**Standard Operating Procedure**

1. Note down the various parameters on the basis of which the different types of markets can be analysed.

For ex :-( Expenses, Competition, Influence area, connectivity). *Refer to the excel file attached(profile\_markets)*.

1. Note down the various types of markets, shops which exists and is necessary for the sustainable environment.

*Refer to excel file attached(Category\_shops).*

1. Classifying the cities into tier 1, tier 2, and tier 3.
2. Look for sources for the data scraping online from profile markets and see what all methods are most appropriate. The sources available must allow the model to replicate itself for different cities.
3. Remember to analyse the quality of information present from the particular source, especially for the tier 3, tier 4 cities.(also certain sites do not give permission to its html content)
4. Google maps are an ideal source for collection of raw data. Certain APIs helps in regulating the data extracted.
5. Quality APIs available

-Google maps APIs available on Google cloud platform

-Maps APIs available on TOM TOM developers.

GCP is a an ideal platform for scraping having various features including query search, nearby search etc but it limits the records to 60 per search and also costs more. On the other hand Tom Tom developers APIs limits to 100 search per query and is economical as well.

8. Make a pipeline in order to scrape data using python script using APIs keys. *Refer to the pipeline attached*

9. Look for inferences from the data.

PIPELINE (Python script)

FLOW OF FUNCTIONS

**Tool Functions**

1. Save function-To save a dataframe as a csv file.
2. Makelist function-To enter a group of things in a list manually.
3. Readfile function- To load a csv or xlsx file from system as a dataframe