

Kripa Shakya

Mac 286 Data Structures

Report

Data Visualization Using Python Programming Language

NYC WiFi Hotspot Locations

In simple terms, “Data” refers to Raw Materials and Visualization refers to “Visualization”. Data is the statistic collection of information and facts made using reasoning and calculations. “Visualization” refers to visual representation. Data visualization is the presentation of data in a pictorial or graphical format. It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns. With interactive visualization, the concept a step further by using technology to drill down into charts and graphs for more detail in it. Data Visualization interactively changing what data sees and how it’s processed. Visual elements like charts, graphs, and maps, data visualization tools provide an accessible way to see and understand trends, outliers, and patterns in data which is more appealing and attractive.

NYC wifi hotspot has been a huge hit. The City of New York has teamed up with private partners to provide free wifi service at the public payphone across the five boroughs at no cost to taxpayers. It started its run from 2015. It replaced the NYC payphone spots and changed it to NYC WiFi Hotspot to follow the internet trend of the 21st Century. Wifi Hotspots are popping up on the sidewalks all over the town. It specifies with accessing the internet for free. It makes it possible to make cell phone calls from all underground NYC subway stations. It has provided many parks across the boroughs offer some semblance of free WiFi. The speed of the internet is 1 Gigabyte per second. WiFi signals has range upto 150 feet. I have particularly found my data from NYC open data. NYC has made their data public which is generated by various New York agencies and public organization for transparency, accountability. It has made it easy and accessible as well. I have used the list of providers, its location, longitude and latitude for the data.

For this research project, I have chose to code in Python programming language. Python is one of the widely and high level programming language which is used for general purposes

and interpreted in a dynamic form. It's design philosophy emphasizes code readability and the syntax allows programmers to express concepts in fewer lines of code than would be possible in languages such as C++, Java, etc. The language enables clear programs on both large and small case. The code written are readable and understandable. Python supports several programming paradigm and object oriented and structured completely. The language is compatible with major platforms and systems. It has a robust and large standard library. For the library, I have imported Pandas. It is an open source BSD licensed library providing high performance. It is one of the major tools that is widely being used specifically generated for data structures and data analysis.

```
hs = pd.read_csv('NYC_wi-fi_Hotspot_Locations.csv')
```

```
hs
```

	OBJECTID	Borough	Type	Provider	Name	Location	Latitude	Longitude	X	Y	...	NTAName	CounDist	Post
0	8641	QU	Free	Spot On Networks	QUEENS BRIDGE - JACOB A. RIIS Settlement House	10-25 41 AVENUE	40.755727	-73.944583	9.996032e+05	214613.274563	...	Queensbridge-Ravenswood-Long Island City	26	1
1	8642	QU	Free	Spot On Networks	QUEENS BRIDGE - JACOB A. RIIS Settlement House	10-43 41 AVENUE	40.755333	-73.944131	9.997285e+05	214469.807003	...	Queensbridge-Ravenswood-Long Island City	26	1
2	8643	QU	Free	Spot On Networks	QUEENS BRIDGE - JACOB A. RIIS Settlement House	10-05 41 AVENUE	40.755751	-73.945166	9.994417e+05	214621.916935	...	Queensbridge-Ravenswood-Long Island City	26	1

```
hs.head()
```

```
providers = list(NYC_Wifi_Hotspot_Locations['Provider'])
```

```
provider_count = {}
```

```
for provider in providers:
...     if provider not in provider_count:
...         provider_count[provider] = 0
...     else:
...         provider_count[provider] += 1
```

```
provider_count
```

```
{'Spot On Networks': 15,
 'ALTICEUSA': 236,
 'Harlem': 127,
 'SPECTRUM': 342,
 'NYCHA': 27,
 'AT&T': 26,
 'LinkNYC - Citybridge': 1730,
 'Downtown Brooklyn': 99,
 'NYPL': 89,
 'BPL': 58,
 'QPL': 64,
 'Manhattan Down Alliance': 35,
 'Transit Wireless': 275,
 'City Tech': 10,
 'Fiberless': 25,
 'Chelsea': 29,
 'Partner': 1}
```

```
NYC_Wi-fi_Hotspot_Locations = NYC_Wi-fi_Hotspot_Locations.ix[:7]
NYC_Wi-fi_Hotspot_Locations.head()
```

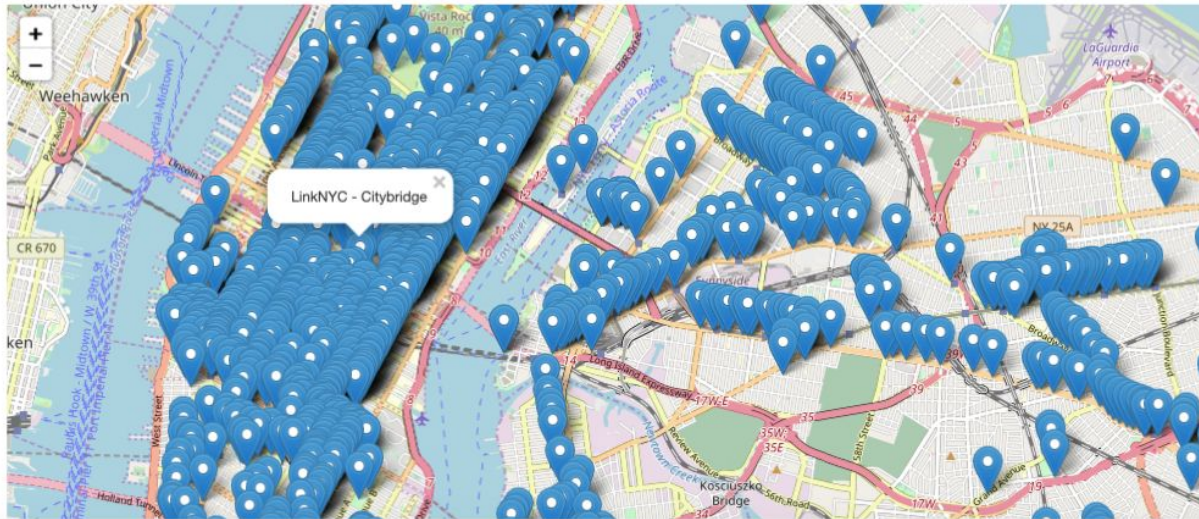
The screenshot of the pictures shows how to retrieve the csv. File and work it into the the program and also imports pandas beforehand. Then on the second pictures the code to find the name of the providers and the of the hotspots it has provided. The largest one can be seen as LinkNYC- Citybridge with 1730 hotspot locations following with transit wireless and so on.

For the data visualization and to mark the locations exactly on the map, I have used Folium library that is supported by python. Folium builds on the data wrangling strengths of the Python ecosystem and the mapping strengths of the Leaflet.js library. Folium makes it easy to visualize data that's been manipulated in Python on an interactive Leaflet map. It enables both the binding of data to a map for choropleth visualizations as well as passing Vincent/Vega visualizations as markers on the map.

```
import folium
```

```
data = hs.head(3205)
```

```
map = folium.Map(location = (40.7128, -74.0060), zoom_start = 10)
for each in data.iterrows():
    marker = folium.Marker(location = [each[1]['Latitude'], each[1]['Longitude']],
                           popup = each[1]['Provider'])
    marker.add_to(map)
display(map)
```



So the pictures above shows the code where folium library is imported. It actually imports the whole world map as a leaflet map. By googling the exact longitude and latitude of New York city and the poles of it, I used it to exactly pinpoint the map of NYC in a frame. Then for the zooming section, I have kept it upto 10. I have used the marker method to point out the exact locations inside the map and called out latitude, longitude, and the provider itself. So when you zoom in closely to the map and pinpoint one area, you would actually see the name of the provider given.

In conclusion, Global Learning is concerned with exploring the interconnections between people and places around the world. So Internet and its usage is a most requirement to stay connected. Hence, NYC has provided us with the facilities of free wifi for its visitors to the local people. Global learning asks us to observe the similarities and differences that exist around our world today and relate these to our own lives. Therefore, These hotspot should also help in more development and it should be provided to most of the parts in the country for its well rise

and urbanizations.

Work Citation:

NYC Open data-

<https://data.cityofnewyork.us/City-Government/NYC-Wi-Fi-Hotspot-Locations/yjub-udmw>

NYC hotspot articles-

<https://www.nycgo.com/articles/wifi-in-nyc>

<https://www.kdnuggets.com/2018/09/visualising-geospatial-data-python-folium.html>

Folium Library-

<https://pypi.org/project/folium/0.1.5/>