Hotel Management Database

Group 10

Group Member：Yue Wu, Haolan Mai, Jiaqing Li, Yucong Li, Ruoxin Xu, Chuwei Peng

Client Description & Objective

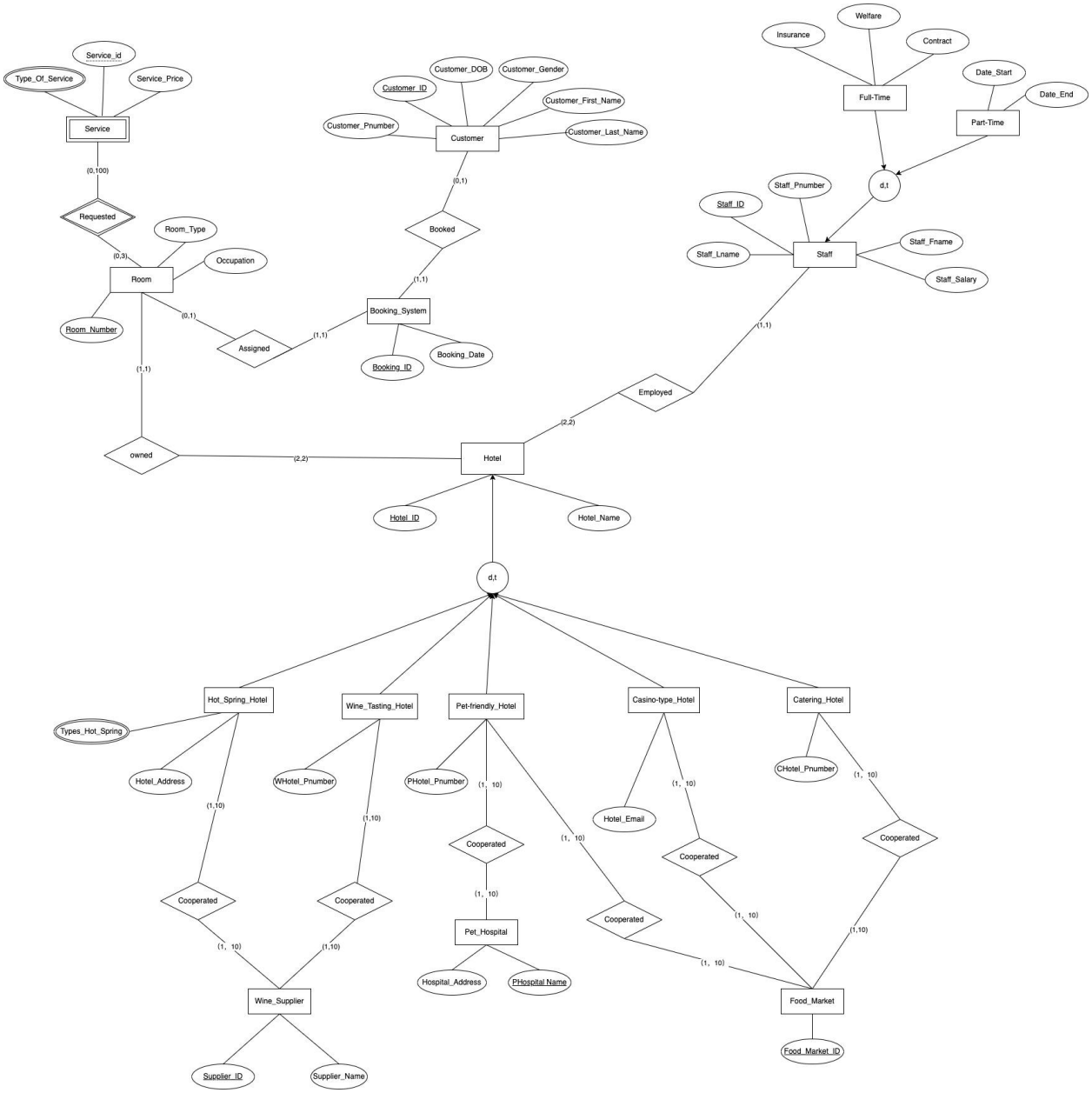
Client: An online travel shopping company

Objective:Make recommendations for the company's customers on Christmas.

Discussion of Data Collection

The hotel is an essential element of our travel, and we are curious about the huge data management system behind hotels. Therefore, we constructed a database named HOTEL. This database collected data from 50 hotels in five different categories, including information about hotel reservations, rooms, staff, and guests. Besides, the database also contains details about the partners associated with these five different types of hotels. Our collection is qualitative based on practical experience and social observation. We think this approach is more subjective. In the process, we need to add our judgment to the data. Therefore, the risk of bias is small. At the same time, we collect data through social observation to determine the dynamics of the situation. This often cannot be measured by other data collection techniques.

Simplified EER



1.Assume that every Pet_Hospital has a distinct Hospital_Name
2. A hotel can have 2 rooms.
3. A room can be owned by a hotel.
4. A hotel has two staff.
5. A staff member can be hired by a hotel.
6. A room can be reserved by 1 and 0.
7. Only one room can be reserved; one customer can make 0 to 1 reservation
8. There can only be one customer per reservation
9. One booking system can only be booked by one customer
10. Let's assume that a service is a weak entity and its parent entity is room.

11.The identity relationship between them is that a service can be requested by 0 to 100 rooms.
12. 1 room can order service 0 to 3 times.
13. The staff has two subclasses, which are divided into full_time and part_time. They are disjoint and total.And we use optionA to convert into the schema.
14. The five hotel types have a disjointed and total relationship with their superclass
15. A hot_Spring_Hotel, or Wine_Tasting_Hotel can work with 1 to 10 Wine_suppliers
16. A Wine_Supplier can cooperated with 1 to 10 Hot_Spring_Hotel and 1 to 10 Wine_Tasting _Hotel
17. One Pet-friendly_Hotel can work with anywhere from 1 to 10 Pet_hospitals
18. A Pet_Hospital can cooperated with 1 to 10 Pet-friendly_Hotel
19. A Pet-friendly_Hotel, or Casino-type_Hotel, Catering_Hotel can work with anywhere from 1 to 10 Food_Markets
20. A Food_Market can cooperated with 1 to 10 Pet-friendly_Hotel, Casino-type_Hotel, or Catering_Hotel
21. We assume Types_Of_Hot_Spring as a multivalued attribute.

Relational Schema Design

1. Hotel (Hotel_ID, Hotel_Name)
    1a. Hot_Spring_Hotel(Hotel_IDˆ1,Hotel_Address)
    1b. Wine_Tasting_Hotel(Hotel_IDˆ1,WHotel_Pnumber)
    1c. Pet_friendly_Hotel(Hotel_IDˆ1,PHotel_Pnumber)
    1d. Casino-type_Hotel(Hotel_IDˆ1,Hotel_Email)
    1e. Catering_Hotel(Hotel_IDˆ1,CHotel_Pnumber)
2. Room(Room_Number, Hotel_IDˆ1,Room_Type, Occupation)
3. Booking_System(Booking_ID,Customer_IDˆ4,Room_Numberˆ2,Booking_Date)
4.Customer(Customer_ID,Customer_Pnumber,Customer_DOB,Cutomer_Gender,Customer_First _Name, Custmer_Last_Name)
5. Staff(Staff_ID, Hotel_IDˆ1,Staff_Pnumber, Staff_Lname,Staff_Fname,Staff_Salary)
    5a.Full-Time(Staff_IDˆ5, Insurance, Welfare, Contract)
    5b.Part-Time(Staff_IDˆ5,Date_Start,Date_End)
6.Food_Market(Food_Market_ID)
7.Pet_Hospital(PHospital_Name, Hospital_Address)
8.Wine_Supplier(Supplier_ID, Supplier_Name)
9.Service(Service_ID, Room_Numberˆ2, Service_Price)
10. Hot_Spring_Cooperated(Hot_Spring_Hotel_IDˆ1a,Supplier_IDˆ8)
11.Wine_Cooperated(Wine_Hotel_IDˆ1b,Supplier_IDˆ8)
12.Pet_cooperated(Pet_Hotel_IDˆ1c,PHospital_Nameˆ7)
13.Pet_cooperated_FM(Pet_Hotel_IDˆ1c,Food_Market_IDˆ6)
14.Casino_cooperated(Casino_Hotel_IDˆ1d,Food Market_IDˆ6)
15.Catering_cooperated(Catering_Hotel_IDˆ1e, Food_Market_IDˆ6)
16.Type_Of_Service(Type_Of_Service,Service_IDˆ9,Room_Numberˆ2)
17.Type_Of_Hot_Spring(Type_Of_Hot_Spring,Hot_Spring_Hotel_IDˆ1a)

Normalization

Our database is in its normal form.

Our database is in the First Normal Form. The definition of the 1NF is all attributes are single-valued. From the EER, we can find that there are two multi-valued attributes, Type_Of_Service and Types_Hot_Spring. In order to normalize them, we have created a relation for each of them.

Our database is in the Second Normal Form. This is because it is in First Normal Form and every non-prime attribute is fully functional depending on the primary key in each relation.

Our database is in the Third Normal Form. Since it is in the Second Normal Form and there are no non-prime attributes in each relation being transitively dependent on the primary key. Every non-prime attribute does not depend on another non-prime attribute.

Our database is in the Boyce Codd Normal Form. Because it is in the Third Normal Formal and every non-superkey cannot imply the superkey in each relation.

MySQL Relationship View

Here shows 10 tables implemented in MySQL Relationship View.

1. Casino_Cooperated

Casino_Hotel_ID	Food_Market_ID
31	F00001
32	F00002
33	F00003
34	F00004
35	F00005
36	F00006
37	F00007

2. Catering_Cooperated

	Catering_Hotel_ID	Food_Market_ID
▶	41	F00001
	42	F00003
	43	F00003
	44	F00004
	45	F00006
	46	F00006
	47	F00007

3. Hot_Spring_Cooperated

	Hot_Spring_Hotel_ID	Supplier_ID
▶	1	S00001
	2	S00002
	3	S00004
	4	S00004
	5	S00005
	6	S00006
	7	S00007

4. Pet_Cooperated

	Pet_Hotel_ID	PHospital_Name
▶	21	Bear
	22	Tiger
	23	Bee
	24	Bear
	25	Dog
	26	Panda
	27	Dinosaur

5. Pet_Cooperated_PM

	Pet_Hotel_ID	Food_Market_ID
▶	21	F00001
	22	F00002
	23	F00003
	24	F00003
	25	F00005
	26	F00006
	27	F00007

6. Wine_Cooperated

	Wine_Cooperated_Hotel_ID	Supplier_ID
▶	11	S00001
	12	S00002
	13	S00003
	14	S00005
	15	S00005
	16	S00006
	17	S00007

7. Hotel

	Hotel_ID	Hotel_Name
▶	1	Hotel01
	2	Hotel02
	3	Hotel03
	4	Hotel04
	5	Hotel05
	6	Hotel06
	7	Hotel07
	8	Hotel08

8. Hot_Spring_Hotel

	Hotel_ID	Hotel_Address
▶	1	0001 berkeley
	2	0002 berkeley
	3	0003 berkeley
	4	0004 berkeley
	5	0005 berkeley
	6	0006 berkeley
	7	0007 berkeley
	8	0008 berkeley

9. Wine_Tasting_Hotel

	Hotel_ID	WHotel_Pnumber
▶	11	2020000001
	12	2020000002
	13	2020000003
	14	2020000004
	15	2020000005
	16	2020000006
	17	2020000007
	18	2020000008

10. Pet_Friendly_Hotel

	Hotel_ID	PHotel_Pnumber
▶	21	3030000001
	22	3030000002
	23	3030000003
	24	3030000004
	25	3030000005
	26	3030000006
	27	3030000007
	28	3030000008

Interesting Queries:

- 1.Find the number of Staff & the number of Room for each type of hotel/room

```
WITH CTE AS
(
SELECT *, 'Hot_Spring_Hotel' AS Hotel_type
FROM Hotel
WHERE Hotel_ID IN (SELECT Hotel_ID FROM Hot_Spring_Hotel)
UNION ALL
SELECT *, 'Wine_Tasting_Hotel' AS Hotel_type
FROM Hotel
WHERE Hotel_ID IN (SELECT Hotel_ID FROM Wine_Tasting_Hotel)
UNION ALL
SELECT *, 'Pet_Friendly_Hotel' AS Hotel_type
FROM Hotel
WHERE Hotel_ID IN (SELECT Hotel_ID FROM Pet_Friendly_Hotel)
UNION ALL
SELECT *, 'Casino_Hotel' AS Hotel_type
FROM Hotel
WHERE Hotel_ID IN (SELECT Hotel_ID FROM Casino_Hotel)
UNION ALL
SELECT *, 'Catering_Hotel' AS Hotel_type
FROM Hotel
WHERE Hotel_ID IN (SELECT Hotel_ID FROM Catering_Hotel)
)
```

```
SELECT Hotel_Type, SUM(Staff_Num) AS Total_Staff, SUM(Room_Count) AS Total_Room
FROM
(
SELECT Hotel_ID, COUNT(Staff_ID) AS Staff_Num
FROM Staff
GROUP BY Hotel_ID) t
LEFT JOIN CTE c
ON c.Hotel_ID = t.Hotel_ID
LEFT JOIN
(
    SELECT Hotel_ID, COUNT(Room_Number) AS Room_Count
    FROM Room
    GROUP BY Hotel_ID
) g
ON c.Hotel_ID = g.Hotel_ID
GROUP BY Hotel_Type
```

Hotel_Type	Total_Staff	Total_Room
Hot_Spring_Hotel	20	20
Catering_Hotel	10	20
Wine_Tasting_Hotel	10	20
Pet_Friendly_Hotel	10	20
Casino_Hotel	10	20

- 2. Which types of hotels/rooms are popular on Christmas according to reservation?

```
SELECT Count(*) As No_of_Reservation,Room_Type
FROM Room
WHERE Room_Number in (SELECT Room_Number FROM Booking_System)
Group By Room_Type
```

No_of_Reservation	Room...
10	Sprin...
8	Wine...
2	Pet_...
2	Cateri...

- 3. Find how many rooms are occupied for each types of hotel/room

```
SELECT Count(Occupation) As No_of_Occupied,Room_Type
FROM Room
Where Occupation = 'Y'
Group By Room_Type
```

No_of_Occupi...	Room...
11	Sprin...
12	Wine...
2	Pet_...
1	Casin...
5	Cateri...

Computation in Python & R

- The occupation rate for each type of hotel
  - There are 20 rooms for each type of hotel, we divide the number of occupied rooms by 20 for each type and then we obtain the occupation rate.

R code and output:

```
library(readr)
occupied <- read_csv("occupied.csv")
occupied['occupation rate'] = occupied['No_of_Occupied']/20
```

No_of_Occupied <dbl>	Room_Type <chr>	occupation rate <dbl>
11	Spring_Room	0.55
12	Wine_Room	0.60
2	Pet_Room	0.10
1	Casino_Room	0.05
5	Catering_Room	0.25

- The ‘Staff/Room’ ratio for each type of hotel
  - There are 20 rooms for each type of hotel, we divide the number of staff by 20 for each type and then we obtain the ‘staff/room’ ratio.

Python code and output:

In [15]:
```
import pandas as pd
import matplotlib.pyplot as plt
data = pd.read_csv('Hotel_Type_final.csv')
```

In [18]:
```
data['Room_Staff_Ratio'] = data["Total_Staff"]/data['Total_Room']
data
```

Out[18]:

	Hotel_Type	Total_Staff	Total_Room	Room_Staff_Ratio
0	Hot_Spring_Hotel	20	20	1.0
1	Catering_Hotel	10	20	0.5
2	Wine_Tasting_Hotel	10	20	0.5
3	Pet_Friendly_Hotel	10	20	0.5
4	Casino_Hotel	10	20	0.5

Explanation:

There are 5 types of hotel: hot spring hotel, wine tasting hotel, pet friendly hotel, casino hotel and catering hotel. Customers can choose different types of hotels according to their travel needs. In addition, we also analyzed two aspects, so as to provide customers with more suitable choices on Christmas.

Firstly, we calculated the ratio of the number of employees to the number of rooms in each hotel. Generally, the higher the ratio, the higher the service quality of the hotel. In addition, we also calculated the occupation rate and booking rate of the hotel. This indicator shows the hotel's Popularity. Some customers like to choose hot hotels, while some want to avoid crowds to spend a quiet holiday, so this indicator can also help customers to choose.

Since all the data is virtual, our data analysis and data visualization is only for illustration. The analysis will be more useful in an actual database.