



# KNIME ve Python Uygulamaları

# içindekiler

- KNIME ve Python Ortamı
- Veri Bilimi Yöntemleri
- Veriyi Tanıma ve Hazırlama
- Veri İşleme (ETL)
- Model Oluşturma (ML)
- Başarı Değerlendirme
- Gerçek Hayat Uygulamaları
- Kapanış & Teşekkürler

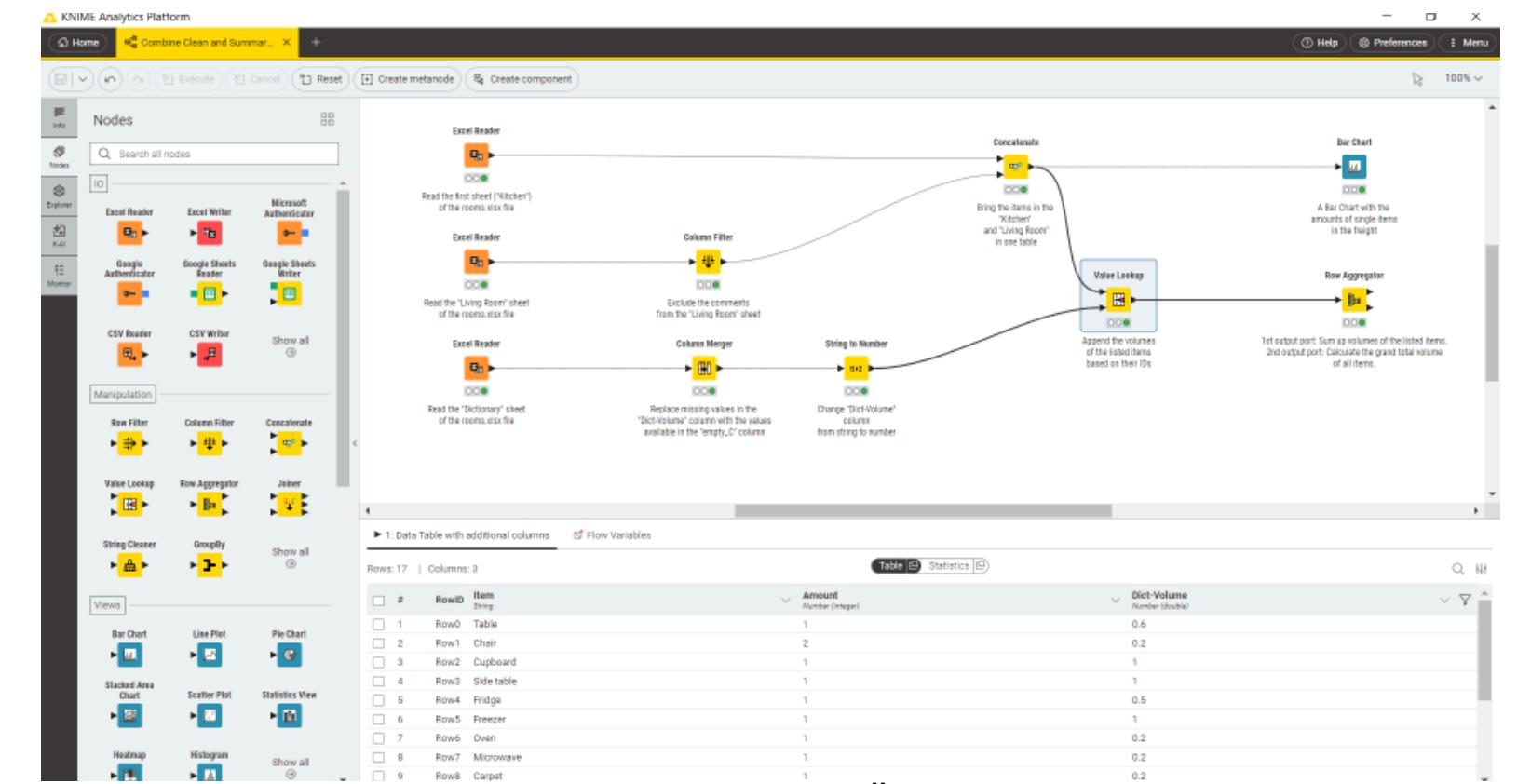
## KNIME

- Sürükle–bırak ile hızlı workflow oluşturma
- Görsellik, kolay kullanım ve otomasyon

## Python

- Esnek dil yapısı
- Geniş kütüphane ekosistemi (NumPy, Pandas, Scikit-learn)
- İleri seviye analiz imkanı

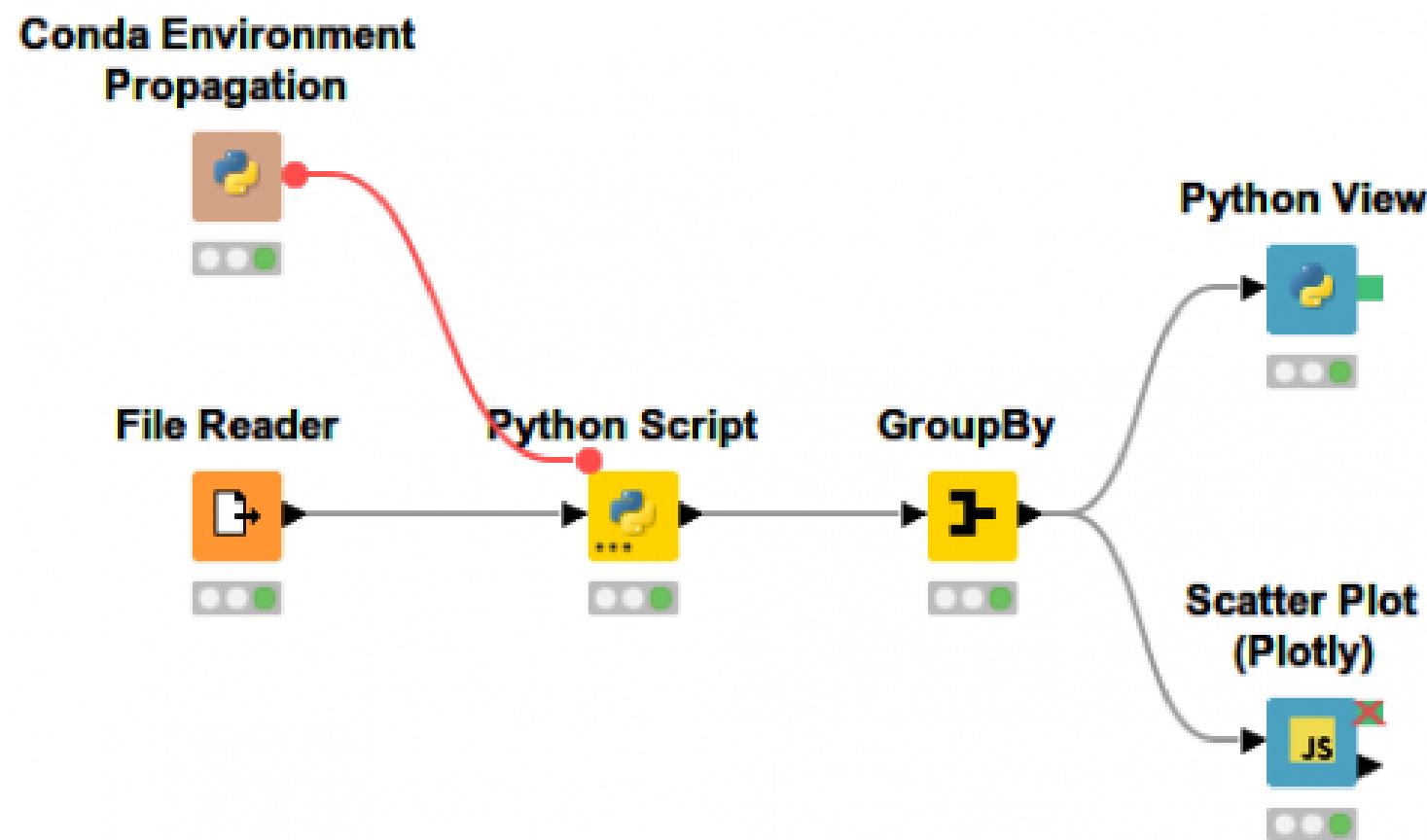
Veri bilimi sürecinde KNIME görselliği sağlarken, Python analitik gücü ekler.



KNIME Workflow Örneği

# KNIME VE PYTHON'UN BİRLİKTE KULLANIMI

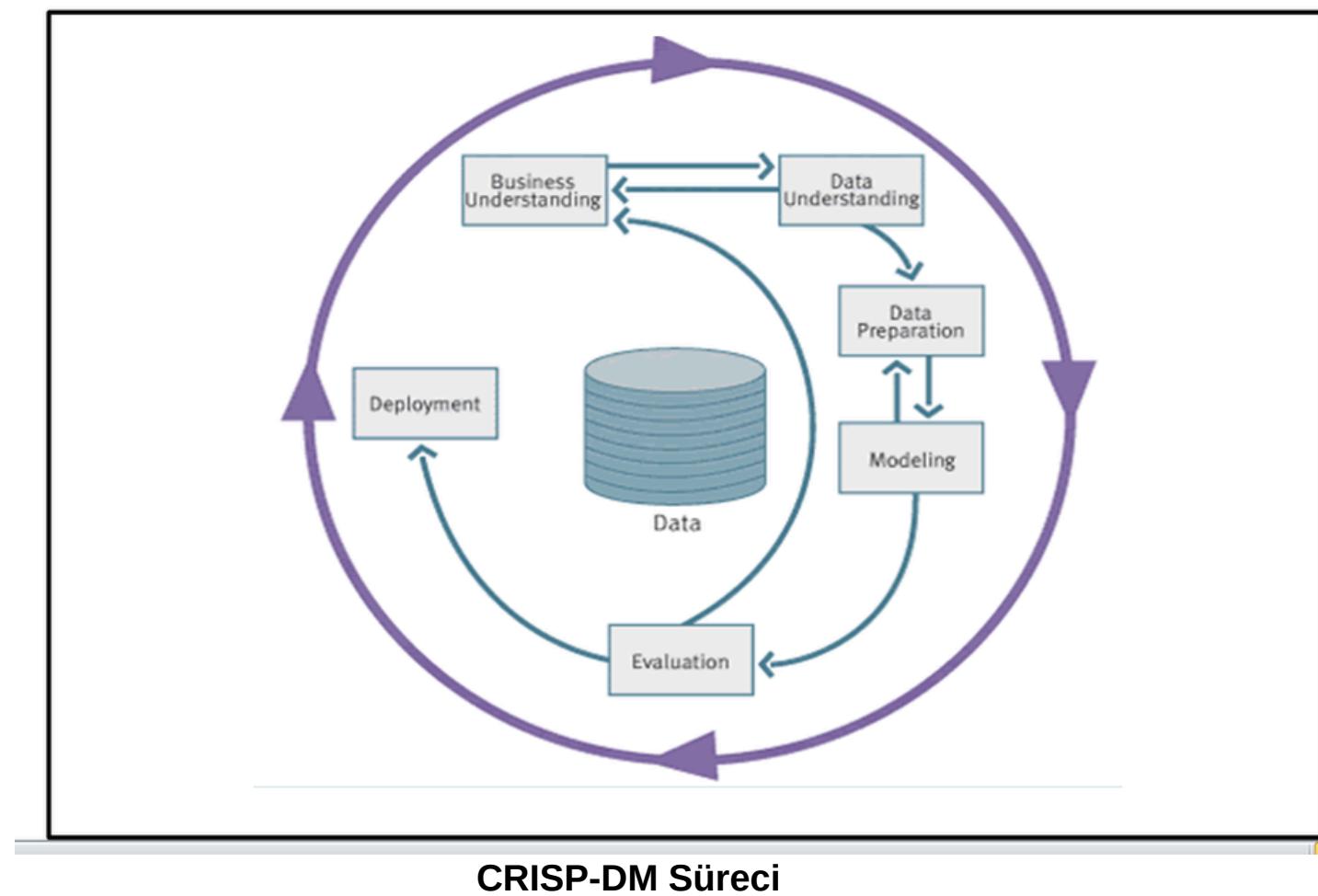
- KNIME → Python Script node ile doğrudan Python kodu çalıştırılabilir
- Büyük veri hazırlıkları KNIME'da, özel algoritmalar Python'da yapılabilir
- Python: Kod esnekliği, KNIME: Görsel raporlama & otomasyon



KNIME – Python Script Node Örneği

# VERİ BİLİMİ YÖNTEMLERİ

- **SEMMA:** Örnekleme (Sample), Keşfetme (Explore), Değiştirme (Modify), Modelleme (Model), Değerlendirme (Assess)
- **CRISP-DM:** Veri madenciliği için yaygın süreç modeli (iş anlayışı → veri hazırlama → modelleme → değerlendirme → dağıtım)
- **KDD:** Knowledge Discovery in Databases → Büyük veri tabanlarından anlamlı bilgi keşfi süreci



# VERİYİ TANIMA VE HAZIRLAMA

## Veriyi Tanıma:

- Veri tiplerinin belirlenmesi (sayısal, kategorik, metin)
- Eksik veya hatalı değerlerin fark edilmesi
- Python: df.info(), df.describe()
- KNIME: Tablo okuyucu, veri önizleme node'ları

## Veriyi Hazırlama:

- Eksik değerleri tamamlama veya çıkarma
- Normalizasyon / standardizasyon
- Kategorik değişkenleri kodlama (One-Hot Encoding)
- Python: pandas, scikit-learn preprocessing
- KNIME: Missing Value, Normalizer node

# Veri İşleme (ETL Süreci)

ETL: Extract (Çıkarma) → Transform (Dönüştürme) → Load (Yükleme)

Extract: Farklı kaynaklardan (CSV, SQL, API) veri çekme

Transform: Veriyi temizleme, dönüştürme, hesaplama ve uygun formata getirme

Load: Veriyi hedef sisteme veya modele yükleme

Python: pandas, SQLAlchemy, pySpark

KNIME: Database Reader, Column Filter, Joiner, ETL node'ları

# Model Oluşturma (ML)

**Amaç:** Veriden öğrenen modeller kurarak tahmin veya sınıflandırma yapmak

**Yöntemler:**

- Regresyon: Sürekli değer tahmini (ör. fiyat)
- Sınıflandırma: Kategorik değer tahmini (ör. spam/ham mail)
- Kümeleme: Benzer grupları keşfetme (ör. müşteri segmentasyonu)

**Python:** scikit-learn, TensorFlow, PyTorch

**KNIME:** Decision Tree, Logistic Regression, K-Means node'ları

```
1 # Basit bir sınıflandırma modeli (Logistic Regression)
2
3 from sklearn.model_selection import train_test_split
4 from sklearn.linear_model import LogisticRegression
5 from sklearn.metrics import accuracy_score
6 import pandas as pd
7
8 from sklearn.datasets import load_iris
9 data = load_iris()
10 X = pd.DataFrame(data.data, columns=data.feature_names)
11 y = data.target
12
13 # Eğitim ve test seti ayırma
14 X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
15
16 # Model oluşturma ve eğitme
17 model = LogisticRegression(max_iter=200)
18 model.fit(X_train, y_train)
19
20 # Tahmin ve doğruluk
21 y_pred = model.predict(X_test)
22 print("Doğruluk Oranı:", accuracy_score(y_test, y_pred))
23
```

# Başarı Değerlendirme (Evaluation)

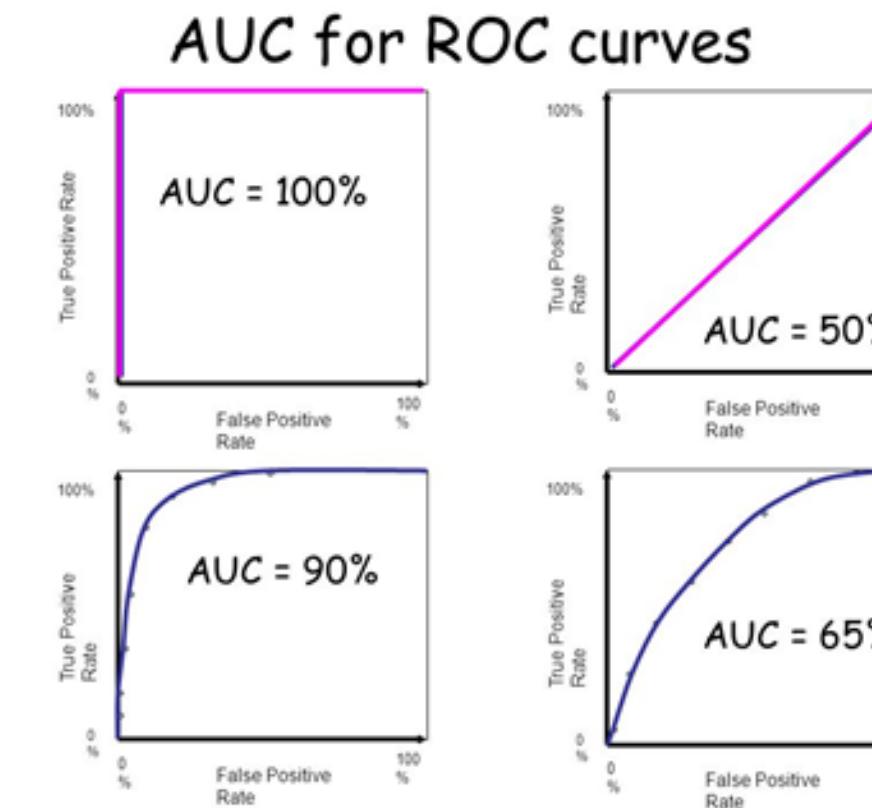
**Amaç:** Modelin ne kadar doğru çalıştığını ölçmek

## Yöntemler:

- Doğruluk (Accuracy): Tüm tahminlerde doğru olanların oranı
- Kesinlik (Precision): Pozitif tahminlerin gerçekten pozitif olma oranı
- Duyarlılık (Recall): Gerçek pozitiflerin ne kadarını yakalayabildiğimiz
- F1 Skoru: Precision ve Recall'un dengeli ortalaması
- ROC-AUC: Modelin ayırcı gücünü ölçer

**Python:** accuracy\_score, classification\_report, roc\_auc\_score

**KNIME:** Scorer, ROC Curve, Confusion Matrix node'ları



# Gerçek Hayat Uygulamaları

## Finans:

- Kredi risk skorlama
- Dolandırıcılık tespiti

## Sağlık:

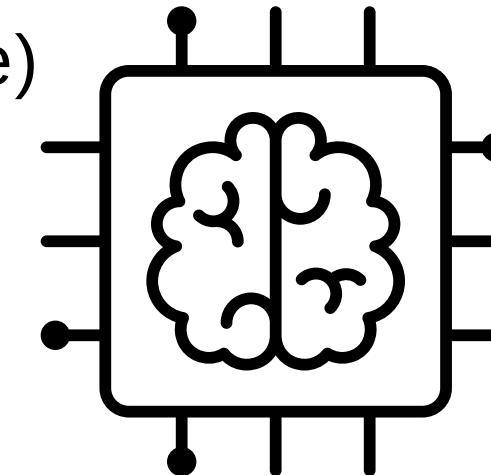
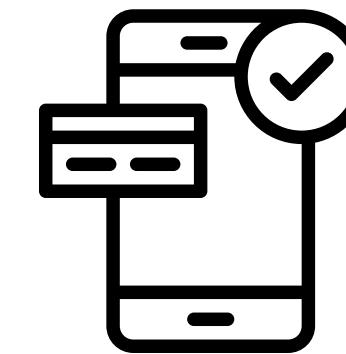
- Hastalık tahmini
- Görüntü işleme ile teşhis

## E-Ticaret:

- Tavsiye sistemleri (ürün önerileri)
- Müşteri segmentasyonu

## Endüstri:

- Tahmine dayalı bakım (Predictive Maintenance)
- Sensör verisi analizi





Teşekkürler