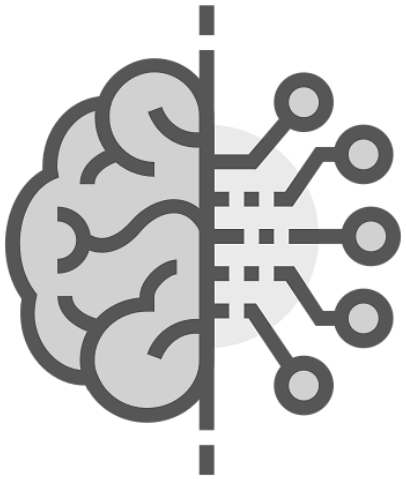
# Artificial Intelligence (Robotics)

Lecture 2: Introduction



Muhammad Rizwan – Corvit Systems

# AI (Robotics)

- **Course Title**:

  AI (Robotics)

- **Course Duration**:

  06 Months

# AI (Robotics)

- **Course Objective**:

  Employable skills and hands on practice for **Artificial Intelligence**, including specialization in **AI Programming**, **Data Science**, **Deep Learning**, **Computer Vision**, **NLP**, **MLOps** & **Robotics**.

# AI (Robotics)

- **Learning Outcome of the Course**:
  - AI for Everyone
  - Algorithmic Programming, Data Analytics
  - Workflow of AI & Machine Learning Projects
  - Deep Learning, Deep Computer Vision
  - Deep Sequence Modeling, NLP, MLOps, Robotics
  - Core Skills

# AI (Robotics)

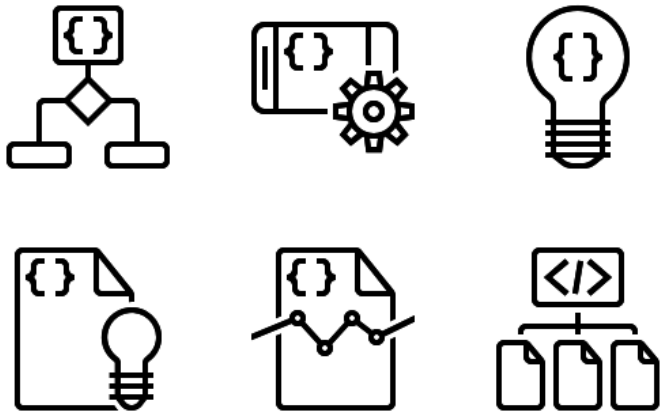
- **Job Opportunities**:
  - Python Developer, Data Analyst, Business Analyst
  - Software Engineer, Machine Learning Researcher
  - Full Stack Data Scientist, Machine Learning Engineer
  - MLOps Engineer

# AI (Robotics)

- **Companies Offering Jobs in the respective trade**:
  - Afiniti, arbisoft, confiz, Careem
  - Folio3, Xavor, i2c, NetSol
  - Kaggle, Crossover, Fiverr, Upwork, freelancer
  - Google, Meta, Netflix, Amazon
  - Uber, Apple, Shopify
  - Research and All Private Institutes …

# AI (Robotics)

- **Course Requirements**:
  - Basic Programming
  - Basic Math, Statistics, Linear Algebra, Derivatives etc.



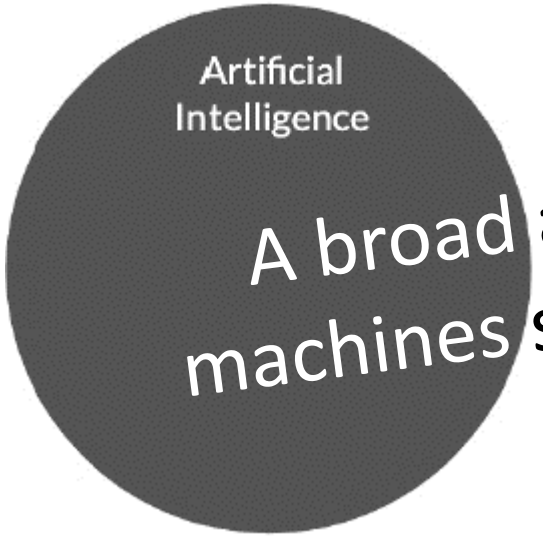
# Artificial Intelligence (Robotics)

AI & Data Science Introduction

# Artificial Intelligence (AI)

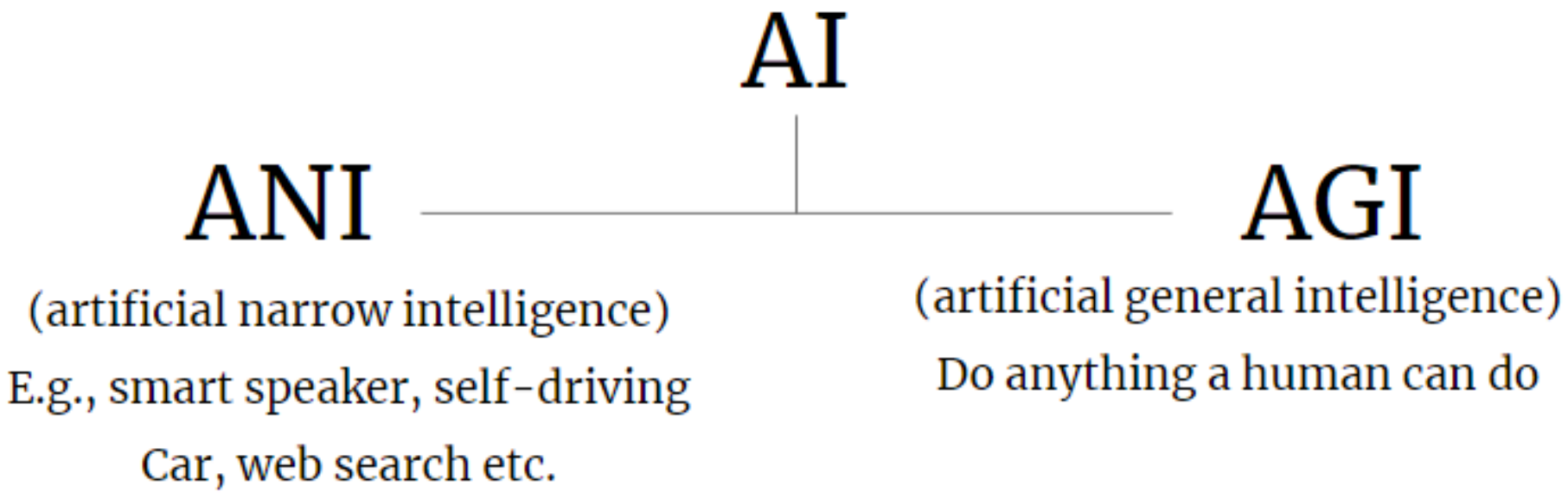
AI – A huge set of tools for making computers behave intelligently

AI is a technique that enables machines to mimic human behavior.




AI is changing the way we work and live

# Demystifying AI

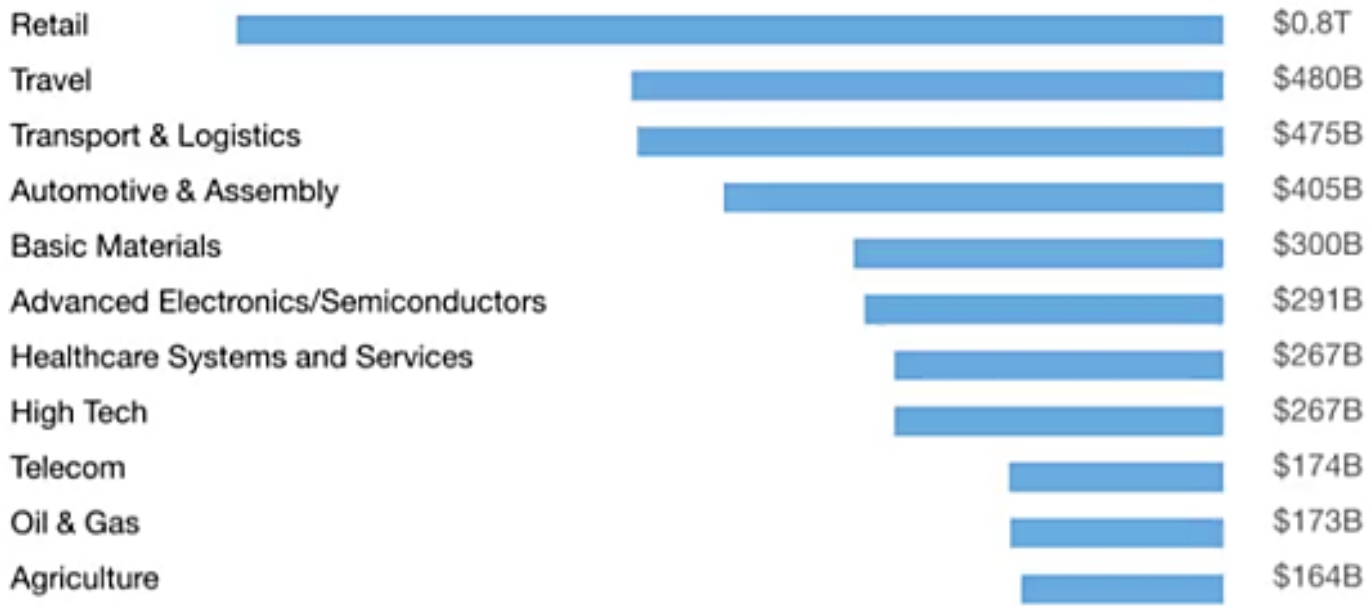



*Any intellect that greatly exceeds the cognitive performance of humans in virtually all domains of interest.*

# AI value creation in all sectors

AI value creation by 2030

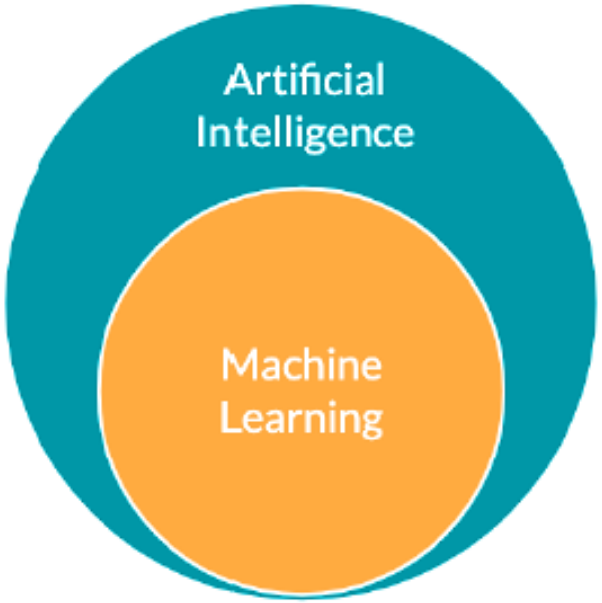
**$13 trillion**



# Machine Learning (ML)

ML – Machine Learning is the most prevalent subset of AI

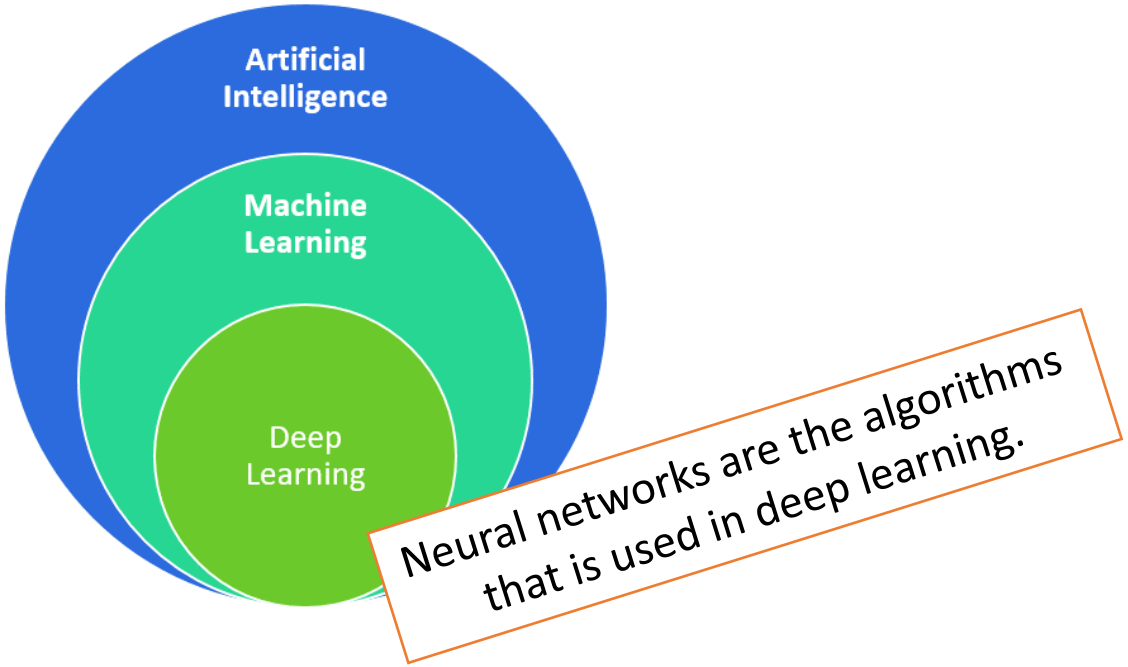
Machine learning is used to make inferences and predictions from data.



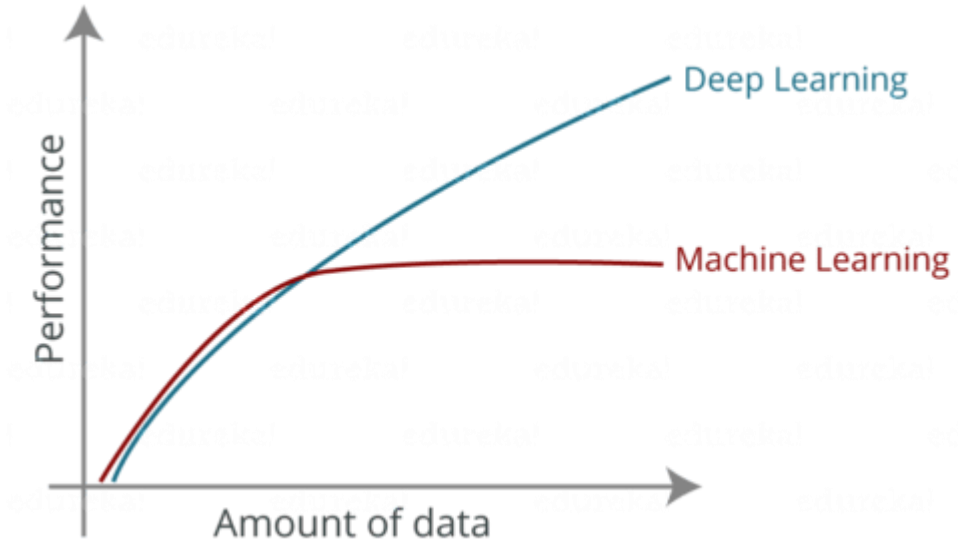

Machine learning uses mathematics and programming techniques.

# Deep Learning (DL)

Deep Learning – Special area of Machine Learning

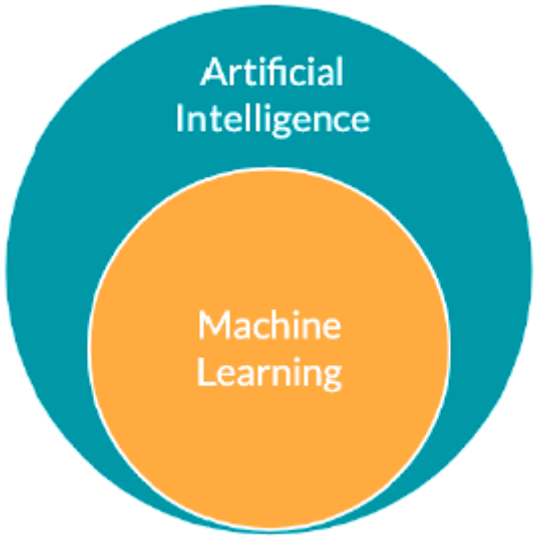



Neural networks are the algorithms that is used in deep learning.



# Data Science

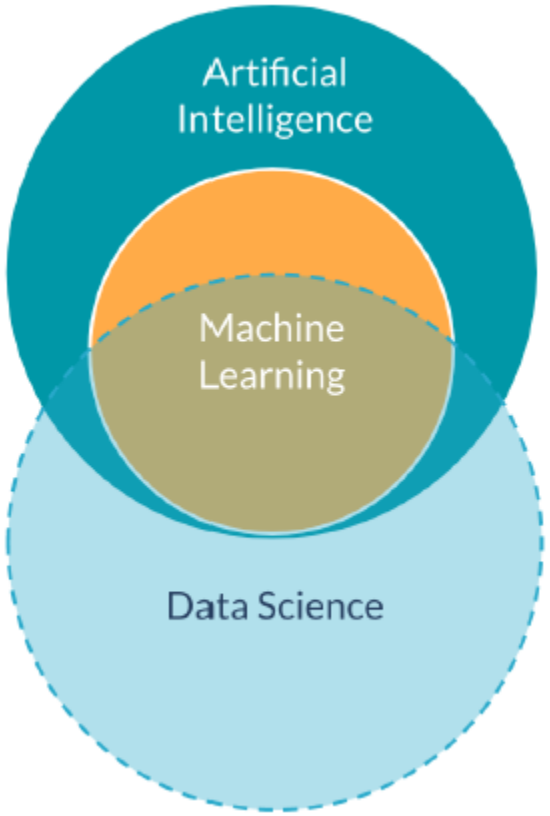
Data science is about discovering and communicating insights from data



# Data Science

**Data science** is about making discoveries and creating insights from data

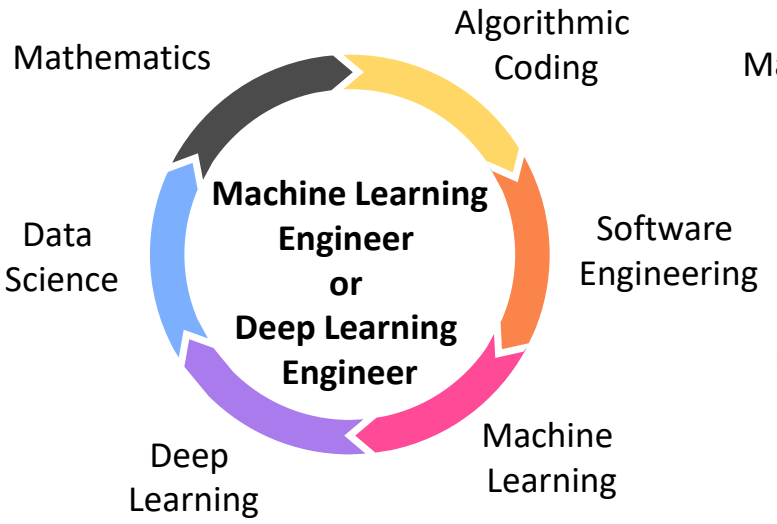
**Machine Learning** is often an important tool for data science work

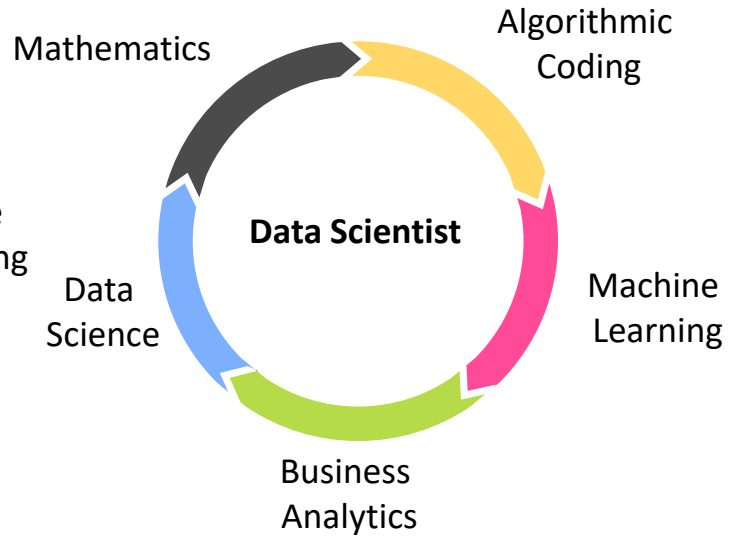


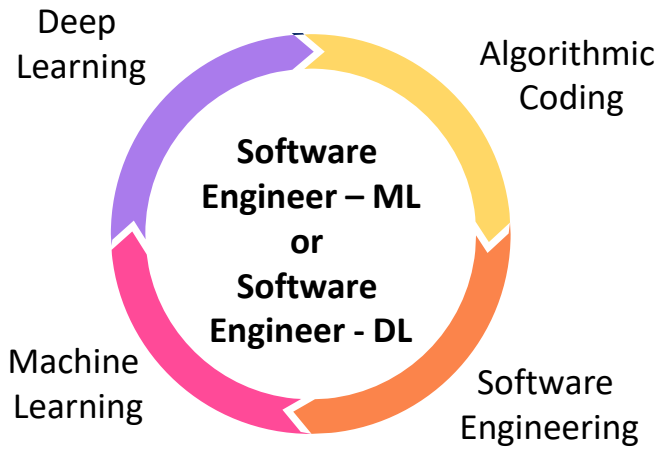
# AI & Deep Learning Specialization

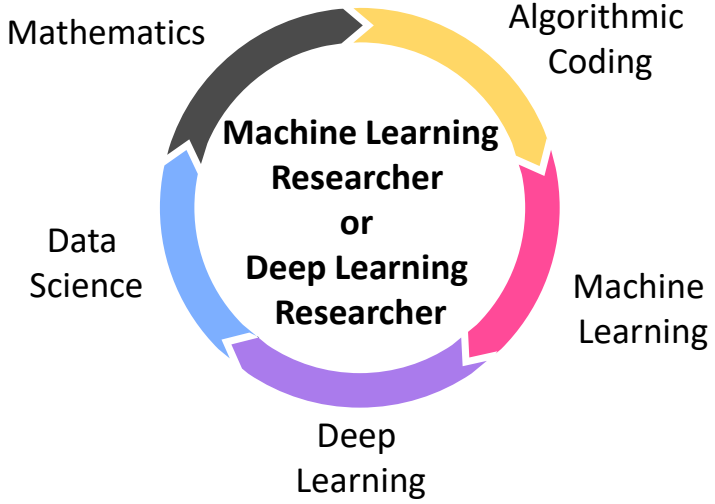
## Computer Vision
## Natural Language Processing

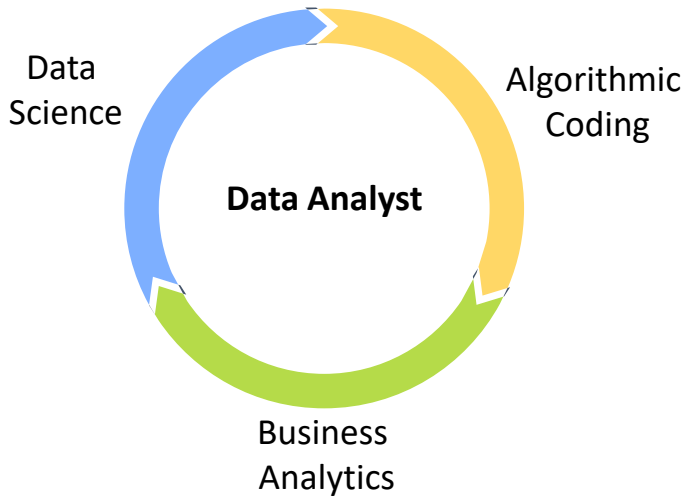
# AI Roles

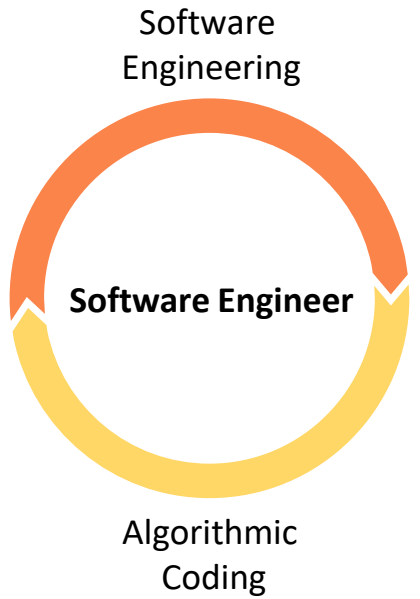


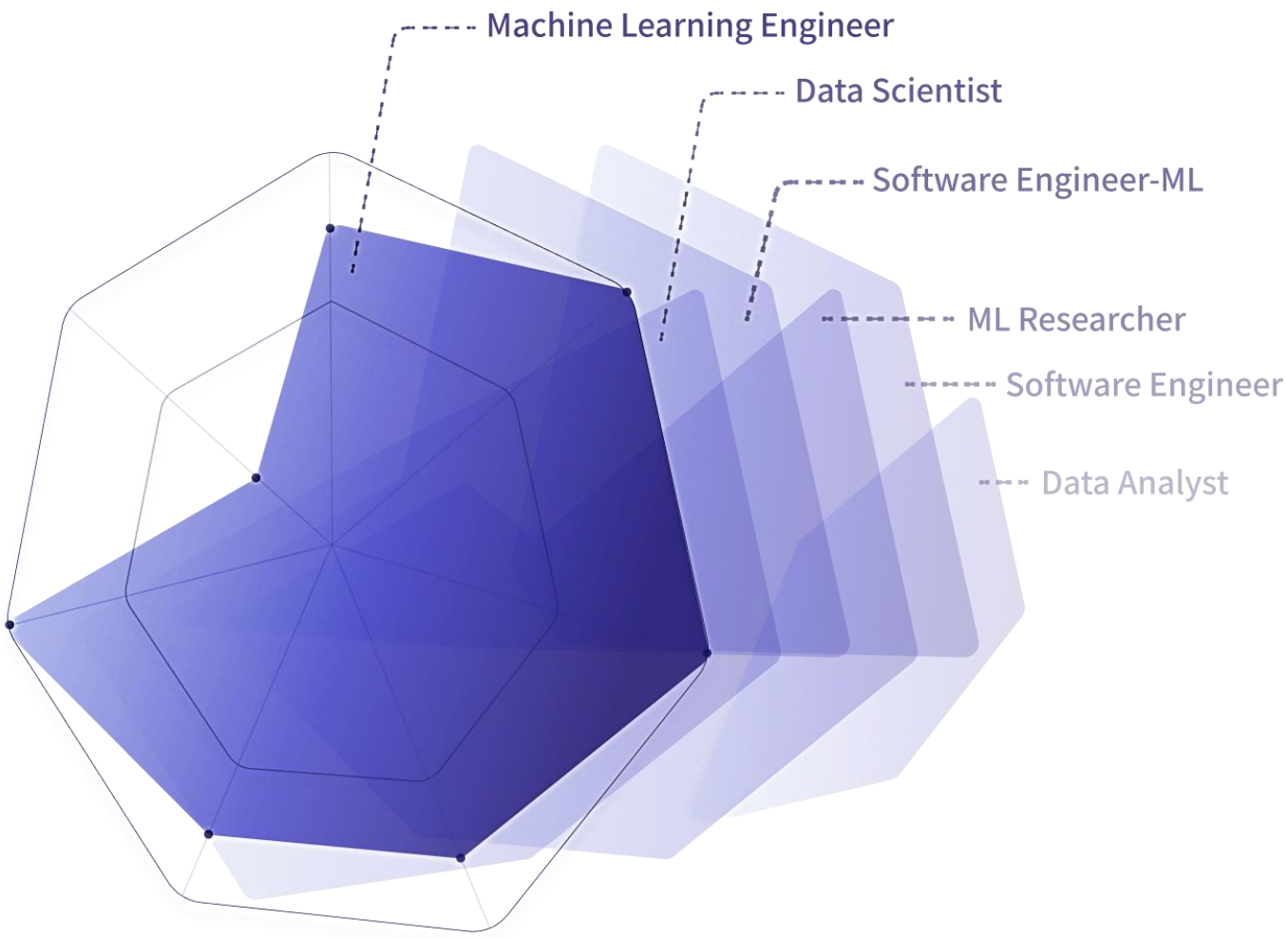




**@mrizwanse**

# Standardized test for ML skills

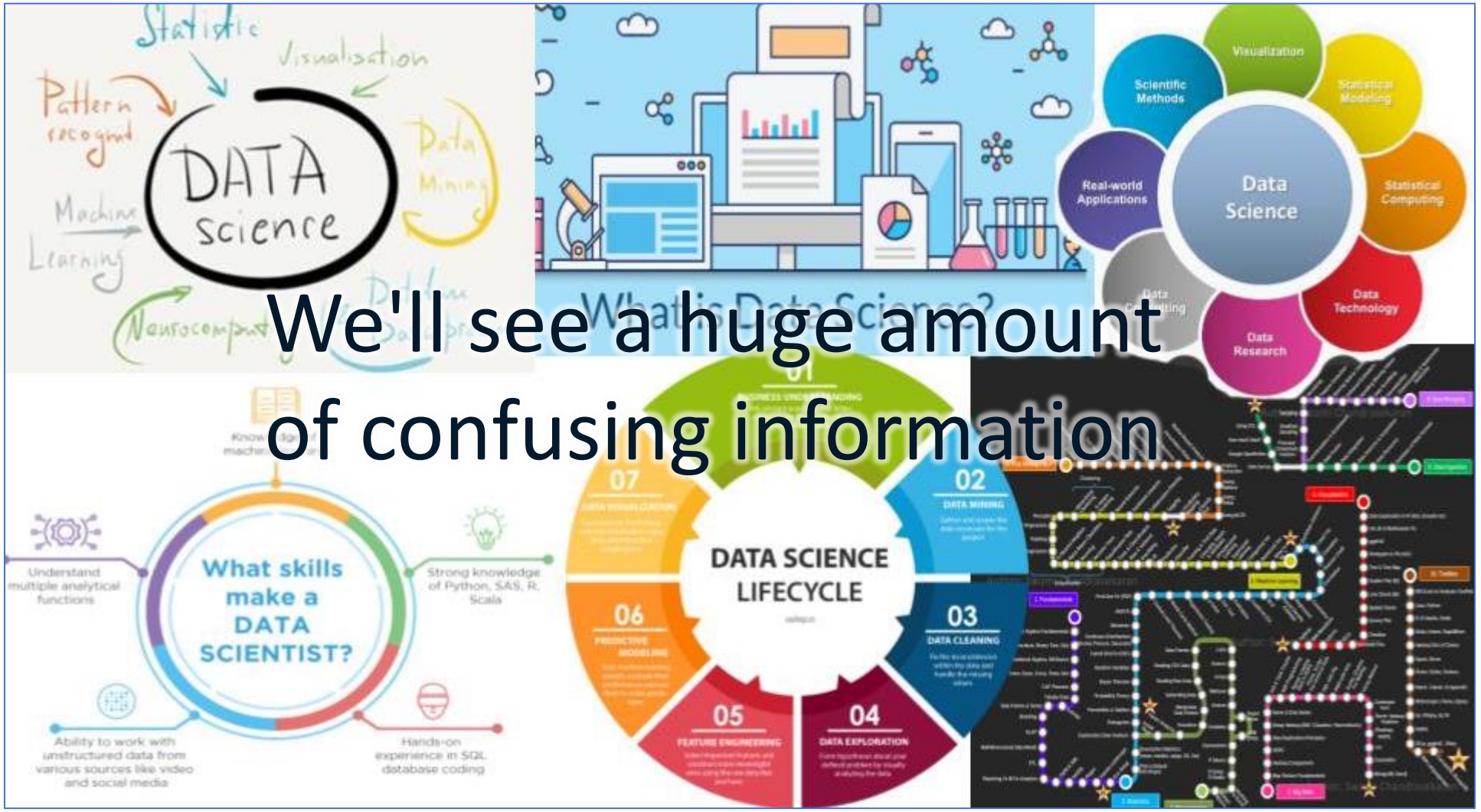


# Data Science

- Introduction to Data Science
- Data Collection and Storage
- Preparation, Exploration and Visualization
- Experimentation and Prediction

# What is Data Science?

# Let’s ask Google – What is Data Science?




We'll see a huge amount of confusing information

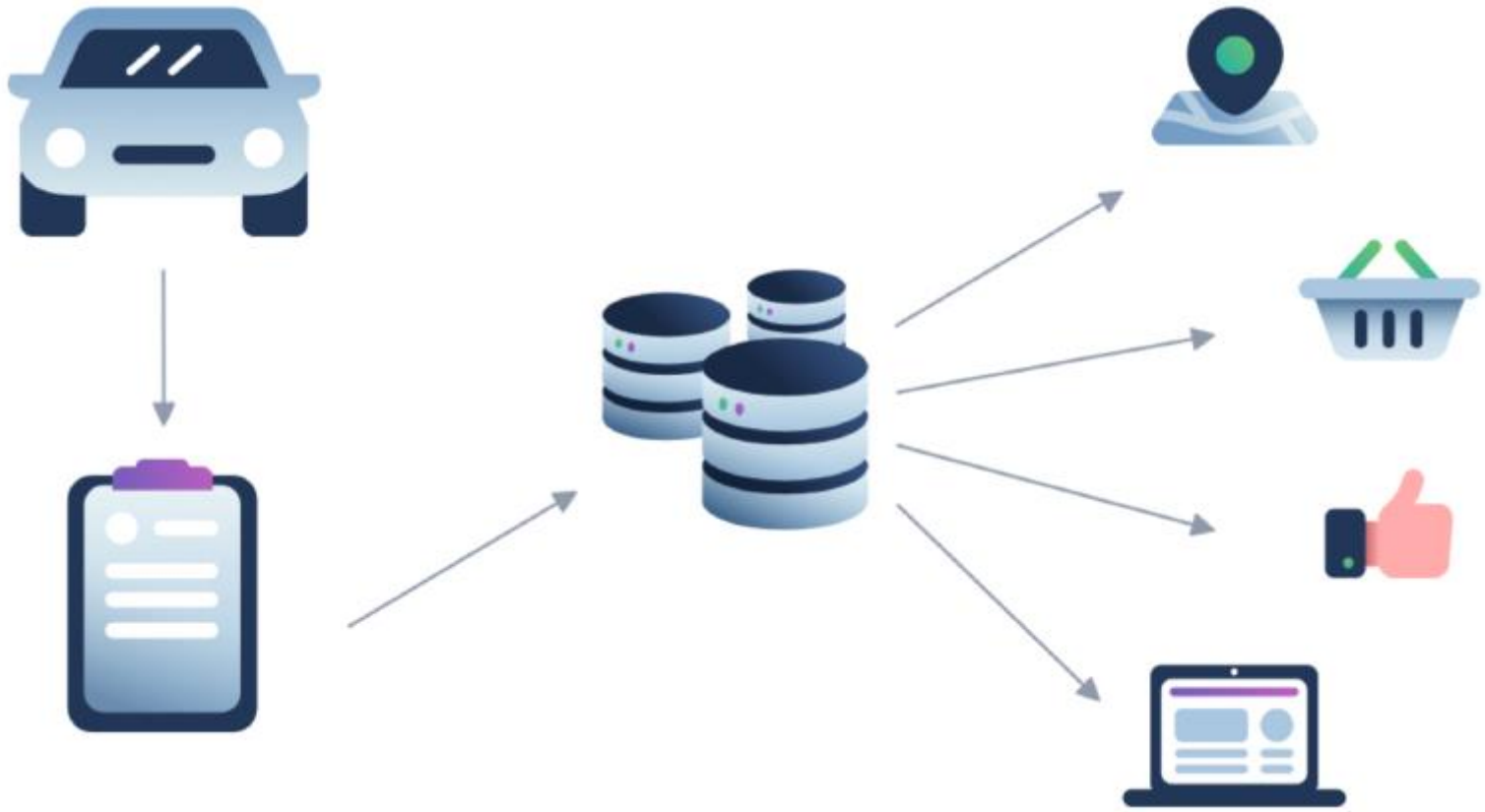
# Making **data** work for you

Use **data** to better describe the present or better predict the future

# What can **data** do?

- Describe the current state of an organization or process
- Detect anomalous events
- Diagnose the causes of events and behaviors
- Predict future events (such as forecasting population size)

# Why now?



# The **Data Science** Workflow

**Data Collection & Storage**

**Data Preparation**

**Exploration & Visualization**

**Experimentation & Prediction**

# **Example**: Customer Segmentation Workflow

*Bilal manages a **data science team** at a subscription-based fast food delivery company. Her team has been asked to investigate loss of customers, also known as customer attrition.*

# Solution – Hint

- Before the data can be explored and analyzed, it must be collected and prepared.
- A line chart is a way of visualizing data.
- Clustering can be used to predict which group an item is most similar to.

# Arrange the steps

- Reformat the delivery date on all entries to be in the same time zone
- Cluster the users into different personas and perform a regression to predict churn for each cluster
- Download the data
- Create a line chart that shows decay in subscriptions by cohort

# Solution

- Download the data
- Reformat the delivery date on all entries to be in the same time zone
- Create a line chart that shows decay in subscriptions by cohort
- Cluster the users into different personas and perform a regression to predict churn for each cluster

# Why?

- Data collection comes first
- Preparation
- Exploration and visualization
- Finally experimentation and prediction

# **Example**: Building a customer service chatbot

*Ahmad and his data team are working on a customer service chatbot. The chatbot is a computer program that uses **data science** to answer basic customer questions via a messenger.*

*The team will use transcripts from over 300,000 customer service interactions to train a chatbot to answer customer questions.*

# Ahmad team will take in the appropriate step of the data science workflow.

- Data Collection and Storage
- Exploration and Visualization
- Experimentation and Prediction

# Ahmad team will take in the appropriate step of the data science workflow.

- Plot the number of conversations vs. the time of day
- Load the transcripts into the data team's database
- Use a **Machine Learning** model to predict possible response for each question
- Collect the timestamps for each transcript
- Create an algorithm that classifies the initial customer question
- Gather customer information for each conversation
- Create a bar chart of the number of conversations of each type

| | |
|---|---|
| **Data Collection and Storage** | **Exploration and Visualization** |
| **Experimentation and Prediction** | |

# Solution – Data Collection and Storage

- Gather customer information for each conversation
- Collect the timestamps for each transcript
- Load the transcripts into the data team’s database

# Solution – Exploration & Visualization

- Plot the number of conversations vs. the time of day
- Create a bar chart of the number of conversations of each type

# Solution – Experimentation and Prediction

- Create an algorithm that classifies the initial customer question
- Use a **Machine Learning** model to predict possible response for each question

# Application of Data Science

# Case studies

- Machine Learning
- Deep Learning
- IoT

# **Case study**: Credit Card Fraud Detection

| Amount | Date | Location | … |
|---|---|---|---|
| 70,000 | 10-Jan-2020 | Lahore | … |
| 90,000 | 15-Sep-2020 | Faisalabad | … |
| 45,000 | 19-Dec-2020 | Lahore | … |
| 93,000 | 25-Jan-2021 | Islamabad | … |
| 20,000 | 28-Mar-2021 | Karachi | … |
| 33,000 | 25-Apr-2022 | Peshawar | … |
| … | … | … | … |

*“What is the probability that this transaction is fraudulent?”*

# What do we need for **Machine Learning**?

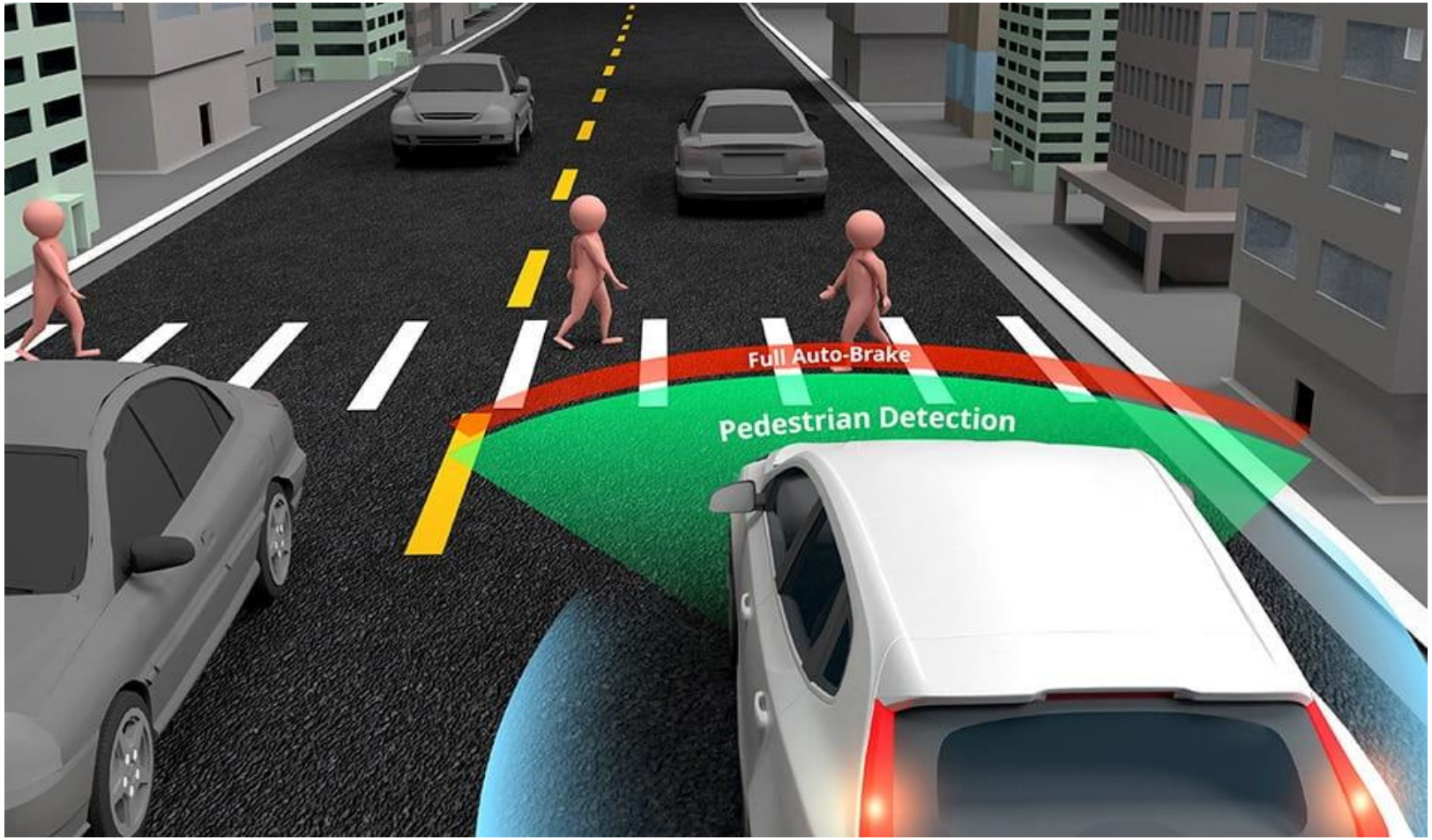
- A well-defined question
  - *“What is the probability that this transaction is fraudulent?”*
- A set of example data
  - *Old transactions labeled as “fraudulent” or “valid”*
- A new set of data to use our algorithm on
  - *New credit card transactions*

# **Case study**: Smart Watch

# Internet of Things (IoT)

- Smart watches
- Internet-connected Home security systems
- Electronic toll collection systems
- Building energy management systems
- Much, much more!

# **Case study**: Image Recognition



# Deep Learning

- Many **Neurons** work together
- Requires much more training data
- Used in complex problems
  - Image Classification
  - Language Learning / Understanding

# **Example**: Assigning Data Science project

*Khalil manages an analytics team and has a few tasks that he hoping to achieve this quarter. The tasks are centered around the following domains:*

- ***Machine Learning***
- ***Deep Learning***
- ***IoT***

*The knowledge to build machine learning and deep learning applications is present within her team. There is another team in the company that is specialized in working with IoT data. Khalil wants to know for which tasks he'll need their help.*

# For each task to select the correct domain

- Automatically summarize text from news articles
- Automatic building cooling using temperature sensors
- Cluster patients by symptoms to help
- Predict ride-sharing prices at a certain time and location based on previous prices
- Flag images that contain a safety violation
- Detect machinery failure with vibration detectors

**Machine Learning** **Deep Learning**

**Internet of Things**

# Solution – Hint

- Clustering can be used to **predict** which group an item is most similar to.
- Text summarization and image classification both require much, much more input data than most data science problems.

# Solution – Machine Learning

- Predict ride-sharing prices at a certain time and location based on previous prices
- Cluster patients by symptoms to help

# Solution – Deep Learning

- Automatically summarize text from news articles
- Flag images that contain a safety violation

# Solution – Internet of Things

- Detect machinery failure with vibration detectors
- Automatic building cooling using temperature sensors

# Data Science roles and tools

Data Engineer

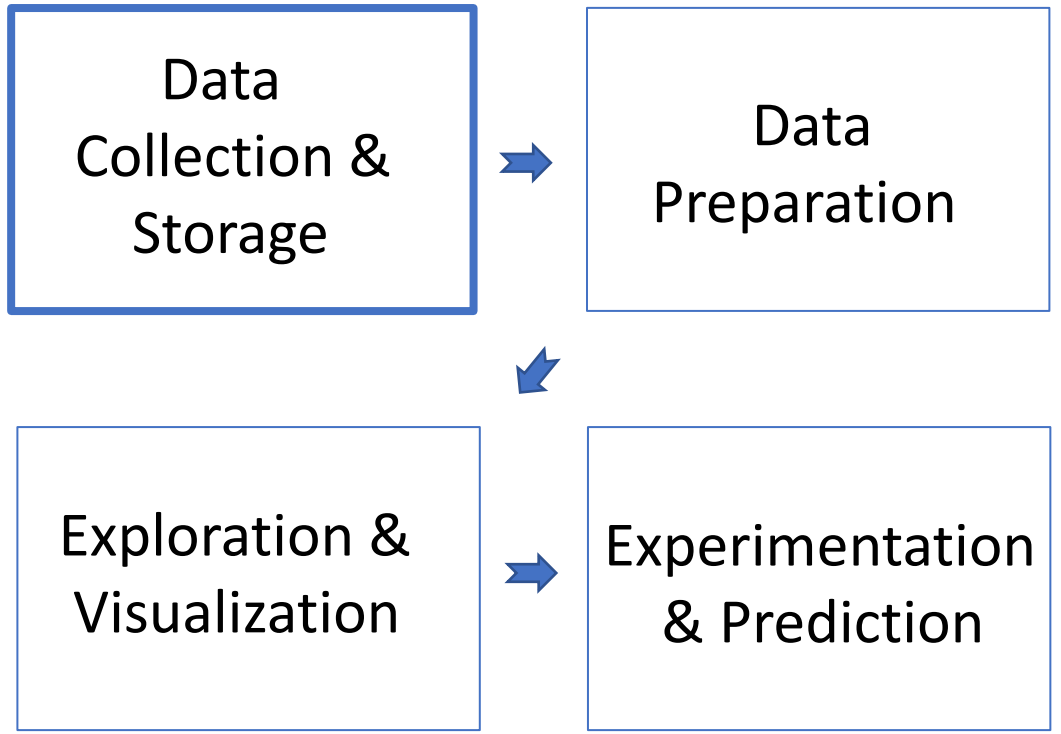
Data Analyst

Data Scientist

Machine Learning Scientist

# Data Engineer

- Information architects
- Build data pipelines and storage solutions
- Maintain data access

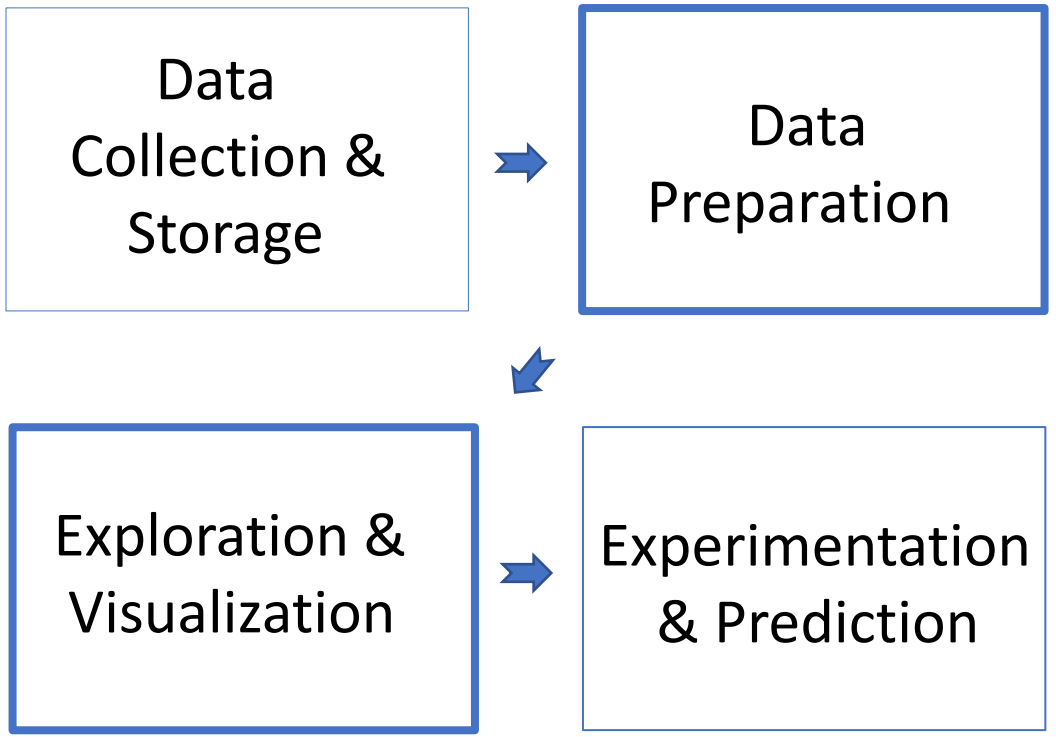


# Data Engineering Tools

- **SQL**
  - To store and organize data
- **Python, Java,** or **Scala**
  - Programming languages to process data
- **Shell**
  - Command line to automate and run tasks
- **Cloud Computing**
  - AWS, Azure, Google Cloud Platform

# Data Analyst

- Perform simpler analyses that describe data
- Create reports and dashboards to summarize data
- Clean data for analysis

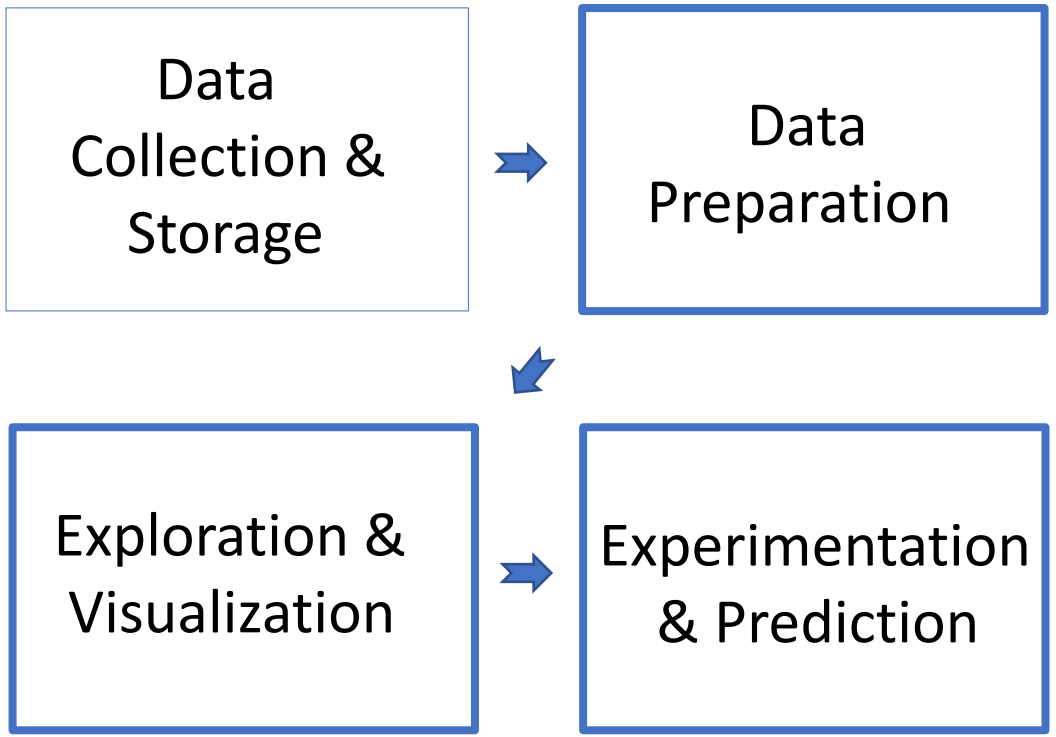


# Data Analyst Tools

- **SQL**
  - Retrieve and aggregate data
- **Spreadsheets (Excel or Google Sheets)**
  - Simple analysis
- **BI Tools (Tableau, Power BI, Looker)**
  - Dashboards and visualizations
- *May have*: Python or R
  - Clean and analyze data

# Data Scientist

- Versed in statistical methods
- Run experiments and analyses for insights
- Traditional Machine Learning

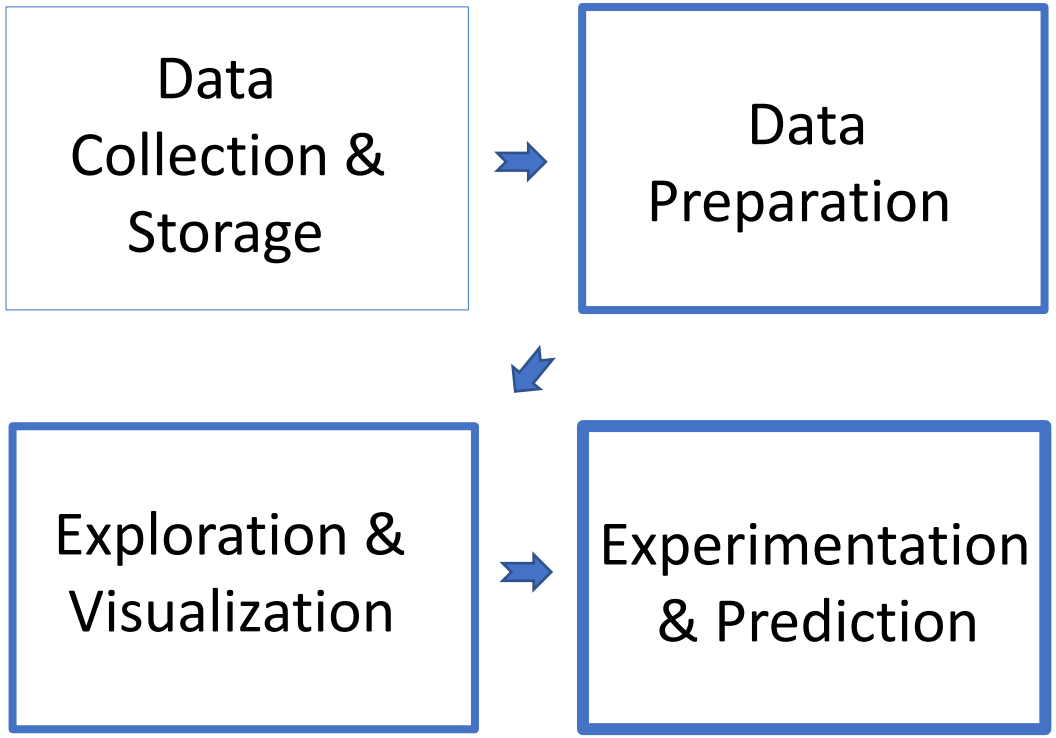


# Data Scientist Tools

- **SQL**
  - Retrieve and aggregate data
- **Python and/ or R**
  - Data Science libraries, e.g., **`pandas`** (Python) and **`tidyverse`** (R)

# Machine Learning Scientist

- Predictions and extrapolations
- Classification
- Deep Learning
  - Image Processing
  - Natural Language Processing



# Machine Learning Tools

- **Python and/ or R**
  - Machine Learning libraries, e.g., Tensorflow or PyTorch or Spark




| Data Engineer | Data Analyst | Data Scientist | ML Scientist |
|---|---|---|---|
| Store and maintain data | Visualize and describe data | Gain insights from data | Predict with data |
| SQL + Python / Java / Scala | SQL + BI Tools + Spreadsheets | Python/R | Python/R |

# Data Analyst Job Posting

**Job Id**: 8675309

**Location**: New York

**Business**: Cool Company

**Job Summary**:

Join our awesome company and do data analysis for us!

**Basic Qualifications**:

- Proficient in SQL for data analysis (including JOINs, WINDOW functions)
- Familiarity with Tableau or Power BI for building dashboards
- ???

# **Example** – Editing a Job post

*Anas is looking to expand her team by hiring a **Data Analyst**. This analyst will be performing ad hoc analyses and building dashboards to track company-wide goals.*

*Anas is working on a job description for a job board. He almost done writing the Basic Qualifications section.*

*Which of the following should he add?*

# Possible Answers

- Some experience with Deep Learning and Neural Networks
- Expert user of Excel or Google Sheets, including VLOOKUP and pivot tables
- Basic proficiency in either Java, Scala or Python for database operations

# Possible Answers

- Some experience with Deep Learning and Neural Networks
- **Expert user of Excel or Google Sheets, including VLOOKUP and pivot tables**
- Basic proficiency in either Java, Scala or Python for database operations

# **Example** – Matching Skills to Job

*Usman manages a Data Science team, and is looking to post some new job listings for a **Data Engineer** and a **Machine Learning Scientist**.*

*Help Usman decide which skill requirements belong with each job.*

# Select each requirements into the correct resume

- Strong Java Skills
- Some higher education in statistics
- Proficient in using Python for prediction and modeling
- Expert at building and maintaining SQL databases
- Experience in using TensorFlow for implementing deep learning architectures

| **Data Engineer** | **Machine Learning Scientist** |
| --- | --- |

# Solution – Data Engineer

- Expert at building and maintaining SQL databases
- Strong Java Skills

# Solution – Machine Learning Scientist

- Proficient in using Python for prediction and modeling
- Experience in using TensorFlow for implementing deep learning architectures

# **Example** – Classifying data tasks

*Rizwan is leading her team's planning session this week. He needs to assign tasks for the coming week and set priorities. Rizwan team is composed of a* ***Data Engineer****, a* ***Data Analyst****, and a* ***Data Scientist****.*

# Assign each task to the current job title

- Update Excel spreadsheet with new graphs
- Give new team members database access
- Train an anomaly detection algorithm
- Create a dashboard for Making team
- Create a new table in the SQL database
- Run a correlation analysis between weather and ice cream sales

| **Data Engineer** | **Data Analyst** | **Data Scientist** |
| --- | --- | --- |

# Solution – Data Engineer

- Create a new table in the SQL database
- Give new team members database access

# Solution – Data Analyst

- Create a dashboard for Making team
- Update Excel spreadsheet with new graphs

# Solution – Data Scientist

- Run a correlation analysis between weather and ice cream sales
- Train an anomaly detection algorithm