Veri Analizi Egitimi (Data Analysis)

# **GÜN - 1**

## **Veri Nedir?**

* Multimedya
* Yapılandırılmış vs. Yapılandırılmamış Veri
* Standart ve Özel Veri Tipleri
* Veri Türleri
* **Uygulama**: Pandas revision + pivot table
* **Uygulama**: Pandas ile temel veri okuma, pd.read\_csv()

## **Veri Görselleştirme**

* Görselleştirme Nedir?
* Tufte’nin 6 Prensibi
* Grafik Türleri
  + Line Plot
  + Scatter Plot
  + Bar Plot
  + Histogram
  + Pie Charts
  + Heatmap
  + Box Plot
  + Kartil Nedir? Nasıl Hesaplanır?
  + Joint Plot
  + KDE(Kernel Density Estimate)
* **Uygulama**: seaborn ile lineplot ve distplotun gosterilmesi

## **Temel İstatistik**

* Betimsel İstatistik Kavramları
* Çarpıklık Kavramı
* Korelasyon ve Korelasyon Matrisi
* Simpsons Paradox
* Anscombe Quartet
* Veri Dağılımı ve Hipotez Testleri
* **Uygulama**: pd.describe

## **Veri Dağılımı**

* Veri ve Dağılım
* Gaussian(Normal) Dağılım
* t-Dağılımı
* Serbestlik Derecesi
* Bernoulli’s Dağılımı
* Exponential Dağılım

# **GÜN - 2**

## **Temel Hipotez Testleri**

* Temel Hipotez testleri
* P değeri
* T testi
* Z testi
* Ki-kare(Chi-Square) Testi
* Hipotez Testinde Hatalar

## **Veri Temizleme**

* 68-95-99.7 Kuralı ve 3 Sigma
* Aykırı (Outlier), Eksik (Missing) ve Duplike (Duplicate) Veriler ve Tespiti
* Z-Skoru
* Eksik değerlerin ele alınması
* Null vs NaN
* Eksik değerler için Pandas Fonksiyonları
* Boyut İndirgeme
* PCA(Principal Component Analysis)
* Collinearity(Çoklu Doğrusal Bağlantı)
* **Uygulama**: pd.corr fakat bagimsiz degiskenler arasindaki analiz gosterilecek

## 

## **Veri Dönüşümü**

* Veri Dönüşüm Teknikleri
  + Round
  + Scaling(Ölçekleme)
  + Label Encoding
  + One Hot Encoding
  + Stack
  + Melt
  + Sort
  + Feature Engineering
* **Uygulama**:astype
* **Uygulama**:round vs
* **Uygulama**:sklearn scalers
* **Uygulama**: pd.get\_dummies, sklearn.LabelEncoder,
* **Uygulama**: sort

## **Veri Azaltma**

* Data aggregation
* Matematiksel Fonksiyonlar ve Aggregation Çeşitleri

# **GÜN - 3 (ML + Proje)**

## **ML Projesi**

* Makine Öğrenmesi
* Veri Analizi ve Makine Öğrenmesinde Kullanılan Araçlar
* Uçtan Uca Makine Öğrenmesi Projesi

## **Pandas Uygulamali Bir Senaryo Seklinde Gerçekleşecek (Proje)**

* Giris: Veri Okunur
* Gorsellestirilir
* Temel Analizi Yapilir
* Eksik Kismi Tamamlanir
* Fazla Kismi Cikarilir
* Korelasyona Bakilir
* Train Test Split
* Modele Verilir
* Ciktisina Bakilir
* Cikis: Geri Yazilir

Things to Consider:

* Day-1 sunumunda boyut indirgeme, pca ve collinearity’nin yerini değiştirsek mi?
* Veriyi iyi tanimak analizin baslangicidir
* Hedefin iyi ayarlanmasi ve ciktinin yani objectivein basta iyi belirlenmesi, sonraki sureclerin tamaminin son derece etkilediginin gosterilmesi gerekiyor.