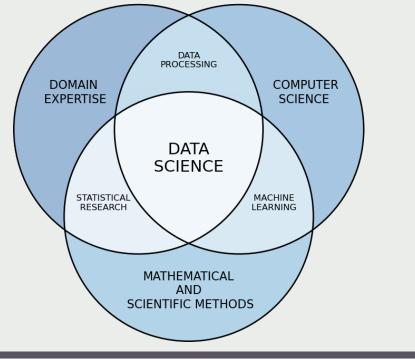
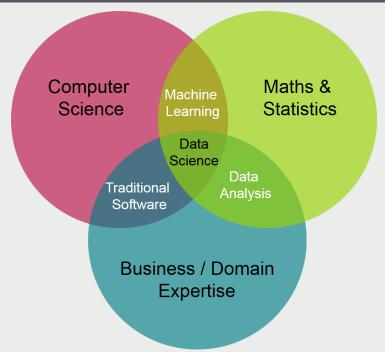
ISE 311 VERİ BİLİMİ

Dr. Öğretim Üyesi Tuğrul TAŞCI

Literatür, Temel Tanım ve Kavramlar

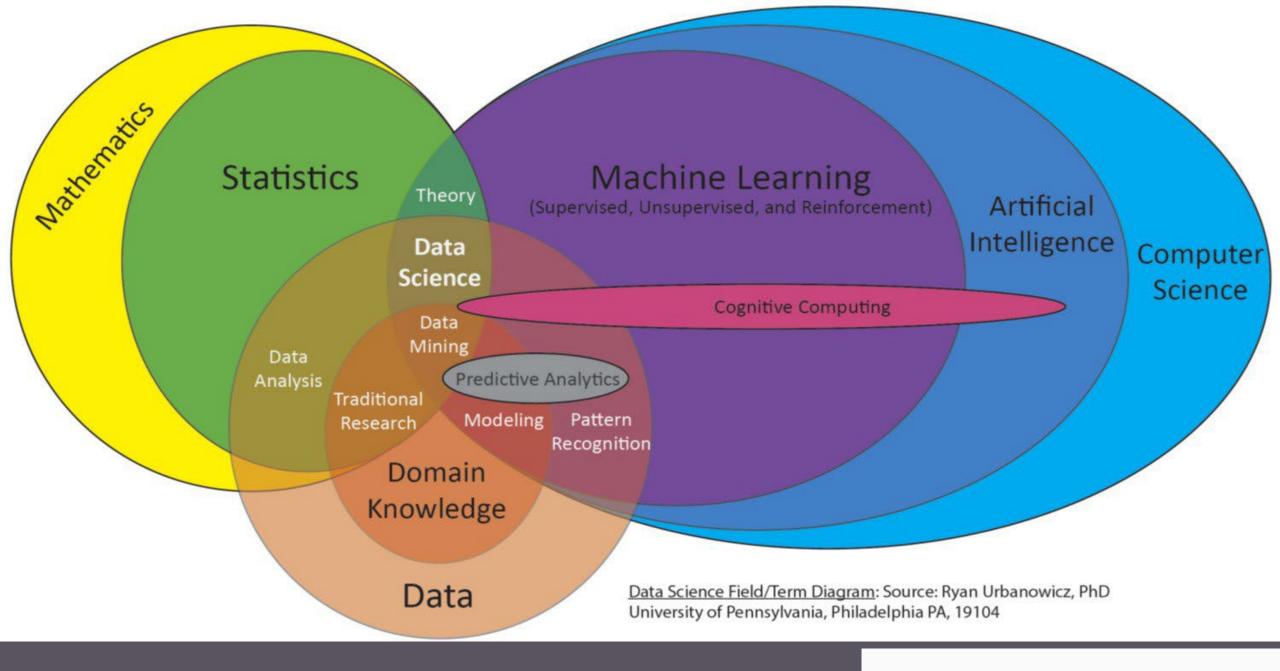


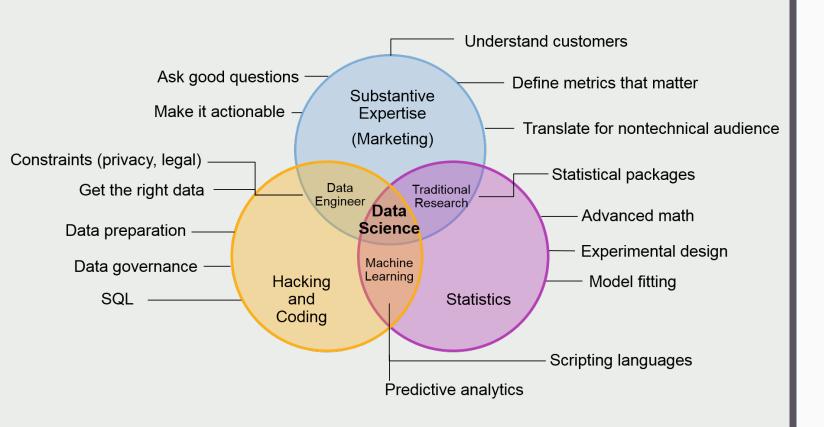


Veri Bilimi Nedir?

Faaliyet alanına ait bilgi ve tecrübe ile bilgisayar bilimleri, matematik ve istatistiğe ait yöntem ve araçlara dayalı olarak kurumsal veri kaynakları üzerinden anlamlı bilgilerin çıkarılması yoluyla iş süreçleri ile ilgili karar alma ve stratejik planlama yapmayı sağlayan disiplinler arası bir bilim ve çalışma alanıdır.

10/10/2022





Veri Bilimi – Alt Disiplinler

Alt Disiplin & Yürütülen Süreçler

Artificial Intelligence (AI)

Human Intelligence Exhibited by Machines

Amazon purchase prediction Smart Email Categorization

Machine Learning (ML)

An Approach to Achieve Artificial Intelligence

Google Maps speed of traffic Facebook facial recognition

Netflix video recommendation

Deep Learning (DL)

A Technique for Implementing Machine Learning

Self-Driving Cars

Speech Recognition Robotics

Data Science

Scientific methods, algorithms and systems to extract knowledge or insights from big data

Data Analysis

Process of inspecting, cleansing, transforming and modeling data

Data Analytics

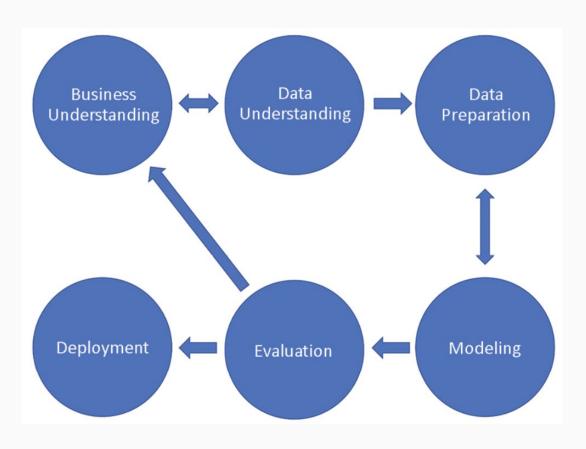
Discovery, interpretation, and communication of meaningful patterns in data

> Data Mining

1950's 1980's 2010's

Data science process Define research goal 1: Setting the research goal Create project charter Data retrieval Internal data e Data ownership 2: Retrieving data External data Errors from data entry Physically impossible values Missing values Data cleansing @ Outliers Spaces, typos, ... Errors against codebook Aggregating data 3: Data preparation Extrapolating data Derived measures Data transformation @ Creating dummies Reducing number of variables Merging/joining data sets Combining data Set operators Creating views Simple graphs Combined graphs 4: Data exploration Link and brush Nongraphical techniques Model and variable selection Model execution 5: Data modeling Model diagnostic and model comparison Presenting data 6: Presentation and automation Automating data analysis

Veri Bilimi Proje Geliştirme Süreci



5 TYPES OF ANALYTICS





What can be improved or automate?

Combines a number of intelligent technologies like artificial intelligence and deep learning to apply human sense to perform certain task.

Perspective Context



What should be done?

Advice business on possible outcome and what they can do. Provide suggestion to solve business problems.

Predictive Insight



What will happen?

Utilizing statistical and machine learning to provide recommendation and provide data about what might happen in the future.

Diagnostic Foresight



Why did happen?

 $\label{prop:eq:explain} \textit{Explain the cause behind the outcome of descriptive analytics}.$

Descriptive Hindsight



What happened?

Gives only information about everything is going well or not in a business without explaining the causes.

Veri Analitiği Tipleri

Seviyeler & Cevaplanan Sorular

THE DATA SCIENCE HIERARCHY OF NEEDS

LEARN/OPTIMIZE

AGGREGATE/LABEL

EXPLORE/TRANSFORM

MOVE/STORE

COLLECT

AI, DEEP LEARNING Python, R, MATLAB,...

Research/Data Scientist

Machine Learning Engineer

A/B TESTING,
EXPERIMENTATION,
SIMPLE ML ALGORITHMS

ANALYTICS, METRICS, SEGMENTS, AGGREGATES, FEATURES, TRAINING DATA SSAS, Oracle Data Mining

Power BI, Tableau, SSIS, SSAS

BI Developer

Data Analyist

CLEANING, ANOMALY DETECTION, PREP

RELIABLE DATA FLOW, INFRASTRUCTURE, PIPELINES, ETL, STRUCTURED AND UNSTRUCTURED DATA STORAGE

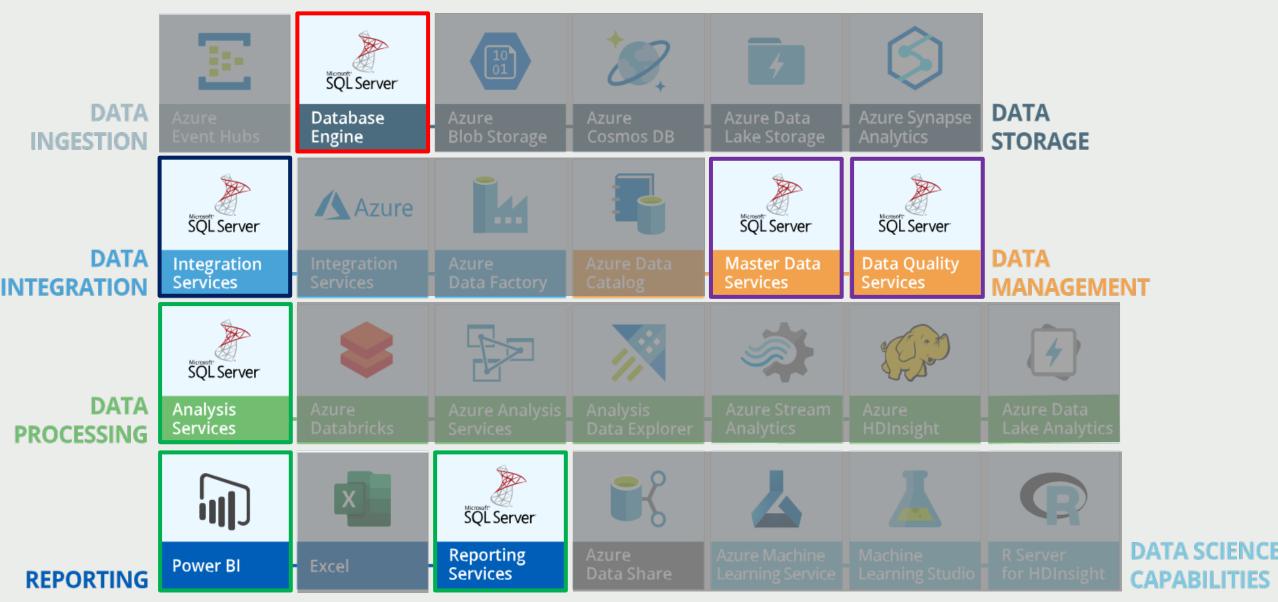
INSTRUMENTATION, LOGGING, SENSORS, EXTERNAL DATA, USER GENERATED CONTENT

SQL Server, SSIS

Data

Engineer

MICROSOFT BUSINESS INTELLIGENCE



DATA SHARING SERVICE

TEŞEKKÜRLER





Dr. Tuğrul TAŞCI ttasci@sakarya.edu.tr