



START HERE

Real Data Analyst Project Kit

Python + PostgreSQL + SQL + Power BI Workflow

Welcome

Thank you for purchasing the Real Data Analyst Project Kit.

This project was designed to help beginners learn a real-world data analytics workflow using:

- Python
- pandas
- PostgreSQL
- SQL
- Power BI

In this project, you will start with messy raw data, clean it using Python, upload the cleaned data into PostgreSQL, analyze business questions using SQL, and build a professional Power BI dashboard.

This project follows the same workflow many real data analysts use in business environments.

What's Included

This project kit includes:

- ✓ Raw messy dataset
 - ✓ Cleaned dataset
 - ✓ Python notebooks/scripts
 - ✓ PostgreSQL setup SQL
 - ✓ SQL business analysis queries
 - ✓ Power BI dashboard file
 - ✓ Dashboard assets/icons
 - ✓ Interview questions & answers
 - ✓ Portfolio guidance documents
 - ✓ SQL cheat sheet
 - ✓ Common errors & fixes guide
-

Recommended Workflow

Follow the project in this order:

Step 1 — Explore the Raw Dataset

Folder:

01_Raw_Dataset

Start by reviewing the messy dataset.

You will notice:

- Missing values
- Duplicate records
- Inconsistent country names
- Invalid or missing dates

This simulates real-world data problems analysts encounter frequently.

Step 2 — Find Data Issues Using Python

Folder:

03_Python_Code

File:

01_find_data_issues.ipynb

This notebook/script demonstrates how to:

- Identify missing values
 - Find duplicates
 - Explore inconsistent values
 - Review date problems
-

Step 3 — Clean the Data

Folder:

03_Python_Code

File:

02_clean_data.ipynb

This notebook/script demonstrates how to:

- Fill missing values
- Standardize country names
- Remove duplicates
- Fix invalid dates
- Export a cleaned CSV file

At the end of this step, you will generate:

cleaned_ecommerce_data.csv

Step 4 — Upload Data into PostgreSQL

Folder:

03_Python_Code

File:

03_upload_to_postgresql.ipynb

This notebook/script demonstrates how to:

- Connect Python to PostgreSQL
 - Create a database connection
 - Upload the cleaned dataset into SQL
-

Step 5 — Create the SQL Table

Folder:

04_SQL_Code

File:

01_create_database_and_table.sql

This SQL script creates the PostgreSQL table used in the project.

1.1_Verify Data and table

This will verify if the data has been uploaded in the table.

Step 6 — Answer Business Questions with SQL

Folder:

04_SQL_Code

File:

02_business_questions.sql

This file contains SQL queries used to answer business questions such as:

- Total revenue
- Top-selling products
- Sales by country
- Monthly sales trends

- Total orders
-

Step 7 — Open the Power BI Dashboard

Folder:

05_PowerBI_Dashboard

File:

ecommerce_dashboard.pbix

This Power BI dashboard demonstrates how to:

- Create KPI cards
 - Build charts
 - Use slicers
 - Create interactive visuals
 - Connect Power BI to SQL or CSV data
-

Step 8 — Review Dashboard Assets

Folder:

06_Dashboard_Assets

This folder contains:

- KPI icons
 - Dashboard graphics
 - Visual assets used in the dashboard
-

Step 9 — Review Interview Questions

Folder:

08_Interview_Questions

This folder contains:

- Common interview questions
- Example answers
- Technical explanations
- Business-focused responses

This section is helpful for:

- Interviews
- Resume preparation
- Portfolio discussions

Important Tools Needed

To complete this project, you should install:

Python

Recommended:

- VS Code
- Jupyter Notebook

PostgreSQL

Includes:

- PostgreSQL database
- pgAdmin 4

Power BI Desktop

Download:

- Microsoft Power BI Desktop

Beginner Notes

If you are new to these tools:

- Follow the project step-by-step
- Run code sections one at a time
- Read comments inside the notebooks/scripts
- Troubleshoot errors slowly and carefully

Remember:

Learning data analytics is a process.

Common Beginner Mistakes

Some common issues include:

- Incorrect PostgreSQL passwords
- Missing Python libraries
- Wrong file paths
- Power BI refresh problems

If you encounter issues, review:

COMMON_ERRORS_AND_FIXES

inside the Project Guide folder.

Portfolio & Career Tips

This project can be used for:

- Resume projects
- LinkedIn projects
- GitHub portfolios
- Interview discussions
- Data analyst practice

For more guidance, review:

HOW_TO_USE_THIS_PROJECT_IN_YOUR_PORTFOLIO

You can find this file in 07_Project_Guide.

Final Thoughts

This project demonstrates a complete data analytics workflow from raw data to dashboard reporting.

By completing this project, you will gain experience with:

- Data cleaning
- SQL databases
- Business analysis
- Dashboard development
- End-to-end analytics workflows

Thank you again for supporting Data Geek Is My Name.

Keep learning and building.