

Battle of the Neighborhoods: New York City

By Hassan Sarhan



WHERE DO I LIVE!?



Introduction

New York is the definition of a global city. People across the world envy those that live there. However New York has its downsides. But we won't let that deter us because we will find a neighborhood that has it all. A neighborhood that is close to trendy venues, has a low crime rate, and is relatively affordable to live in. We will find it using data science.

Livability Index

$$\textit{Livability Index} = \textit{norm}(\textit{Venues}) + \textit{norm}(-\textit{rent}) + \textit{norm}(-\textit{crimes})$$

The “Livability Index” is a term I coined to calculate how livable a neighborhood is based on its cost of living, its popularity, and its safety. We must normalize all the values to give equal weight to each of them and the rent and the crime data must be negative because they have a negative effect on the livability index.

```
In [176]: NYC.sort_values(by = 'Livability Index', axis = 0, ascending = False)
```

Out[176]:

	Neighborhood	Rent	Num_of_Crimes	Borough	Latitude	Longitude	Num of Venues	Livability Index
6	Woodside	1950	1	Queens	40.746349	-73.901842	80	0.067118
31	Brooklyn Heights	3456	1	Brooklyn	40.695864	-73.993782	100	0.040235
11	Jackson Heights	2222	2	Queens	40.751981	-73.882821	80	0.033391
25	Clinton Hill	2671	3	Brooklyn	40.693229	-73.967843	96	0.027437
40	Civic Center	4241	2	Manhattan	40.715229	-74.005415	100	-0.018090
			2	Brooklyn	40.599519	-73.998752	47	-0.020923
			1	Queens	40.728974	-73.857827	44	-0.022139
			3	Manhattan	40.707107	-74.010665	100	-0.027460
			2	Queens	40.784903	-73.843045	41	-0.030876
			1	Brooklyn	40.576825	-73.965094	42	-0.031658
			3	Manhattan	40.867684	-73.921210	58	-0.037297
32	Park Slope	3513	2	Brooklyn	40.672321	-73.977050	74	-0.042107

Let's see the winner

Woodside, Queens

With a low cost of living, high number of venues, and a low crime rate Woodside might be the best neighborhood in NYC



Out[180]:

	Neighborhood	Rent	Num_of_Crimes	Borough	Latitude	Longitude	Num of Venues	Livability Index	Rank
44	Chelsea	4359	9	Staten Island	40.594726	-74.189560	104	-0.159482	3
43	Chelsea	4359	9	Manhattan	40.744035	-74.003116	104	-0.159482	3
39	Financial District	4005	3	Manhattan	40.707107	-74.010665	100	-0.027460	4
45	Greenwich Village	4415	3	Manhattan	40.726933	-73.999914	100	-0.047118	6
29	Prospect Heights	3373	3	Brooklyn	40.676822	-73.964859	79	-0.044749	7
31	Brooklyn Heights	3456	1	Brooklyn	40.695864	-73.993782	100	0.040235	8
32	Park Slope	3513	2	Brooklyn	40.672321	-73.977050	74	-0.042107	18
48	Upper West Side	4654	20	Manhattan	40.787658	-73.977059	100	-0.410235	20
47	West Village	4524	5	Manhattan	40.734434	-74.006100	100	-0.047118	6
41	Morningside Heights	4314	2	Manhattan	40.807070	-73.956940	100	-0.047118	6

Adding Formal Rankings

These rankings were obtained from niche.com

Avg. Rent: \$4359

of Crimes: 9

of Venues: 104

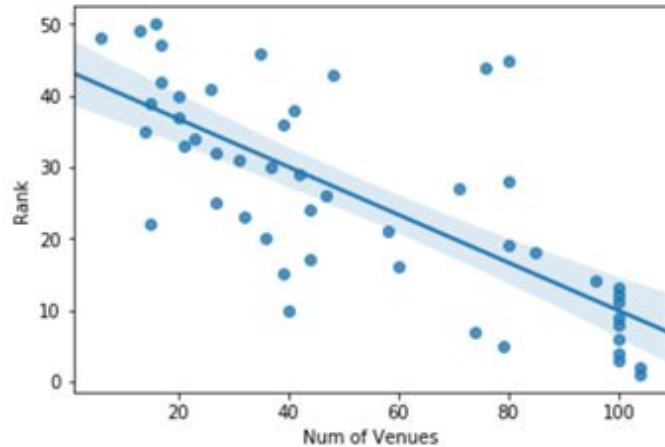
Chelsea, Manhattan



Num of Venues vs. Rank and Rent vs. Rank

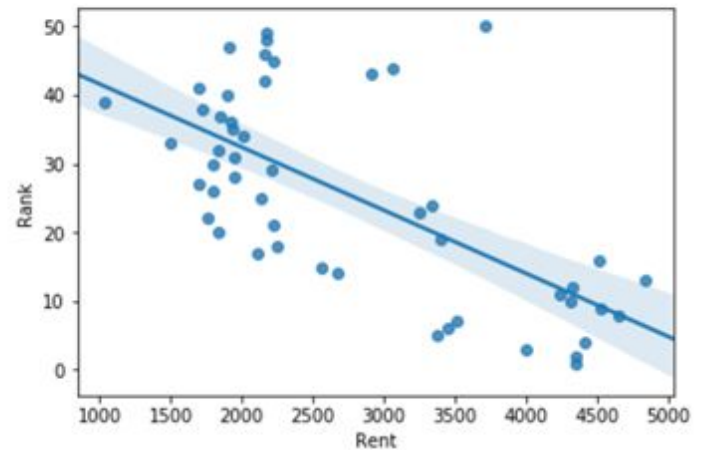
```
[170]: sns.regplot(NYC_Ranked2['Num of Venues'], NYC_Ranked2['Rank'])
```

```
[170]: <matplotlib.axes._subplots.AxesSubplot at 0x1bf8a22ade>
```



```
[168]: sns.regplot(NYC_Ranked2['Rent'], NYC_Ranked2['Rank'])
```

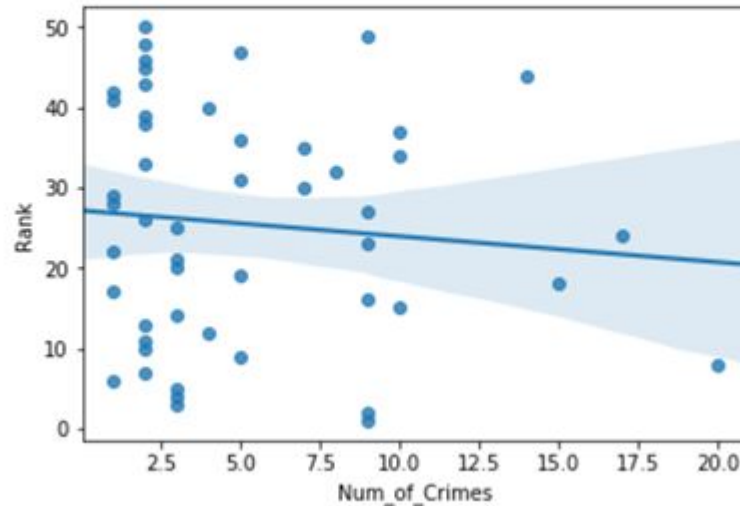
```
[168]: <matplotlib.axes._subplots.AxesSubplot at 0x1bf90a621>
```



Num of Crimes vs. Rank

```
In [169]: ▶ sns.regplot(NYC_Ranked2['Num_of_Crimes'], NYC_Ranked2['Rank'])
```

```
Out[169]: <matplotlib.axes._subplots.AxesSubplot at 0x1bf90abcef0>
```



Observations

We can see that the number of crimes has little effect on the rank of the neighborhood. This leads me to believe that the crime rate across NYC is low and is not a major factor in deciding where to live

Predictive Modeling and Evaluation

A linear model can be defined that predicts the ranking of a neighborhood based on the number of venues in a neighborhood, number of crimes in a neighborhood and the average rent of a neighborhood. The linear model used is multiple linear regression. It is not very accurate however with an r-squared score of ~ 0.72 , this is most likely due to the low amount of training data available

An aerial photograph of the New York City skyline at sunset. The Empire State Building is the central focus, with its spire reaching towards a sky filled with soft, orange and purple clouds. Other skyscrapers are visible on either side, and the city's dense grid of buildings extends to the horizon. The water of the harbor is visible in the distance.

**Which neighborhood
would you live in?**