

DERS İZLENCESİ (SYLLABUS)

VERİ BİLİMİNE GİRİŞ

Dersin Adı: Veri Bilimine Giriş

Öğretim Elemanı: Dr. Ömer Gökdaş

Dönem: 2025 / 2026 (Bahar)

Haftalık Ders Saati: 11.30 – 14.30 (Çarşamba)

Ders Dili: Türkçe

1. Dersin Tanımı

Bu ders, veri bilimine uygulamalı bir giriş sunar. Amaç; Python ekosistemini kullanarak veriyi okuma, temizleme, görselleştirme, keşifçi analiz (EDA) yapma ve temel makine öğrenmesi iş akışını (train/test, leakage, metrikler) kavratmaktadır. Ders uygulamaları ağırlıklı olarak Google Colab ortamında yürütülür.

2. Öğrenme Çıktıları

- Python ile NumPy/Pandas kullanarak veri setlerini yükler, dönüştürür ve analiz eder.
- Eksik veri (NaN), grup işlemleri (groupby/pivot), birleştirme (merge) gibi veri hazırlama adımlarını uygular.
- Matplotlib/Seaborn ile temel ve istatistiksel görselleştirmeler üretir, grafik yorumlar.
- Kaggle veri setleri ile EDA yapar ve bulguları raporlar.
- scikit-learn ile train/test ayrimı, leakage kavramı, regresyon ve sınıflandırma için temel metrikleri uygular.
- SQL (SQLite) ve temel API veri çekme yaklaşımını veri bilimi iş akışına bağlar.
- Uçtan uca mini proje akışını yönetir (Veri → EDA → Görselleştirme → Modelleme → Sunum).
-

3. Haftalık Ders Planı

Hafta	Konu Başlığı	İçerik ve Uygulama (özet)
1	Veri Bilimine Giriş, Python Temelleri (Colab)	Veri bilimi süreci, ders akışı, Colab ile çalışma düzeni, temel Python yapı taşları
2	NumPy + temel istatistik + lineer cebir	Array mantığı, vektör/matrís işlemleri, temel istatistik (ortalama-std), lineer cebire giriş
3	Pandas-1 (okuma, filtreleme, select)	CSV okuma, DataFrame/Series, select/loc/iloc, temel filtreleme
4	Pandas-2 (NaN + groupby/pivot + merge)	Eksik veri stratejileri, groupby, pivot, merge/join ile tablo birleştirme

5	Görselleştirme (Matplotlib, Seaborn)	Çizgi/sütun/dağılım grafikleri, temel grafik okuryazarlığı, doğru yorum-pratik
6	Kaggle Tanıtım, EDA (Pandas + Görselleştirme)	Kaggle üzerinden veri seti seçimi, EDA akışı, ilk bulgular
7	SQL + API	SQLite ile temel SQL; API'den veri çekmeye giriş
8	Vize	Kod okuma, çıktı yorumlama, veri manipülasyonu ve temel kavramlar
9	ML mantığı + leakage + train/test (scikit-learn)	Gözetimli öğrenme mantığı, doğru ayırım, leakage örnekleri, temel pipeline düşüncesi
10	Linear Regression + metrik (scikit-learn)	Regresyon modeli, değerlendirme metrikleri, yorumlama ve hata analizi
11	Classification + F1/Confusion (scikit-learn)	Sınıflandırma mantığı, confusion matrix, precision/recall/F1, karar eşikleri
12	Kaggle yarışması	Mini challenge akışı: veri → EDA → baseline model → raporlama
13–14	Proje sunumları	Proje çıktılarının sunumu: problem tanımı, EDA, görseller, modelleme, sonuçlar

4. Yazılım ve Araçlar

- Platform: Google Colab (önerilen) / Jupyter Notebook
- Python kütüphaneleri: NumPy, Pandas, Matplotlib, Seaborn, scikit-learn
- Veri kaynakları: Kaggle veri setleri, simülasyon/ornek veri setleri
- Veri tabanı: SQLite (SQL temelleri)
- API: Temel düzeyde veri çekme ve Pandas'a aktarma yaklaşımı

5. Sınav ve Değerlendirme Politikası

- Vize: Kod yazdırılmaktan çok kod okuma, çıktı yorumlama, kavram ve iş akışı değerlendirmesi
- Dönem projesi: Veri seti seçimi (Kaggle vb.), EDA + görselleştirme + temel modelleme + sunum

6. Kaynaklar

- Eğitmen ders notları (Jupyter Notebooks): <https://github.com/ogokdas/Veri-Bilimine-Giris>
- W3Schools: <https://www.w3schools.com/>
- Kaggle Datasets: <https://www.kaggle.com/datasets>