COVID-19 Vaccination Database Project

Project Overview
This project demonstrates the end-to-end design and implementation of a relational database system using the Our World in Data COVID-19 vaccination dataset. The goal was to clean and organise real-world data, apply database design principles, and extract meaningful insights through SQL queries. **Steps Undertaken**
- **Data Acquisition:** Collected CSV files from the Our World in Data (OWID) GitHub repository, which provide vaccination statistics globally, including doses administered, vaccine types, and demographic data.
- **Database Design:** Created an Entity-Relationship Diagram (ERD) and defined clear assumptions for the schema.
- **Normalisation:** Applied First, Second, and Third Normal Forms (1NF, 2NF, 3NF) to eliminate redundancy and ensure data integrity.
- **Schema Implementation:** Developed the schema in SQLite, including primary keys, foreign keys, and integrity constraints.
- **Data Import:** Populated the database (Vaccinations.db) with updated CSV data after cleaning and restructuring.
- **SQL Queries:** Implemented and executed multiple queries to analyse vaccination progress, including:
• Countries with daily vaccination rates above average.
• Countries with cumulative doses above average.
• Vaccine types used per country.
• Maximum vaccinations by reporting source.
• Weekly vaccination comparisons for Australia, Germany, England, and France.
Key Deliverables
- Entity-Relationship Model and assumptions.
- Normalised relational schema.
- SQL scripts for schema creation and queries.
- SQLite database file (Vaccinations.db).
Frameworks Applied
- Database Normalisation (1NF, 2NF, 3NF).
- Relational Database Theory (ERD, schema mapping).
- SQL Standards (DDL, DML, joins, aggregates).
Skills Demonstrated
- Data modelling and database design.
- SQL schema creation, queries, and optimisation.
- Data cleaning, import, and integrity enforcement.
- Communication of technical work in a professional format.