CS 327E Milestone 11 due Sunday, 05/03.

Create an Airflow DAG that automates the data processing pipeline you built for your secondary dataset.

Required functionality:

- DAG creates a new BQ dataset named `<source>_workflow_staging` to store the staging tables.
- DAG creates a new BQ dataset named `<source>_workflow_modeled` to store the modeled tables.
- DAG loads the CSV files for your dataset into `<source>_workflow_staging` and runs through the series of SQL and Beam transformations, writing the modeled tables to `<source>_workflow_modeled.`
- DAG executes dependent tasks in proper sequence.
- DAG executes independent tasks in parallel.
- DAG is implemented in a standard Python file named `<source>_workflow.py`

Not in-scope:

- DAG copies the CSV files into GCS.
- DAG creates the database views for reporting.

Testing and verification:

- DAG must produce the same end-results as the Dataflow jobs and/or SQL scripts from Milestone 10.

CS 327E Milestone 11 Rubric
Due Date: 05/03/20

Create file `<source>_workflow.py` that implements the transforms you built for your secondary dataset. The DAG should run operations in a decently efficient manner (operations that don't depend on one another should run concurrently, etc.) **-100** `<source>_workflow.py` does not exist in the repository **-50** `<source>_workflow_staging` dataset does not exist in BQ **-50** `<source>_workflow_modeled` dataset does not exist in BQ **-20** `<source>_workflow_staging != <source>_staging` **-20** `<source>_workflow_modeled != <source>_modeled` **-15** for each missing Dataflow job dependency **-15** for each missing SQL dependency **-10** each task that runs in parallel with its dependency **-10** each task that runs after another task it does not depend on	100

`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**