**Data:**

Data is provided by the NYC Department of Information Technology and Telecommunications. The dataset consists of records for every public WIFI hotspot (ones provided by or in partnership with the city) in New York City. It contains over 2500 records overall.

Below are the field we are Considering:

OBJECTID

Borough

Type

Provider

Name

Location

Latitude

Longitude

X

Y

Location\_T

Remarks

City

SSID

SourceID

Activated

BoroCode

BoroName

NTACode

NTAName

CounDist

Postcode

BoroCD

CT2010

BCTCB2010

BIN

BBL

DOITT\_ID

Location (Lat, Long)

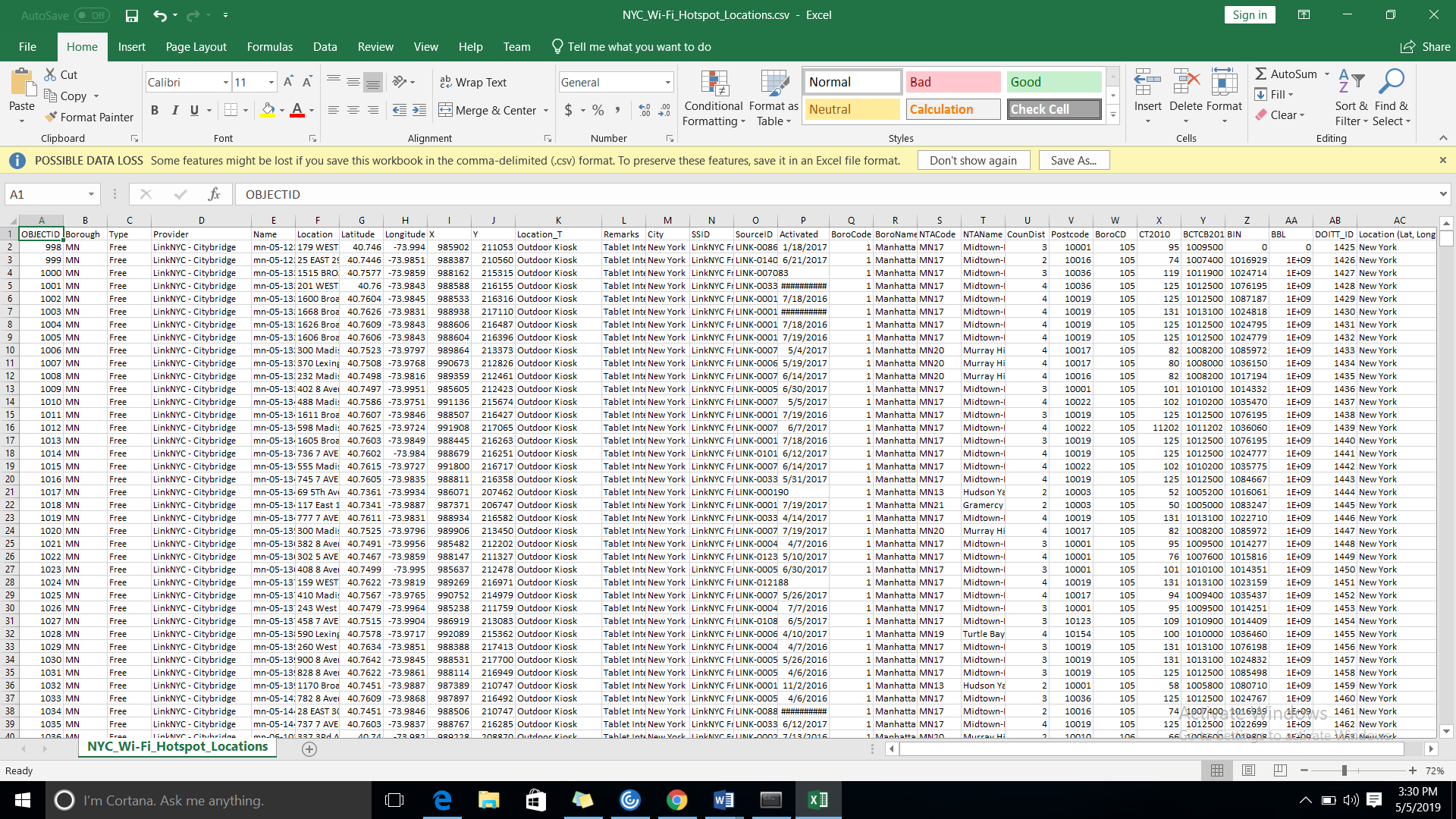
All the above field might not ne required for the Analysis. EDA will be performed on the Entire Dataset and will come up with the required variable which are important and relevant. The Dataset available will be used with the Foursquare location Data

We are trying to find an answer for below questions:

Does free public WiFi tend to cluster around certain (more affluent) areas?

Who are the free WiFi providers, and where do they do it?

Below is the snapshot of the Source Data which will be used for Analysis :



The k-means clustering algorithm is used to form clusters of WiFi usage based on latitude and longitude data associated with specific providers.