**Retail Orders Analysis using Python and SQL**

**By: Hussain Muslim**

**Introduction**

This project analyzes retail orders to derive actionable insights using Python and SQL. The goal was to understand sales trends, top-performing products, and profit patterns, showcasing end-to-end data analysis workflow.

**Tech Stack**

* **Python**: Data extraction and preprocessing.
* **SQL**: Data modeling and querying.
* **Libraries**: Kaggle, pandas.

**Dataset**

* **Name**: Retail Orders
* **Source**: Kaggle
* **Description**: This dataset contains sales records for a fictional retail business, including order details, shipping modes, customer segments, and product categories.

**Workflow**

1. **Data Acquisition**:
   * Downloaded the dataset using the Kaggle API in Python.
   * Extracted and loaded the CSV into a SQL table for analysis.
2. **Data Modeling**:
   * Created the df\_orders table in SQL with fields like order\_id, ship\_mode, region, sale\_price, and profit.
3. **Analysis**:
   * Queried for key metrics:
     + Total sales by region and product category.
     + Top 10 revenue-generating products.
     + Average discount by category.
4. **Insights**:
   * Identified regions with high-profit margins.
   * Highlighted product categories driving the most revenue.

**Key SQL Queries**

1. **Top 10 Revenue-Generating Products**:

SELECT product\_id, SUM(sale\_price) AS total\_revenue

FROM df\_orders

GROUP BY product\_id

ORDER BY total\_revenue DESC

LIMIT 10;

1. **Profit Analysis by Region**:

SELECT region, SUM(profit) AS total\_profit

FROM df\_orders

GROUP BY region

ORDER BY total\_profit DESC;

Thank you.

**LinkedIn:** [**https://www.linkedin.com/in/hussainmuslim/**](https://www.linkedin.com/in/hussainmuslim/)

[**Hussainsmk1@gmail.com**](mailto:Hussainsmk1@gmail.com)