Low Level Design (LLD)

Analysing Expenditure

Written By:	Sohel Datta
Last Revised Date:	05-01-2025

Document Version Control:

Version	Date	Author	Description
1.0	04-01-2025	Sohel Datta	Initial draft of the document.
1.5	05-01-2025	Sohel Datta	Architecture & Architecture Description appended and updated

Contents

1. Introduction

1.1. Why this Low-Level Design Document?

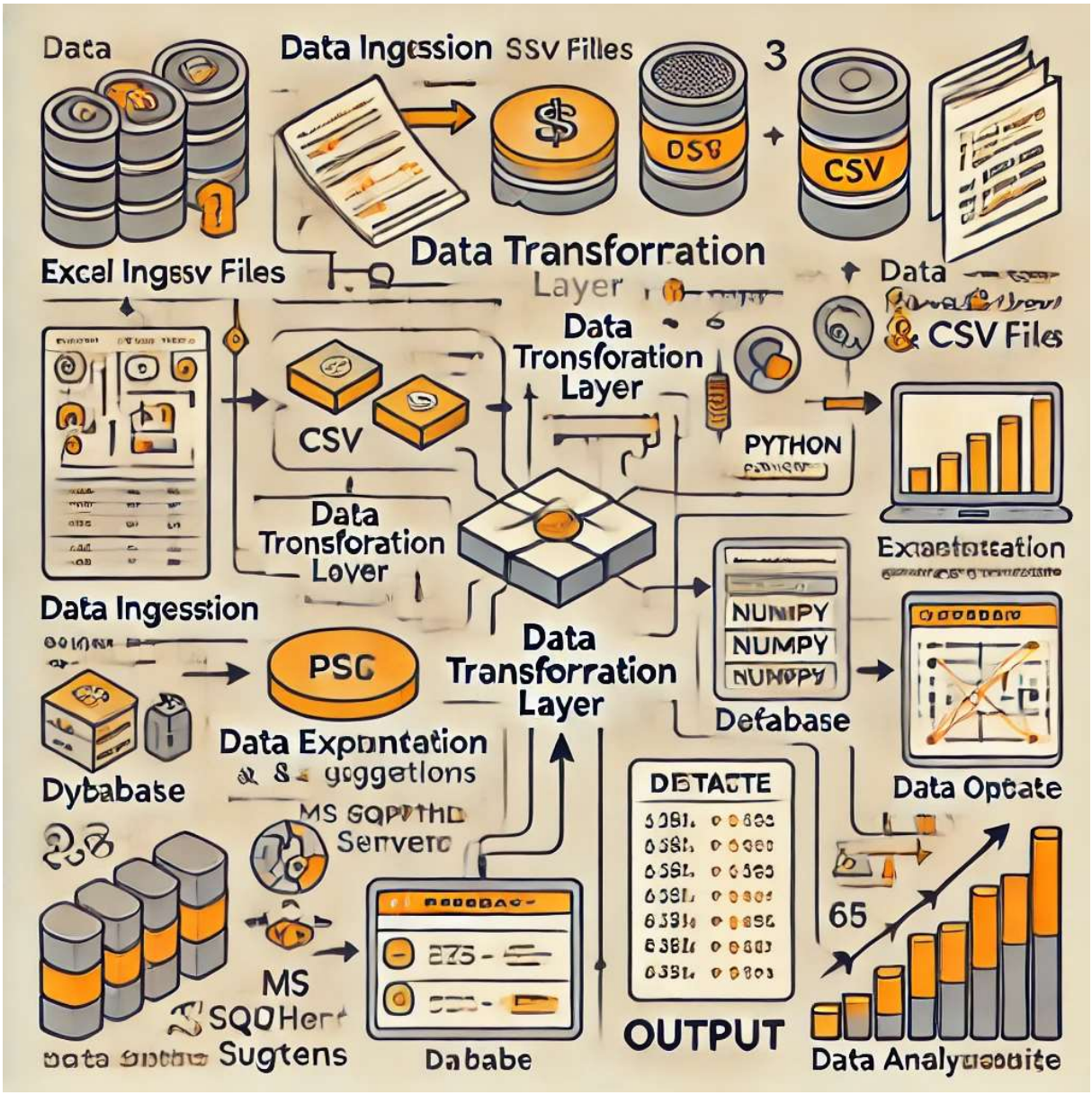
This LLD document serves as a detailed technical reference for implementing the expenditure analysis system. It specifies the components, methods, and workflows necessary for ensuring the successful deployment of the project.

1.2. Scope

The scope of this document includes:

- Data ingestion, transformation, and loading into a SQL server.
- Analytical models to identify cost patterns.
- Integration of tools for data visualization and reporting.

2. Architecture



3. Architecture Description

3.1. Data Description

- Source Files**: Salary Data.xlsx, Zomato Schema.xlsx, DA Assignment.xlsx.**
- Attributes**: Employee ID, Expenditure Category, Amount, Date, Vendor.**

- Volume**: Approximately 1GB of data.**

3.2. Data Transformation

- **Steps**:
  - Parse Excel files to extract structured data.
  - Handle missing values using imputation techniques.
  - Normalize fields (e.g., converting currency to USD).
  - Create derived columns (e.g., Monthly Expenditure).
- **Tools**:
  - Python: Pandas, NumPy.
  - Jupyter Notebook for testing transformations.

3.3. Data Insertion into Database

➢ **Database**: Microsoft SQL Server.

➢ **Table Schema**:

Field Name	Data Type	Description
EmployeeID	INT	Unique identifier for employees.
Category	VARCHAR (50)	Expenditure category.
Amount	FLOAT	Expense amount.
Date	DATE	Transaction date.
Vendor	VARCHAR (50)	Vendor associated with expens

3.4. Connection with SQL Server

Library: pyodbc.

Connection String:

```
connection = pyodbc.connect(
    'DRIVER={ODBC Driver 17 for SQL Server};'
    'SERVER=DESKTOP-SD7QIA4\\Sohel;'
    'DATABASE=ExpenditureDB;'
    'Trusted_Connection=yes;'
)
```

Insert Query:

INSERT INTO Expenditures (EmployeeID, Category, Amount, Date, Vendor)

```
VALUES (?, ?, ?, ?, ?)
```

3.5. Export Data from Database

- Export aggregated data for visualization using SQL queries.
- Output formats: CSV, JSON, or direct API integration for tools like Power BI.

3.6. Deployment

- Environment:
  - **Local**: Windows 11, Python 3.10, SQL Server 2019.
  - **Production**: Cloud-hosted SQL Server, Dockerized ETL pipeline.
- **Steps**:
  - Automate the pipeline with a scheduler (e.g., CRON, Airflow).
  - Deploy a dashboard using Power BI/Tableau.

4. Unit Test Cases

Test Case ID	Test Scenario	Expected Outcome
TC_01	Parse Excel file	Successfully parse and extract data
TC_02	Handle missing values	Missing values replaced with default values
TC_03	Normalize currency fields	Currency fields converted to USD
TC_04	Insert data into SQL database	Data inserted into ExpenditureDB without errors
TC_05	Query data for monthly expenditure	Accurate aggregated results

Test Case ID	Test Scenario	Expected Outcome
TC_06	Export data to CSV	CSV file generated with correct formatting