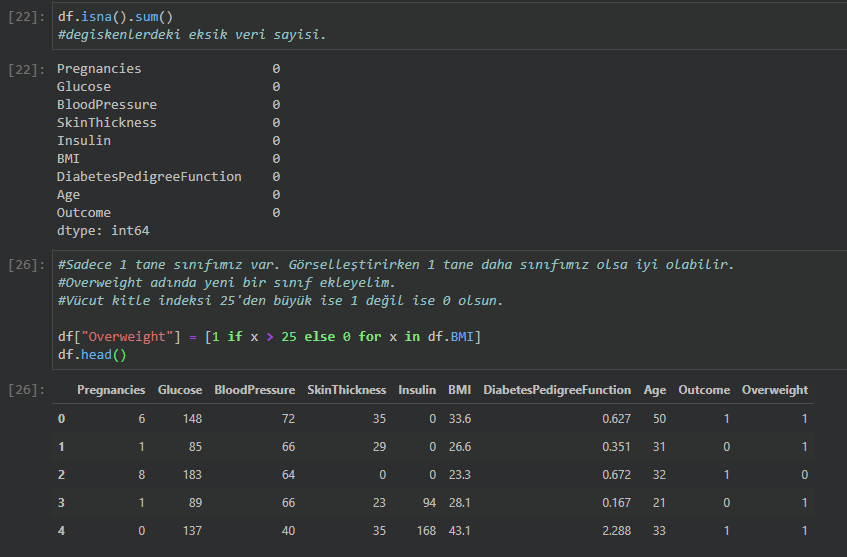
**Outcome** categoric, diğer değişkenler ise numeric veri.









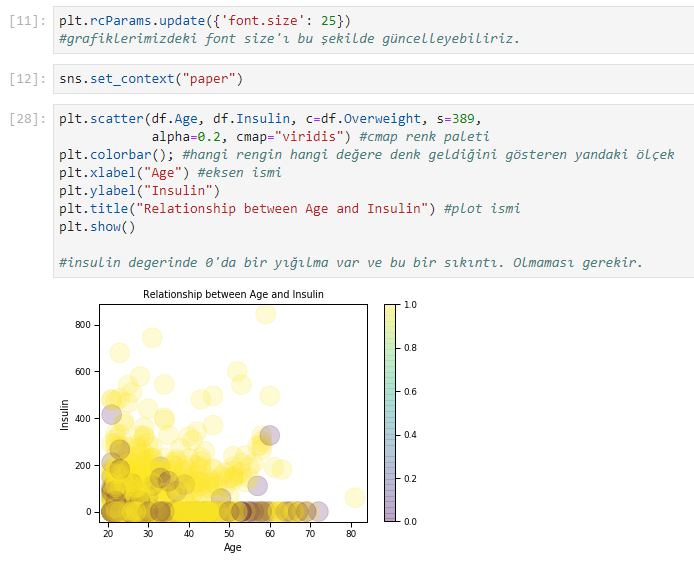


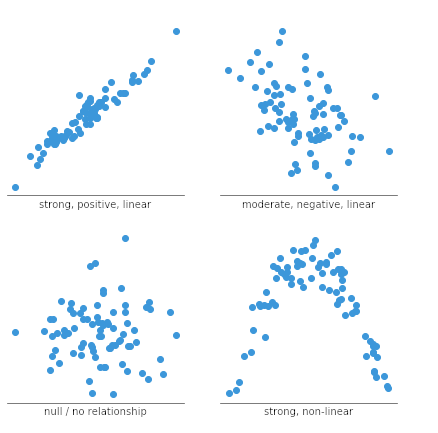
## Veri Görselleştirme

### Relational Plots with Matplotlib

**Relational Plots** iki tane değişkenin arasındaki ilişkiyi gösteren grafiklerdir.

* **Scatter Plot**: İki değişken arasındaki ilişkinin dağılımını veri noktalarıyla gösterir.
* **Lineplot**: İki değişken arasındaki ilişkiyi sürekli gösterir. Veri noktaları birbirine çizgilerle bağlıdır. (Zaman serilerinde kullanılır.)
* **s parametresi:** marker boyutu
* **c parametresi:** marker rengi, hangi değişkeni tuttuğu da yazılabilir.
* **alpha:** marker opaklığı

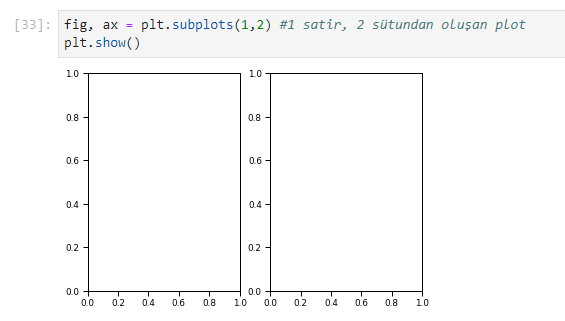


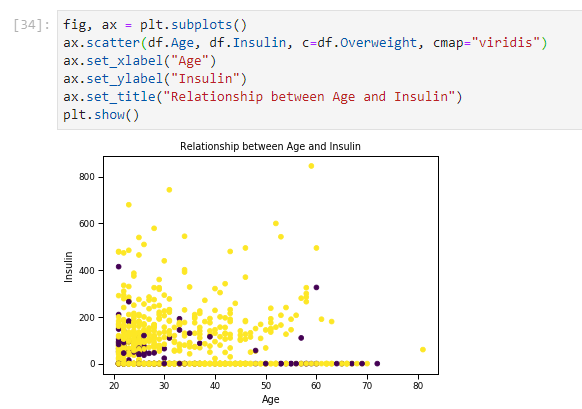


### Scatter plot with Subplots

subplot'ı bir plottan iki tane küçük plot çıkarıyoruz gibi düşünebiliriz.

fig, ax = plt.subplots(): figure ve axes object oluşturur. figure’de her şey var, axes data’yı tutuyor.

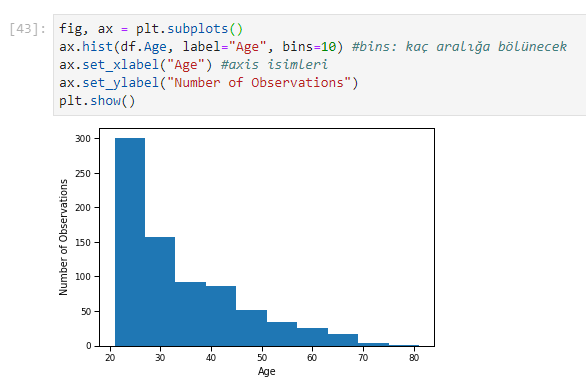


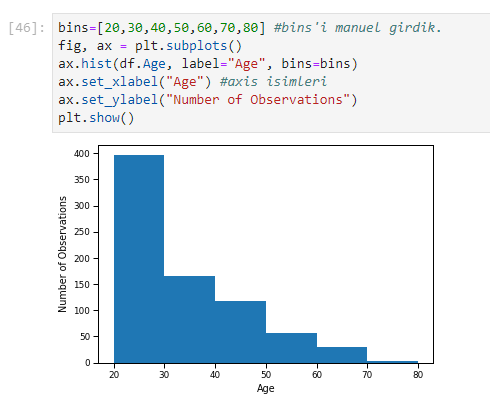


Categorical Plots with Matplotlib

#### Histogram

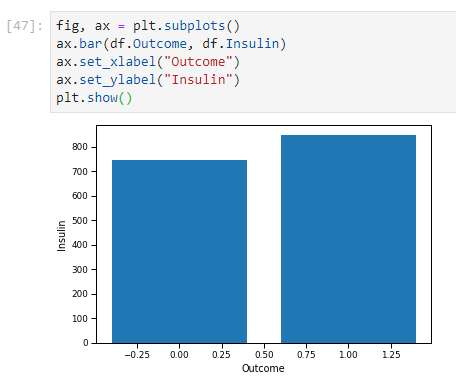
Numerik ya da kategorik verilerde dağılımı yorumlamamıza yardımcı olur.



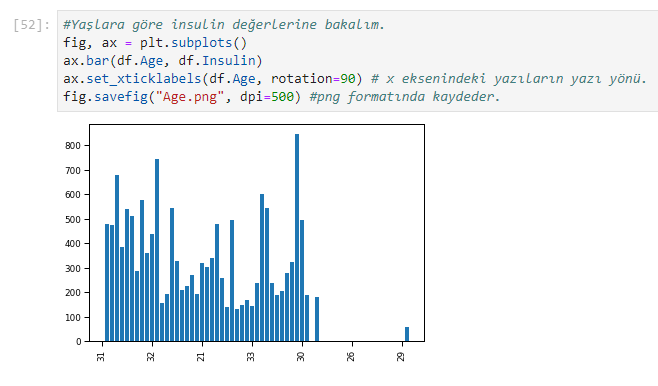


#### Bar Plot

Kategorik verilerin özelliklerine bakmamızı sağlar.



### Figure Kaydetme

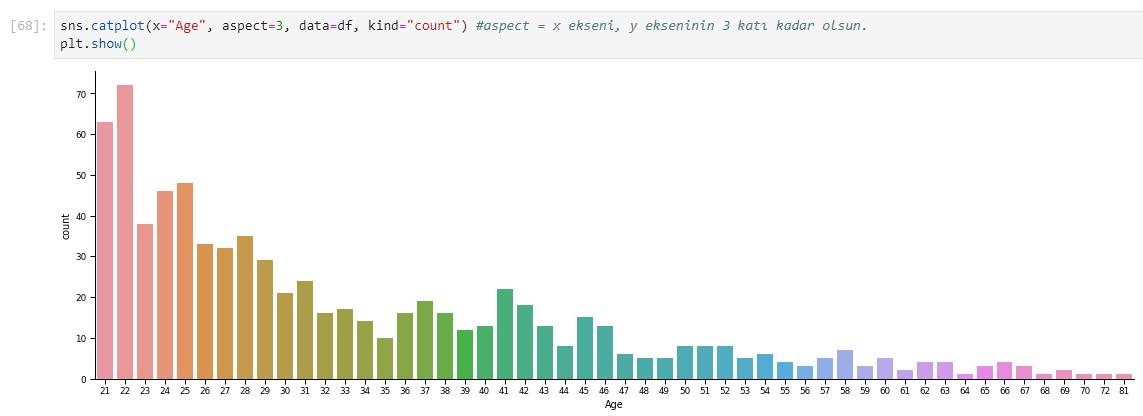
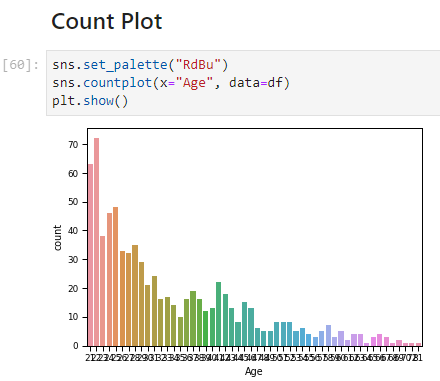


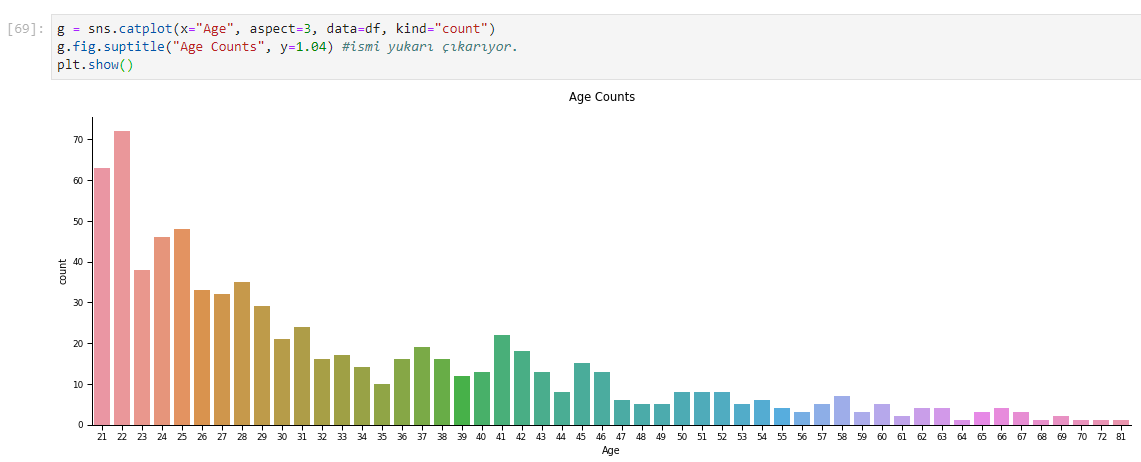
* **fig.savefig(“Age.png”):** kayıp olmadan kaydeder, yüksek kalitelidir ama çok hafıza tutar
* **fig.savefig(“Age.jpg”, quality=50):** websitesine konulabilir
* **fig.savefig(“Age.png”, dpi=200):** dots per inch, dense rendering
* **fig.set\_size\_inches([5,3]):** aspect ratio

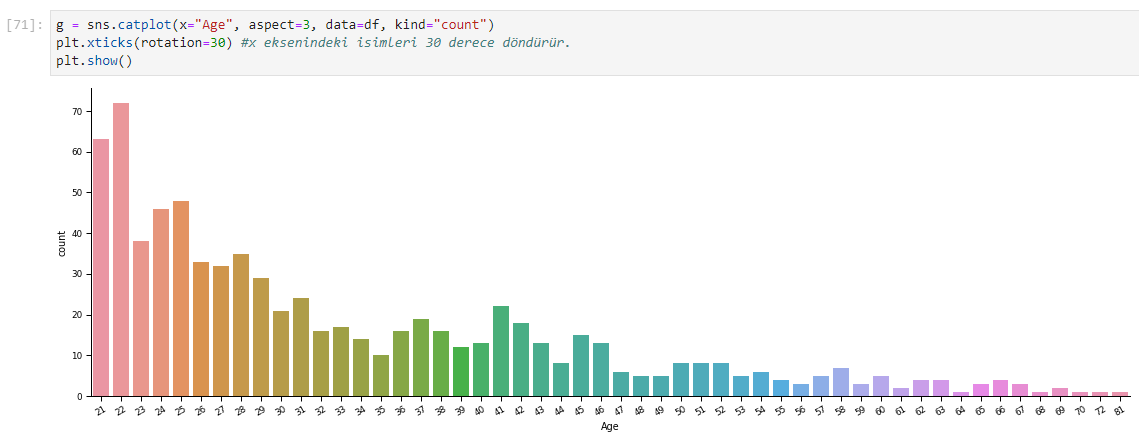
### Seaborn

* **FacetGrid** (relplot(), catplot()) subplot'lar oluşturabilir.
* **AxesSubplot**(scatterplot, countplot) bir tane plot oluşturur.

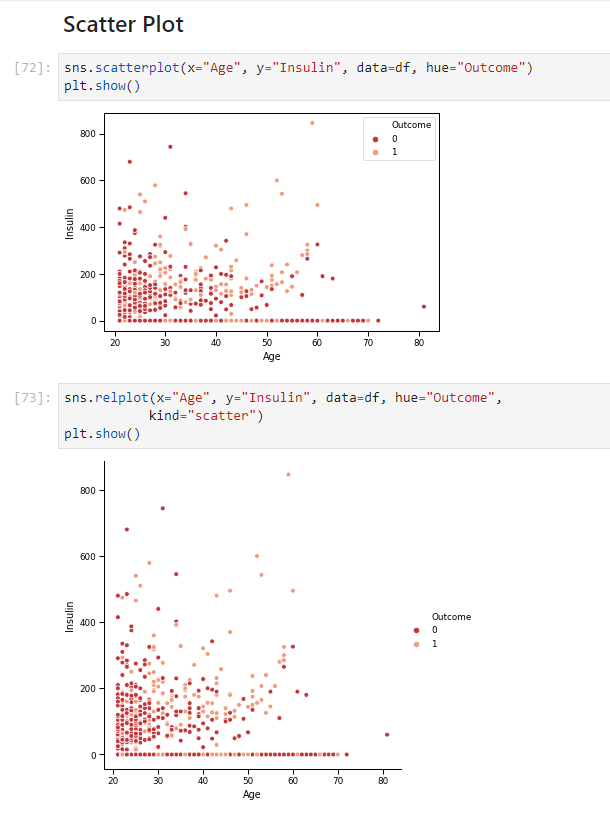
#### Count Plot & Cat Plot



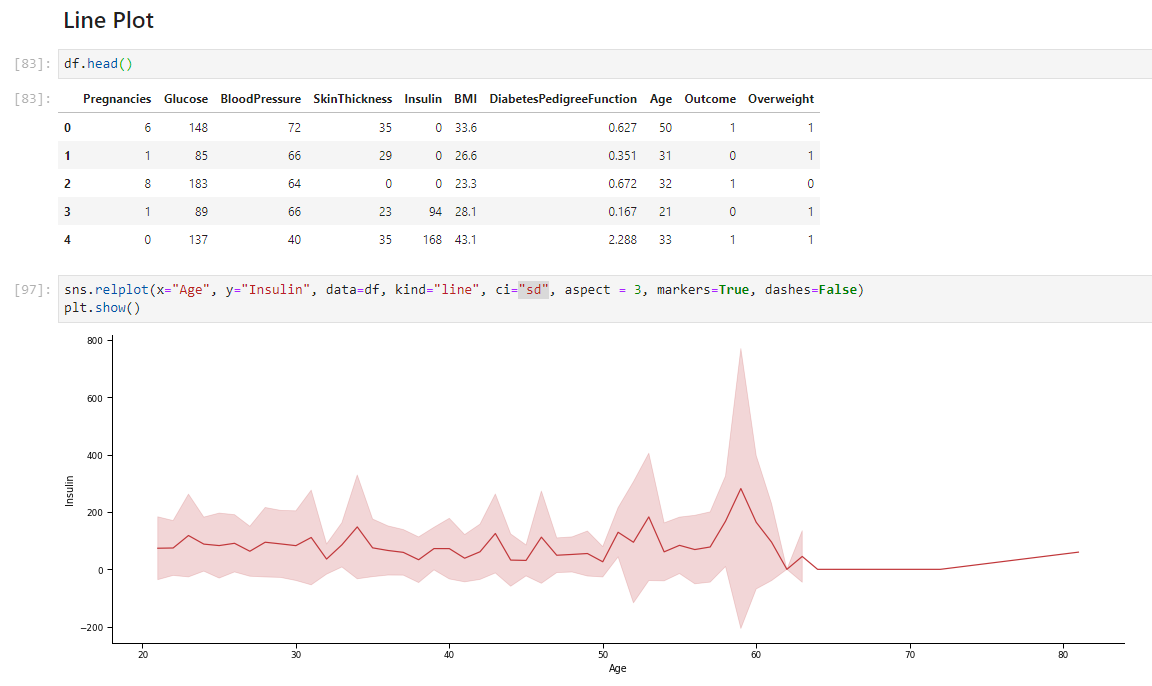


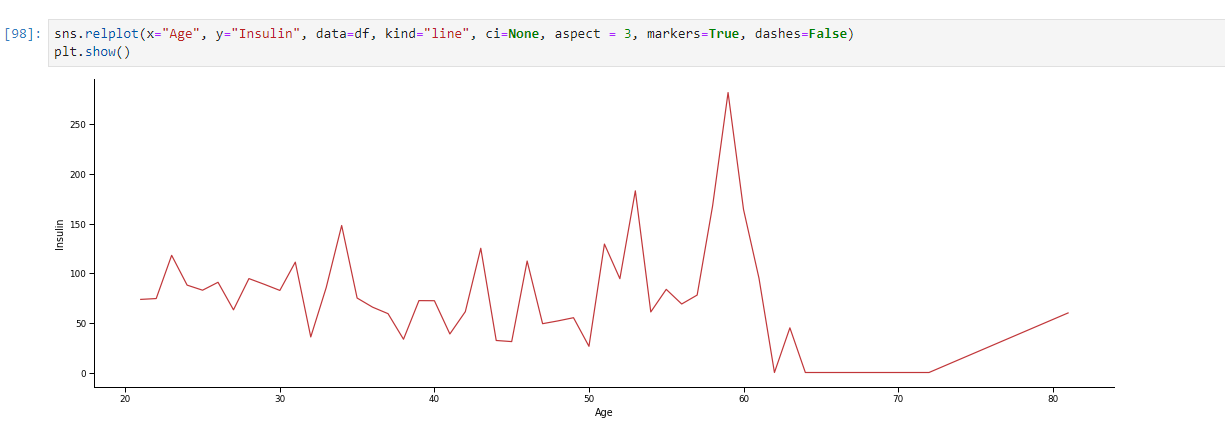


#### Scatter Plot

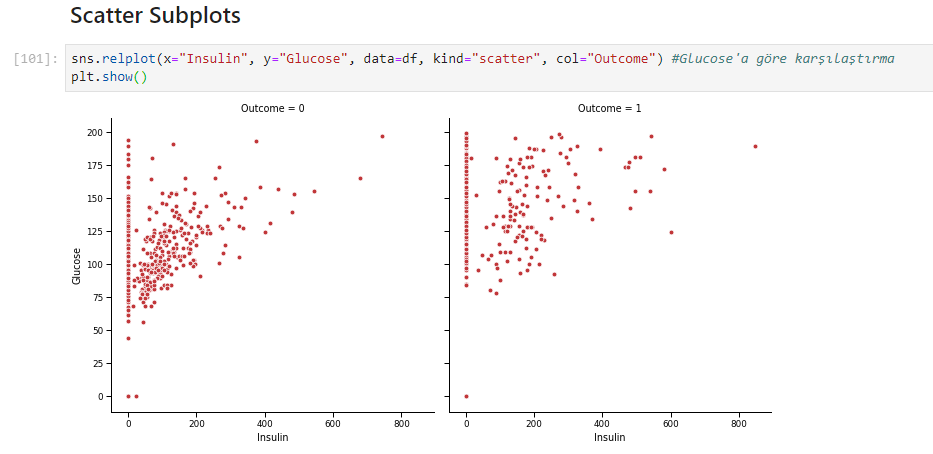


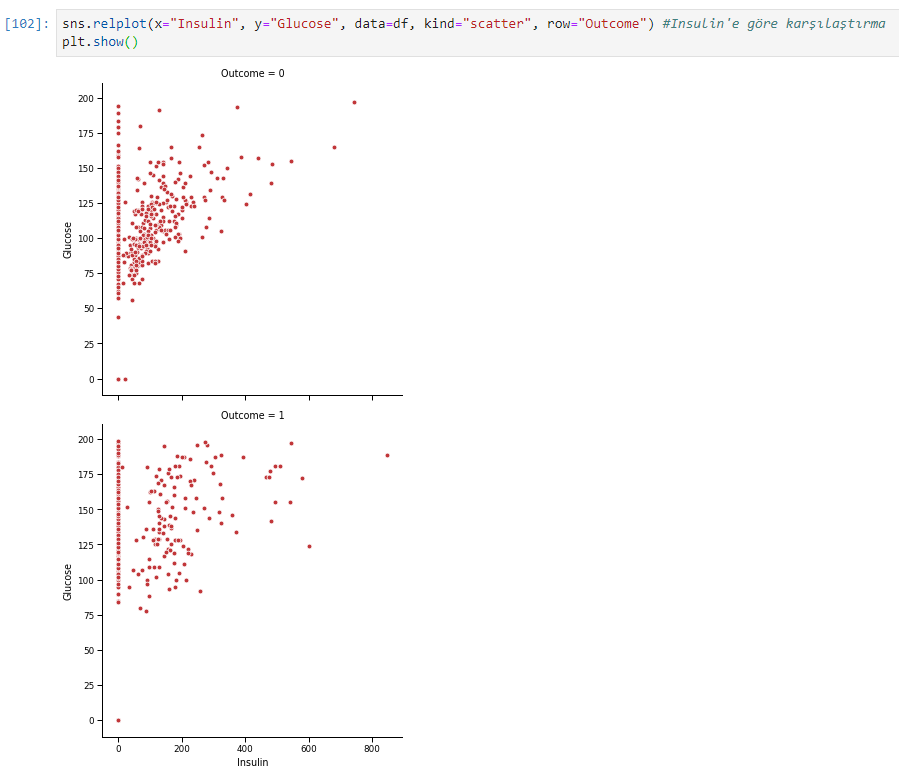
#### Line Plot

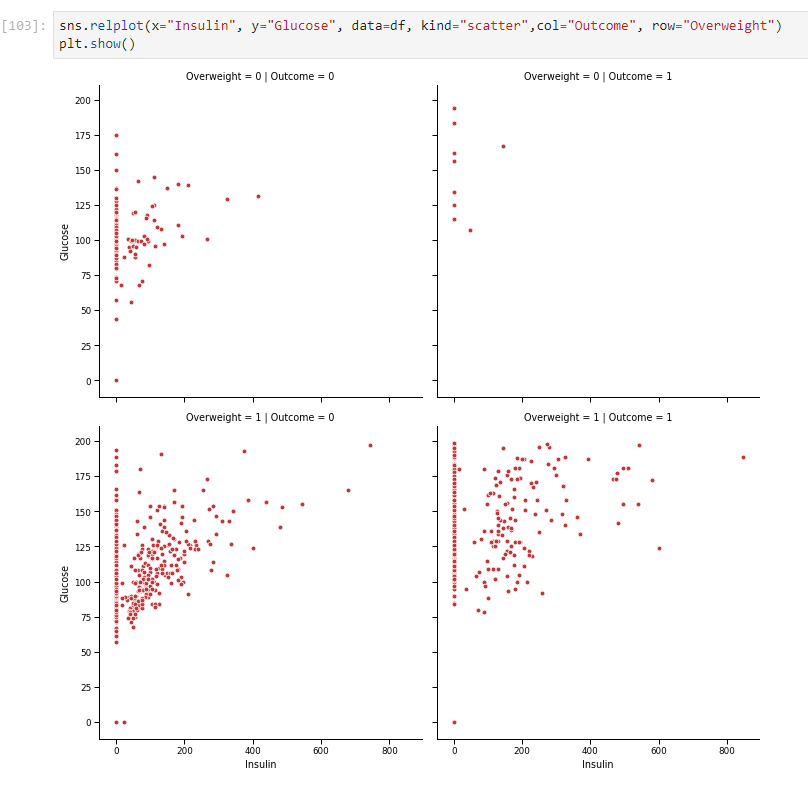




#### Scatter Subplots



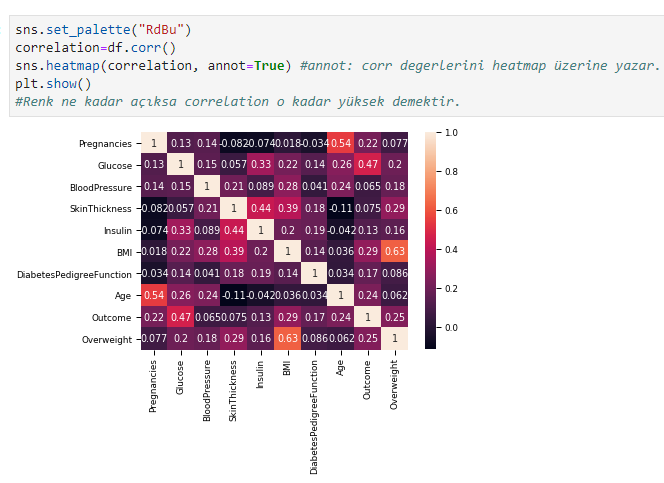




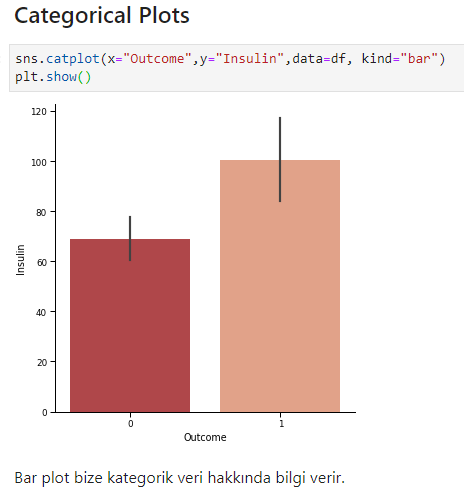
#### Heatmap

Öznitelikler arasındaki ilişkiye, korelasyona bakmamızı sağlar.

Korelasyon ne kadar iyiyse makine öğrenmesi modelimiz o kadar düzgün çalışır.



#### Categoric Plot



#### Box Plot

