Documentation Report: Analysis of NYC Neighborhood Financial & Social Data

Sahar Ahmadi

Introduction

For this project, I chose the Neighborhood Financial Health Digital Mapping and Data Tool dataset from open data, which contains data on NYC neighborhoods, including poverty rates, various financial health indicators, demographics, median incomes, and social indicators. The dataset includes 52 columns and 386 rows. This is an object-oriented Python program that cleans and prepares my dataset for further analysis. At the end, this program saved my cleaned dataset as a new CSV file.

This topic matters because financial health at a neighborhood level can help inform policy decisions, resource allocation, community support programs, policymakers, and city planners allocate resources effectively and design programs that target underserved areas.
The main research questions for this project are:

- How do IndexScores vary across NYC boroughs and neighborhoods?
- Which neighborhoods in NYC have the highest and lowest financial health scores?

Methods

For the data cleaning and processing part this data set I used Python. The main Python libraries that this program used are Pandas for data cleaning and manipulation, re for extracting patterns such as PUMA codes. OS for file handling and validation, in addition, this program also used web scraping by using BeautifulSoup library for extract public headlines related to NYC poverty and financial conditions. and finally Matplotlib for data visualization (bar charts, histogram). This project also has text mining techniques by tokenizing neighborhood names,

removing non-alphabetic characters, and performing frequency analysis to identify common keywords correct data type to ensure accurate analysis. The project used Object-Oriented Programming (OOP) through a custom data-cleaning class by creating a Neighborhood class to encapsulate neighborhood attributes and methods for categorizing scores and income, file handling**,** data processing with Pandas to prepare the dataset for analysis and insights. For data analysis part this program grouped data by boroughs, calculated averages, and extracted key subsets (e.g., neighborhoods with IndexScore > 3). Finally, for visualization plotted bar charts and histograms to illustrate patterns and distributions.

Analysis and Findings

Average IndexScore by Borough

Borough	Avg_IndexScore	Avg_Income	TotalOutcome
Manhattan	5.317228	47216.4	76.859381
Staten Island	4.438161	33380	7.665787
Queens	3.605581	29391.57143	2.570436
Brooklyn	3.275751	28471.44444	-22.717395
Bronx	1.956922	21453.9	-70.426484

According to our findings Manhattan has the the highest average Index Score (5.32) and highest total outcome (76.86), indicating it is the strongest-performing borough in terms of financial health. In another hand Bronx has the lowest average Index Score (1.96) and a significantly negative total outcome (-70.43), highlighting financial struggles in this borough. Finally, Queens, Brooklyn, and Staten Island fall in the middle range, with moderate Index Scores and outcomes, showing some neighborhoods are doing well while others not. For insight part I will say there is a clear geographic disparity in neighborhood financial health, with

Manhattan leading and Bronx lagging, suggesting resource allocation and policy interventions could be borough-specific.

Neighborhoods with IndexScore > 3

Neighborhoods with Index Score > 3:

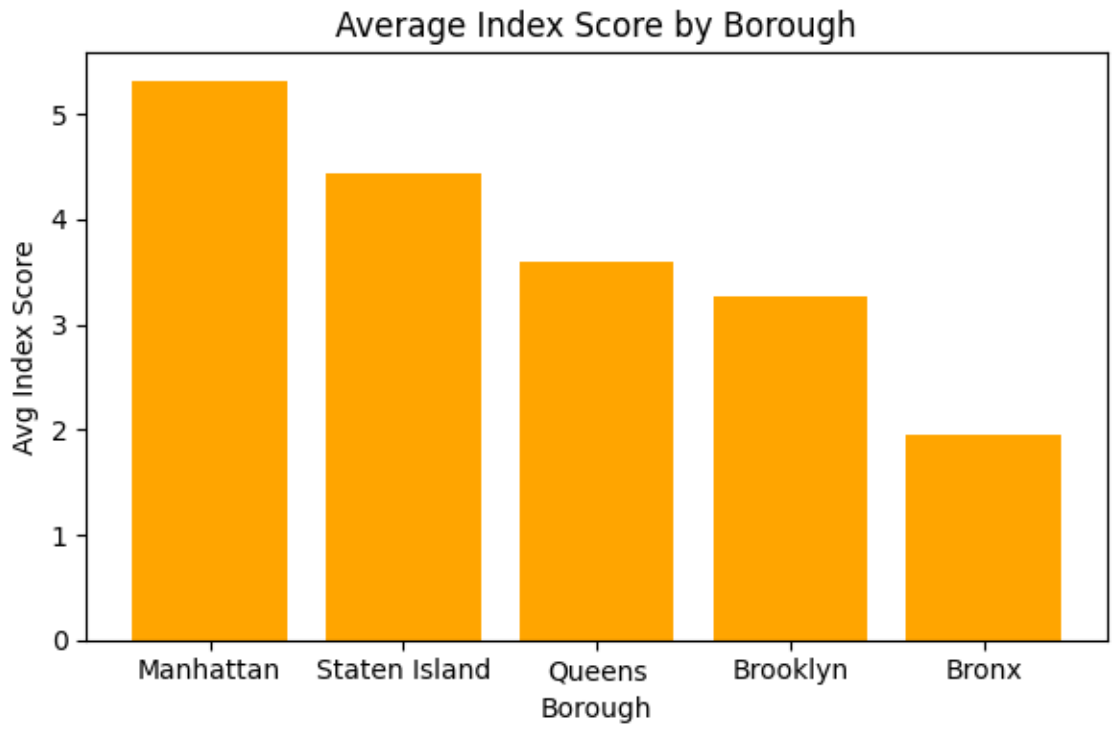
	Neighborhoods	IndexScore	Median_Income
0	Riverdale Fieldston & Kingsbridge	3.970856	30437
1	Wakefield Williamsbridge & Woodlawn	3.535724	26136
2	Coop City Pelham Bay & Schuylerville	3.993143	31573
14	Upper East Side	6.628084	71933
15	Upper West Side & West Side	5.895308	65844
16	Chelsea Clinton & Midtown Business District	9.240777	65905
17	Murray Hill Gramercy & Stuyvesant Town	8.787774	73067
19	Battery Park City Greenwich Village & Soho	10.000000	79181
20	New Springville & South Beach	4.587394	35413
21	Port Richmond Stapleton & Mariners Harbor	4.288928	31347
22	Greenpoint & Williamsburg	4.385768	35889
25	Brooklyn Heights & Fort Greene	6.444394	44313
26	Park Slope Carroll Gardens & Red Hook	7.200655	57393
27	Crown Heights North & Prospect Heights	3.955711	29895
30	Canarsie & Flatlands	4.458754	32210
31	East Flatbush Farragut & Rugby	3.529227	27248
34	Bay Ridge & Dyker Heights	3.305963	30858
36	Flatbush & Midwood	3.583589	25708
37	Sheepshead Bay Gerritsen Beach & Homecrest	4.059021	29307
40	Astoria & Long Island City	4.901631	31658
43	Bayside Douglaston & Little Neck	4.620964	37432
44	Queens Village Cambria Heights & Rosedale	4.255354	32961
45	Briarwood Fresh Meadows & Hillcrest	4.097157	30239
47	Forest Hills & Rego Park	5.390182	41029
48	Sunnyside & Woodside	3.644233	31245
49	Ridgewood Glendale & Middle Village	3.565296	31420
50	Richmond Hill & Woodhaven	3.097412	26357
51	Jamaica Hollis & St. Albans	3.490817	25946
52	Howard Beach & Ozone Park	3.418566	27016
53	Far Rockaway Breezy Point & Broad Channel	3.272490	25920

According to the above analysis, Neighborhoods with high Index Scores or greater than 3 are in Manhattan (e.g., Battery Park City, Chelsea Clinton & Midtown, Upper East Side), reflecting stronger economic conditions and higher median incomes. Some neighborhoods in Queens and Brooklyn also have index greater than 3, though their median incomes but generally lower than Manhattan. Bronx and some areas, although they have an index greater than 3 but still

show that even moderately performing neighborhoods may face income challenges. For the insight, I will say financial health at the neighborhood level correlates strongly with median income and possibly other socioeconomic factors.

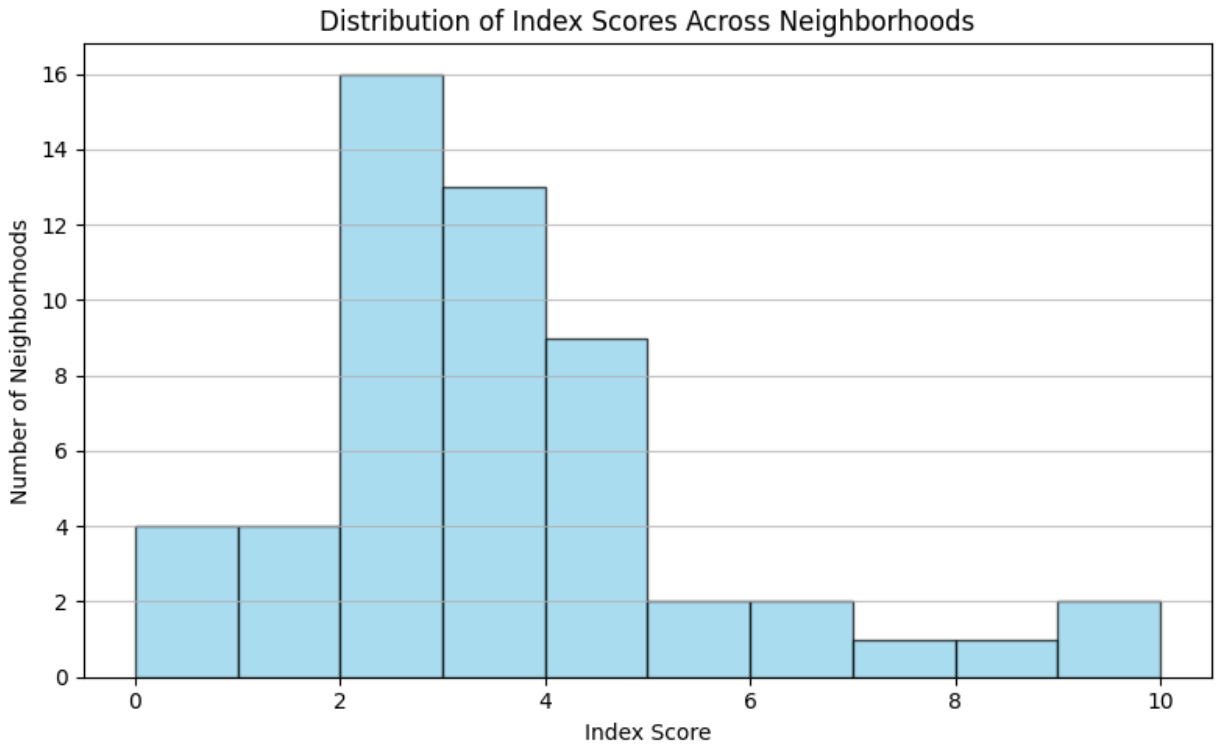
Visualizations

Bar Chart – Average IndexScore by Borough



Clearly shows Manhattan has the highest average Index Score compared to other boroughs.

Histogram – Distribution of IndexScores Across Neighborhoods



Histogram shows the distribution of Index Scores across all neighborhoods, highlighting that most neighborhoods fall between 2–5, it means that most NYC neighborhoods have moderate financial health, rather than extremely strong or extremely weak outcomes. For the insight I will say financial health is uneven but concentrated around the middle extreme cases are less common.

Business or Practical Implications

These findings are valuable because for example policymakers can target low-scoring neighborhoods which in this finding was Bronx to improve financial stability, education access, employment, and community services with programs. Other thing about these findings is for businesses that they can use this data to identify areas for investments, partnerships, or

community programs. Finally, community organizations can use these insights to prioritize outreach and resources efficiently.

Limitations

- Sample size: After data cleaning and removing missing and duplicate values reduced the number of usable records, which may limit generalizability.
- Other external factors: This data set does not include other external factors such as crime rates, housing cost in those areas which makes it harder to do further analysis.

Conclusion

This analysis revealed significant disparities in financial health across NYC neighborhoods and boroughs such as Manhattan, Queens, Brooklyn, Bronx. High-income neighborhoods tend to have stronger financial health scores that in here regarding our findings was Manhattan, while lower-income areas are at risk like Bronx. For the next steps I could include much larger dataset because after data cleaning all we will be left with small sample size. Maybe will add multi-year data, add additional data sets with external variables such as crime rates, housing cost in those areas, and developing predictive models to forecast financial health changes at the neighborhood level.

References

Consumer. "Neighborhood Financial Health Digital Mapping and Data Tool." Cityofnewyork.us, 2 May 2022, data.cityofnewyork.us/Business/Neighborhood-Financial-Health-Digital-Mapping-and-/r3dx-pew9/about_data. https://data.cityofnewyork.us/Business/Neighborhood-Financial-Health-Digital-Mapping-and-/r3dx-pew9/about_data

Visual Studio Code. https://code.visualstudio.com/