

Project Report: Data Analysis and Visualization Using SQL, Python, and Power BI (**E-commerce Return Rate Reduction Analysis**)

1. Introduction

The objective of this project is to demonstrate the complete data workflow—starting from data extraction using SQL, performing data manipulation and analysis using Python, and finally visualizing insights through an interactive Power BI dashboard.

2. Tools & Technologies Used

- SQL (MySQL/Microsoft SQL Server): For data storage and query operations
- Python: For data analysis, processing, and integration with SQL
- Power BI: For interactive dashboard creation and data visualization
- Libraries in Python:
 - mysql.connector: To connect with the SQL database
 - pandas: For data manipulation
 - matplotlib & seaborn: For visualizations (if used)

3. Project Workflow

Step 1: SQL Database Design and Management

- Created and managed a relational database using SQL.
- Developed multiple tables with proper relationships and constraints.
- Inserted sample data to simulate real-world scenarios.
- Performed operations like SELECT, JOIN, GROUP BY, ORDER BY, etc., to retrieve meaningful insights.

Step 2: Python Integration with SQL

- Established a connection between Python and SQL using mysql.connector.
- Retrieved data from SQL tables and loaded it into Pandas DataFrames.
- Performed data cleaning, transformation, and exploratory data analysis in Python.
- Script saved in a Jupyter Notebook (Project_1.ipynb) which includes:
 - Database connection setup

- Query execution
- Data fetching and processing
- Optional charting using Python libraries

Step 3: Power BI Dashboard Creation

- Imported the cleaned and transformed dataset into Power BI.
- Designed a professional dashboard to visualize key metrics.
- Used Power BI features like:
 - Slicers
 - Pie charts, bar graphs, and line charts
 - Drill-down capabilities
- The Power BI file (Project_5.pbix) contains a dynamic and user-friendly dashboard for decision-making.

4. Outcomes & Learnings

- Gained practical knowledge of working with relational databases.
- Learned how to integrate Python with SQL to automate data pipelines.
- Developed hands-on experience in creating insightful dashboards using Power BI.
- Improved understanding of the full data lifecycle—from raw data to business insights.

5. Conclusion

This project showcases the integration of data engineering and analytics skills. By connecting SQL with Python and visualizing the outcomes in Power BI, a comprehensive data pipeline was built. This can be effectively used in real-time business environments to drive data-driven decisions.

6. Attachments

- Project_1.ipynb: Python Jupyter Notebook file showing SQL connection and data processing
- Project_1 Python work using SQL Database.pdf: Documentation of Python-SQL work
- Project_5.pbix: Power BI Dashboard file