

Team Project: IT Startup Case Study

Introduction

Your team is the executive committee of a fictitious IT startup company, which is providing a platform for restaurants, service industry in general, to publish their menus, promotions, ads, etc... The company has been so successful that a well-designed database to tract access to its platform, to monitor usage, and to develop business insights have been turning out to be increasingly crucial.

The company's business model is roughly based on revenue from monthly subscription fees paid by the clients, in return providing access to different platforms (including the one hosted by you), as a result increasing the visibility of the clients to the general public. There are three subscription levels and your employees act as agents for more than 15,000 clients.

Project Functional Requirements (FRs)

A database model needs to be established with schemas, normalized tables, relationships, queries, and business intelligence analysis. In particular, below are the itemized functional requirements.

FR1 (Human Resources): Add a third department, Operations, to the Department table. Add executives' data (three individuals including yourselves) into the Employee table (i.e. new rows 11, 12, 13). Position field will be 'Executive'. Add another column (called Head) to the Employee table and assign each department to one executive through Dept_ID. *Specifically, if an employee is a head of a department, his/her Head field will be Dept_ID of that department (i.e. 1,2,3) otherwise it will be zero* (All these steps must be performed in your SQL script).

FR2 (Data Model): In addition to the existing tables (**Employee** and **Department**) create and populate the following tables: **Client**, **View**, **Pricing**, **TypeClient**, and **AgentRegion**¹. **Be extra careful about data types.** PK-FK pairs' data types in different tables should be the same². Assign primary keys for each table. Create their Entity Relationship (ER) diagram in Management Studio. Note, data for the other new tables is provided in the Excel file above in this module. (Your script must include: CREATE TABLE commands for all tables (including the Client and View) and BULK INSERT command for Client and View. You can populate the other tables with simple copy/paste from Excel, otherwise provide the SQL code.)

FR3 (Queries): Write the following retrieval queries in your script. * means format as a number with two decimals and ** means format as a currency.

FR3.Q1: Top ten Spas & Salons *names* that have the highest views.

FR3.Q2: All clients whose *names* **start OR end** with the term 'Grill', along with their *cities*, *subscription fees***, and *number of views*.

FR3.Q3: Count of client types (Arts & Entertainment, Bakery, Restaurant, etc.) with their average views per client* and average subscription fees per client** sorted with respect to average views per client in descending order.

FR3.Q4: Cities (along with their regions) for which total number of views for non-restaurant clients are more than 15 (sorted in a descending order of total number of views).

¹ Note that this is an associate table between Client and Employee tables.

² For instance, *Planno* field in the Pricing table is named as *Pricing* field in the Client table. Similarly, *ClientID* in the Client table is *ID* in the View table. Ponder over fields by examining tables' data and identify those connections.

FR3.Q5: Number of clients and average views per client* with respect to the hosts in a descending order of average views.

FR3.Q6: Number of clients, their total fees**, total views, and average fees per view** w.r.to regions, sorted in descending order of average fees per views.

FR3.Q7: All views (all columns) that took place after October 15th, by Kindle devices, hosted by Yelp from cities where there are more than 10 clients. Also add the name of the client (as a first column) and city of the client (as a second column) for each view.

FR3.Q8: All non-executive employee full names in the first column, number of their regions, number of their clients, and number of views for those clients in columns 2, 3, and 4, respectively.

FR4 (Business Intelligence): Retrieve the corresponding data through two queries (provide in your script) then make the following computations and visualization in Excel.

FR4.BI1: Is there a correlation between price paid and number of views for clients? Comment in Excel.

FR4.BI2: Create a chart with average number of views per day (during the month of October) in the vertical and hours of the day (0 to 23) in the horizontal axis. Apply a simple regression model (i.e. linear, polynomial up to order 3, exponential, etc.) on the chart via Excel trendlines. Which regression model has the best explanatory power in terms of the measure of goodness of fit (i.e. R^2 value)? Only keep the best regression equation on the chart. Is there a prominent pattern?

FR5 (General Rules):

- Write all SQL code in a single script.
- Write all team members' names in the comments.
- Write functional requirement codes and descriptions in the comment outs.
- Do not change field (column) and table names.
- For subqueries, use indentation. Your subqueries should work separately inside the parenthesis.
- Follow our best style for not losing minor points.
- Deliverables:
 - ER Diagram (.pdf)
 - Complete SQL script file (.sql)
 - Excel for Business Intelligence analysis with only two sheets (.xlsx)
- The grading rubric:
 - FR4.BI1 and FR4.BI2 are 10 points each
 - All other FR's are 8 points each
 - Total 100 points
- Submit your work through Canvas / Assignments / Team Project
 - One submission per team
 - Write team member names in both Canvas comments and in the file
 - Due date: Check it via Canvas Assignments