**Brainstorming**

Use this document to start defining your research.

* Brainstorm initial ideas
* List of topics you need to look for (papers)
* Start looking at Data
  + Have a list of available data
  + Have a list of data that would be good to gain access to (we want to do this sooner than later

Step1: Get Twitter API Access:

My project aims to explore and analyse tweets about United Nations' Sustainable Development goals.

It is an open-ended project to see what people think and which of the 17 goals do they support the most.

It can help us access the importance of each goal across the world.

The Twitter API and data will help me fetch the right data for analysis. It will help me look for the correct keywords associated with the United nation's Sustainable development goals, in turn helping me to receive specific filtered data.

My project will mostly use counts, proportions and filtering to analyse Twitter data.

I will also try to use Natural Language Processing techniques to deep dive well into the data which can help me to receive better information.

Outcomes will be shared as simple graphs or counts in PDF or Powerpoint presentations. If I'm able to produce extraordinary results which can help people in the world, I might decide to try to build a research paper about my analysis and approach.

Can View:

In the meantime, you can get a head start by learning about the Twitter API by browsing our [docs](https://developer.twitter.com/en/docs), [tutorials](https://developer.twitter.com/en/docs/tutorials), and [community forums](https://developer.twitter.com/en/community).

About Twitter data <https://blog.hubspot.com/website/how-to-use-twitter-api#:~:text=What%20is%20the%20Twitter%20API,stands%20for%20Application%20Programming%20Interface>.

Code to pull Data: https://www.askpython.com/python/examples/extracting-tweets-using-twitter-api

Step 2: Look at Data

* SDGs in general - Rajani
  + Used Keywords - #SDGs,#SDG, #SustainableDevelopmentGoals
  + Pros
  + Cons
* SMB - David
  + Used Keywords- #SupportSmallBusinesses,#SupportSmallBusiness,#localbiz
  + Pros
  + Cons
  + Info : Useful link <https://twitter.com/search?q=Support%20Small%20Business%20(businesses%20OR%20business)%20(%23SupportSmallBusiness)&src=typed_query>

**Tweets we need :**



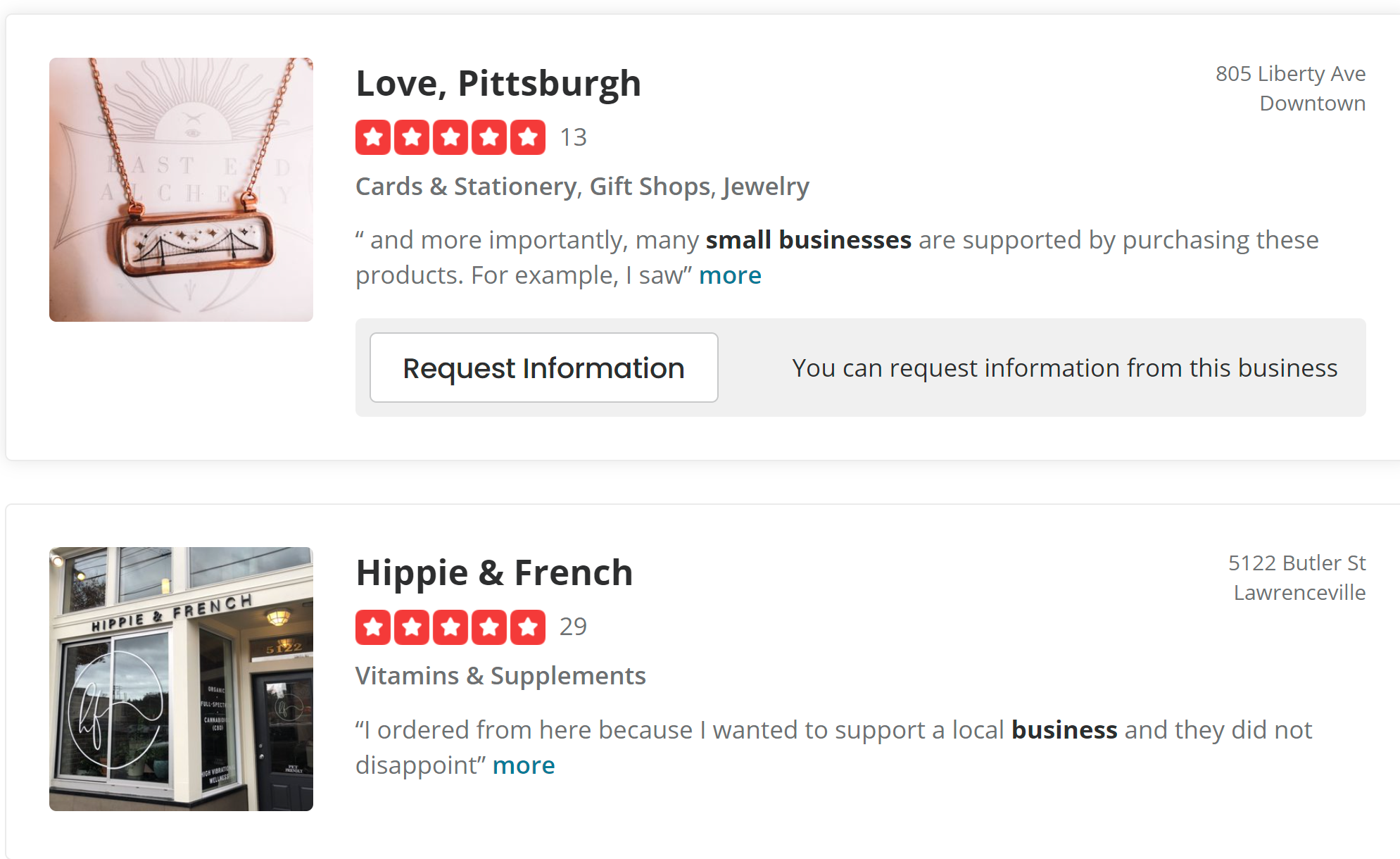
**Tweets we need to filter (Self promote)**





**Can use Yelp reviews to extract small biz info per city:**

<https://www.yelp.com/search?find_desc=Small+Businesses&find_loc=Pittsburgh%2C+PA>



Pulled data (4300+ rows): <https://drive.google.com/file/d/1-FfyL3NIelrfRQ-SmZx5P2cjv9FrgwpB/view?usp=sharing>

Based on content, it has many retweets and more than half of data is duplicate. However, based on user and tweets, there are only 5 duplicate rows

Also, the data has to be cleaned for non-English sentences?

RT @khairulaming: Thread bisnes makanan untuk KL dan Selangor ðŸ‘‡ðŸ». Feel free to promote your food business kat setiap daerah yg dibahagikanâ€¦

Multiple non-English characters (Itâ€™s, e!ðŸŒˆðŸ, etc) need to be removed and the content needs to be cleaned.

Only 35 geotags (<1%) found in the data. Alternative thoughts about the next steps on where to move.