



Data Science Life Cycle



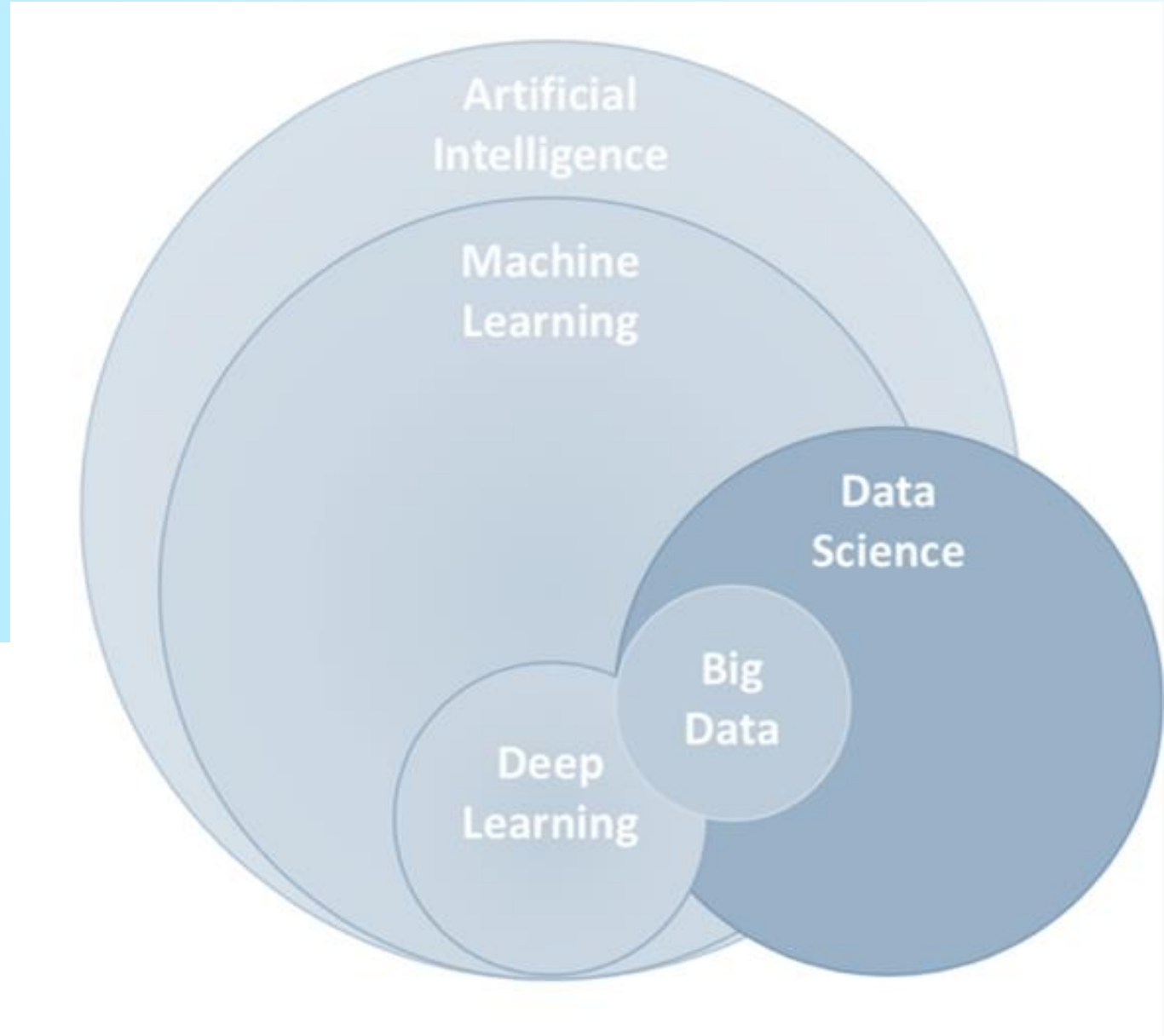
- **Introduction**
- **Business understanding**
- **Data understanding**
- **Data preparation**
- **Data Modeling**
- **Model Evaluation**
- **Model Deployment**

## OBJECTIVES

# DATA SCIENCE

---

Data science is the field of study that combines domain expertise, programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data. Data science practitioners apply machine learning algorithms to numbers, text, images, video, audio, and more to produce artificial intelligence (AI) systems to perform tasks that ordinarily require human intelligence. In turn, these systems generate insights which analysts and business users can translate into tangible business value.



# DIFFERENT TYPE

---

## - Data Analyst

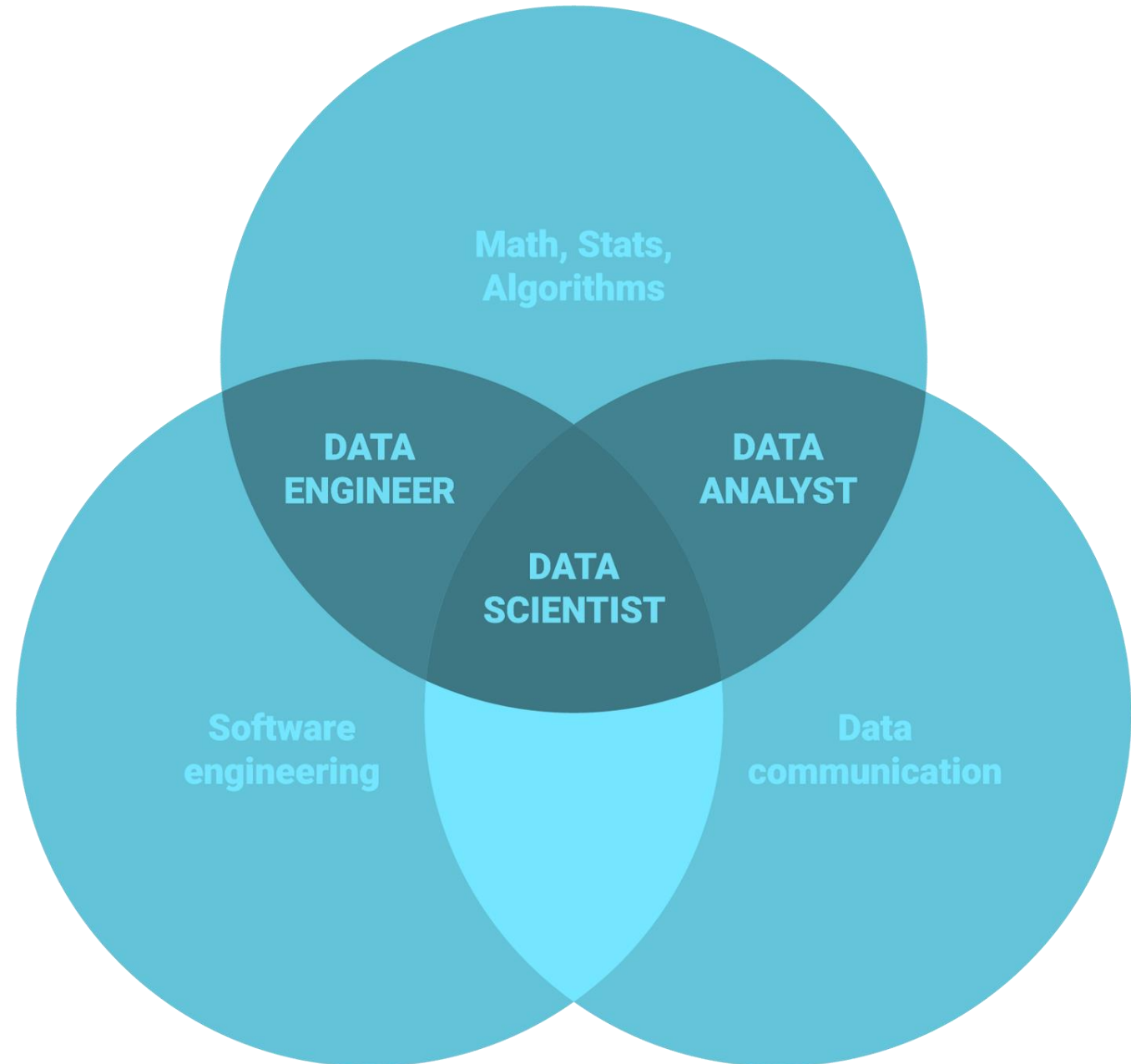
The process of the extraction of information from a given pool of data.

## - Data Engineering

*person who specializes in preparing data for analytical usage*, also involves the development of platforms and architectures for data processing.

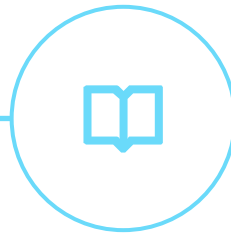
## - Data Scientist

the field of study that combines domain expertise, programming skills, and knowledge of mathematics and statistics to extract meaningful insights from data.

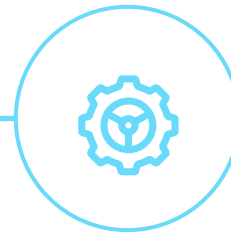




**Business  
understanding**



**Data  
understanding**



**Data  
preparation**

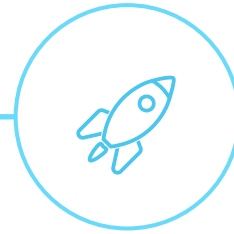
**LIFE CYCLE**



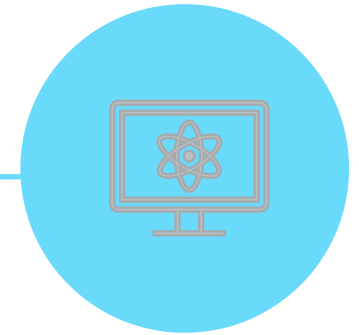
**Modeling**



**Evaluation**



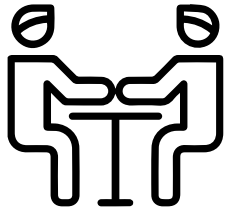
**Deployment**



**LIFE CYCLE**

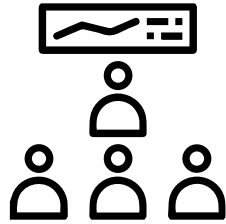
# Business Understanding

---



Understanding Business  
Requirements

---



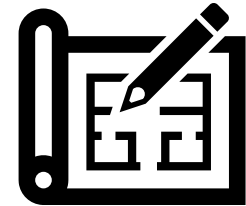
Analyzing supporting  
information

---



Converting to data  
mining problem

---

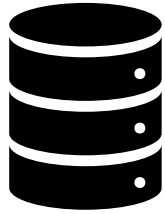


Preparing a preliminary  
plan

---

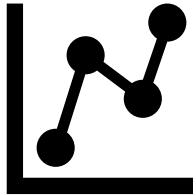
# Data Understanding

---



Data Collection

---



Data Properties

---



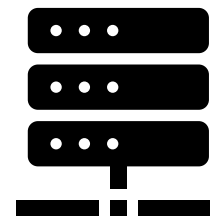
Data Quality

---



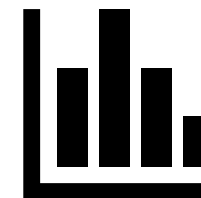
# DATA PREPARATION

---



Final Dataset Selection

---

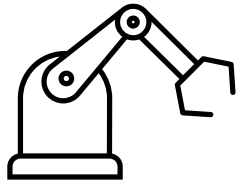


Prepare the data

---

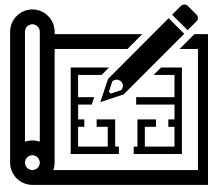
# Data Modeling

---



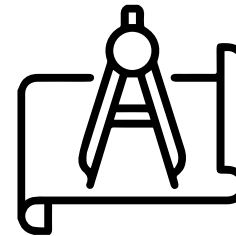
**Selection and Creation**

---



**Testing Plan**

---



**Parameter Tuning/Testing**

---

# TOOLS

Tools used in Data manipulation

# DATA SCIENCE

# TOOLS

## ANACONDA DISTRIBUTION

Most Trusted Distribution for Data Science

### ANACONDA NAVIGATOR

Desktop Portal to Data Science

### ANACONDA PROJECT

Portable Data Science Encapsulation

### DATA SCIENCE LIBRARIES

#### Data Science IDEs



#### Analytics & Scientific Computing



#### Visualization



#### Machine Learning



...and many more!

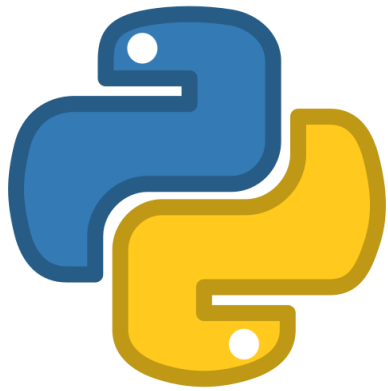


Data Science Package & Environment Manager

# ANACONDA

## TOOLS

---



**Python**



**Colab**

**THANK YOU**