

Sri Lanka Institute of Information Technology

**Data Warehousing and Business Intelligence**

**(IT3021)**

Assignment 2

**Hotel Reviews**

Submitted by :

**IT19374666**

**Dissanayake Adikaramge D.M.**

1. **Data Source**

I selected a hotel review data set for the assignment which includes the ratings by the reviewers, number of reviews seen of a review by the reviewees, the commission paid to the reviewer, the relevant hotel details, and the reviewer details.

This hotel review data set comprises of 10,000 records of hotel reviews over 16 years from 2002 – 2018.

The data set has hierarchies in hotel location such as country-> province -> city and in reviewer entity has user\_province -> user\_city hierarchy.

**Following is the ER Diagram for the chosen data set.**

Diagram

Description automatically generated

Data set was downloaded from the following link :

<https://www.kaggle.com/datafiniti/hotel-reviews>

**Dimensional Model**

**Diagram

Description automatically generated**

For the data warehouse of the review data set ,I implemented a star schema. DimHotel,DimReviewer , DimReviewDetails and DimDate are dimensions and Review is the fact table in the data warehouse.

I merged the hotel and hotel location tables as shown in the relational model above together to create a single dimension table named DimHotel as shown in the above dimensional model to develop it to a star schema.

Further I implemented the DimHotel dimension as a **slowly changing dimension**.

**Grain** : Details of a review on a specific hotel by a reviewer .

**Assumptions:**

I decided hotel dimension is a slowly changing dimension assuming hotel name, categories ,primary categories and location details can be changed over time.

1. **SSAS Cube implementation**
2. Initially I added the ‘DataWarehouse\_Hotel\_Reviews’ database as the data source in the SSIS analysis service.
3. Then I created the data source view as follows;

Diagram

Description automatically generated

1. I created a cube for the data source using all the dimensions (DimDate, DimReviewer, DimHotel and DimReviewDetails) and FactReview as the measure group table with all the measures in the Fact table as measures.
2. Then I added a hierarchy for the DimDate as Year>Month>Date and a hierarchy for the DimHotel as Province > City >Name.

Graphical user interface

Description automatically generated

Graphical user interface

Description automatically generated

1. I created a KPI to measure the review counts to be greater than 15.

Graphical user interface, text, application, email

Description automatically generated

1. I created a role ‘Administrator’ giving all the access and permission.

Graphical user interface, text, application, email

Description automatically generated

1. I deployed the cube as a SSAS database.

Graphical user interface, text, application, email

Description automatically generated

1. **Demonstration of OLAP operations**
2. Initially I connected Excel PowerPivot to the cube using MDX query.

I used the following MDX query for the OLAP drill down and roll up operations.

Graphical user interface, text, application

Description automatically generated

Then I successfully loaded the query data to the PowerPivot backend.

Graphical user interface, text, application, email

Description automatically generated

**Drill down**

I created an pivot Table to analyze count of views, count of reviews and average rating for Hotel where we can drill down from the Country to Province to City to Hotel.

Graphical user interface, application

Description automatically generated

Following is the table that created for the drill down operation.

**Graphical user interface, application, table

Description automatically generated**

**Roll up.**

I used the above created table to perform the roll up operation for the hotel review details analysis.

Table

Description automatically generated

We can analyze province wise review details for hotels using above table.

**Slice**

1. I initially connected Excel Power View to the cube.

Graphical user interface, text, application

Description automatically generated

1. Then I added a pie chart to analyze province wise review counts and stacked bar chart to analyze hotel wise review counts for 2015 and 2016.
2. Finally, I added a slicer to analyze primary category wise review counts.

Chart, pie chart

Description automatically generated

**Dice**

1. I initially connected PowerPivot to the implemented cube.

Graphical user interface, application

Description automatically generated

1. I used Hotel Dimension to analyze review count, view count based on the selected provinces as below and Date Dimension to analyze the same measures from year 2013-2018.

Table

Description automatically generated with medium confidence

**Pivot**

1. From the connected cube I implemented the following pivot table.

It depicts the review count and view counts based on year and province.

Table

Description automatically generated

1. I changed the year row to column for better analysis of the measures.

Graphical user interface, application, table

Description automatically generated

1. **SSRS Reports**

**Report with a matrix**

1. In report builder ,I created a data source connection to ‘HotelReviewDataWarehouse’ which is embedded to the report.
2. Then I created a data set using the following SQL query.

Text

Description automatically generated

The following matrix report demonstrates province wise average review rating and view count for years( 2015-2018); year demonstrates in columns and province in rows.

Chart

Description automatically generated

**Paramiterized report**

1. Initially I created a new data set from following SQL query.

**Text

Description automatically generated**

1. Then I added a new data set for Province details.

Text

Description automatically generated

1. Next I added a data set for Hotel details.

Text

Description automatically generated

1. I added a parameter for Province list as follows to accept multiple values.

Graphical user interface, text, application, email

Description automatically generated

1. Then I added a parameter for Hotel list as follows to accept multiple values.

Graphical user interface, text, application, email

Description automatically generated

Following report consists of a bar char for hotel wise average rating for 2017 and 2018 and an area graph for hotel wise views when province and hotel names are passed as parameters.

Graphical user interface, application

Description automatically generated

**Drill-down report.**

1. Initially I created a new data set from following SQL query.

Text

Description automatically generated

Following report depicts hotel wise average review rating and view count which is drilled down from province for 2015-2018.

**Table, calendar

Description automatically generated with medium confidence**

**Drill-through report.**

1. Initially I created a new data set from following SQL query for the level 1 report.

**Text

Description automatically generated**

1. Then I created the following report with a bar chart to depict province wise average rating and an area graph to depict province wise view count.

Graphical user interface

Description automatically generated

1. I created a data set from following SQL query for the level 2 report.

**Text

Description automatically generated**

The following level 2 report demonstrates the hotel wise average rating and hotel wise view counts in an area graph.

Graphical user interface, application, Teams

Description automatically generated

1. Then I created a parameter for the provinces.
2. Finally, I added an action in the level 1 report to navigate to the level 2 report as follows;

**Graphical user interface, application

Description automatically generated**