

DATA SCIENCE BOOTCAMP

Session 1: Data Science Terms | Jupyter Notebook Demo

SESSION 1 OUTLINE

- **DATA SCIENCE TERMINOLOGIES**
- **MACHINE LEARNING AND ITS PROCESS**
- **SUPERVISED VS UNSUPERVISED
LEARNING**

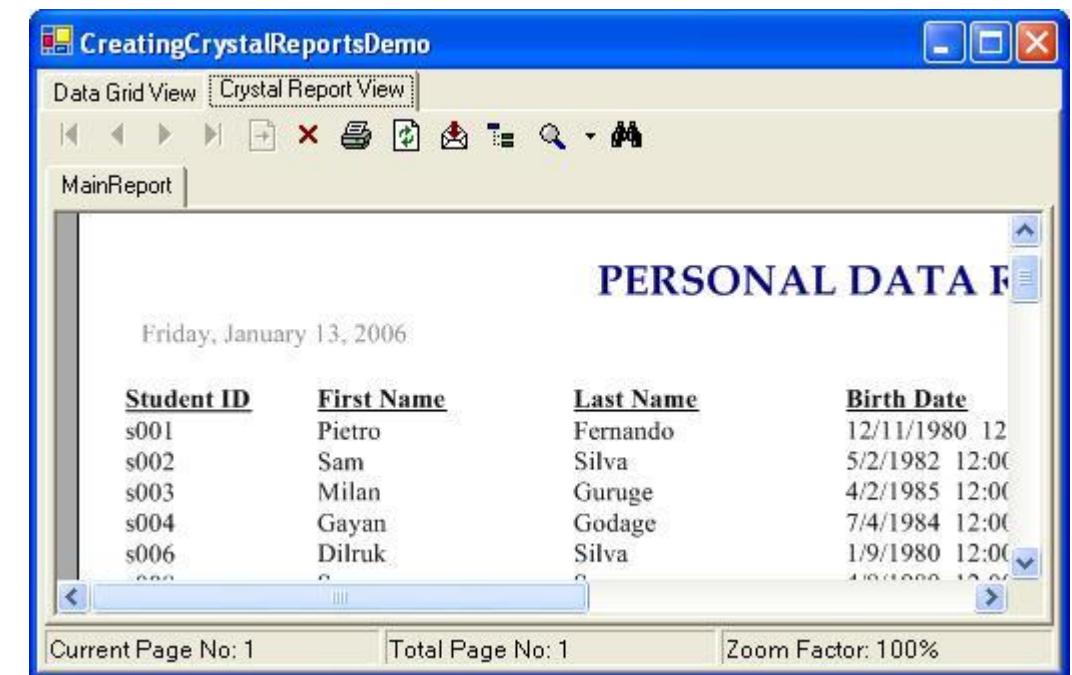
SESSION 1 OUTLINE

- **DATA SCIENCE TERMINOLOGIES**
- **MACHINE LEARNING AND ITS PROCESS**
- **SUPERVISED VS UNSUPERVISED
LEARNING**

DATA SCIENCE TERMINOLOGIES: REPORTS

The system reporting tool gives you the ability take periodic full, or partial, snapshots of the system, and save them for future reference.

DATA SCIENCE TERMINOLOGIES: REPORTS



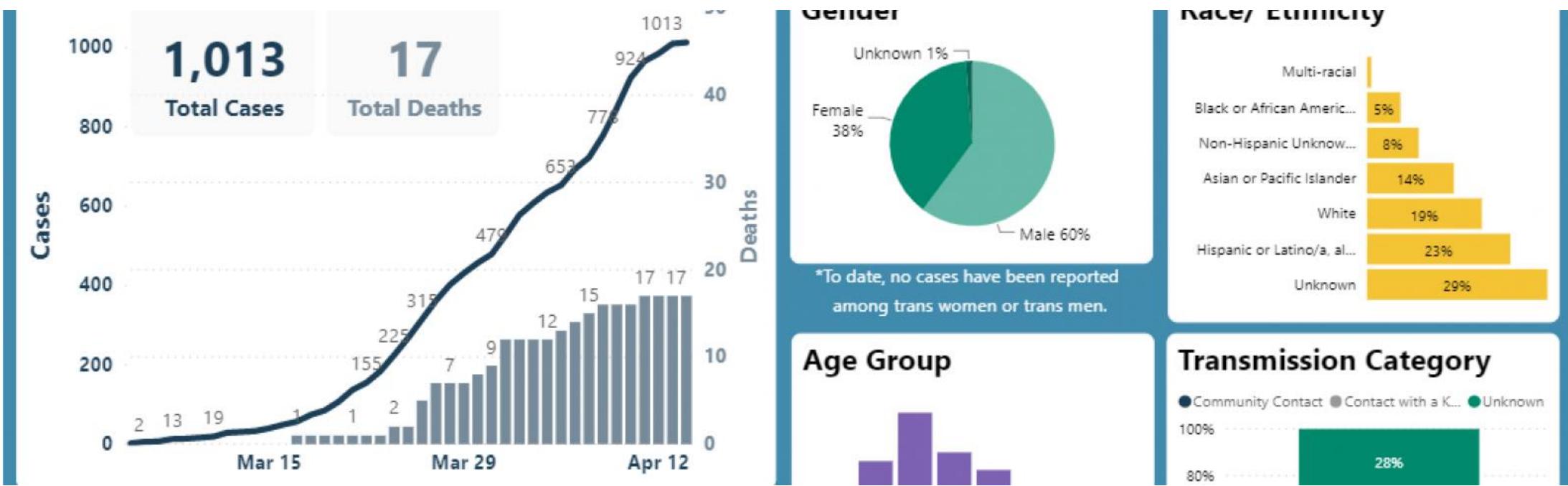
DATA SCIENCE TERMINOLOGIES: ANALYTICS

Analytics is the scientific process of discovering and communicating the meaningful patterns which can be found in data.

DATA SCIENCE TERMINOLOGIES: ANALYTICS

Analytics is the scientific process of discovering and communicating the meaningful patterns **which can be found in data.**

DATA SCIENCE TERMINOLOGIES: ANALYTICS



DATA SCIENCE TERMINOLOGIES: REPORTS VS ANALYTICS

ADMIN PANEL

Sales Report

From date: February 01 2018 To date: March 19 2018

Search

Order No.	Customer Name	Cashier Name	Order Type	Payment	Date
51	Zehab	Zehab	Collection	Cash	March 19 2018
21	03338613392	Zehab	Dine In	Cash	March 19 2018
31	Pasha 0333 861...	Zehab	Dine In	Cash	March 19 2018

Total Amount: 4132.5 Generate Report



From historical data to insights via Dashboard



DATA SCIENCE TERMINOLOGIES: BIG DATA

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so large size and complexity that none of traditional data management tools can store it or process it efficiently.

DATA SCIENCE TERMINOLOGIES: BIG DATA

Big Data is a collection of data that is huge in volume, yet growing exponentially with time. It is a data with so **large size and complexity that none of traditional data management tools** can store it or process it efficiently.

DATA SCIENCE TERMINOLOGIES: 6Vs BIG DATA

The six Vs of big data

Big data is a collection of data from various sources, often characterized by what's become known as the 3Vs: *volume, variety and velocity*. Over time, other Vs have been added to descriptions of big data:

VOLUME	VARIETY	VELOCITY	VERACITY	VALUE	VARIABILITY
The amount of data from myriad sources. 	The types of data: structured, semi-structured, unstructured. 	The speed at which big data is generated. 	The degree to which big data can be trusted. 	The business value of the data collected. 	The ways in which the big data can be used and formatted. 

DATA SCIENCE TERMINOLOGIES:

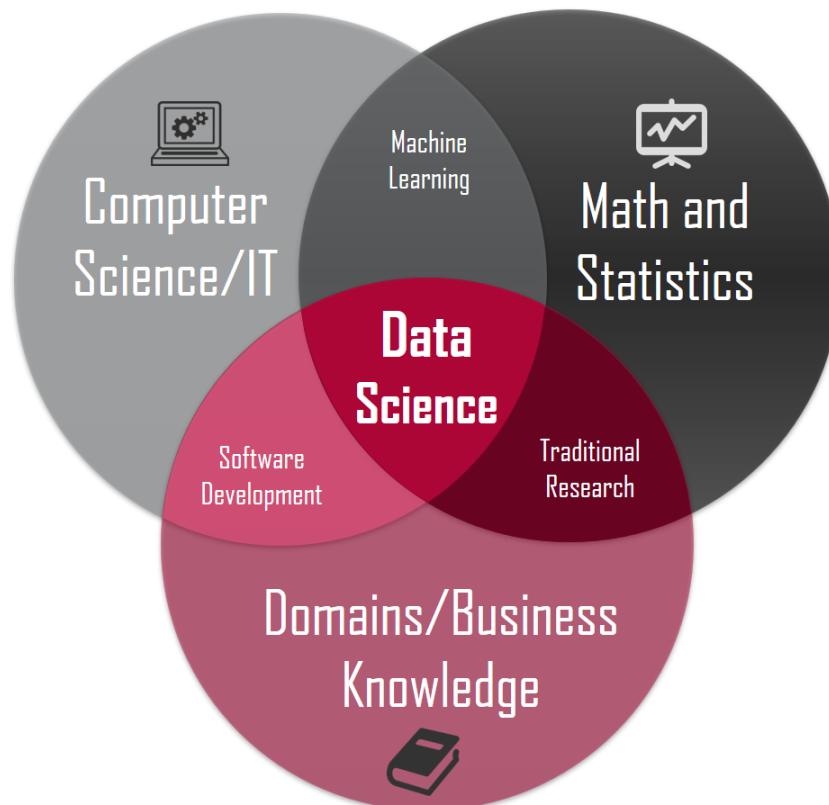
DATA SCIENCE

Data science is a field of applied mathematics and statistics that provides useful information based on large amounts of complex data or big data.

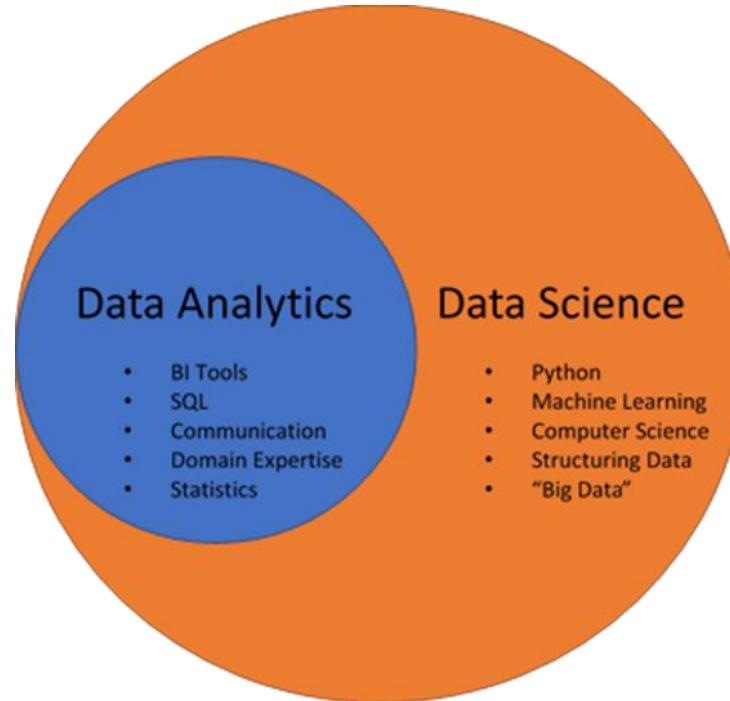
DATA SCIENCE TERMINOLOGIES: DATA SCIENCE

Data science is a field of applied **mathematics and statistics** that provides useful information based on large amounts of complex data or big data.

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE



DATA SCIENCE TERMINOLOGIES: DATA SCIENCE VS ANALYTICS



ANALYTICS

INSIGHTS FROM THE PAST



DATA SCIENCE

UTILIZING DATA FROM THE PAST FOR FUTURE

Google Images: "analytics vs data science"

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE VS ANALYTICS

DESCRIPTIVE ANALYTICS

WHAT HAPPENED?

PREDICTIVE ANALYTICS

WHAT WILL HAPPEN?

DIAGNOSTIC ANALYTICS

WHY IT HAPPENED?

PRESCRIPTIVE ANALYTICS

WHAT SHOULD HAPPEN?

ANALYTICS

INSIGHTS FROM THE **PAST**

DATA SCIENCE

UTILIZING DATA FROM THE PAST FOR **FUTURE**



DATA SCIENCE TERMINOLOGIES: DATA SCIENCE VS ANALYTICS

DESCRIPTIVE ANALYTICS

SITE DOWN!!
(WHAT HAPPENED?)

DIAGNOSTIC ANALYTICS

SITE TRAFFIC – MANY SERVICE CONNECTED
(WHY IT HAPPENED?)

ANALYTICS
INSIGHTS FROM THE **PAST**



PREDICTIVE ANALYTICS

NOTIFICATION ALERT: SITE A WILL GO DOWN
(WHAT WILL HAPPEN)

PRESCRIPTIVE ANALYTICS

RECOMMENDED ACTION: CHECK THE SERVICE
CONFIGURATION LIMIT
(WHAT SHOULD HAPPEN?)

DATA SCIENCE

UTILIZING DATA FROM THE PAST FOR **FUTURE**



Google Images "Cellsit Internet"

DATA SCIENCE TERMINOLOGIES: ANALYTICS – BUSINESS INTELLIGENCE

Business intelligence combines business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations make more data-driven decisions.

DATA SCIENCE TERMINOLOGIES: ANALYTICS – BUSINESS INTELLIGENCE

Business intelligence combines business **analytics**, **data mining**, **data visualization**, **data tools and infrastructure**, and best practices to help **organizations** make more data-driven decisions.

DATA SCIENCE TERMINOLOGIES: ANALYTICS – BUSINESS INTELLIGENCE

Business intelligence combines business analytics, data mining, data visualization, data tools and infrastructure, and best practices to help organizations make more data-driven decisions.

Tools: Tableau, PowerBI, Google Data Studio

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – STATISTICAL METHODS

Statistical methods are mathematical formulas, models, and techniques that are used in statistical analysis of raw research data.

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – STATISTICAL METHODS

Statistical methods are mathematical formulas, models, and techniques that are used in statistical analysis of raw research data.

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – STATISTICAL METHODS

Demo: Excel

$$y=mx+b$$

Where:

m = slope

b = intercept

x = independent

y = dependent

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – MACHINE LEARNING

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – MACHINE LEARNING

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms **to imitate the way that humans learn**, gradually improving its accuracy.

MOVE **OOK** **BACK**

JANE ♀

HP

TARZAN ♂

HP

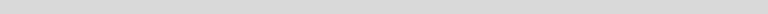
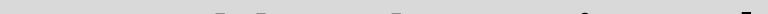
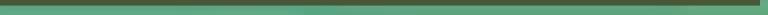
What will TARZAN do?
Biology 101 REVISITED! COMPUTER
TARZAN do?

► Give a ROCK.
Give a FLOWER.

MOVE **AOK** **BBACK**

JANE ♀

HP



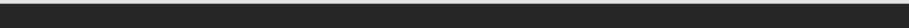
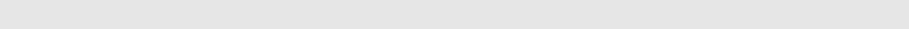
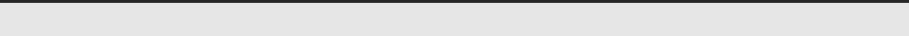
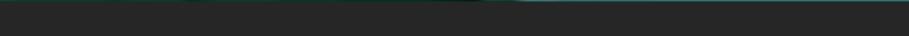
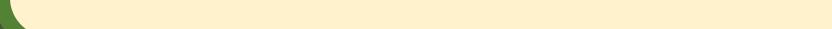
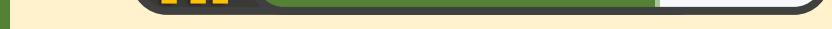
TARZAN does something that gives her a surprise!

▶ Give a ROCK.
Give a FLOWER.



TARZAN ♂

HP



+MOVE AOK BACK

JANE ♀

HP



TARZAN ♂

HP

~~TARZAN~~ (COMPUTER) gives her
TARZAN do?

Give a ROCK.
► Give a FLOWER.

+MOVE AOK BACK

JANE ♀

HP



TARZAN! Come here! What does TARZAN do?



TARZAN ♂

HP

▶ Catch a SNAKE
Give a CUTE RABBIT.

MOVE **AOK** **BBACK**

JANE ♀

HP



TARZAN ♂

HP

TARZAN (COMPUTER) gives her
TARZAN do?

Catch a SNAKE
► Give a CUTE RABBIT.

+MOVE AOK BACK

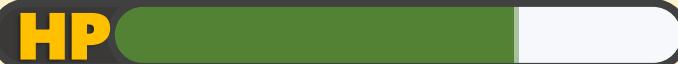
JANE ♀

HP



TARZAN ♂

HP



TARZAN! yells at
TARZAN do?

► YELLS
SMILE



MOVE **OOK** **BACK**

JANE ♀

HP



TARZAN ♂

HP

TARZAN (COMPUTER) smiles at
TARZAN do?

YELLS
► **SMILE**

+MOVE ↗OK ⏪BACK

TARZAN learns
how to win
JANE's heart



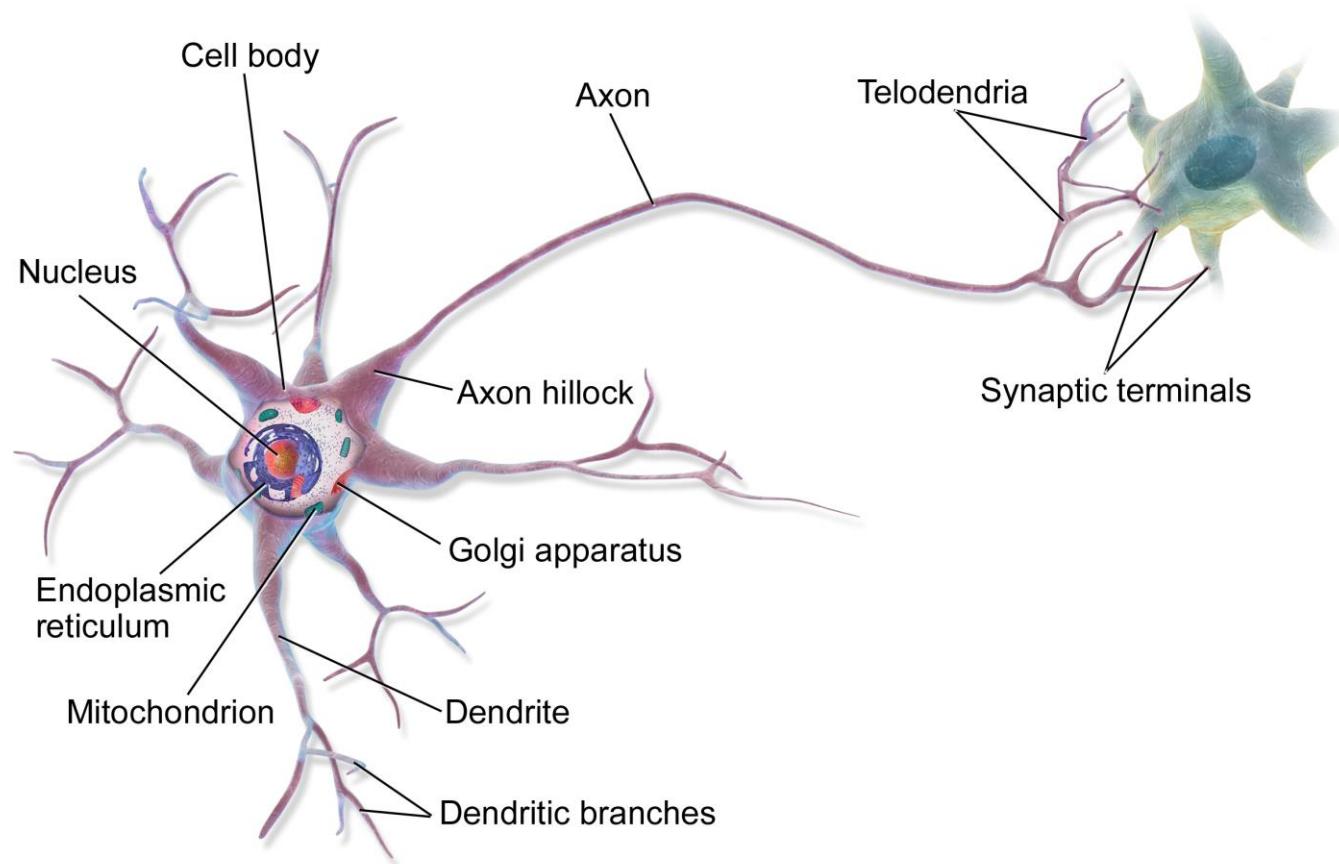
DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – DEEP LEARNING

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain called artificial neural networks.

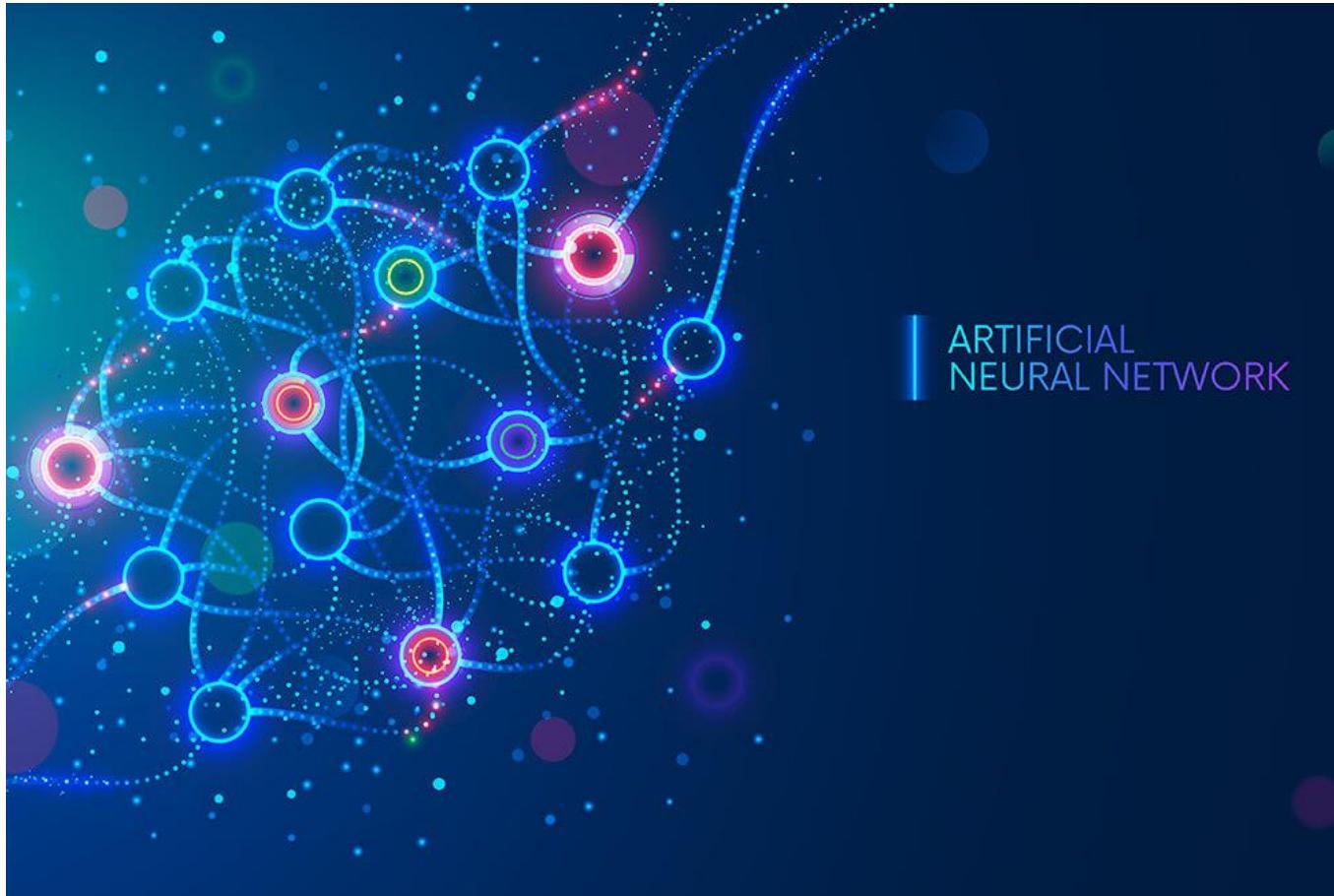
DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – DEEP LEARNING

Deep Learning is a subfield of machine learning concerned with algorithms inspired by the **structure and function of the brain** called artificial neural networks.

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – NEURAL NETWORKS



DATA SCIENCE TERMINOLOGIES: DATA SCIENCE – NEURAL NETWORKS



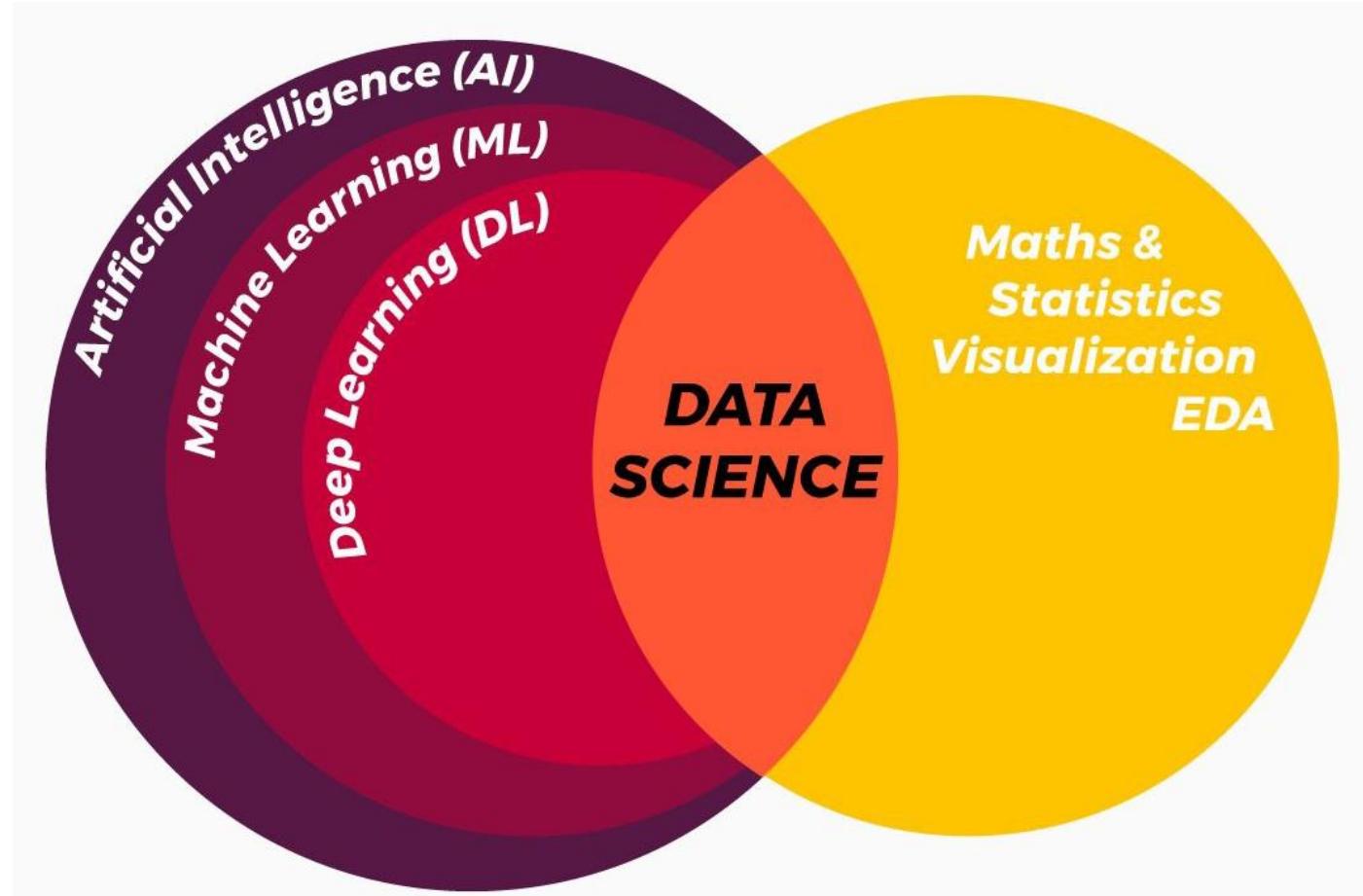
DATA SCIENCE TERMINOLOGIES: DATA SCIENCE VS ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE VS ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI), the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings

DATA SCIENCE TERMINOLOGIES: DATA SCIENCE VS ARTIFICIAL INTELLIGENCE





JUPYTER: JULia, PYthon, and R

CARE AND SHARE!

DATA SCIENCE BOOTCAMP

Session 1: Data Science Terms | Jupyter Notebook Demo