Milestone 11 due Friday, 04/26.

Part 1:

1. Implement your cross-dataset queries using the transformed tables from Milestone 10:
   - Develop the **6 interesting queries** which you described in `CROSS-DATASETS.txt`.
   - At least 4 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`).

2. Create the 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 6 charts to your existing Data Studio dashboard.
   - Take a screenshot of your dashboard and save it as `dashboard-v3.png`.

Part 2:

In this milestone, you will plan your Airflow workflow which will be implemented as part of Milestone 12. The workflow will consist of the SQL and Beam transforms you wrote to normalize your datasets. You will need to choose between the transforms for `dataset1` or `dataset2` since we won't have the time to create a workflow for both datasets.

- Decide which dataset to create a workflow for (i.e. choose between `dataset1` and `dataset2`).
- List each task in the workflow. For example, `create_Student_table` and `transform_Student_table` are two tasks belonging to the college dataset workflow.
- Specify the dependencies between the tasks of the workflow. For example, the `transform_Student_table` task is dependent on `create_Student_table` task.
- Save your workflow descriptions (tasks and dependencies) in a file called `WORKFLOW.txt`

CS 327E Milestone 11 Rubric
Due Date: 04/26/19

Part 1 - Create file `cross-dataset-queries.sql` that run the queries discussed in the previous milestone. Comment each query with the function they perform. **-30** no `cross-dataset-queries.sql` in repository, or missing queries **-5** each missing or erroneous query, up to **-30** **-5** each missing or incorrect comment, up to **-30** **-5** each query not on a transformed table, if there are less than 4 Create file `dashboard-v3.png` that visually displays the data returned by your queries. **-30** no `dashboard-v3.png` present **-5** each missing query, up to **-30** **-5** each chart without a title, up to **-30**	60
Part 2 - Create file `WORKFLOW.txt` detailing your workflow for transforming a specific dataset. Include the dataset, the process details, and any dependencies. **-40** no `WORKFLOW.txt` in repository **-5** each missing task in workflow, up to **-40** **-5** each missing dependency, up to **-40**	40

`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**