Data

The data of this project comes from multiple sources.

- ## Neighborhood

The data for neighborhoods in Paris was extracted by web scraping using BeautifulSoup library for Python. The neighborhood data is scrapped form the Wikipedia page.

```
data = requests.get("https://en.wikipedia.org/wiki/Category:Suburbs_of_Paris").text

soup = BeautifulSoup(data,'html.parser')

neighborhoodList = []

for row in soup.find_all("div", class_="mw-category")[0].findAll("li"):
    neighborhoodList.append(row.text)

paris_df = pd.DataFrame({"Neighborhood": neighborhoodList})

paris_df.head()
```

- ## Geocoding

```
def get_latlng(neighborhood):
    lat_lng_coords = None
    while(lat_lng_coords is None):
        g = geocoder.arcgis('{}, Paris, France'.format(neighborhood))
        lat_lng_coords = g.latlng
    return lat_lng_coords

coords = [ get_latlng(neighborhood) for neighborhood in paris_df["Neighborhood"].tolist
```

- ## Venue Data

Venue data is found by passing the required parameters in the Foursquare API and creating a data frame to contain all the details.

```
for lat, long, neighborhood in zip(paris_df['Latitude'], paris_df['Longitude'], paris_c

  url = 'https://api.foursquare.com/v2/venues/explore'

  params = dict(
  client_id=CLIENT_ID,
  client_secret=CLIENT_SECRET,
  v='20180323',
  ll='40.7243,-74.0018',
  query='coffee',
  limit=1
  )
resp = requests.get(url=url, params=params)
data = json.loads(resp.text
results = requests.get(url).json()["response"]['groups'][0]['items']


    for venue in results:
        venues.append((
```