Data Warehousing and Data Mining  
Report

Marriott International Hotel

By

|  |  |  |
| --- | --- | --- |
| Mr. Ruksat | Soladda | 6188021 |
| Ms. Preechaya | Wannabhoom | 6188065 |
| Mr. Teekawin | Kirdsaeng | 6188077 |
| Mr. Podchawat | Worathanakul | 6188104 |
| Ms. Wanwisa | Laowsiriwong | 6188135 |

A Report Submitted in Partial Fulfillment of

the Requirements for

ITCS453 Data Warehousing and Data Mining

Faculty of Information and Communication Technology

Mahidol University

2/2020

Table of Content

|  |  |
| --- | --- |
|  | Page |
| Overview | 3 |
| Business Process | 3 |
| Issue | 3 |
| Business Requirement for Data Warehouse | 3 |
| Objective and Scope | 4 |
| Expected Benefit | 4 |
| Data Source | 4 |
| ER-Diagram | 4 |
| Star Schema | 5 |
| ETL Process Description | 8 |
| Analysis and Visualization Report | 10 |
|  |  |
|  |  |

Overview

Marriott International Inc. is an American hospitality company founded by J. Willard Marriott in 1957. Throughout the year Mr. Marriott keeps having new ideas and developing their business. Today's Marriott International is the largest hotel chain in the world by number of available rooms. It has 30 brands with 7,484 properties in 131 countries and territories around the world over 1,400,693 rooms. The headquartered in Bethesda Maryland. However, we will focus on the Marriott Hotel in Thailand. According to the Marriott website there are 9 branches all over Thailand; Bangkok, Hua-Hin, Phuket, Phuket Island and Rayong. They also opened restaurants and bars to attract more customers that didn’t use the hotel service.

Business Process

Our customers come from online booking and walk-in customers. For walk-in customers they will have the reception on helping them book the room. If customers do the booking online the room will be reserved immediately. After the customer pays for the payment the financial will approve that the customer has paid then the customer will receive an invoice. For customers that are interested in eating at the restaurant they reserve a seat there. For the restaurant there will be a restaurant employee to provide comfort ability to customers and for customers to enjoy their meals. Moreover, the customer can choose a maximum of 3 rooms per booking. The customer will have to tell us their check in date and checkout date.

Issue

Nowadays, there is a lot of tourism in Thailand. So, many hotels need the best online system to support tourists to reply to the customers' searches with the necessary and updated information. On the other hand, the hoteliers also need information about their organization's competitive set, which implies having access to information about the clients, competitors, and all the stakeholders associated with the hospitality activity. There is a lot of data in the hotel such as account, customer information, hotel booking information, employee information and many more. So, it will be huge data and hard to manage data. Moreover, the information data of each department does not match to other departments because some departments have not updated yet.

Business Requirements for data warehouse.

* Show the total number of booked rooms of each branch for each year.
* Show the most and least booked room type.
* Show the total number of customers in each branch for each year.
* Show the total number of room sales for each branch.
* Show the total night stay of customers for each branch.
* Show the total number of reservations in a restaurant.

Objectives and Scope

* To collect information that can be used for analyzing the popularity of each branch, managing sales, and to check about customers interested in room types.
* To predict the months of the majority of customers visited in each branch.
* To visualize the sales rate for the manager to analyze the total income.

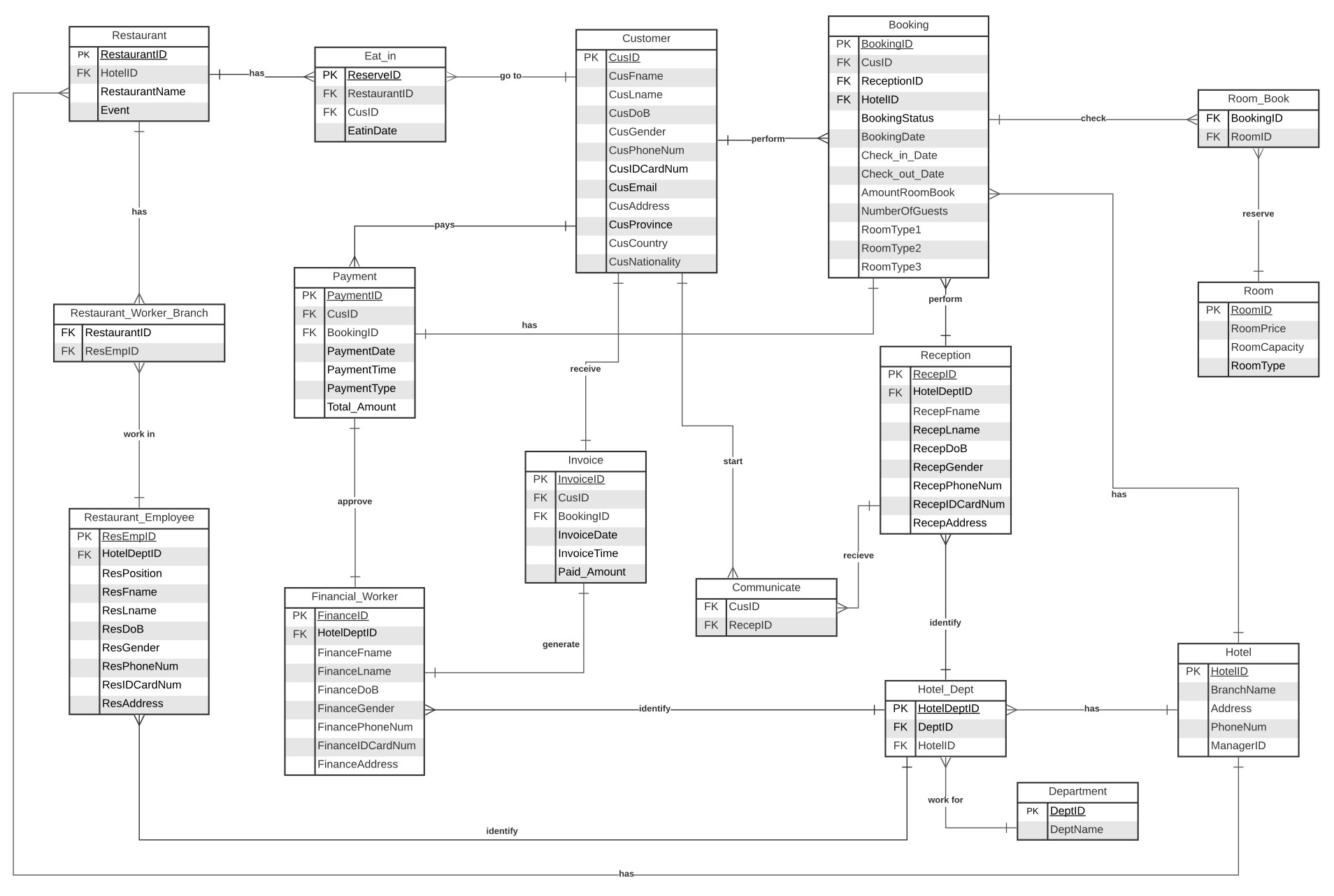
Expected benefits

* To make easier decisions about promotional advertising marketing plans, sales, risk and strategies.
* In order to keep track of the hotel data and know what and when the data was collected.
* Get an in-depth look at the most popular rooms and the most popular dates that customers book the most in each branch

Data Source

Since there are 9 branches in Thailand we will focus on the data in each branch. The data of the hotel and restaurant name address and other information are real. However, the data of customers or employees working for the hotel are generated because of privacy. We decided to use a data generator to generate it.

ER - Diagram



Star Schema

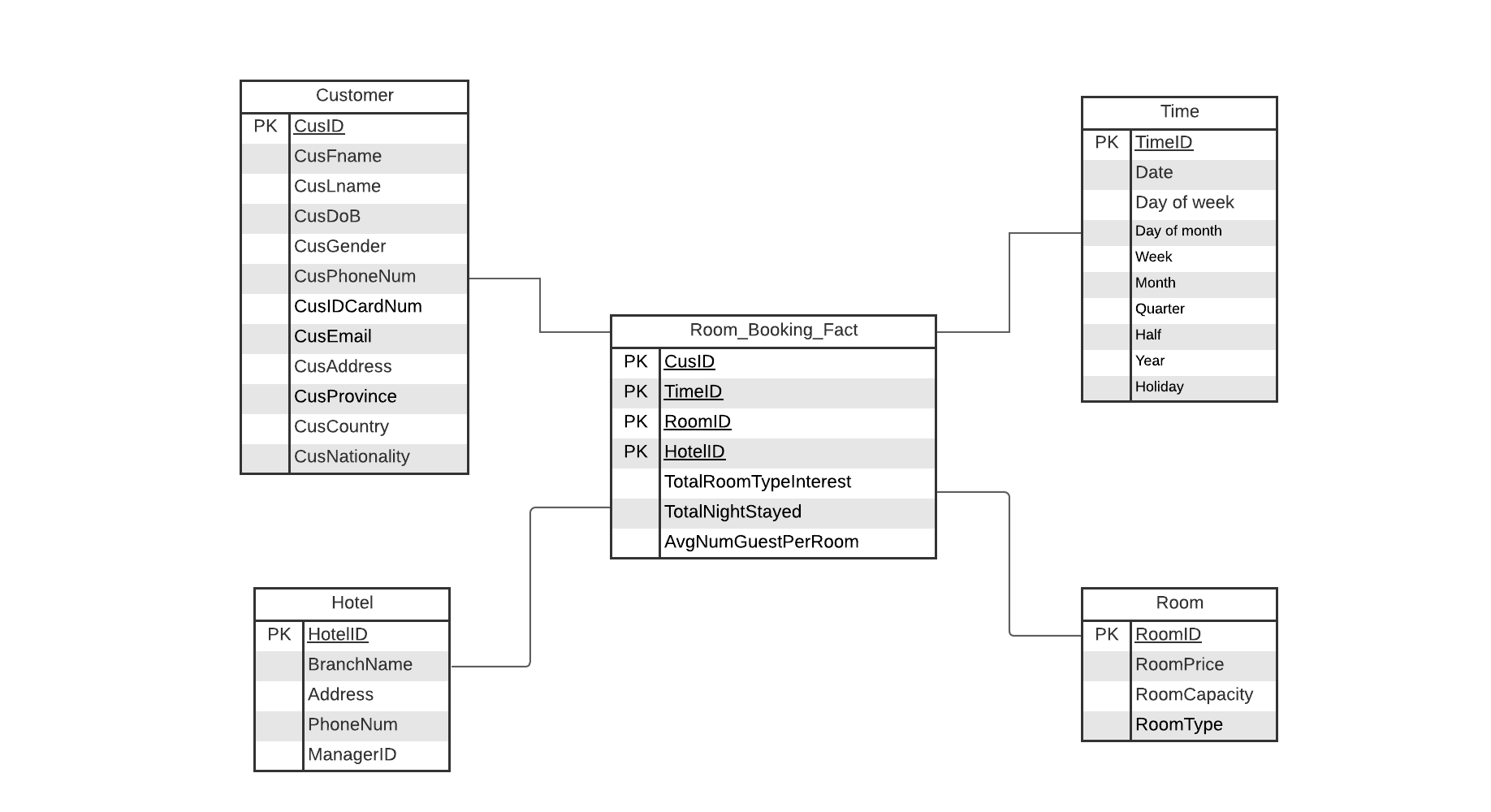
**Room Booking Fact Table**

Dimension:

1. Customer
2. Hotel
3. Time
4. Room

Fact:

1. TotalRoomTypeInterest
2. TotalNightsStayed
3. AvgNumGuestPerRoom



The Room Booking Fact table helps to know which month and day has the average number of guests per room, type of room that customer interest and total nights stayed. The average number of guests per room, type of room interest, night stayed can be derived from customer, hotel, time and room. It helps to analyze the demand of the type of room that customers want to use in each situation, the average number of guests per room, and the number of the customers that stayed many nights. The total room type interest is to check what kind of room the customer are most interested in. The total night stayed is to see how many nights does customer stayed in the hotel from different branches. For the average number of guests per night is to check how many customers come to stay in one room.

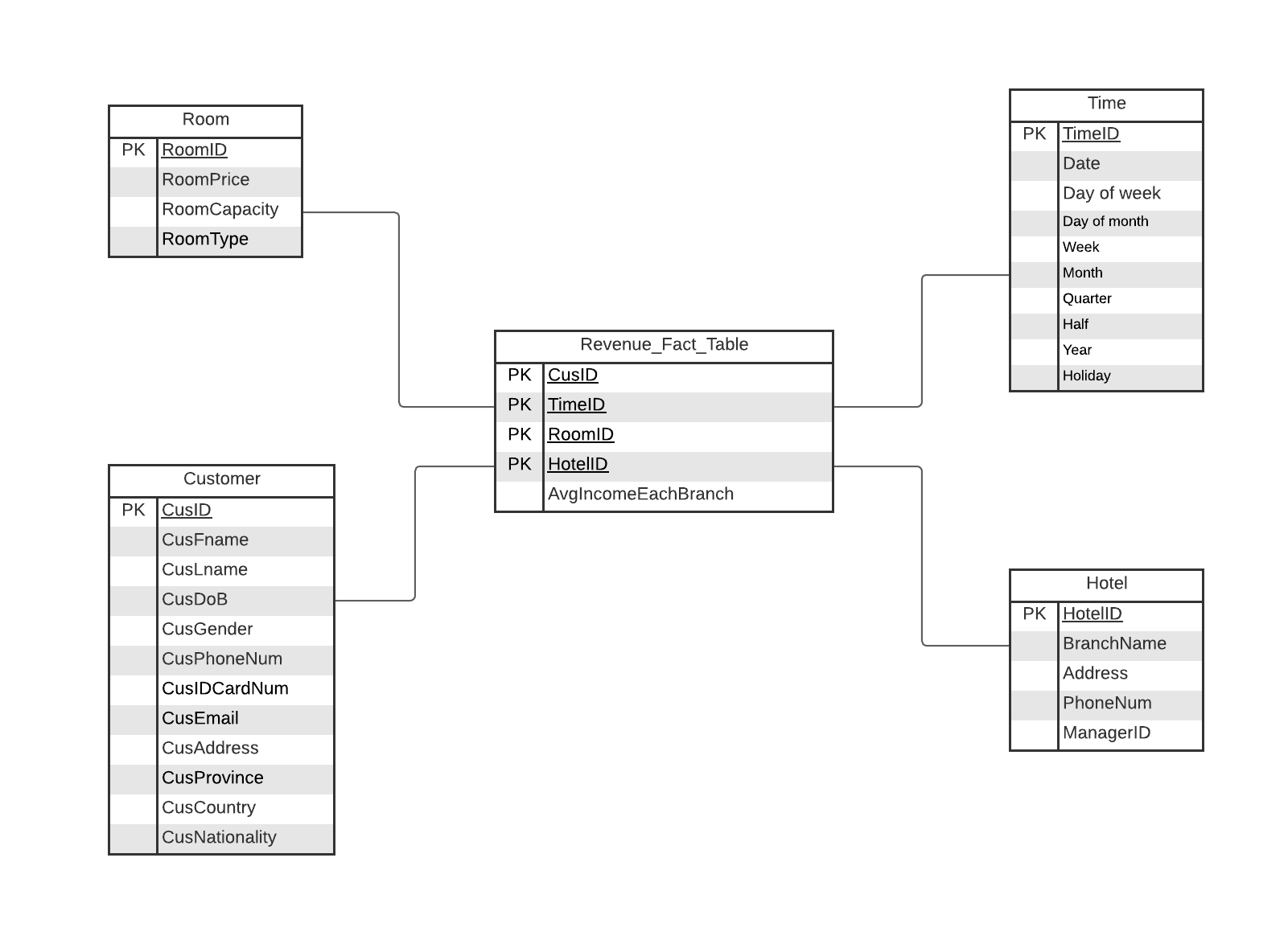
**Revenue Fact Table**

Dimension:

1. Customer
2. Hotel
3. Time
4. Room

Facts:

1. AvgIncomeEachBranch



The revenue fact table helps to know the income in this segment is the primary income, it helps to find out how many sales each store is in each room and what the average monthly gross income is. The average income each branch of can be derived from Room, Customer, Hotel, Time. This will help analyze the sales rates and revenue of each branch.

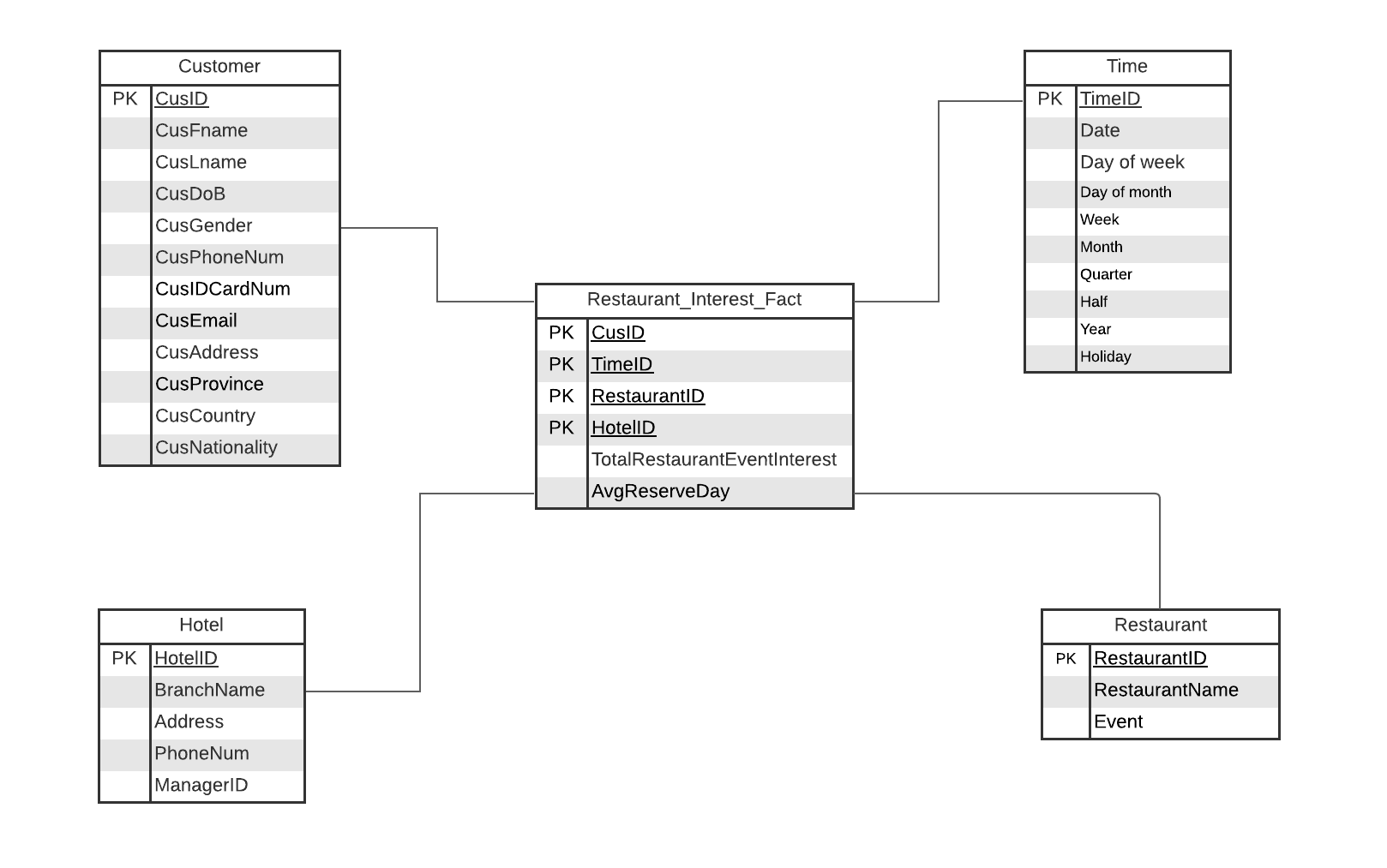
**Restaurant Interest Fact**

Dimension:

1. Restaurant
2. Time
3. Customer
4. Hotel

Fact:

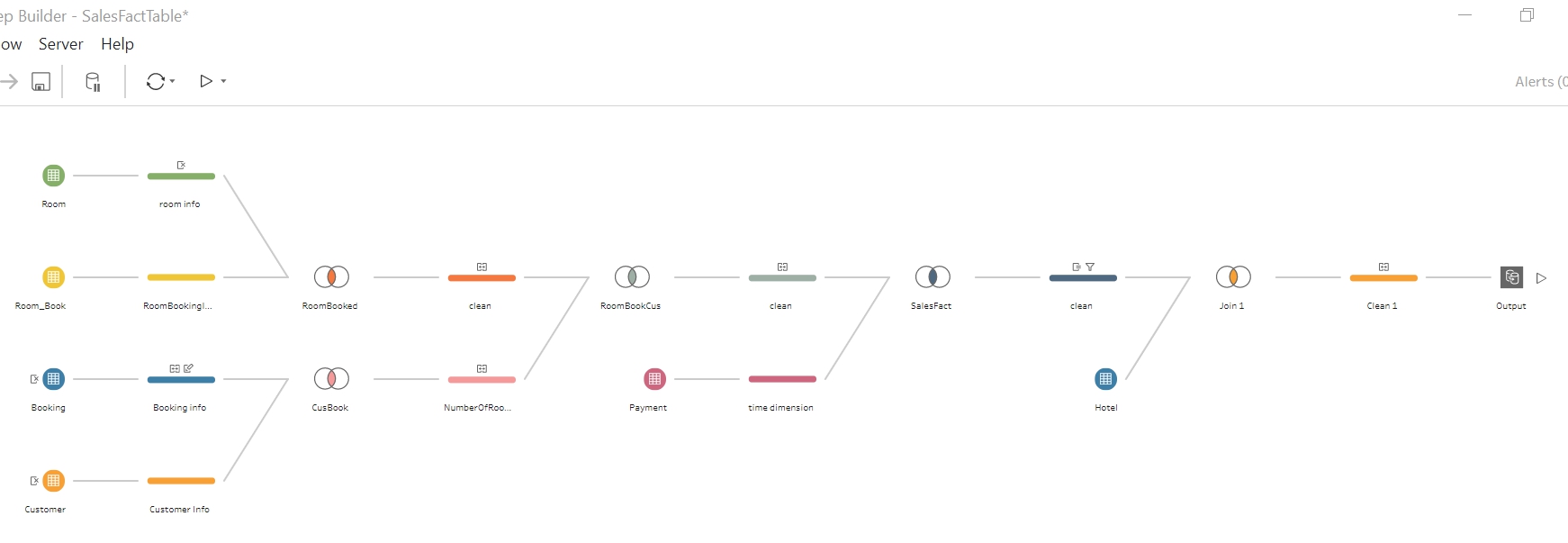
1. TotalRestaurantEventInterest
2. AvgReserveDay



The Restaurant Interest Fact table helps to know what kind of restaurant that customer is interested in. The total restaurant can be derived from customer, hotel, time and restaurant. It helps to analyze the amount of the customer that uses the restaurant, the total of the restaurant type that customer interest, the average reserve day. The total restaurant event interest is to see how many customers are interested to eat in a specific restaurant room which in each room the event won’t be the same if it is the same branch. The average reserve day is to check what day customers are interested in coming in the restaurant to be able to prepare enough food.

ETL Process Description

**Revenue Fact Table**



Hotel dimension:

It extracts the data from the Hotel table, including HotelID, BranchName, Address, PhoneNum, and ManagerID.

Room dimension:

It extracts the data from the Room table, including RoomID, RoomPrice.

Customer Dimension:

It extracts the data from the Customer table, including CusID, CusFname,CusLname.

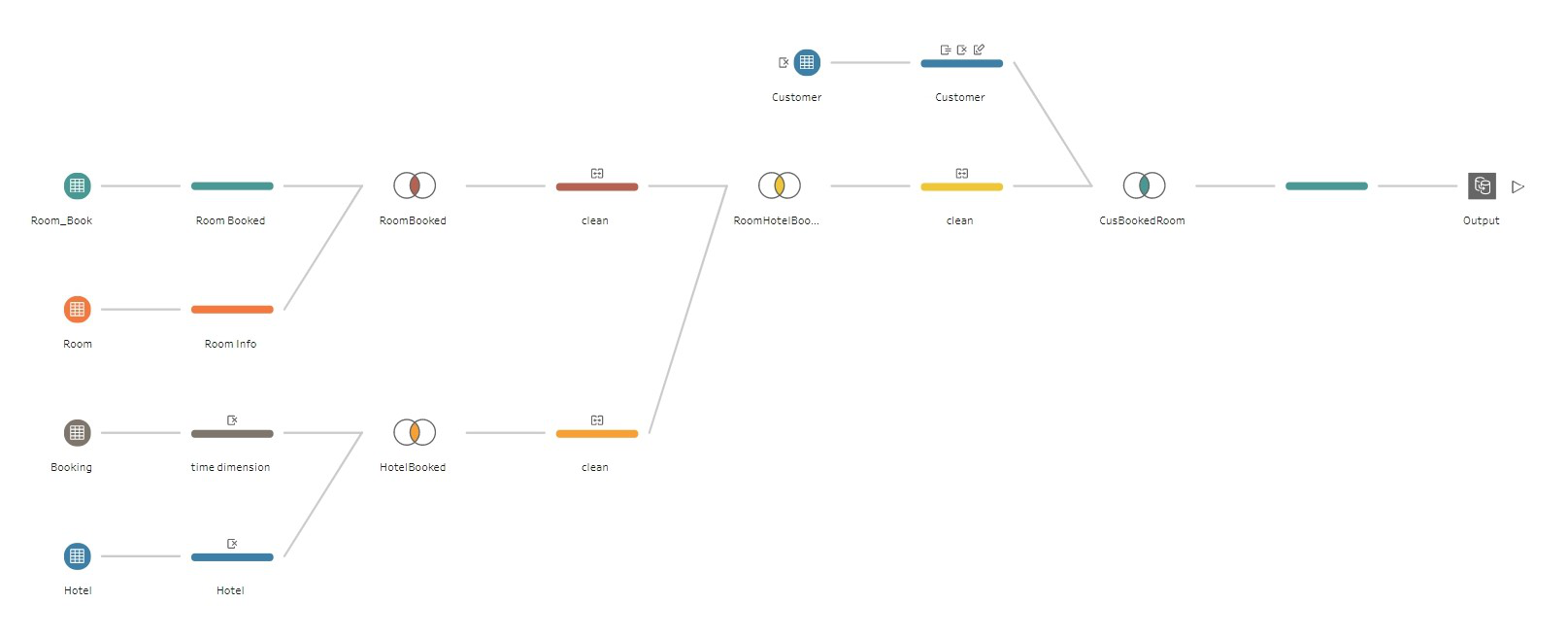
Time dimension:

It extracts the date from the PaymentDate in the Payment table.

Fact:

The avg sales each branch fact is extracted from the attribute TotalAmount in the Payment table.

**Room Book Fact Table**



Hotel dimension:

It extracts the data from the Hotel table, including HotelID, BranchName, Address.

Room dimension:

It extracts the data from the Room table, including RoomID, RoomPrice, RoomCapacity, and RoomType.

Customer Dimension:

It extracts the data from the Customer table, including CusID.

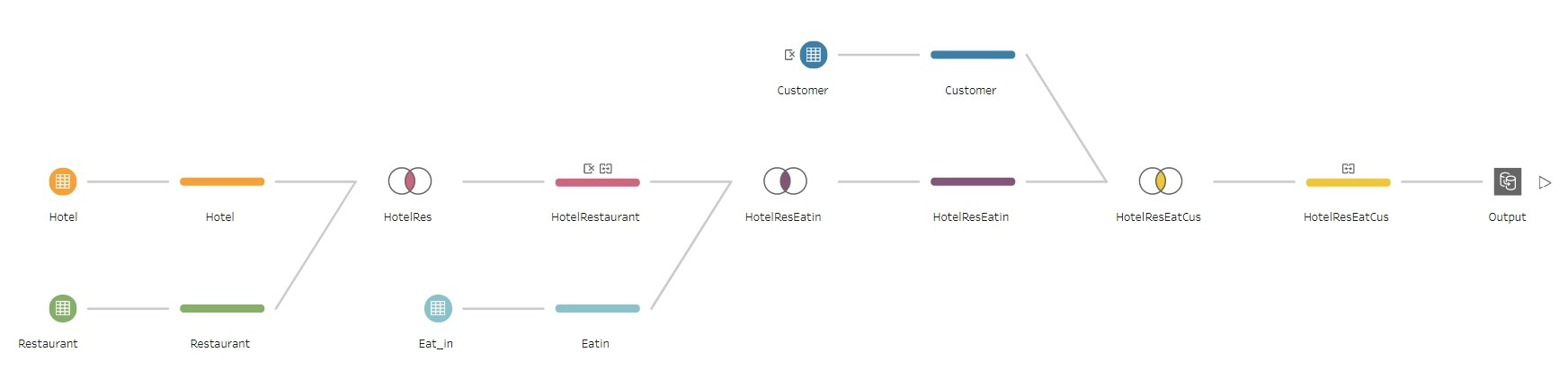
Time dimension: Booking

It extracts Booking date from the checkin and checkout in the Booking table.

Fact:

The NumOfOccupiedRoom fact extracted from the attribute branch and booking in the hotel, booking table. TotalRoomTypeInterest fact extracted from the attribute Roomtype, Room\_Book in the room table. TotalNightsStayed fact extracted from the attribute Check\_in\_Date, Check\_out\_Date, BranchName in the booking and hotel table.

**Restaurant Interest Fact Table**



Hotel dimension:

It extracts the data from the Hotel table, including HotelID, BranchName, Address. Restaurant dimension:

It extracts the data from the Restaurant table, including RestaurantID, RestaurantName, ResDeptID, and Event.

Customer Dimension:

It extracts the data from the Customer table, including CusID, CusFname, CusLname.

Time dimension: Eat in

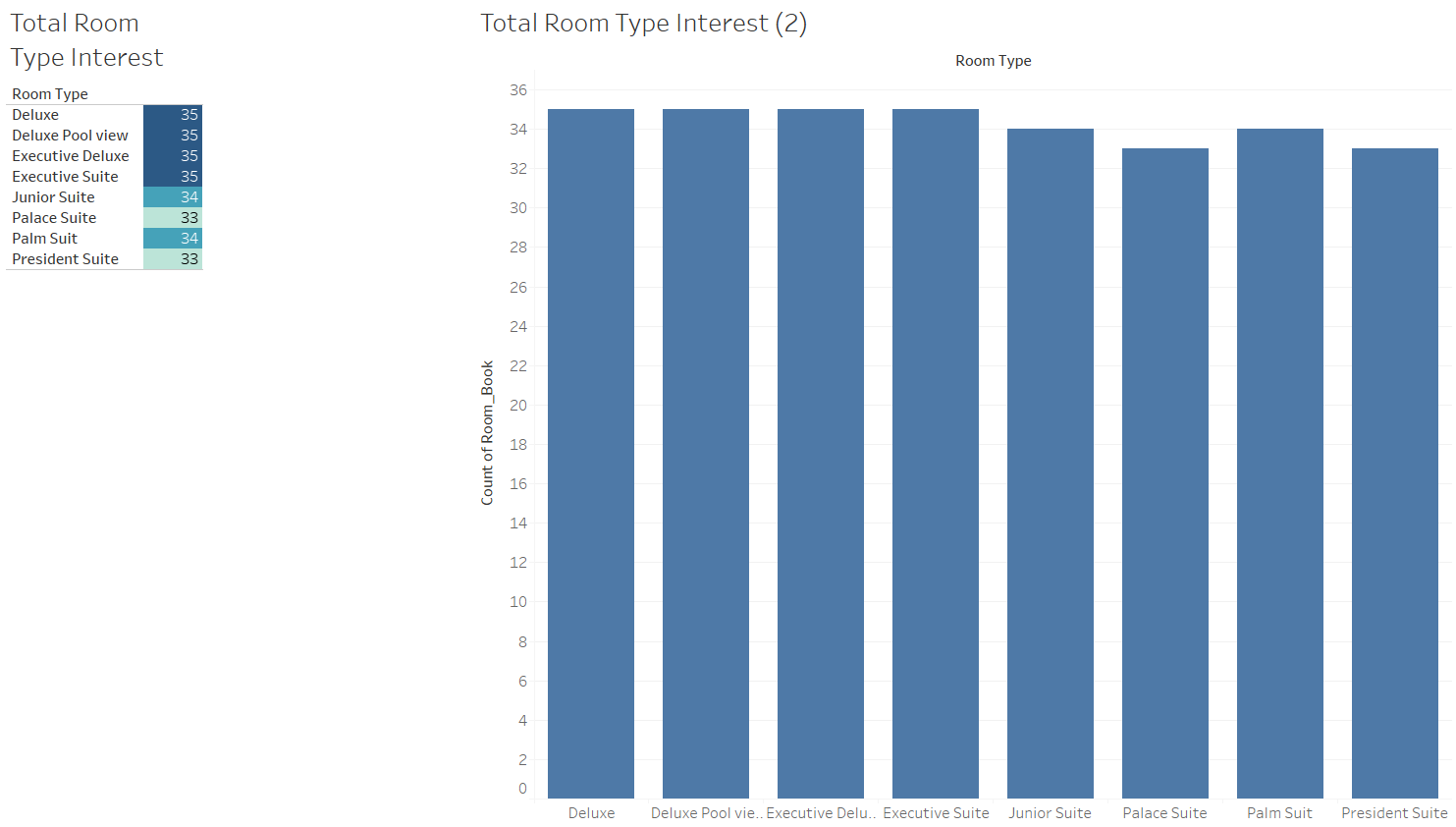
It extracts the date from EatIndate in the EatIn Table

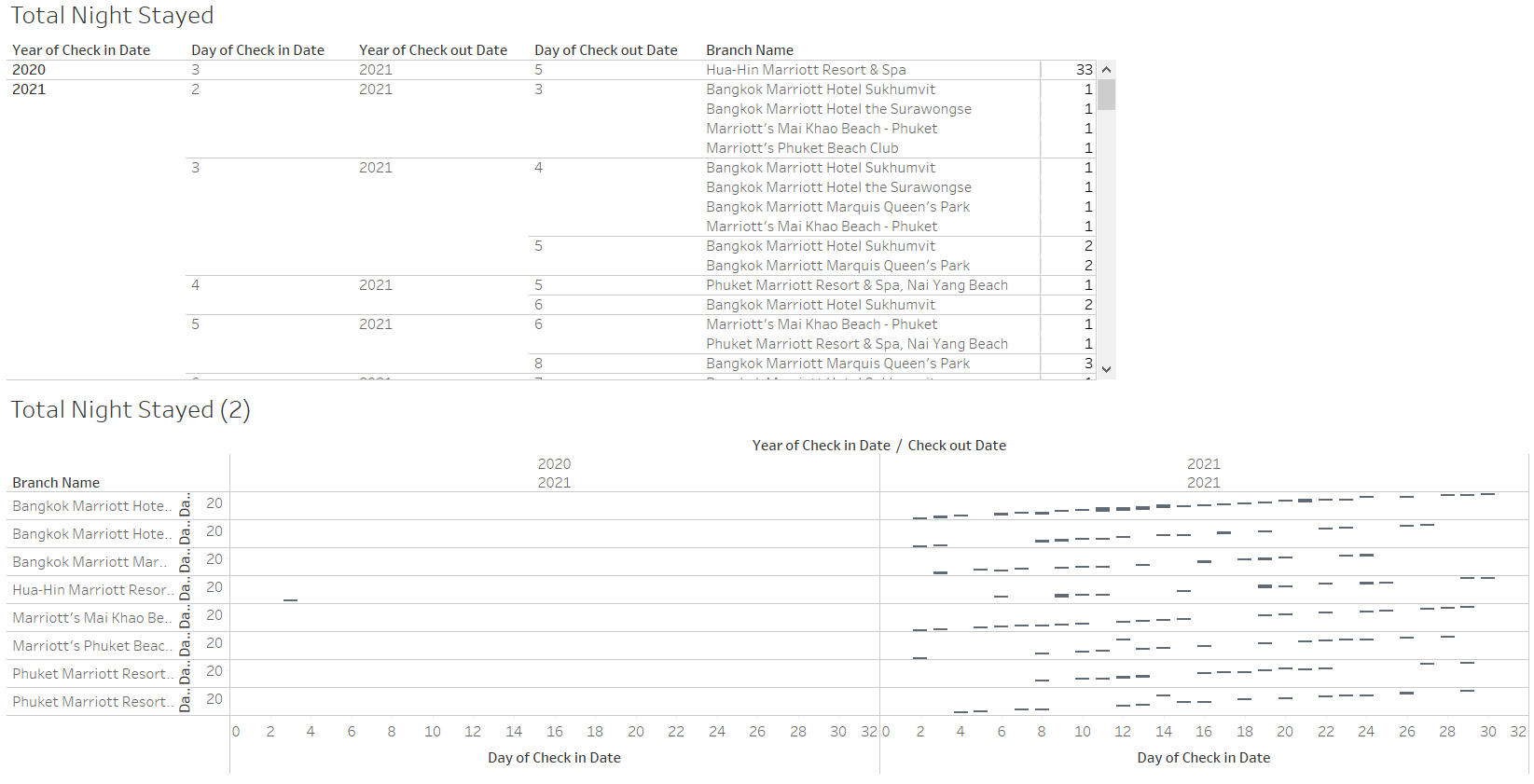
Fact:

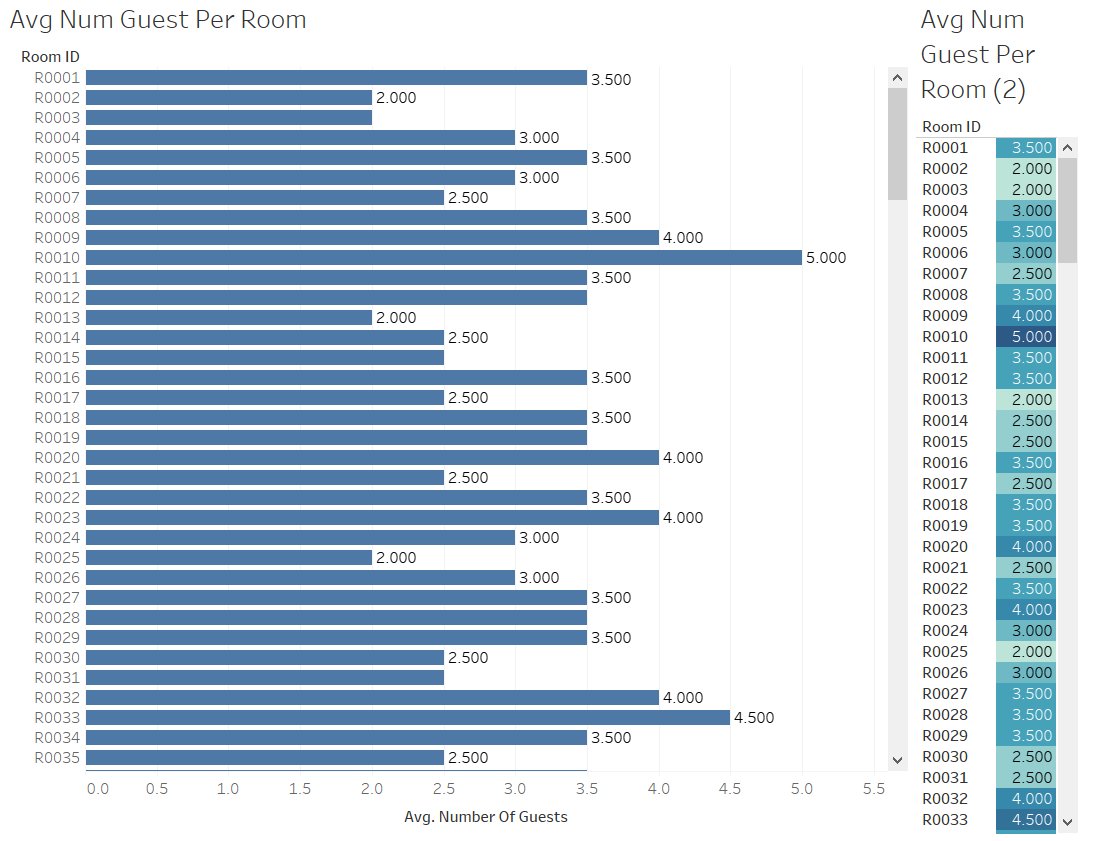
TotalRestaurantEventInterest fact extracted from the attribute Event in the restaurant table. Num\_Of\_Reserve\_Table fact extracted from the attribute RestaurantName, RestaurantsID, ReserveID in the restaurant and Eat\_in table.

Analysis and Visualization Reports

**Room Book Fact Table**



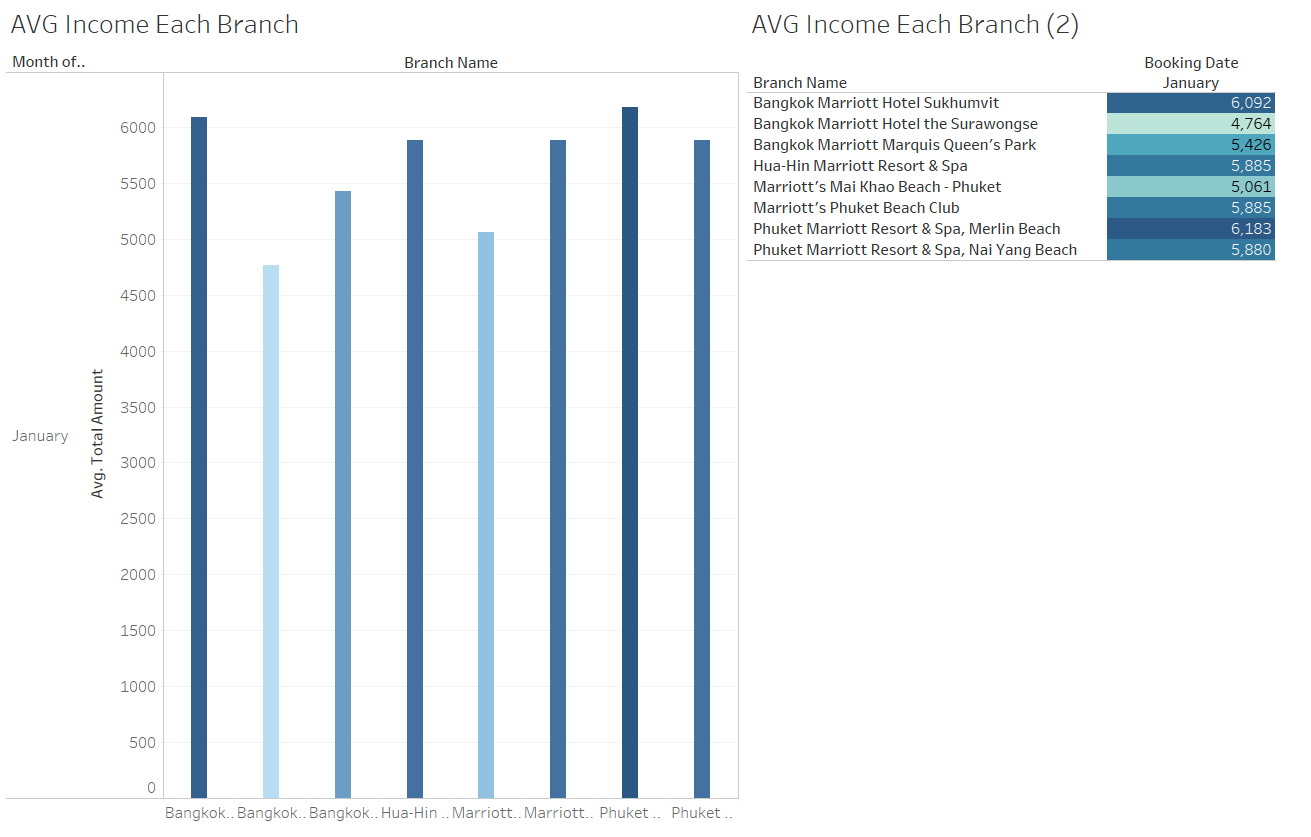




The visualized data from this story are derived from room booking star schema which made up for the purpose to analyze various types of room including the type of room that customer interest, the total night stayed in each branch and the average number of the guest in each room. This data can answer the question about which type of room that customer wants to stay, the number of rooms on night stay and the average number of guests per room in each branch. Furthermore, the manager can use this graph to anticipate the type of room that customer is most interested in, the day that customer wants to check in and check out ,and the average number of guests per room in each branch. These graphs can answer company questions as follows.

* Which type of room is the customer most interested in?
* Which type of room is the customer least interested in?
* What is the average duration of customer stayed night?
* What day of the month is the customer most checkin?
* Which is the average guest per room?

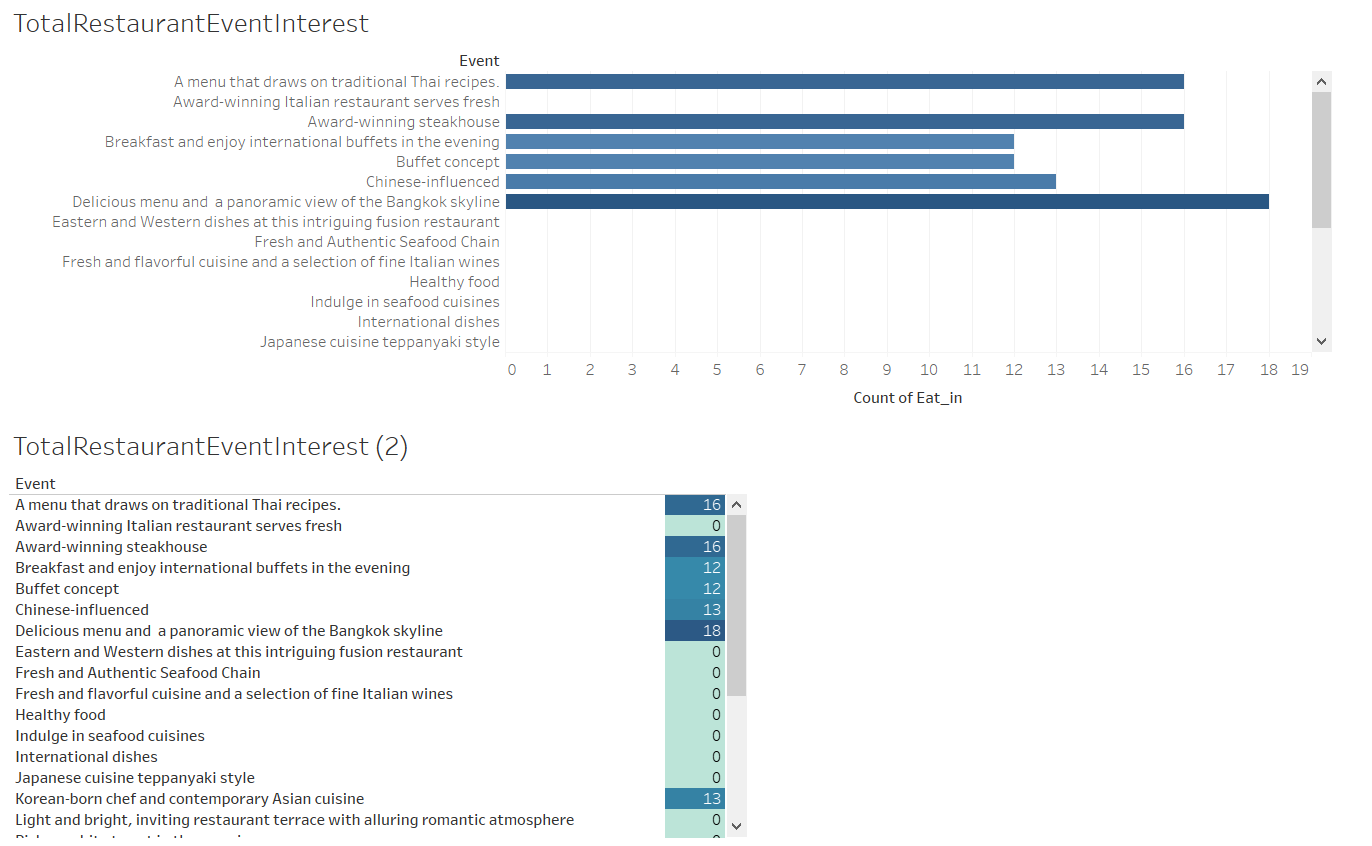
**Revenue Fact Table**

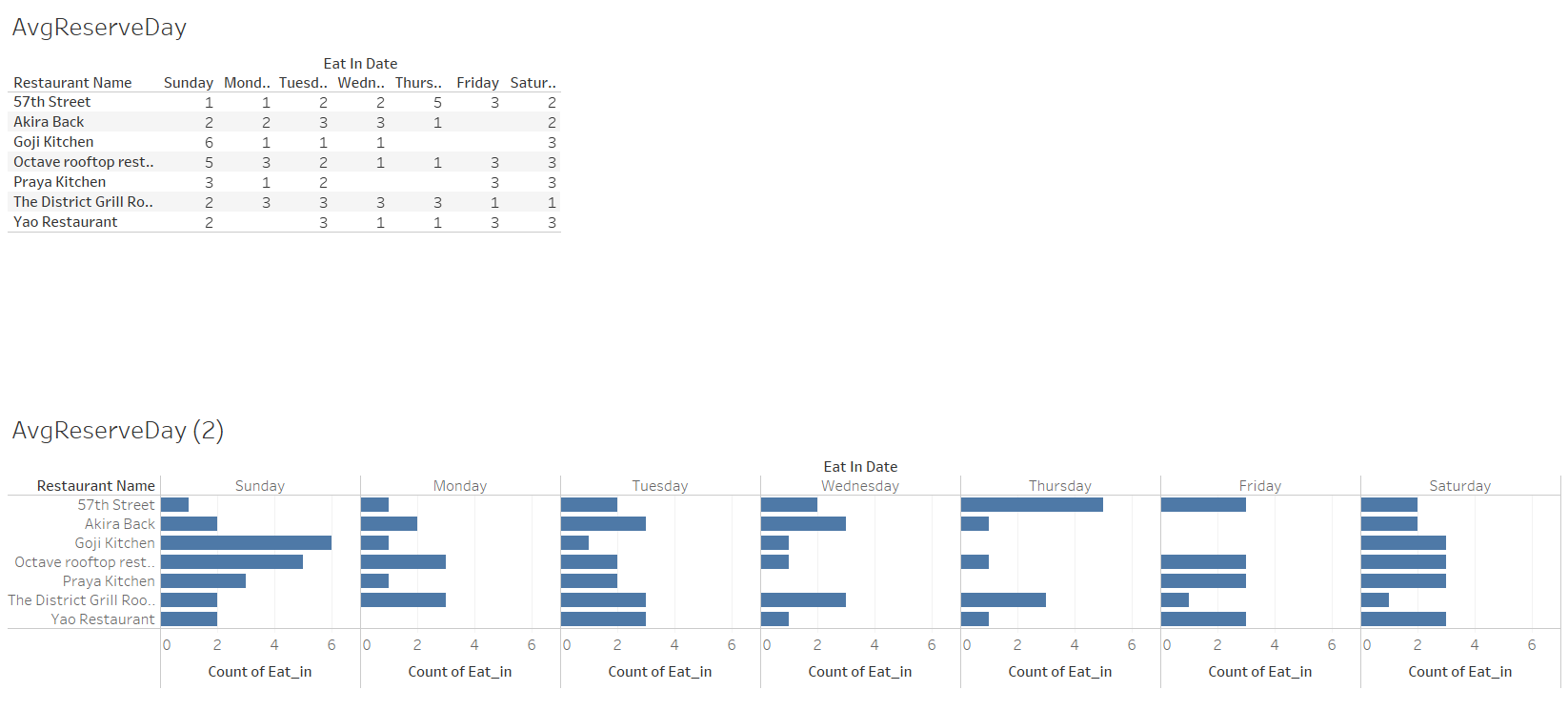


The visualized data from this story are derived from the revenue star schema which made up for the purpose to analyze the average income in each branch. This data can answer the company question about which branch has the most income. Furthermore, the manager can use these graphs to anticipate the average of income in each branch. These graphs can answer company questions as follows.

* Which branch has the most sales?
* Which branch has the least sales?
* Which is the average total sales for each branch?

**Restaurant Interest Fact Table**





The visualized data from this story are derived from the restaurant interest star schema which made up for the purpose to analyze the restaurant event interest and the average reserve per each day as well. This data can answer the company questions about which event that the customer mostly interests and the average restaurant has the customer reserve per day. Furthermore, the manager can use these graphs to anticipate the type of the event that customer interest and the demand in each restaurant per day. These graphs can answer the company question as follows.

* Which event is the most interesting?
* Which restaurant has the least reserve table per day?
* Which restaurant has the most reserve table per day?
* Which event is the least interesting?
* How many times is service in each restaurant?