

Data

Following data sources will be used to examine the business problem:

1. New York City Dataset.

Description: This file was created as a guide to New York City's neighborhoods that can be found on a website, "New York: A City of Neighborhoods." Link: https://geo.nyu.edu/catalog/nyu_2451_34572

Best estimates of label centroids were established at a 1:1,000 scale, but are ideally viewed at a 1:50,000 scale. This dataset will provide the addresses of neighborhood of NYC in json format. An extract of the json is as follows:

```
{'type': 'Feature',
'id': 'nyu_2451_34572.306',
'geometry': {'type': 'Point',
'coordinates': [-74.08173992211962, 40.61731079252983]
},
'geometry_name': 'geom',
'properties': {'name': 'Fox Hills',
'stacked': 2,
'annoline1': 'Fox',
'annoline2': 'Hills',
'annoline3': None,
'annoangle': 0.0,
'borough': 'Staten Island',
'bbox': [-74.08173992211962,
40.61731079252983,
-74.08173992211962,
40.61731079252983]}}
```

2. Foursquare API.

Description: Foursquare API, a location data provider, will be used to make API calls to retrieve data about venues in different neighborhoods.

Link: <https://developer.foursquare.com/docs>

Venues retrieved from all the neighborhoods are categorized broadly into "Arts & Entertainment", "College & University", "Event", "Food", "Nightlife Spot", "Outdoors & Recreation", etc. An extract of an API call is as follows:

```
'categories': [{'id': '4bf58dd8d48988d110941735',
'name': 'Italian Restaurant',
'pluralName': 'Italian Restaurants',
'shortName': 'Italian',
'icon': {'prefix': 'https://ss3.4sqi.net/img/categories_v
2/food/italian_',
'suffix': '.png'}},
'primary': True}],
'verified': False,
```

```
'stats': {'tipCount': 17},  
'url': 'http://eccorestaurantny.com',  
'price': {'tier': 4, 'message': 'Very Expensive', 'current'  
cy'
```