

SCENERIO 1

WC1 DIMENSIONAL MODEL

1. Business Requirement: The data to be ingested and transformed in this project will provide extensive coverage on daily product usage of clients (e.g. over- or under- usage) and provide the ability to generate relevant BI reports via Tableau for various teams internally.

2. Granularity: Lowest level is Day level
Therefore it will be: Per day, Per User, Per Account

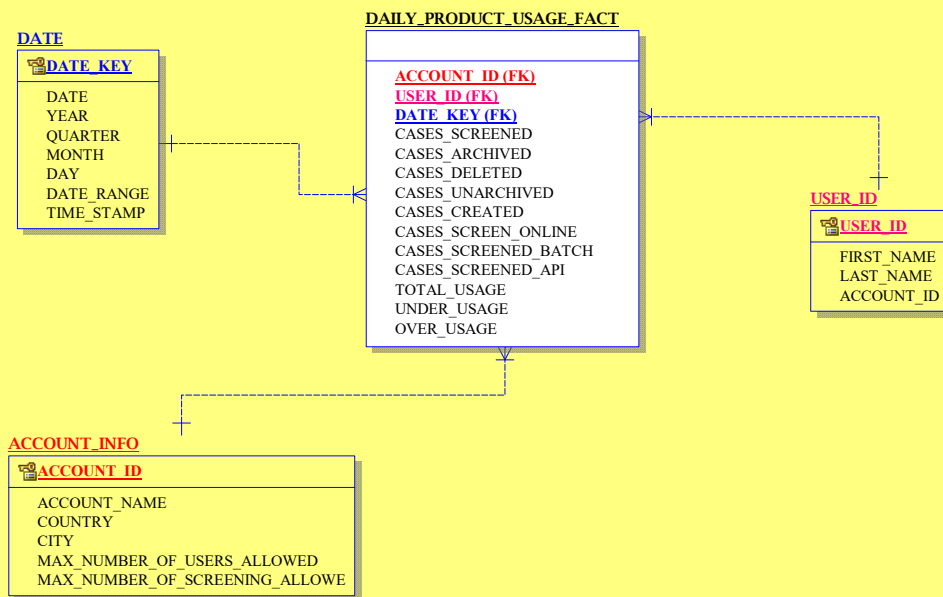
3. Fact and Dimension Tables:

- Fact Table: Daily Usage Summary
- Dimension Tables: User Information, Account Information, Date Dimension

4. Fact Table Design: DateID, User ID, Account ID, Cases Created, Cases Screened, Cases Screened - Online, Cases Screened - Batch, Cases Screened - API, Cases Archived, Cases Deleted, Cases Unarchived, Total Usage, Under Usage, Over Usage

5. Dimension Table Design:

- User Dimension: User ID, First Name, Last Name
- Account Dimension: Account ID, Account Name, Country, City, Max Number of Users Allowed, Max Number of Screenings Allowed
- Date Dimension: Date ID, Date, Year, Quarter, Month, Day, Date Range, Time Stamp



SCENERIO 2

Approach 1: Enhancing the Existing Design

1. Granularity Enhancement:

Modify the existing data ingestion process to capture data at a daily level rather than just monthly. This will provide more granular insights into usage patterns.

2. User and Usage Details:

Enhance the User Dimension table to include more detailed user information. Collect additional usage details, such as the type of screening (online, batch, API) and other relevant metrics.

3. Data Availability:

Adjust the data pipeline to ensure that data for the current month becomes available sooner, perhaps on a daily basis, rather than waiting until the 2nd of the following month.

4. Additional Metrics:

If users require additional metrics beyond what was initially designed, add those metrics to the Fact Table.

Pros:

- Maintains a single implementation, reducing complexity.
- Provides a more comprehensive view of usage data.
- Allows users to access more detailed and up-to-date information.

Cons:

- May require significant changes to the existing infrastructure.
- Could increase the complexity of the ETL processes and data management.
- Users may need to adapt to the updated dashboards and reporting.

Approach 2: Supporting Two Concurrent Implementations

1. Parallel Implementations:

Maintain the existing EDW solution with its limitations to serve users who are accustomed to it. Create a new, separate EDW solution based on the design from Interview Scenario #1 to meet the advanced needs of users.

2. Data Integration:

Develop a mechanism to synchronize data between the two implementations, ensuring consistency in data that overlaps.

3. User Segmentation:

Identify which users or teams require the advanced features and which can continue using the existing solution.

Pros:

- Existing users are not disrupted, and they can continue using the familiar interface.
- Advanced users get access to enhanced features and more detailed data.
- Flexibility to adapt to varying user needs.

Cons:

- Maintenance and synchronization between the two implementations can be complex.
- Increases infrastructure and operational costs due to maintaining two separate solutions.
- Potential confusion for users who need to choose between the two options.