

Milestone 9 due Friday, 12/07.

1. Update your ERD to reflect the newest version of all your tables:
 - Diagram should represent all latest tables in `dataset1` and `dataset2`.
 - Entities should specify the field names, data types, and keys of each table.
 - Diagram should visually identify the dataset that each table belongs to (e.g. use a different background color for each dataset).
 - Name your updated ERD file `ERD-v3.pdf`.

2. Implement your cross-dataset queries using the transformed tables from Milestones 6 and 8:
 - You should end up with 4 interesting queries that span `dataset1` and `dataset2`.
 - At least 2 queries must use a table that was transformed through Beam.
 - Wrap each query in a SQL view using the `CREATE VIEW` command.
 - Create each view in `dataset1`.
 - Copy the SQL into a `cross-dataset-queries.sql` file.
 - Add a short comment above each SQL statement to describe the query. Comments should begin with a `--` (e.g. `--this is a legal comment in SQL`).

3. Create data visualizations:
 - Open [Data Studio](#)
 - Create a Data Source inside Data Studio for each SQL view.
 - Create a chart in Data Studio that visualizes the data in a compelling way.
 - Add the 4 charts to your existing Data Studio dashboard.
 - Take a screenshot of your dashboard and save it as `dashboard-v4.png`.

4. Verify queries, views, and charts:
 - re-run queries and views developed in previous milestones.
 - fix any broken queries and views you find.
 - update the appropriate `.sql` files with code fixes.
 - open Data Studio report and fix any broken charts on your dashboard.

CS 327E Milestone 9 Rubric

Due Date: 12/07/18

<p>Part 1 - ERD should match the schema of BigQuery tables, and relations should be finalized.</p> <ul style="list-style-type: none"> -30 no ERD-v3.pdf in repository, or missing tables -5 missing field names, missing data types, or missing keys (max. -15 for each table) -20 tables are not color-coded by dataset -5 per incorrect schema 	30
<p>Part 2 - Build views in BigQuery of your queries created from Milestone 7, and create them under dataset1. Save the queries into cross-dataset-queries.sql.</p> <ul style="list-style-type: none"> -40 no cross-dataset-queries.sql in repository, or missing queries -5 each missing comment, up to -20 -10 each missing view from dataset1, up to -40. -10 each of 2 queries that don't use transformed tables 	40
<p>Part 3 - Create data visualizations for your SQL views using Data Studio, and save them in a dashboard-v4.png file.</p> <ul style="list-style-type: none"> -30 no dashboard-v4.png in repository, or missing charts -5 missing titles/labels from charts, up to -20 	30
<p>submission.json submitted into Canvas. Your project will not be graded without this submission. The file should have the following schema:</p> <pre>{ "commit-id": "your most recent commit ID from Github", "project-id": "your project ID from GCP" }</pre> <p>Example:</p> <pre>{ "commit-id": "dab96492ac7d906368ac9c7a17cb0dbd670923d9", "project-id": "some-project-id" }</pre>	Required
<p>Total Credit:</p>	100