Subscription Box Analytics – Coursework Documentation

1. Business Description:  
The service allows users to subscribe to boxes containing products based on their interests (e.g., books, food, cosmetics). A one-time purchase or subscription for a specific period is available. Each user can specify their interests and assign an importance value from 1 to 10. Boxes are assembled with products, delivered to users, and accompanied by payments, reviews, and delivery statuses.

2. Database Structure:

2.1. OLTP:  
Contains the following entities:  
- Users  
- Interests  
- Users\_Interests (bridge between users and interests)  
- Boxes  
- Items  
- Box\_Items (items in the box)  
- Subscriptions  
- Payments  
- Deliveries  
- Reviews

2.2. OLAP:  
- Dimensions:  
 - dim\_Date  
 - dim\_Interest  
 - dim\_Customer (with SCD Type 2)  
 - dim\_Box  
- Facts:  
 - fact\_Subscription  
 - fact\_Delivery\_Performance  
- Bridge table:  
 - bridge\_Customer\_Interest

3. Purpose of OLTP and OLAP:

- OLTP – collects all business data and is used for daily operational queries.

- OLAP – through fact tables, you can determine which boxes are most popular, how activity changes over time, and how efficiently delivery is performed.

5. Script Execution Architecture:  
- Intermediate staging tables are used.  
- Data is loaded using the COPY command from CSV files located in the PostgreSQL system folder.  
- After loading, data is transferred from stg\_\* to the main OLTP tables.  
- From OLTP, data is moved to the OLAP structure through an ETL process.

7. How to Run Scripts:

7.1. How to Launch OLTP:  
- Transfer all CSV files to the PostgreSQL system folder (import) using the terminal.  
- Copy all scripts from files 1.1–1.5 into PgAdmin and execute each script sequentially.

Note:  
- File 1.5 contains queries for OLTP.

7.2. How to Launch OLAP:  
- Copy all scripts from files 1.1–1.4 into PgAdmin and execute each script sequentially.

Note:  
- File 1.4 contains queries for OLTP.

6. Tools:  
- PostgreSQL (pgAdmin)  
- Power BI Web  
- Excel (for storing exported CSVs)  
- Python (to merge all CSV files into one Excel file)  
- Terminal

7. Visualization (Power BI):  
Visual reports were created for:  
- Subscriptions over time  
- Delivery volume and delays  
- User interests  
- Box ratings  
Data models were configured, table relationships were established, and slicers were added for filtering by date, interest, and gender.