**Data Warehousing Major Project**

**Group No. 4**

**Name: Roll No.**

**Nikhil Sachdeva 2016061**

**Raj Kamal Yadav 2016076**

**Rajat Bansal 2016260**

**“Problem Statement: Suggest hotels for users based on their recommendations and preferences.”**

**DIEM Schema Representation:**

**Step I: Building the DIEM Schema**

1. **The first decision is the choice of place.**
   1. **Uncertainties**
      1. Hotel Brand-Value
   2. **Action**
      1. Accessibility
   3. **Objective**
      1. Maximize Leisure Options
      2. Maximize Value for Expenditure
   4. **Object**
      1. :

Accessibility(User, Date, Choice of Place, Occupancy)

* + 1. :

Hotel Brand-Value(User, Date, Star Rating, Quality of Services, Reviews Rating)

* + 1. :
       1. Leisure Options
       2. Value for Expenditure

1. **The second decision is determining the expected overall expenditure based on the choice of Hotel.**
   1. **Uncertainties**
      1. Hotel Facilities
      2. Room Facilities
      3. Hotel Policies
      4. Dining Facilities
   2. **Action**
      1. Preferences
   3. **Objective**
      1. Minimize Expenditure
      2. Maximize Satisfaction
   4. **Object**
      1. :

Preferences(User, Date, Reservation Cost, Transport Cost, Dining Options, Leisure Activities, Period Of Stay, Payment Options)

* + 1. :

Hotel Facilities(User, Date, Swimming Pool, Breakfast Included, Parking, Order-In, Dining Included)

Room Facilities(User, Date, Single Room, Double Room, Deluxe Suites, Attached Bathrooms, Hostel Room)

Hotel Policies(User, Date, Pets, Check-In/Out, Health-Care Options, Cancellation Charges, Medical Charges)

Dining Facilities(User, Date, Breakfast, Lunch, Dinner, Bar)

* + 1. :
       1. Expected Expenditure
       2. Satisfaction

**Step II Business Intelligence Elicitation**

1. **Accessibility**(User ID, Choice of Place, Occupancy, Date)

**Derived attributes:**

Choice of Place = Location, Weather Conditions

**Yields**

**Accessibility**(User ID, Location, Weather Conditions, Occupancy, Date)

1. **Preferences**(User ID, Reservation Cost, Transport Cost, Services Requested, Dining Options, Leisure Activities, Period Of Stay, Payment Options, Date)

**Derived attributes:**

Payment Options = Mode of Payment, Discount

**Yields**

**Preferences**(User ID, Reservation Cost, Transport Cost, Services Requested, Dining Options, Leisure Activities, Period Of Stay, Mode of Payment, Discount, Date)

**Step II Choice Information Elicitation**

All the Oobj are considered in this case:

**UserSatisfiability (User, Date, User Rating)**

**GOM4DW Schema**

1. **Hotel Brand-Value(User, Date, Star Rating, Quality of Services, Reviews Rating)**



1. **Accessibility(User, Date, Location, Weather Conditions, Occupancy)**



1. **Preferences(User, Date, Reservation Cost, Transport Cost, Dining Options, Leisure Activities, Period Of Stay, Mode of Payment, Discount)**



1. **Hotel Facilities(User, Date, Swimming Pool, Breakfast Included, Parking, Order-In, Dining)**



1. **Room Facilities(User, Date, Single Room, Double Room, Deluxe Suites, Attached Bathrooms, Hostel Room)**



1. **Hotel Policies(User, Date, Pets, Check-In/Out, Health-Care Options, Cancellation Charges, Medical Charges)**



1. **Dining Facilities(User, Date, Breakfast, Lunch, Dinner, Bar)**

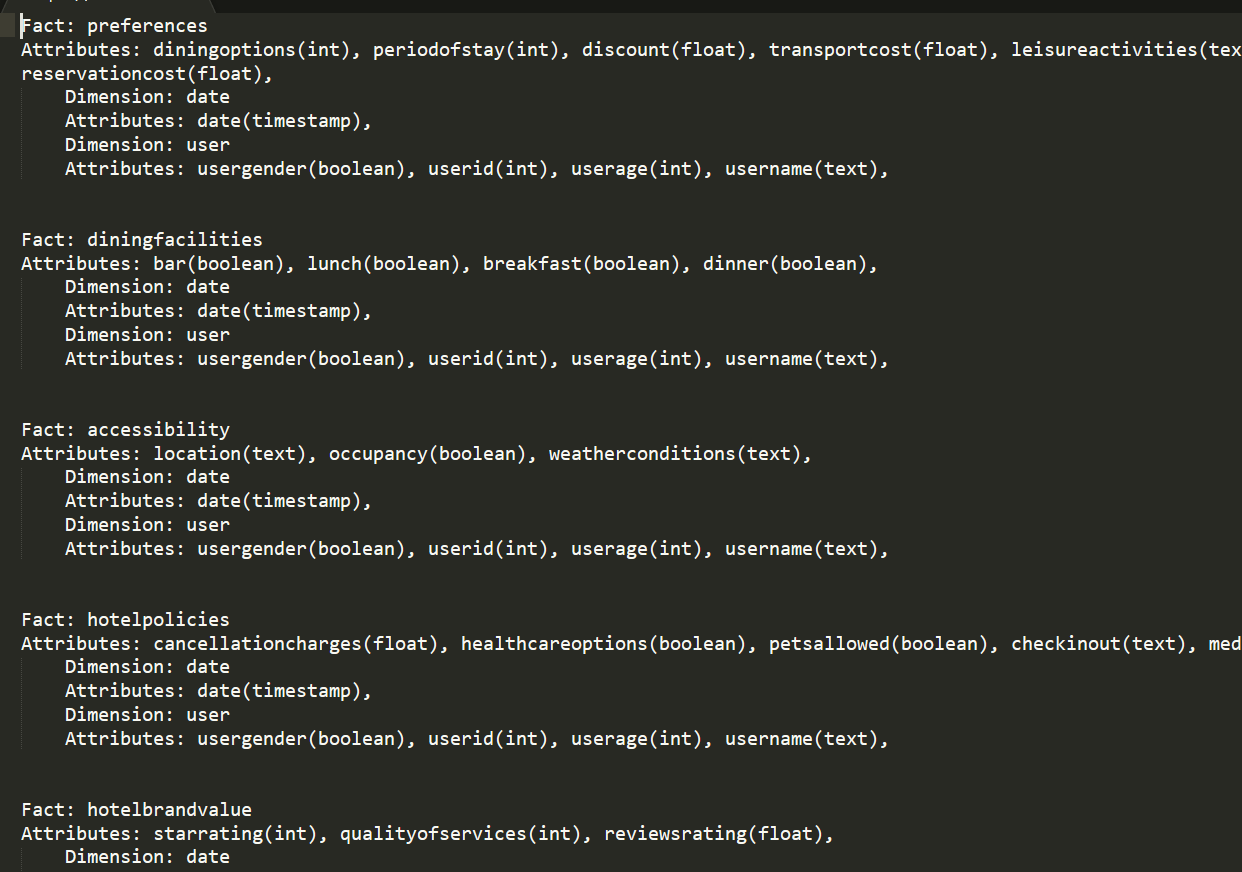


1. **User Satisfiability(User, Date, UserRating)**



**GOM4DW to ROLAP**

**Star Schema**

****

**Different Types of Facts and Dimensions**

**Dimensions**

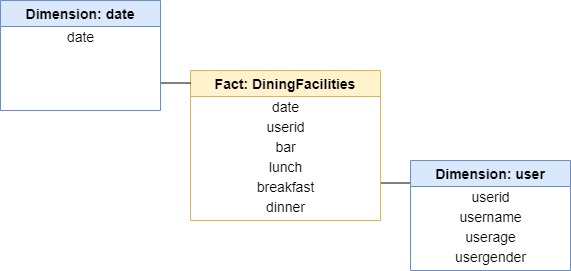
Date: Degenerate Dimension, Conformed Dimension

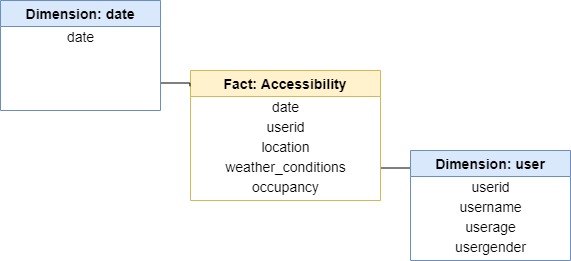
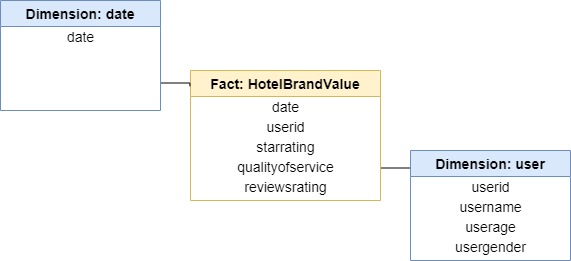
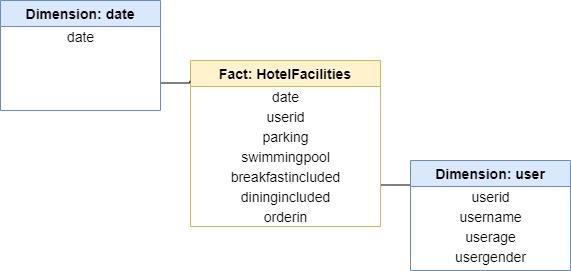
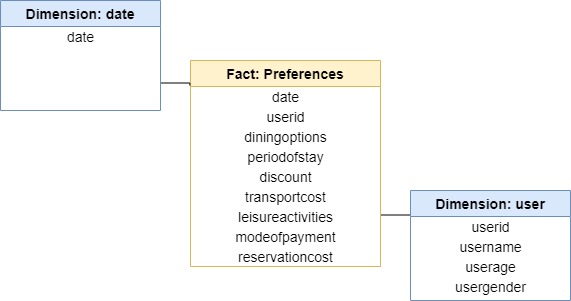
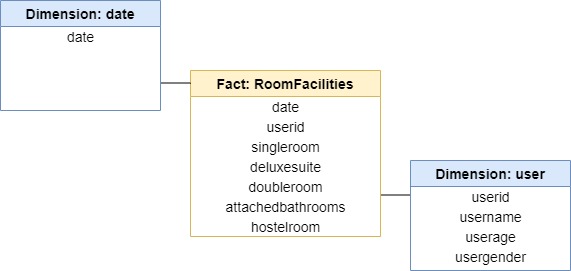
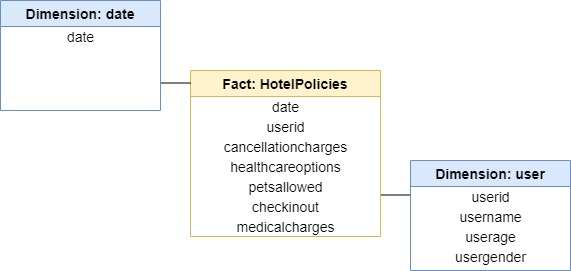
User: Casual Dimension, Conformed Dimension

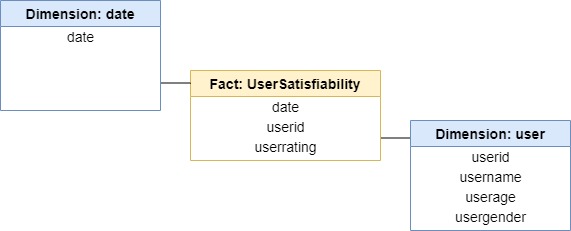
Since, Date and User Dimensions are shared by all Fact tables, they behave as conformed dimensions. Further, Date also acts as a Degenerate Dimension as its only attribute and primary key, date, is taken as an attribute into their corresponding fact tables. Whereas for User Dimension, it acts as a Casual Dimension because its userid attribute is taken as a primary in each of the Facts tables. Changing the user will directly affect the facts too.

**Facts**

|  |  |
| --- | --- |
| Transaction Fact | UserSatisfiability, Preferences, DiningFacilities, HotelPolicies, RoomFacilities, HotelFacilities, HotelBrandValue, Accessibility |
| Reason being, we are taking the input from the user for each of the above facts. Moreover, all the inputs are being inputted into the facts table continuously by individual users. In the following diagrams, you can see that each fact table constitutes the primary keys of each of their connected dimension tables. | |

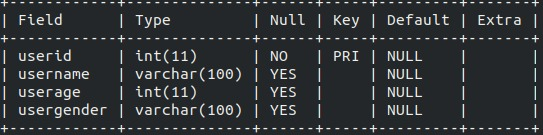
****

****

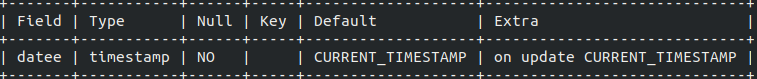
****

**SQL Table Schemas**

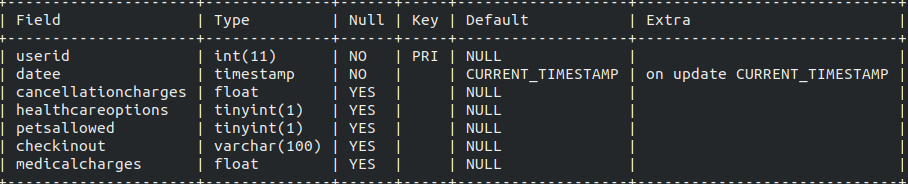
1. User Details



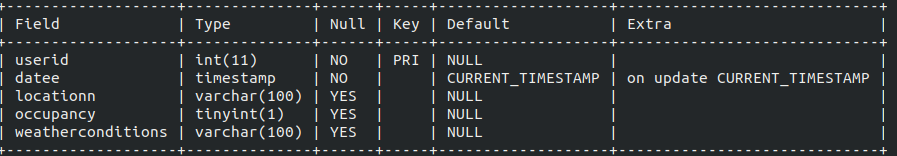
1. Date



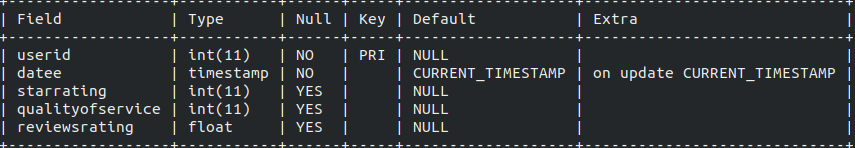
1. Hotel Policies



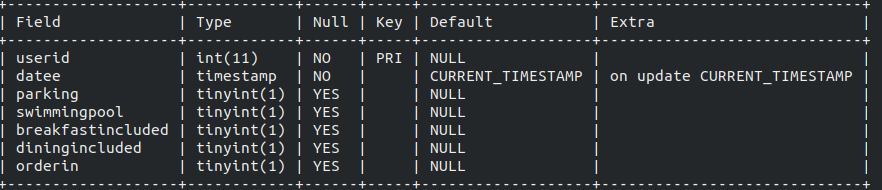
1. Accessibility



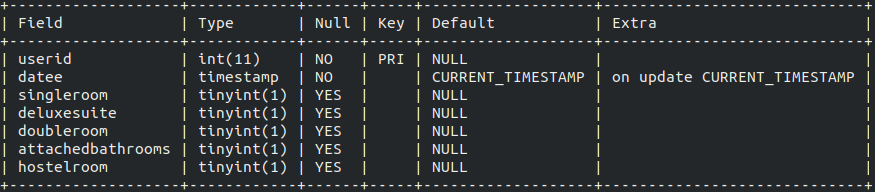
1. HotelBrandValue



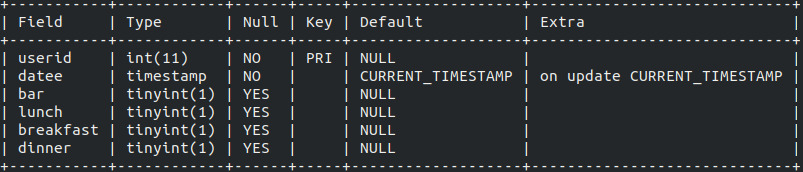
1. HotelFacilities



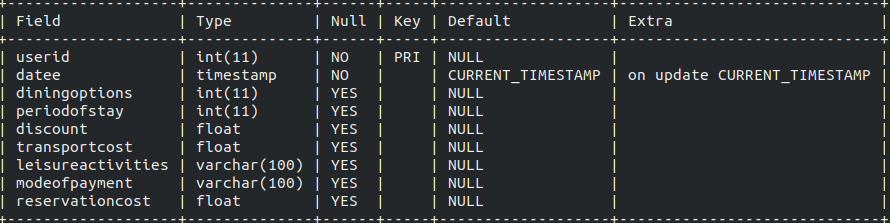
1. RoomFacilities

****

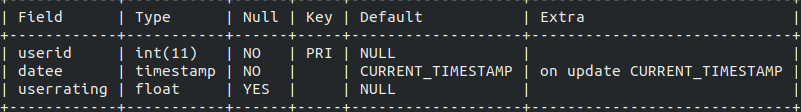
1. DiningFacilities

****

1. Preferences

****

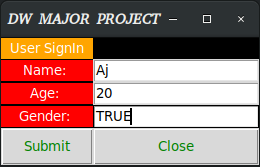
1. UserSatisfiability

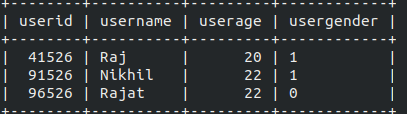
****

**ETL Process**

App Login:

1. When a user logs in for the first time, its record is created in the UserDetails table.

****

****

1. After logging in, the user can access the following options.

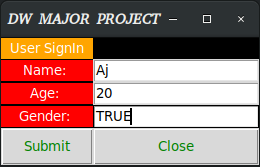
****

**Type-1 Query**

To show demonstration of Type 1 queries, we have edited the UserDetails table by clicking on the “Edit Credentials” option:

****

Then the same login page is shown, but now it’s within the scope of the original user. After submitting, whatever changes are reflected in the original table:

****

The query (it is an example query, when we change something from (name, age, gender)):

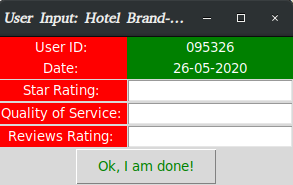
**“Update UserDetails Set name ="+ *str*(name.get()) +", age ="+ *str*(age.get())+", gender="+ *str*(gen.get())+ " where userid =" + userid +");”**

**Type-2 Query:**

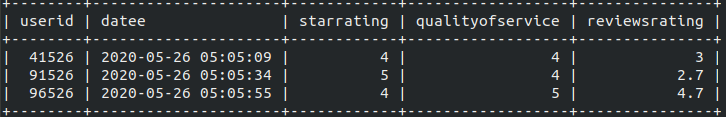
To show the demonstration for type-2 query, we have access following options (in gray background):



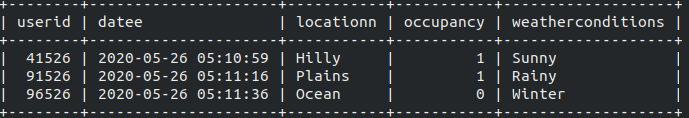
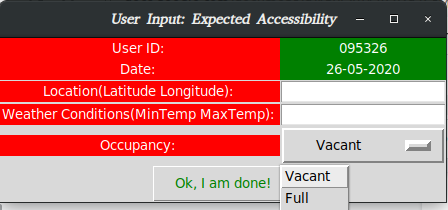
After clicking Hotel Brand Value, we get following dialog, where the user can enter their data:

****

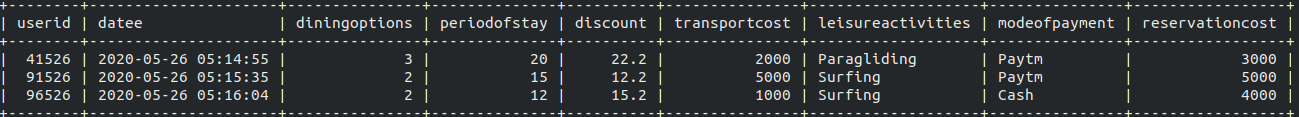
Then, following changes are made in the respective table:

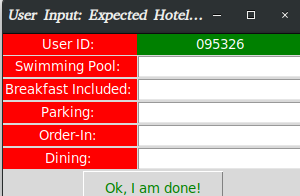
****

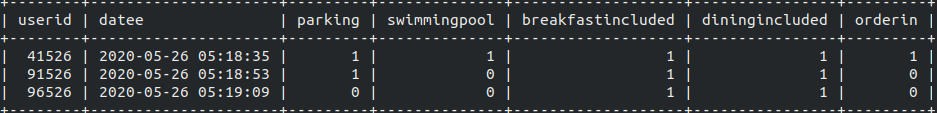
Same goes for other options:

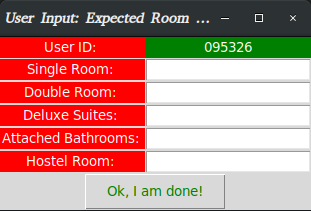
****

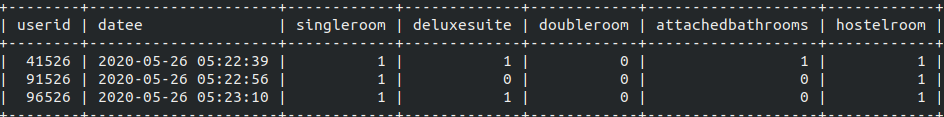
****

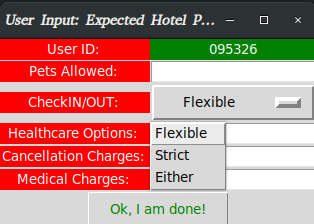
****

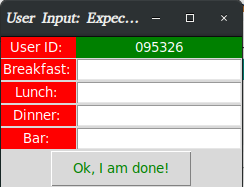
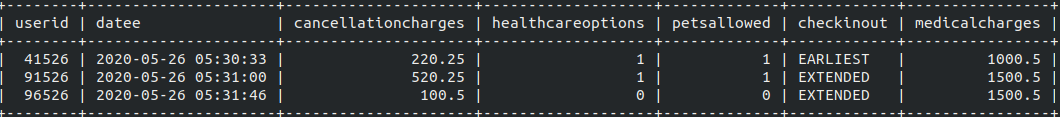
****

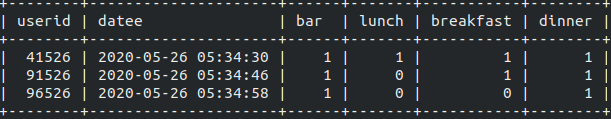
****

****

****

****

****

****

The Query for for inserting into the respective Fact tables (Shown for one of the options):

**"INSERT INTO HotelBrandValue(userid, datee, starrating, qualityofservice, reviewsrating) VALUES (" + userid + "," + "NOW()," + *str*(starrating) + "," + *str*(qos) + "," + *str*(reviewsrating) + ");"**