Introduction to Data Science

Unlocking the Power of Data

MohammadReza **EskandariNasab**

PhD Student at Utah State University

06/06/2024

What is Data Science?

- **Definition**: "Data Science is the discipline of extracting knowledge and insights from structured and unstructured data using scientific methods, processes, algorithms, and systems."

- **Key Fields**: Statistics, Machine Learning, Computer Science.

- **Goal**: To make data-driven decisions and predictions.
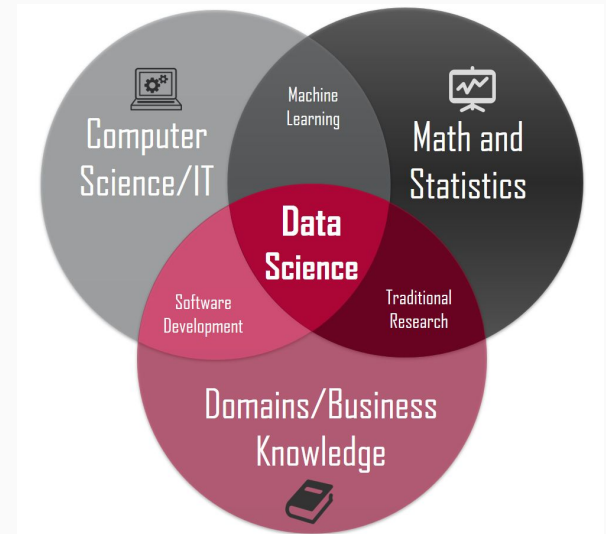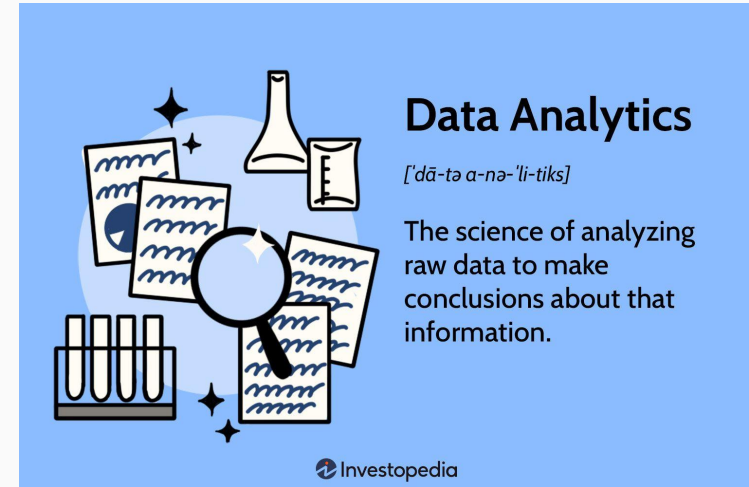
The Data Science Process

Stages:

1. **Data Collection:** Gathering raw data.

2. **Data Cleaning:** Removing inaccuracies and filling missing values.

3. **Data Analysis:** Applying statistical methods and algorithms.

4. **Data Visualization:** Creating graphs and charts.

5. **Model Deployment:** Using models in real-world applications.
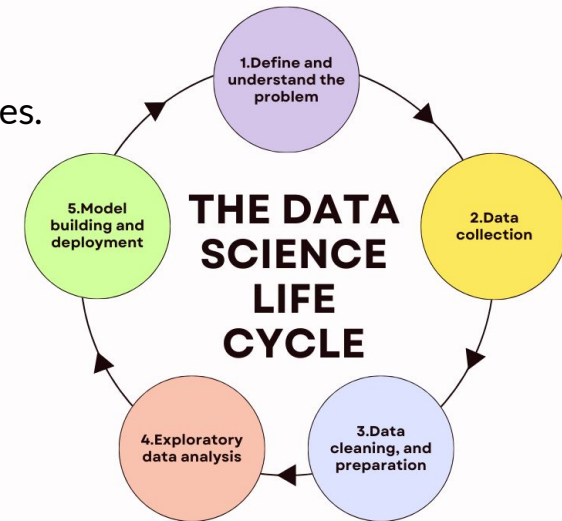
Key Components of Data Science

1. **Data Understanding:** Knowing the types and nature of data.

2. **Data Manipulation:** Techniques for processing data (e.g., sorting, merging).

3. **Data Storage:** Use of databases and data warehouses.

4. **Importance of Data Quality:** Impact on analysis accuracy.

5. **Machine Learning:** Builds models to predict future data or automate decisions.

Data Analytics

['dā-tə ə-nə-'li-tiks]

The science of analyzing raw data to make conclusions about that information.

Investopedia

Tools and Technologies

Programming Languages: Python, R

Libraries and Frameworks:
    Python: Pandas, NumPy, scikit-learn, TensorFlow
    R: ggplot2, dplyr, caret

Development Environments: Jupyter Notebook, RStudio

Applications of Data Science

Healthcare: Disease prediction, medical imaging.

Finance: Credit scoring, algorithmic trading.

Retail: Customer segmentation, inventory management.

Transportation: Route optimization, demand forecasting.

Challenges in Data Science

Data Privacy: Ensuring personal data is protected.

Data Quality: Dealing with poor or incomplete data.

Keeping Up-to-Date: Staying current with new methods and technologies.

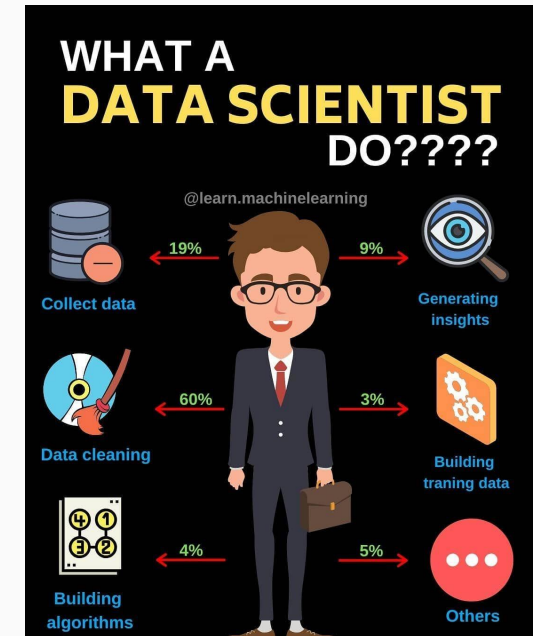**Skill Diversity:** Need for a broad range of skills.

Common Data Science Problems

Title: Introduction to Data Science

MohammadReza **EskandariNasab** | PhD Student at Utah State University

UtahStateUniversity

Becoming a Data Scientist

Education: Background in computer science, mathematics, or domain-specific fields.

Skills: Statistical analysis, programming, problem-solving.

Pathways: Academic degrees, online courses, bootcamps.

Title: **Introduction to Data Science**

MohammadReza **EskandariNasab** | PhD Student at Utah State University

Conclusion and Resources

Summary: Data Science is a powerful tool for insight and innovation across many fields.

Further Learning:

- **Online Platforms:** Coursera, Udemy, edX
- **Books:** "Data Science for Business" by Foster Provost, "Pattern Recognition and Machine Learning" by Christopher Bishop
- **Communities:** Kaggle, Stack Overflow, GitHub

https://github.com/samresume/USU-Workshop

Title: Introduction to Data Science

MohammadReza EskandariNasab | PhD Student at Utah State University

UtahStateUniversity

Thank You for Your Attention

MohammadReza **EskandariNasab**
PhD Student at Utah State University

06/06/2024