NYC Housing Pulse (2015–2025)

Tracking Affordability and Market Breadth Across Boroughs

Harsh Golani & Abhishek Agarwal

Group: InsightLoop

DS-GA 1007 — Programming for Data Science

Fall 2025

New York University

Abstract

This project analyzes a decade of NYC neighborhood sales (2015–2025) to measure how prices and *accessibility* evolved across the five boroughs. Using official NYC Department of Finance datasets—Annualized Neighborhood Sales Summaries (2015–2024) and Rolling Sales (2025 YTD)—we assemble a clean, reproducible panel and introduce two interpretable indicators: the Affordability Index, a proxy for entry-level affordability, and Market Breadth, the share of neighborhoods participating in year-over-year price growth. Deliverables include a unified dataset, a modular Python codebase, a concise analytical notebook with publication-quality figures, and a short presentation aligned with DS-GA 1007 evaluation criteria.

1 Introduction and Motivation

New York City's housing market transformed markedly over the past decade: a low-rate expansion in the mid-2010s, COVID-era disruptions, and a tightening cycle beginning in 2022. These shifts affected not only price levels and transaction volumes but also *who* benefited from appreciation and *which* neighborhoods remained within reach for entry-level buyers.

Typical public analyses emphasize citywide medians or short windows. This project adopts a neighborhood-level view and asks two practical questions: How has affordability evolved by borough, and how widespread has price growth been? We construct a longitudinal dataset and compute two transparent, defensible metrics—Affordability Index and Market Breadth—to capture both price levels and the *breadth* of market strength.

2 Dataset

We rely exclusively on official NYC Department of Finance sources:

- **Annualized Neighborhood Sales Summaries (2015–2024):** yearly Excel files per borough summarizing neighborhood statistics by building-class category (e.g., 1–3 family, condo, co-op, mixed-use). Core fields include *Neighborhood*, *Building Class Category*, *Number of Sales*, and *Average/Median Sale Price*.
- **Rolling Sales (2025 YTD):** transaction-level files for the most recent 12 months, one per borough, containing *BOROUGH*, *NEIGHBORHOOD*, *BUILDING CLASS CATEGORY*, *SALE PRICE*, and *SALE DATE*.

After cleaning and harmonization, the combined panel covers more than 400 neighborhoods across the five boroughs. The primary working file is `nyc_housing_pulse_2015_2025.csv`, described in Table 1.

3 Methodology

Implementation is in Python (Pandas, NumPy, Matplotlib) with a modular structure (`src/nyc_sales/`) to ensure clarity and reproducibility.

Ingestion and harmonization. We parse 55 Excel files (2015–2024 annualized summaries; 2025 rolling sales), auto-detect header rows (some sheets contain title blocks), standardize column names,

Table 1: Canonical schema for the merged neighborhood–year panel.

Column	Description
`borough`	Borough (Manhattan, Brooklyn, Queens, Bronx, Staten Island)
`neighborhood`	NYC Department of Finance-defined neighborhood label
`building_class_category`	High-level use/class grouping
`num_sales`	Annual (or YTD) transaction count
`avg_sale_price`, `median_sale_price`	Mean/median neighborhood sale price
`year`	Calendar year (2015–2025)

and coerce price fields to numeric types. Older `.xls` years use the legacy reader, while newer `.xlsx` files use `openpyxl`. Rolling sales are aggregated to the neighborhood–year level to align with annual summaries.

Cleaning and features. We remove invalid or zero-price transactions, trim repeated headers, and normalize neighborhood strings. Where available, min/median/max statistics are retained; otherwise, they are recomputed consistently. Year-over-year changes and identifiers (e.g., `borough_code`) are added.

Custom metrics. We define two interpretable measures:

$$\textsc{Affordability Index}_{b,t} = \mathrm{Q}_{0.25}\Big(\{\overline{P}_{n,b,t} : n \in \mathcal{N}_{b,t}\}\Big), \tag{1}$$

$$\textsc{Market Breadth}_t = \frac{1}{|\mathcal{N}_t|} \sum_{n \in \mathcal{N}_t} \mathbf{1}\Big(\overline{P}_{n,t} - \overline{P}_{n,t-1} > 0\Big), \tag{2}$$

where $\overline{P}_{n,b,t}$ is the neighborhood mean (or median) price for neighborhood n in borough b and year t, $\mathcal{N}_{b,t}$ is the set with sufficient observations, and $\mathcal{N}_t$ is the citywide set for year t.

Visualization. We produce publication-quality graphics: borough trajectories (2015–2025), Affordability Index by borough, Market Breadth citywide, and a focused 2025 snapshot highlighting top and bottom neighborhoods. All figures employ consistent scales, descriptive titles, and color-blind-safe palettes.

4 Research Questions

1. How have neighborhood prices evolved across boroughs between 2015 and 2025?
2. Which boroughs experienced the steepest declines in entry-level affordability?
3. Was the post-COVID rebound broad-based (high Market Breadth) or concentrated in select neighborhoods?
4. Where does 2025 YTD stand relative to pre-COVID and prior-cycle peaks?

5 Expected Outcomes and Deliverables

- **Unified dataset:** `nyc_housing_pulse_2015_2025.csv`, cleaned and ready for analysis.
- **Modular codebase:** clear ingestion, harmonization, metric, and visualization modules.
- **Analytical notebook:** a concise narrative containing 4–6 key figures and takeaways.
- **Presentation:** a 6–8 slide deck summarizing findings for a general audience.

6 Relevance and Broader Impact

By emphasizing affordability and the *distribution* of gains, this project contributes to understanding where housing pressures are intensifying, which neighborhoods remain comparatively accessible, and whether recoveries are inclusive or concentrated. The framework is transparent and readily extensible to future years or comparable datasets in other cities.

Tools and Implementation Details

- **Language/stack:** Python 3 — Pandas, NumPy, Matplotlib.
- **Environment:** Jupyter Notebooks for analysis; Git/GitHub for version control.
- **Data source:** NYC Department of Finance Annualized Neighborhood Sales Summaries (2015–2024) and Rolling Sales (2025 YTD).

References

- NYC Department of Finance. *Property Annualized Sales Update.* nyc.gov/finance/property-annualized-sales-update
- NYC Department of Finance. *Property Rolling Sales Data.* nyc.gov/finance/property-rolling-sales-data
- NYC Open Data Portal. *Property Sales (Historical and Current).*