

Milestone 7 due Friday, 03/15.

1. Update your ERD to reflect the following:
  - state of modeled tables post Beam transforms.
  - field names, data types, and keys (PK, FK) for each entity.
  - relationships between entities.
  - save the file as `ERD-v3.pdf`
  
2. Develop aggregate queries in SQL on `dataset1` as follows:
  - 8 SQL queries with aggregation.
  - At least 1 aggregate function per query.
  - At least 2/5 aggregate functions among the 8 queries.
  - At least 5 `GROUP BY` clauses among the 8 queries.
  - At least 3 `HAVING` clauses among the 8 queries.
  - Copy the SQL code into a file named `aggregate-queries.sql`.
  - 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:
  - Choose your 4 most interesting aggregation queries.
  - Create a BQ view for each query.
  - Open [Data Studio](#)
  - Create a Data Source in Data Studio that accesses each view.
  - Create a chart in Data Studio that visualizes the data in a compelling way.
  - Add the 4 charts to a single Data Studio report (aka dashboard).
  - Take a screenshot of your dashboard and save it as `dashboard-v1.png`.

CS 327E Milestone 7 Rubric

**Due Date: 03/15/19**

<p>Upload an updated ERD that finalizes your table schema after Beam transformations have been applied.</p> <ul style="list-style-type: none"> <li>-40 ./ERD-v3.pdf not found in repository</li> <li>-20 ERD missing a primary key for each table</li> <li>-10 each relation missing a foreign key, or each foreign key without an associated relation</li> <li>-10 missing data types on a table</li> </ul>	40
<p>Create data visualizations for your datasets and save them in dashboard-v1.png. These data visualizations should represent a view.</p> <p>The image should contain 3 charts made from Data Studio, with a relevant title for each one describing the dataset.</p> <ul style="list-style-type: none"> <li>-20 ./dashboard-v1.png not found in repository</li> <li>-10 each missing chart, up to -20</li> <li>-5 each missing title, up to -20</li> <li>-5 each chart representing a table, up to -20</li> </ul>	20
<p>Create a file aggregate-queries.sql containing 5 queries involving aggregation. Two should involve the use of a GROUP BY clause. Each SQL query should be preceded by a comment describing its function.</p> <ul style="list-style-type: none"> <li>-40 ./aggregate-queries.sql not found in repository</li> <li>-5 each missing aggregate statement, up to -40</li> <li style="padding-left: 20px;">If all statements use the same aggregate function, you will only receive half credit</li> <li>-5 each missing or incorrect comment, up to -25</li> <li>-5 each missing GROUP BY clause of the 5, up to -25</li> <li>-5 each missing HAVING clause of the 3, up to -15</li> </ul>	40
<p>submission.json submitted into Canvas. Your project <b>will not</b> 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><b>Total Credit:</b></p>	<b>100</b>