Improve your client's Social Media strategy

openclassrooms.com/en/projects/improve-your-clients-social-media-strategy/assignment

🕐 50 hours

This project was recently improved! If you started working on it before June 17th, 2022, you have the option of continuing to use the **previous version**. You can also choose to use the new and improved version. If you wish to continue on the previous version, you can specify "version before 06/17/22" on your deliverable.

As a freelance Data Analyst, you have landed a 3-month contract at Datazine, a French **online magazine company.** You have been hired to help with their social media strategy and to provide advice on the potential launch of an English version of their magazine.

Subscribers can sign up for the online magazine to engage with the content. Readers can also connect to Datazine's Facebook page, which gives Datazine an opportunity to analyze how readers from across the world engage with Datazine.

Datazine wants to better understand the readers that engage via Facebook, and the staff ultimately hopes to use this information to encourage new subscribers to the online magazine.

On your first day, the Head of Communications sends you an email to brief you on your first project.

Subject: First Project - Querying media analytics using SQL

Hi,

Welcome to Datazine. I am Alexandra and I lead Communications at Datazine.

We want to **evaluate the effectiveness of our magazine's Facebook strategy.**

Following is the information that we need to analyze:

- Key indicators, such as location, gender, age, and language of our users.
- The best time of day, and the best day of week to publish Facebook content.
- Assess the feasibility of launching a US English version.

Datazine downloads all their media analytics data from Facebook's server but since these are raw numbers, they need someone who can draw insights from these data files.

Like most organizations, Datazine also wants to set up a **database** where they can keep a record of all such social media analytics.

Please could you set up the database containing the Facebook data and query the required insights from the database?

I've attached the dataset and a brief with the important Facebook marketing terminology which appear in the data files.

We can also set up some time to discuss further details.

Regards,

Alex

Attachment:

- <u>Data files extracted from Facebook servers</u>
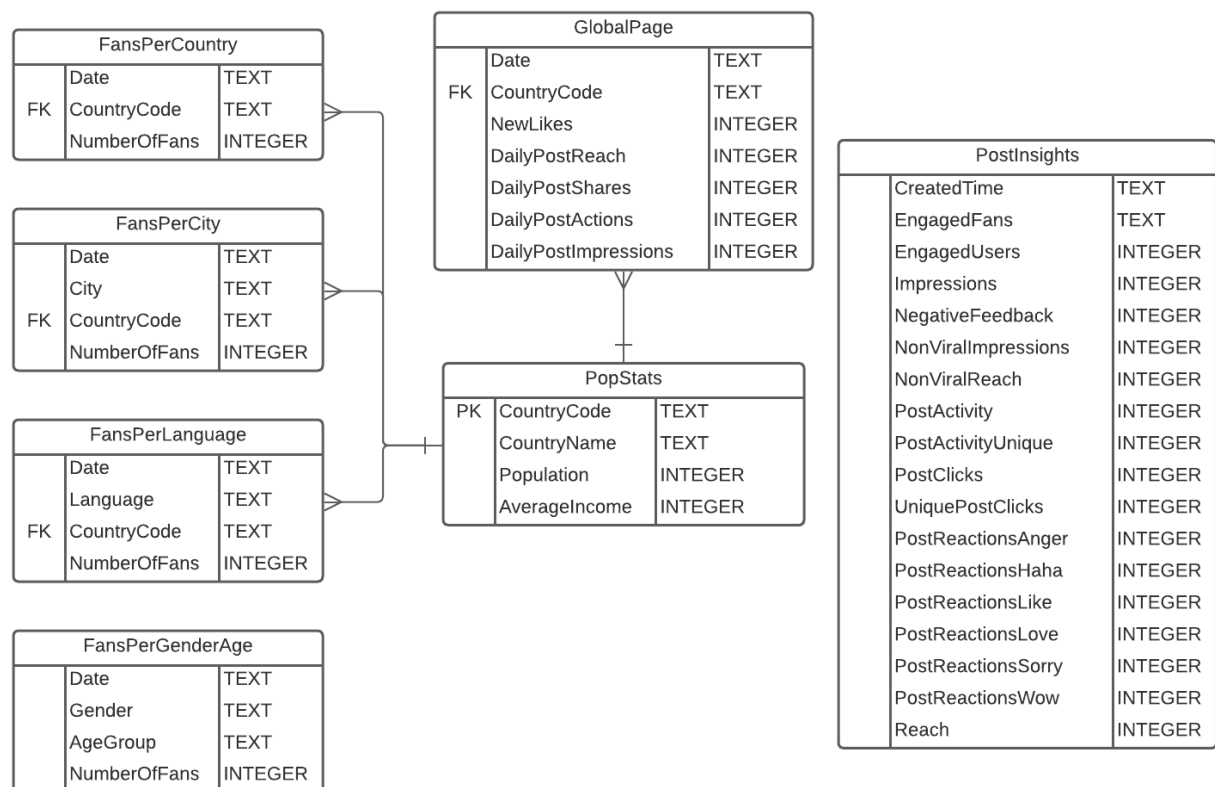- <u>Brief - marketing terminology</u>

After your discussions with Alexandra, you detailed out what needs to be done in the following <u>specifications document</u> which you'll have to complete.

Your tasks

Task 1 - Loading CSV files into a Database

Files don't offer a scalable solution and it's hard to query & filter data from files. So you should load the data files into a database.

Each data file corresponds to a table in the database you will create. The following Entity Relationship Diagram (ERD) represents the structure of the final database.

FansPerCountry

	Date	TEXT
FK	CountryCode	TEXT
	NumberOfFans	INTEGER

FansPerCity

	Date	TEXT
	City	TEXT
FK	CountryCode	TEXT
	NumberOfFans	INTEGER

FansPerLanguage

	Date	TEXT
	Language	TEXT
FK	CountryCode	TEXT
	NumberOfFans	INTEGER

FansPerGenderAge

Date	TEXT
Gender	TEXT
AgeGroup	TEXT
NumberOfFans	INTEGER

GlobalPage

	Date	TEXT
FK	CountryCode	TEXT
	NewLikes	INTEGER
	DailyPostReach	INTEGER
	DailyPostShares	INTEGER
	DailyPostActions	INTEGER
	DailyPostImpressions	INTEGER

PopStats

PK	CountryCode	TEXT
	CountryName	TEXT
	Population	INTEGER
	AverageIncome	INTEGER

PostInsights

CreatedTime	TEXT
EngagedFans	TEXT
EngagedUsers	INTEGER
Impressions	INTEGER
NegativeFeedback	INTEGER
NonViralImpressions	INTEGER
NonViralReach	INTEGER
PostActivity	INTEGER
PostActivityUnique	INTEGER
PostClicks	INTEGER
UniquePostClicks	INTEGER
PostReactionsAnger	INTEGER
PostReactionsHaha	INTEGER
PostReactionsLike	INTEGER
PostReactionsLove	INTEGER
PostReactionsSorry	INTEGER
PostReactionsWow	INTEGER
Reach	INTEGER

Datazine ERD
Link to pdf

In order to set up the database for the project, you need to **create** a new **database** and then inside that database, create **a table for each of the files** in the provided dataset using the CREATE commands.

Download the data extracted from Facebook and create tables in SQLite (or any other RDBMS of your choice) according to the above ER diagram. Note the primary and foreign keys indicated in the diagram, which you will need to add to your database structure.

After successful creation of these tables, import (or load) the data from CSV files to their dedicated table.

If you are using SQLite, here is a Quick Start Guide to get you started.

Task 2 - Query data using SQL to answer questions

With the data loaded in SQL tables, you can now start querying data.

Write SQL statements to answer all business questions listed by the Facebook community manager in the specification document.

You are required to use SQL for this work, but you may choose whichever RDBMS you like. For beginners, we recommend SQLite.

Deliverable

1. The completed specification document, to be used as a visual aid for your project presentation.

To make it easier for your work to be reviewed by the jury, upload your project deliverable to the platform in a zip folder called 'Project_title_LastName_FirstName'.

Use the following naming convention for your deliverable:

LastName_FirstName_1_specification_document_mmyyyy

For example, the deliverable could be named: Smith_Mary_1_specification_document_042022

Project Presentation

You'll be presenting your answers and to an assessor who will play the role of Alexandra (Head of Communications).

During the oral presentation, your assessor will play the role of Alexandra, Head of Communications. The assessor will challenge your decisions, so be prepared to defend your work. The presentation will be structured as follows:

- Presentation of deliverables (15 minutes)
    - Describe the different tables you have created (Task 1).
    - Present the answers to the different business questions, as per the specification document (Task 2).
    - Connect to your database and run three random queries among those you've created (Task 2).
- Discussion (10 minutes)
        Playing the role of Alexandra, the assessor will ask you questions about your methodology and your deliverables.
- Debrief (5 minutes)
        At the end of the sessions, the assessor will stop playing the role of Alexandra so that you can debrief together.

Your presentation should last 15 minutes (+/- 5 minutes). Respecting presentation time requirements is important in professional environments. In consequence, if your presentation is under 10 minutes or over 20 minutes, you may be asked to redo the assessment.

Skills

- 💡
    Retrieve and analyze data using SQL queries

- 💡
    Apply relational algebra operations using SQL

- 💡
  Prepare the database and dataset using SQL

- 💡
  Apply the concepts of the relational model using SQL