CLASS PROJECT

APAN 5310 - SQL & RELATIONAL DATABASES

COLUMBIA UNIVERSITY
School of Professional Studies

Scope

- Assume you are a team of database experts hired by a client

- Your client either (a) has a lot of archived, unstructured data that they need to have organized for easy access and analysis or (b) they need a database for a new app/system/service they are about to offer

- Your client requires two levels of access to the data:

    - Analysts: write and execute SQL code as well as access the database through Python/R

    - Managers/C-executives: high-level overview through visualizations and interactive dashboards that automatically update when new data is stored in the database

SPS

Requirements

- Based on instructor guidance your team will either have to propose a scenario and have it approved or select from the provided list of scenarios

- Find data relevant to the scenario, course instructors will review the data and approve. There is no specific limit but small datasets or pre-defined relational datasets are not appropriate.

- Develop a relational schema and load data (Python/R can automate the loading process)

- Demonstrate the benefits of the relational database by writing at least 10 complex queries that lead to valuable client insights

- Build interactive dashboards in Metabase/R Shiny/Bokeh

COLUMBIA UNIVERSITY
School of Professional Studies