CS 327E Milestone 7 due Sunday, 03/15.

1. Update your ERD to reflect the Beam tables in your modeled dataset:
   - The diagram should **omit** any modeled tables which were replaced by Beam tables.
   - Each entity should be represented by its collection of field names, data types, and primary key (PK).
   - Each child entity should also contain a foreign key (FK).
   - The diagram should include valid relationships between entities.
   - Name the ERD file `<source>_erd_beam.pdf` where `<source>` is the source of your `dataset1`.

2. Practice writing SQL queries with aggregations:
   - Create a new Jupyter notebook named `<source>_analysis.ipynb` where `<source>` is the source of your `dataset1`.
   - Write **6** queries with at least 1 aggregate function per query.
   - At least 3 queries must use a `JOIN` clause.
   - At least 4 queries must use a `GROUP BY` clause.
   - At least 2 queries must use a `HAVING` clause.
   - At least 4 queries must use an `ORDER BY` clause.
   - Add a short Markdown comment above each SQL statement to describe the function of the query.

3. Create data visualizations:
   - Choose **2** of your most interesting aggregate queries.
   - Create a database view for each one from your `<source>_analysis.ipynb` notebook. The view must have a descriptive name and a prefix of `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 both charts to a single Data Studio report (aka dashboard).
   - Download your dashboard as a PDF file and save it as `dashboard-v1.pdf`.

CS 327E Milestone 7 Rubric
Due Date: 03/15/20

Upload an updated ERD that finalizes your table schema after Beam transformations have been applied. -**40** `./<source>_erd_beam.pdf` not found in repository **-20** ERD is missing one or more entities **-20** ERD is missing one or more primary keys **-10** ERD is missing one or more foreign keys **-10** ERD is missing or incorrect relationship between entities	40
Create a Jupyter notebook 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. **-30** `./<source>_analysis.ipynb` not found in repository **-5** each missing aggregate statement, up to **-30** If all statements use the same aggregate function, you will only receive half credit **-5** each missing JOIN clause of the 3, up to **-15** **-5** each missing GROUP BY clause of the 4, up to **-20** **-5** each missing HAVING clause of the 2, up to **-10** **-5** each missing ORDER BY clause of the 4, up to **-20** **-5** each missing or incorrect comment, up to **-30**	30
Create data visualizations and download them as `dashboard-v1.pdf`. These visualizations should represent the results from the two BQ views. The file should contain 2 charts made from Data Studio, with a relevant title for each one describing the dataset. **-30** `./dashboard-v1.pdf` not found in repository **-10** each missing chart, up to **-20** **-5** each missing title, up to **-10** -**5** each chart created from a BQ table instead of a BQ view, up to **-10**	30

`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**