# **Project Title: Python and SQL Integration for Data Analysis**

## **Description**

This project demonstrates how to integrate Python with SQL for performing data retrieval, storage, and analysis. Using Python libraries like pandas and sqlite3, data is fetched from an SQL database, manipulated in Python, and visualized using matplotlib. The goal is to automate the process of analyzing order statuses and derive insights into order frequency.

## **Prerequisites**

Ensure you have the following installed:

* **Python 3.6 or later**: You can download it from [Python's official website](https://www.python.org/downloads/).
* **SQLite3**: Most Python installations come with SQLite built-in.
* **Libraries**:
  + pandas (pip install pandas)
  + sqlite3 (comes with Python)
  + matplotlib (pip install matplotlib)

## **Project Structure**

The project consists of the following steps:

1. **Database Connection**:
   * A connection is established to an SQLite database containing order data using Python’s sqlite3 module.
2. **Data Retrieval**:
   * SQL queries are executed using Python to fetch data for analysis.
3. **Data Cleaning**:
   * The dataset is cleaned, handling missing values, removing duplicates, or filtering based on specific conditions.
4. **Data Visualization**:
   * A bar chart is generated using matplotlib to represent the frequency of different order statuses.
5. **Analysis and Results**:
   * Based on the data visualization, it was observed that over 65% of the orders were marked as "Delivered."
6. **Conclusion**:
   * The integration of SQL with Python enhances data analysis capabilities by allowing efficient data retrieval, manipulation, and visualization. It also facilitates automation of database tasks.

### **Benefits of Python-SQL Integration:**

* **Data Retrieval**: Fetch specific data for analysis.
* **Data Storage**: Store processed data into databases.
* **Data Manipulation**: Update or delete records based on analysis.
* **Automation**: Automate data-driven workflows and reports.

## **How to Run the Project**

1. Clone the repository or download the project files.
2. Ensure all dependencies are installed by running
3. Open the Jupyter notebook (.ipynb file).
4. Run each cell to execute the project step-by-step.