LEARNING PROGRESS REVIEW

WEEK 1

Syntax Group

TABLE OF CONTENT

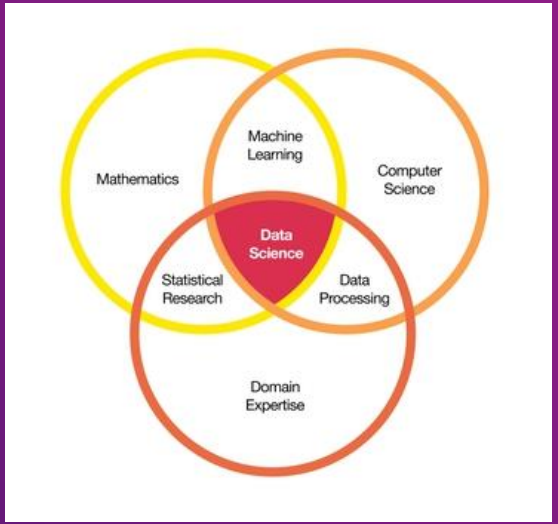
01

INTRODUCTION AND DATA SCIENCE METHODOLOGY

let's get started!

WHAT IS DATA SCIENCE?

is a skill to cultivate or process data to get insights or prediction algorithms with the combination between Mathematics, Computer Science, and Business Understanding or Domain Knowledge



Data Science Interdisciplinary Field
Source: exposit.com

WHAT ARE THE DATA SCIENTIST WORK?

DECISION SUPPORT

Report and Dashboard

SCORING

Algorithm

RECOMMENDATION

Problem Solving

CLASSIFICATION

Further Category Prediction

FORECASTING

Prediction

DETECTION

Anticipation

HARD SKILLS FOR A DATA SCIENTIST?

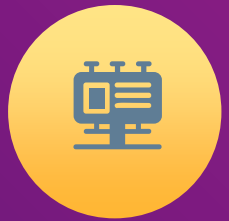
SQL & DATABASE

APPLIED STATISTICS

PYTHON & R

DATA VISUALIZATION & BUSINESS INTELLIGENCE

MACHINE LEARNING & DEEP LEARNING



SOFT SKILLS?

DATA STORYTELLING

CRITICAL THINKING

COMMUNICATION

TEAMWORK

DATA SCIENCE METHODOLOGY

1. Data Processing
2. Discovering Insights
3. Building Prediction Algorithms
4. Framework Data Science Methodology

DATA PROCESSING

DEFINE THE PURPOSE OF DATA PROCESSING

01

02

DATA COLLECTION

03

DATA COMBINED

04

DATA ANALYSIS

DISCOVERING INSIGHTS

DESCRIPTIVE ANALYSIS

DIAGNOSTICS

STATISTICAL TEST

BUILDING PREDICTION ALGORITHMS

- Based on the relationship between the data.
- Classical statistics, machine learning, and human-centered modeling approaches.
- Can be implemented in website-based applications, mobile, and ERP systems.

FRAMEWORK DATA SCIENCE METHODOLOGY

- Business Understanding
- Analytics Approach
- Data Requirements
- Data Collection
- Data Understanding
- Data Preparation
- Modelling
- Evaluation
- Deployment
- Feedback

02

INTRODUCTION TO DATA & DATABASE

let's get acquainted

WHAT IS DATA?

Data is a collection of facts obtained from measurements or observations that provide an overview of a situation. They are 2 types of data, quantitative and qualitative.

DATA BASIS

is an organized collection of structured information, or data, typically stored electronically in a computer system. The advantages of using a database are:

- Speed
- Space
- Security
- Shareability
- Standardization
- Accuracy
- Consistent

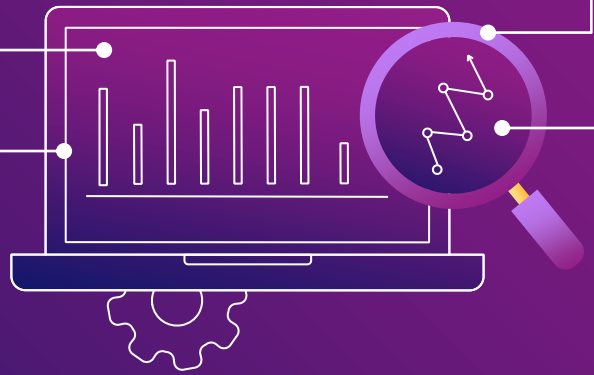
RELATIONAL DATABASE MANAGEMENT SYSTEM (RDBMS)

"...is a type of database management system (DBMS) that uses a relational model for organizing and managing data."

RDBMS DATABASE

MySQL

Oracle DB



PostgreSQL

SQL Server

THE ADVANTAGES OF RDBMS

DATA DEFINITION

DATA RETRIEVAL

DATA UPDATION

USER ADMINISTRATION

FUNCTION OF RDBMS

DATA DICTIONARY

DATA STORAGE

DATA TRANSFORMATION AND PRESENTATION

SECURITY MANAGEMENT

MULTIUSER ACCESS CONTROL

BACKUP AND RECOVERY MANAGEMENT

DATA INTEGRITY

DATABASE COMMUNICATION

INTERFACES

TRANSACTION MANAGEMENT

TYPES OF RDBMS

OPERATIONAL DATABASE

JSON and XML

RELATIONAL DATABASE

MySQL, PostgreSQL, and Oracle DB.

DATA WAREHOUSE

SQL Server and Teradata

END-USER DATABASE

SQLite

DISTRIBUTED DATABASE

AWS

STRUCTURED QUERY LANGUAGE (SQL)

is a programming language used for managing and manipulating relational databases

DATA DEFINITION LANGUAGE (DDL)

CREATE; RENAME;
ALTER; DROP

DATA MANIPULATION LANGUAGE (DML)

SELECT; INSERT;
UPDATE; DELETE

DATA TYPES IN SQL

NUMERIC

TINYINT; SMALLINT;
MEDIUMINT; INT

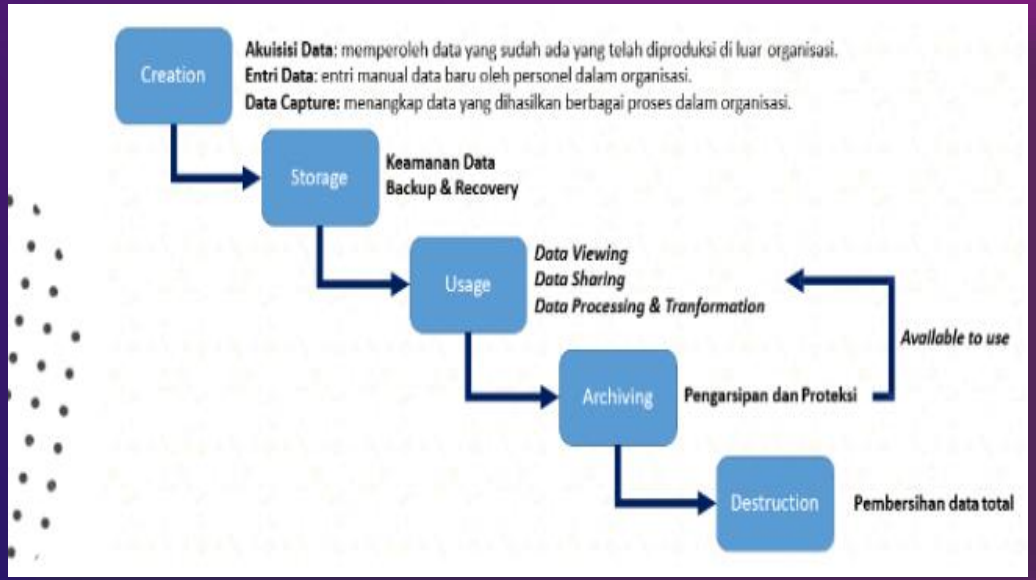
DATE AND TIME

DATE; TIME;
DATETIME; YEAR

STRING

CHAR; VARCHAT;
TEXT

DATA LIFECYCLE MANAGEMENT (DLM)



Data Lifecycle Management Step
Source: digitalskola.com

DEFINITION

Data lifecycle management is a model of organizing the cycles and flow of data within a system

ADVANTAGE

- Meet data needs
- Data security or protection

DBEAVER

One of the popular SQL clients which is free and user-friendly, it's also desktop-based.

SYNTAX GROUP

Aldrich Alfatera Unpapar

Wa Ode Sukmasarny Musdigaswati

Patma Oktaviana

Otniel Sukma Priyambodo

Jason Tadeus

Thank You