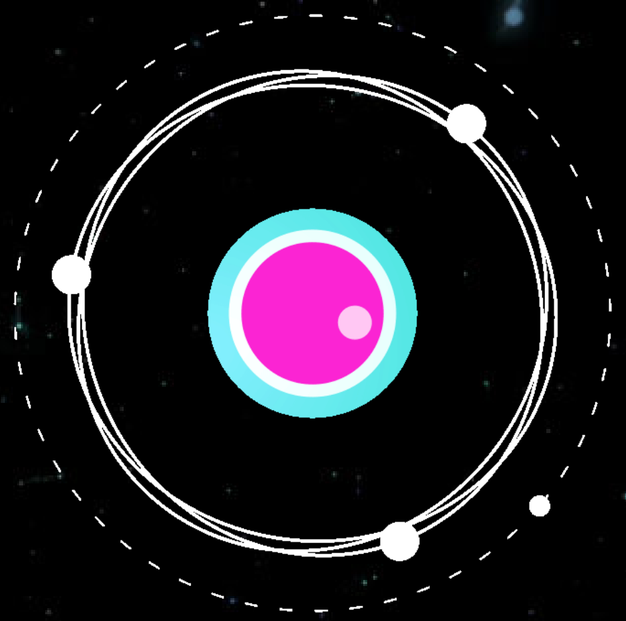
DATA CELL TRAINGING ROADMAP

INPT

CIT



MADE BY:

TAHA ELGHABI

PLAN

OVERVIEW, DATA ROADMAP

DATA COLLECTION

EXPLORATORY DATA ANALYSIS

CLEANING, FEATURE ENGINEERING

VISUALIZATION

PLAN

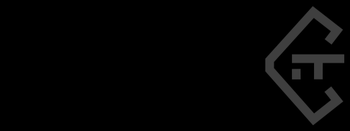
INTRO TO MACHINE LEARNING

MODEL BUILDING

HYPERPARAMETERS OPTIMIZATION

MODEL EVALUATION

MLOPS



DATA CELL TRAINING S1: AN OVERVIEW

INPT

CIT

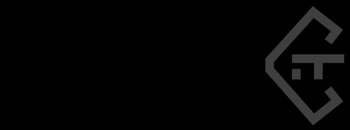
MADE BY:

TAHA ELGHABI

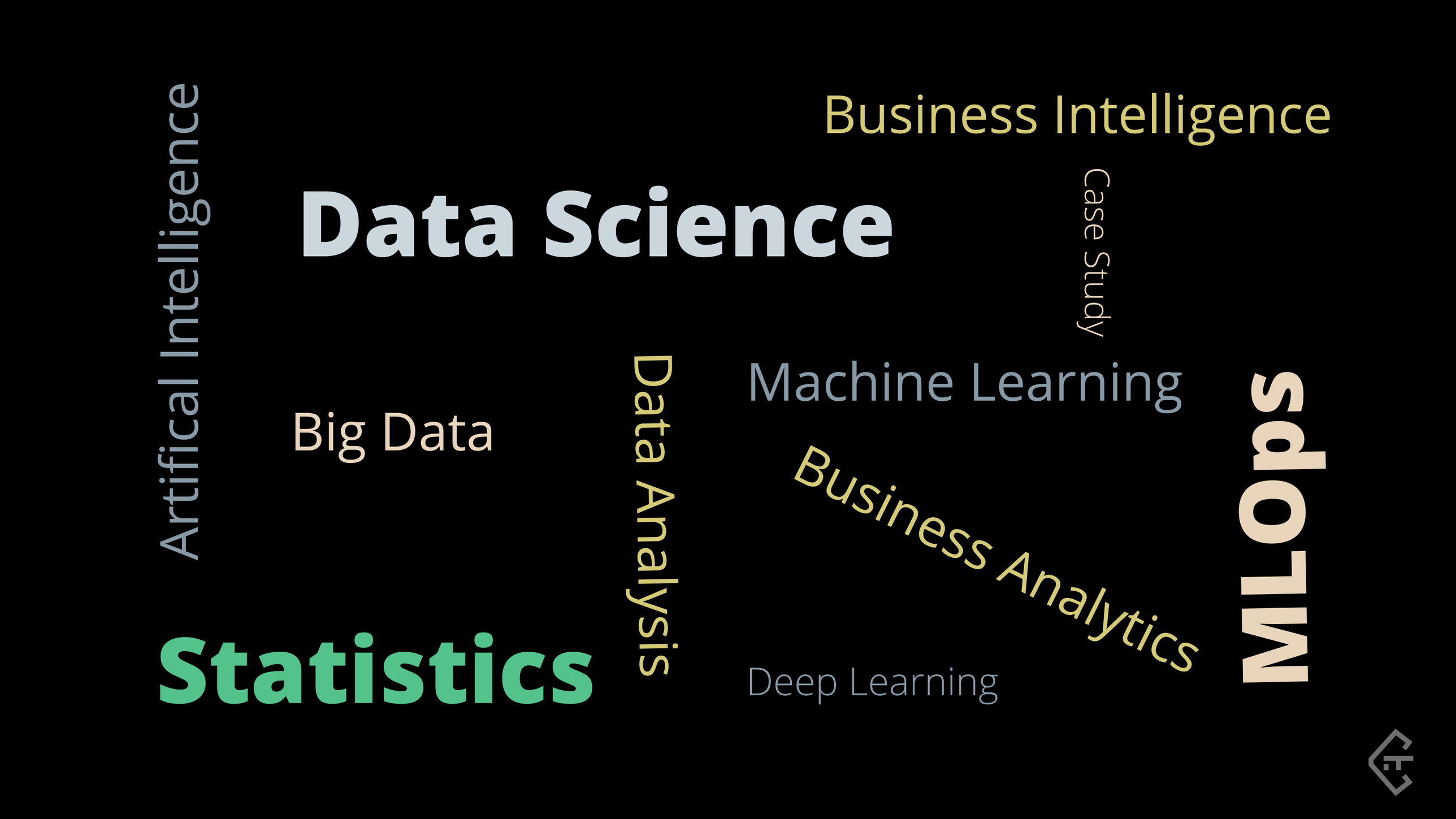
OVERVIEW
DATA SCIENCE ROADMAP

PLAN

- Data Science buzzwords
- What the heck is Data Science
- Workflow of a Data Science Project
- Tools that we'll be using
- Some tips and resources



Data Science buzzwords


Artifical Intelligence
Business Intelligence
Data Science
Case Study
Machine Learning
Big Data
Data Analysis
Business Analytics
MLOps
Statistics
Deep Learning

What the heck is Data Science

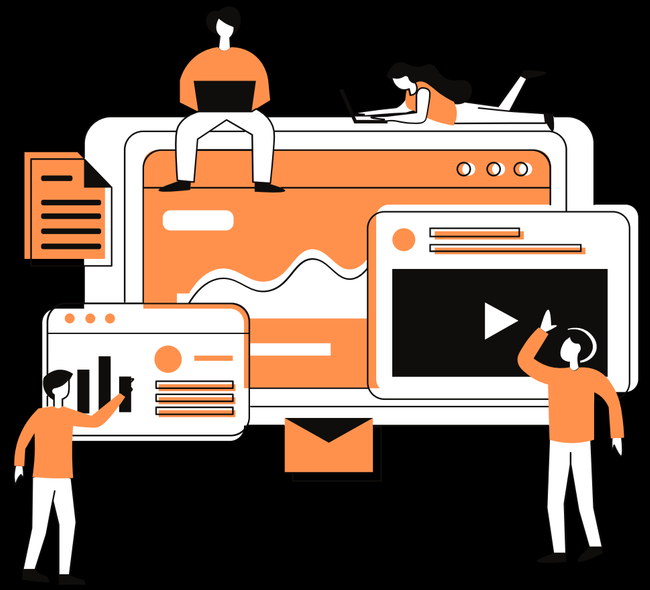
What is Data?

Data is a bunch of facts,
such as numbers, words,
measurements...
or could be a descriptions
of things used to better
understand a subject

Types of Data

Structured

Unstructured

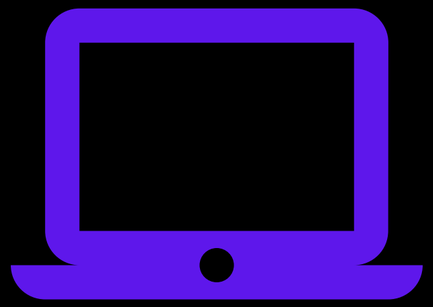


"It all starts with a **Dataset**"

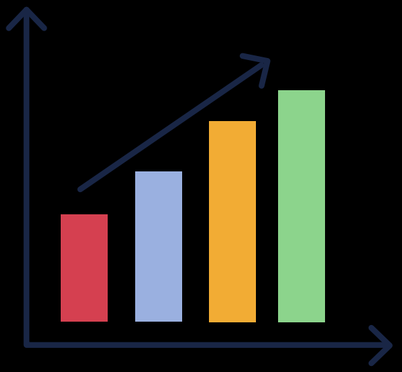
- Unknown Data Scientist

	PassengerId	Survived	Pclass	Name	Sex	Age	SibSp	Parch	Ticket	Fare	Cabin
0	1	0	3	Braund, Mr. Owen Harris	male	22.0	1	0	A/5 21171	7.2500	NaN
1	2	1	1	Cumings, Mrs. John Bradley (Florence Briggs Th...	female	38.0	1	0	PC 17599	71.2833	C85
2	3	1	3	Heikkinen, Miss. Laina	female	26.0	0	0	STON/O2. 3101282	7.9250	NaN
3	4	1	1	Futrelle, Mrs. Jacques Heath (Lily May Peel)	female	35.0	1	0	113803	53.1000	C123
4	5	0	3	Allen, Mr. William Henry	male	35.0	0	0	373450	8.0500	NaN

Computer Science



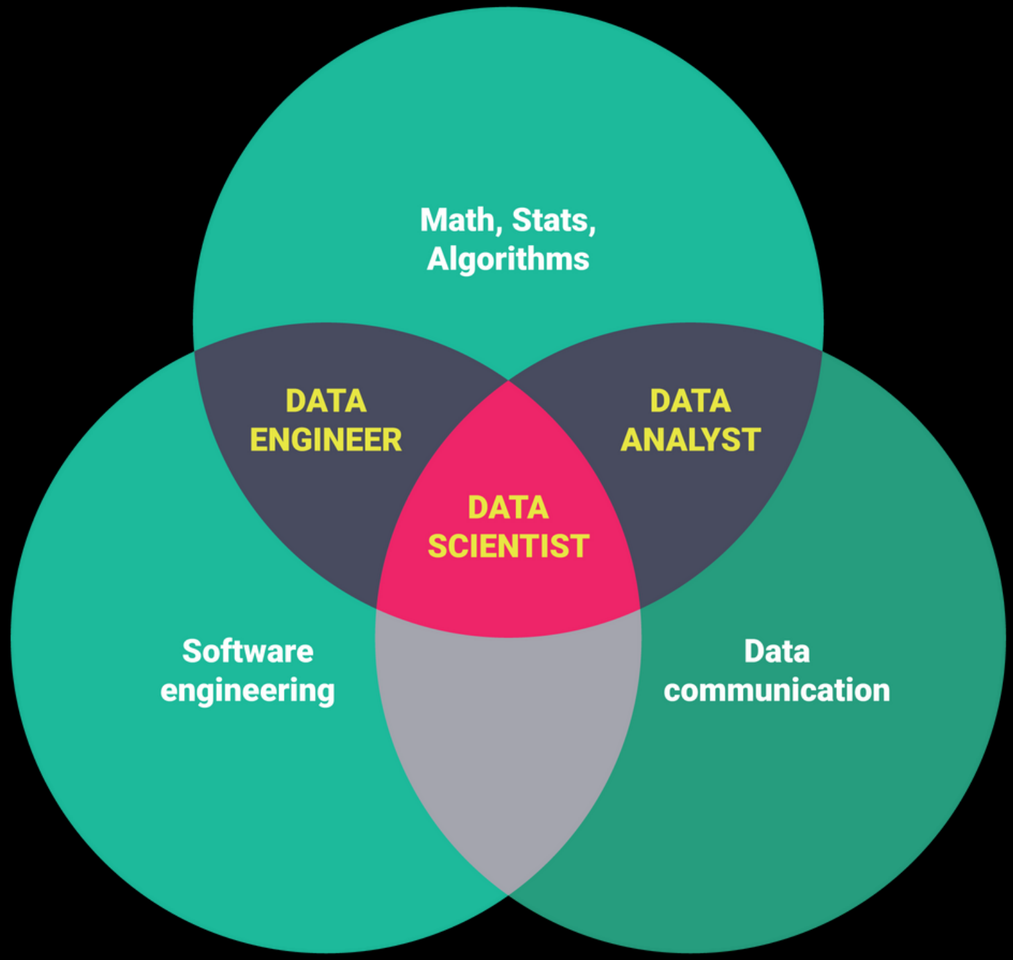
Math & Statistics



Data Science

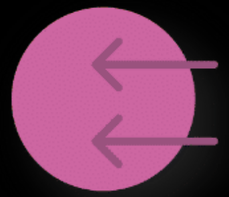
Domain Knowledge

Data Scientist, Engineer, Analyst ?



Workflow of a Data Science Project

Data Science Lifecycle



Data Acquisition

Data Pre-processing

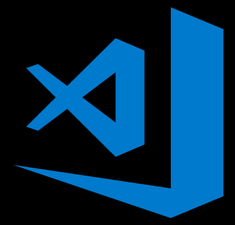
Model Building

Pattern Evaluation

Knowledge Representation

Tools that we'll be using

IDEs and Languages



Visual Studio Code

Some tips and resources

kaggle

ASK Question It's Free