



1. ETL Process:

* Extract: Extract data from CSV files (my new and old datasets).
* Transform: Normalize the data by cleaning it, creating dimension tables (e.g., Product, Customer, Store) and a fact table (Sales).
* Load: Load the transformed data into MySQL using Python MySQL connector.

2. Dimensional Modeling (Star Schema):

* Create Dimension Tables: Define tables to store descriptive data
* Create Fact Table: Create a Sales Fact Table to store transactional data like quantity sold, total sales, etc., with foreign keys to dimension tables.

3. Stored Procedures:

* Create Stored Procedures in MySQL: Write procedures to generate monthly reports (e.g., sales, revenue, average sales).

4. Load Data into Power BI/Tableau:

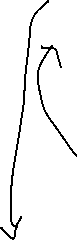
* Connect MySQL to Power BI/Tableau: Use connectors to link the database to your visualization tools.
* Visualize: Create dashboards and analyze sales trends, product performance, and store comparisons.

In Summary:

* ETL: Extract, Transform, and Load data into MySQL. ETL or ELT?
* Star Schema: Create dimension and fact tables.
* Stored Procedures: Generate reports in MySQL.
* Power BI/Tableau: Visualize data in interactive dashboards.

Sales Report based on Month and year and product :

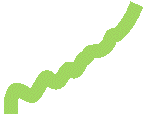
|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Product | Store | Month | Year | Quantity | Total Sales | Average sales |
| Clothing | Miami, FL | 4 | 2024 | Random | Random | Random |
| Home Applicances | Miami, FL | 4 | 2024 | Random | Random | Random |



Dataset 1

Dataset 2 generated by AI

Last step Visualize it

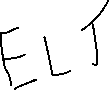


TO DO LIST :  
Start ERD

Walmart Dataset (two dataset one real life and one generated by AI) (10,000 rows with different types of category)



Mysql (normalising and creating tables )



Python (preprocessing my csv and loading the dataset)



Dimensional (Dimension and Fact table in my SQL) Star



Stored procedure to generate monthly sales report

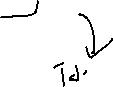


Tableau / Power BI