

Introduction to Data Management - HW2: DDL Script Assignment

Team Number: 20

Team Members: Abhinav Sharma (ass2575), Allie Touchstone (awt529), Aritra Chowdhury (ac79277), Avery Shepherd (ams9694), Harsh Mehta (hdm564), Vishal Gupta (vg22846)

Problem Statement

We want to build a database that effectively implements the ERD and each table has proper relationships. The aim is to create an organized, efficient and fast database so that we're able to reduce redundancies and ensure entities are easy to query for analysis.

Proposed Solution

Deletion of Indexes, Sequences and Tables

The beginning of our code drops all of the sequences and tables later created. If running the code for the first time, there is an expected error for dropping the tables. It is written like this so the script can be run over and over again without any problems.

Creation of Sequences and Tables

According to the given ERD, we then created tables. We defined the data types for the features in each table and added the required constraints for example, defining the foreign key and the primary key it is referring to. These constraints will help preserve the integrity of our database. We generated sequences to assign default incremental values to all primary keys like Customer_ID, Reservation_ID, etc.

Data Seeding

We tested our constraints initially by passing data in accordance with the constraints thereby ensuring that they were working as expected (for example, we added Foreign Keys in accordance with the incremental Primary Keys in referenced tables). Then we further tested our implementation by inserting the given values into the dataset.

Creation of Indexes

Indices help speed the retrieval of data. Apart from creating indices for foreign keys that aren't primary keys we create two additional indices for better usability. The first is *Customer_Name_ix*, a combination of customers' first and last name from the *Customer* table. This index could be imperative while understanding which customers are repeat customers and thereby more valuable for hotels' business. The second is *Room_Floor_Room_Number_ix* which similarly is a combination of Floor and room number from the *Room* table. We expect this index to be of use when analyzing unused capacity of hotels at any given point in time

Assumptions

We have assumed primary keys are both unique and not null. Secondly, if a Check_Out_Date is not given, the Reservation_Total is assumed to be the charge for one night.

Conclusion and Next Steps

We have designed the database in accordance with the ERD provided while using sequences to allow easier primary key generation and auto-increment. We've defined a limited number of indices so that we have high utility with little to no impact on speed. The next key step would be maintaining and updating the indices in our database as data volume increases and priority of which analysis are of key interest change over time.