

Milestone 7 due Friday, 10/25.

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-dataset1-beam.pdf`

2. Develop aggregate queries in SQL on `dataset1` as follows:
 - 6 SQL queries with aggregation.
 - At least 1 aggregate function per query.
 - At least 2/5 aggregate functions among the 6 queries.
 - At least 4 `GROUP BY` clauses among the 6 queries.
 - At least 3 `HAVING` clauses among the 6 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 3 most interesting aggregate queries.
 - Create a BQ view for each query. Assign each view a descriptive name and prefix each view with `v_` (e.g. `v_Highest_Nominated_Movies`).
 - Open [Data Studio](#)
 - Create a Data Source that accesses each view. You'll need one Data Source per view.
 - Create a chart in Data Studio that visualizes the data in a compelling way.
 - Add all 3 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: 10/25/19

<p>Upload an updated ERD that finalizes your table schema after Beam transformations have been applied.</p> <ul style="list-style-type: none"> -40 ./ERD-dataset1-beam.pdf not found in repository -20 ERD missing a primary key for each table -10 each relation missing a foreign key, or each foreign key without an associated relation -10 missing data types on a table 	40
<p>Create a file aggregate-queries.sql containing 6 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"> -30 ./aggregate-queries.sql not found in repository -5 each missing aggregate statement, up to -30 <li style="padding-left: 20px;">If all statements use the same aggregate function, you will only receive half credit -5 each missing or incorrect comment, up to -25 -5 each missing GROUP BY clause of the 4, up to -20 -5 each missing HAVING clause of the 3, up to -15 	30
<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"> -30 ./dashboard-v1.png not found in repository -10 each missing chart, up to -30 -5 each missing title, up to -15 -5 each chart representing a table, up to -15 	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

