CS 378 Project 5, due Thursday, 10/17.

In this project, we will build the mart layer of the warehouse, a layer that sits on top of the intermediate layer and takes the intermediate tables as input. The purpose of the mart layer is to source dashboards and reports that help business leaders understand the health of their business and decide where to make investments. The mart layer is arguably the most important one in the warehouse because it directly impacts the end users.

The mart layer consists of multiple marts. Each one is typically designed to meet the reporting needs of a different business unit. Marts can take on several shapes: they can be represented as a traditional star schema, a snowflake schema or even as standalone reporting tables.

Given the short timeline of this project, we will build our mart layer as a collection of standalone reporting tables.

Objectives
- Imagine the personas of your end users and what reporting needs they may have that can be answered by your warehouse data. Come up with 3-5 business questions that will drive the design of your reporting tables. Make note of these questions in your Colab notebook.
- Implement 10 reporting tables to answer the questions that you came up with. Note that some questions may translate into a single reporting table while others could produce 2-4 reporting tables.
- Each reporting table should join multiple intermediate tables and aggregate the results in a meaningful way.
- As a whole, your reporting tables should be derived from at least 2/3 your intermediate tables and from 100% of your data sources.

Implementation Details
- Create a mart area in BigQuery and populate it with your resulting objects (star schema or standalone views). **Do NOT mutate the intermediate area**. All mutations should be applied only to the mart area of the warehouse. Name your BQ dataset `[your-domain]-mrt` where `mrt` is short for mart. For example, air_travel_mrt.
- Develop a Colab notebook that implements the 10 reporting tables. Name your notebook `5-[your-domain]-mrt.ipynb`.
- Document your business questions as short Markdown comments in your notebook using section headers.
- Document any non-trivial design choices, especially if you came up with broad business questions that you want to explore from different angles.
- As you craft your queries, make note of any outstanding data issues that you encounter (e.g. missing values, etc.) as Markdown comments in your notebook.
- Choose descriptive names for your reporting tables.

- Use lowercase letters for table names and column names in this layer.
- Create a new folder in your repo and name it `project5`. Store your notebook for this project in the `project5` folder. You do not need to create an ERD.
- Create a submission.json file and upload it to Canvas by the deadline. Only one person per group needs to do this step.

CS 378 Project 5 Rubric
Due Date: 10/17/24

`5-[your-domain]-mrt.ipynb` is thorough and meets all requirements
    -10 for each empty or missing reporting table in mart dataset
    -7 for each reporting table created without joins or aggregation
    -7 mart data as a whole is missing a data source
    -7 mart data as a whole does not span 2/3 of intermediate tables
    -5 notebook is missing one or more business questions
    -5 notebook lacks Markdown annotations and is hard to follow
    -5 did not follow naming convention for dataset, tables or columns
  -80 missing file
	100

`submission.json` submitted into Canvas. Your project **will not** be graded without this submission. The file should have the following schema:
 	
`{`
    `"commit-id": "your most recent commit ID from Github",`
    `"project-id": "your project ID from GCP"`
`}`
 	
Example:
 	
`{`
    `"commit-id": "dab96492ac7d906368ac9c7a17cb0dbd670923d9",`
    `"project-id": "some-project-id"`
`}`
	Required

Total Credit:	**100**