

# Week 9: Milestone 1

## Final Project

- **Milestone 1** (Lab 7) - Upload data for Austin and two other cities into BigQuery, and prepare an ERD for one of the cities
- **Guidelines for Milestone 1:**
  - Data hosted as buckets on Google Cloud Storage
    - `gs://utcs-spr2018-datasets/`
  - Submissions require both a Git SHA and a project ID
    - Project ID: your GCP Project ID, found on your Google Cloud dashboard under “Project Info”
- **BigQuery** - Advantages
  - Divides query into “sub-queries” to run the query in parallel (at the same time) by passing sub-queries down a tree-like structure
  - No concept of “keys” - queries are assumed to be correct
  - Stores data by columns
  - Supports standard SQL
- **BigQuery** - Disadvantages
  - Little support for data updating and deleting
  - Limit to a number of concurrent queries
  - No B-Tree Indexing - BigQuery is designed for large data retrieval