ALL data should have been collected.
Your team should perform a preliminary analysis of the data
and attempt to answer one question relevant to your project proposal which you will submit as a
pull request.

Main Progress in Deliverable 1 stage:

- ☑ Specific APIs selected and tested into useful and not useful for data gathering.
- ☑ Data of restaurants in Boston area collected and filtered.
- ☑ Dataset established and integrated from several sources.
- ☑ Data utilization plan for next stage made for project goals

Plans for next stage:

- ☐ Clean the dataset for deeper level
- ☐ Start to make use of the dataset to answer questions in project summary

Progress in details

1. We use Google Place API and Yelp API to obtain restaurants in Boston Area, and also some data features include:

- ratings (from 0 to 5)
- price_level (from 0 to 5)
- reviews
- business status
  - open
  - close
  - temporarliy closed
- address
- opening hours(day & week)

Note: we also tried below 2 popular delivery APIs, but neither of them give us access:

- Grubhub: only for signed-in restaurants, cannot sign up for an account.
- Uber Eats: "Your application currently does not have access to any scopes. Please contact your Uber Business Development representative or Uber point of contact to request access." + Please do not share this document or API endpoint details with anyone who is not authorized to have access.

2. We utilize safegraph to get the location name. After step 1, we have get each resteraunt's name and we can join the csv files with safegraph dataset on the same restaurant name to get full detailed-information about this resteraunt.

- visit counts (pv)
- visitor counts (uv)
- visits by day
- popularity by day/hour

- time
  - which means we can filter data by time

3. From above data we get, we can solve the problem

- How many businesses are open? How many are temporarily closed? (based on yelp and google places data)
- What hours are they operating under?
- How much activity is happening at different types of businesses by restaurant type (e.g. Italian vs. Thai vs. pizza, etc.), by neighborhood or zip code? (use safe graph data)? Do hours of operation seem to impact overall activity/ success?

But now we are still working on these big datasets, so now we cannot get the exact number of these problems.