

1 Data

1.1 Data sources

The data for this project will be obtained from Foursquare as well as <https://data.cityofnewyork.us> where some open data about New York is available. The data from Foursquare will be used to identify whether a certain park has good coffee-shops nearby based on their rating on Foursquare. Unrated coffee-shops will be considered as poor quality coffee-shops and thus won't be considered as competition for the mobile barista. Also, coffee-shops not within walking distance of 500 metres from a specific park will also be ignored.

The data about park locations will come from the open data portal mentioned above. The coordinates of parks in NYC are given a geojson file format, which fits well for the analysis we intend to perform. Large parks are divided into smaller segments in that dataset, so they will be joined to form a "single" park.