## Actions

1. Get project up on GitHub - DONE
2. Add CA CSV in git - DONE
3. Creation of notebooks
   1. ET (ExtractTransform)
   2. L (Load)
4. Git Set Up
   1. Resources Folder - DONE
   2. Rdme - DONE
   3. gitignore - DONE
   4. Report (as a markdown) - DONE
5. Change numbers to INT

Use COVID data: <https://covidtracking.com/data/api>

Use both API and CSV

Linda: California - CSV

John: New York – API

Missing casesViral

DeJuan: Set up, manage GitHub

Emory: SQL, Slack Channel ETL project. Group 5; create ERD, Entity Relationship Document

Siddarth: Create Report outline/framework

## Data to keep

1. Columns to copy
2. A: Date – keep format. **Primary Key**
3. B: State – keep
4. E: deathIncrease – keep (deaths)
5. I: hospitalizedCurrently – keep (daily\_Hospitalization)
6. L: inIcUCurently – keep (Icu\_Hospitalized)
7. U: postiveCasesViral – keep
8. V: positiveIncrease – keep (Positive\_case\_increase
9. AF: totalTestResults – keep
10. AG: totalTestResultsIncrease – keep (Test\_Result\_Total)

## Type of Transformation

1. Copy of columnar data
2. Rename keys
3. Joining data sets
4. Cleaning: Nulls; treat as zeros “0”, using Python