**Why New York’s Housing Market Still Dominates**

A data-driven look at what drives price, affects affordability, and creates opportunity for investors

**Introduction**

New York City’s housing market is one of the world’s most watched and most complicated asset markets: extreme density, enormous wealth concentration, varied property types, and intense local regulation all collide here. Imagine a city where a single square foot can be worth as much as a small car — that’s New York. Today I’ll show you how space, location and short-term supply swings make that happen, and why those mechanics matter for anyone who rents, saves, or dreams of owning a place in the future. The goal of this project is simple and ambitious: to explain what makes NYC’s housing market strong and resilient, to identify which property and market attributes drive price, to segment the city into meaningful market types, and to produce clear, actionable recommendations grounded in your cleaned transaction-level data and the annual market aggregates you provided.

**What I aim to solve:** Which structural factors and market dynamics most explain NYC price levels and trends? In other words: when a property sells for a lot of money in NYC, what explains it — the physical features, the neighborhood, the market-wide supply picture, or other factors?

**Key Research Questions**

* Which property attributes (sqft, beds, baths, property type) most strongly predict sale price?
* How have median sales & list prices and inventory changed over time, and what story do those trends tell?
* How does value density vary spatially across NYC (by neighborhood / sub locality)?
* Can neighborhoods be clustered into distinct market types with different investment/policy implications?
* How do supply-side signals (inventory, months of supply, new listings) relate to price dynamics?
* Given the above, what are clear, data-driven recommendations for buyers, sellers/investors, and policymakers?

**What is real estate, why it matters, and why NYC is the right case study?**

Real estate is simply defined as land and permanent structures built on it plus the bundle of rights associated with ownership and use. It's both a consumption good (shelter, space) and a financial asset (stores wealth, collateral).

Why it matters:

* For households: housing is the primary store of wealth and the largest expense in budgets.
* For investors: property returns, yield, and capital appreciation are core portfolio considerations.
* For cities and policymakers: housing affects labor mobility, inequality, tax bases, and social stability.

**Why NYC?**

* High transaction volume + many property types across a compact area allows powerful comparative analysis
* Extreme price dispersion and sharp neighborhood-level differences make drivers visible and actionable.
* The market is globally integrated (investor flows matter), but also locally constrained (zoning, supply restrictions), yielding clear testable hypotheses about supply, demand, and price.

**Why not Kenya?**

We retain Kenya as a recurring contrast to illustrate scale and social differences — but we did not perform the same deep, transaction-level analysis for Kenya because:

* Kenya’s transaction data for housing is less centralized and more fragmented between formal and informal markets, making a fully comparable cleaned transactions dataset unavailable without additional collection.
* The structure of ownership and the role of informal transactions, self-built housing, and non-registered sales complicate direct comparison.

For scale context: Kenya GDP per capita ≈ $2,206.1 (2024, World Bank); poverty headcount ≈ 39.8% (2022, KNBS) — this underscores the housing market in Kenya, which is often primarily about access and basic shelter rather than a large store-of-wealth asset as we see in NYC. Even though the capital city Nairobi is an emerging real estate hub, especially with all the recent high-rise developments and road projects, it is far too limited in scope of data to use for storytelling.

**1) Distribution of House Sale Prices — the market’s shape and what it hides**

**Insights:**

* When you look at the price distribution you don’t just see numbers — you see a social story. The chart shows that most transactions sit well below the headline ‘average’ prices, and then there’s a long right tail of very expensive sales. In plain terms: a handful of ultra-expensive deals (penthouses, big building sales) push up the average and make the market look richer than it is for most people.
* This does not mean the market isn’t expensive, as seen by the large number of sales past $1 Million USD, but rather, for the average consumer, you are still likely to find an affordable place if you look across neighborhoods.

Why this matters: policies, conversations, and headlines that quote the mean price make housing seem more affordable for the “typical” person than it is — that’s misleading. For young renters and small families/couples, the median (the middle of the pack) is a far better reality check of what housing looks like on the ground.

What this means: When you hear “the city’s housing is worth $X,” ask “is that the average or the median?” — the median tells the real story for most people.

**2) Average Price vs Square Footage — how space really pays**

**Insights:**

* Bringing back the topic of neighborhoods, it seems most have a similar average sqft. However, the Bronx has the lowest average sqft alongside the lowest average price. This corresponds to our previous analysis as to how majority of sales in NYC are within this average price.
* Price rises with size, but not linearly. Small increases in square footage at the lower end (think studio → one-bed) often translate to big bumps in price, whereas moving from a big apartment to a much bigger apartment yields smaller percentage increases. In other words, the first few square feet give you the biggest bang for the buck.

Why this happens: the market values functional space — a bedroom or second bathroom can change who can live there (from single occupant to couple or family), and demand for those utility jumps is high. Meanwhile, ultra-large units are rarer but don’t attract proportionally more buyers per extra square foot.

If you’re renting, think carefully whether a marginally larger unit is worth the jump in rent. If you hope to buy later, small improvements that increase usable space (open a wall, better layout, add a storage solution) often increase value more than cosmetic upgrades.

**3) Price per Sqft in Each Neighborhood — how location rewrites value**

**Insights:**

* This is where the city’s personality becomes obvious: a square foot in one neighborhood can cost two or three times what it costs in another. Neighborhoods with top average price per sqft are not only expensive — they’re consistently expensive across many sales (tight medians), while others show wide ranges, meaning some lucky buyers pay far less than others in the same area.
* This explains why the data shows our median as well as average prices to be high across the entire city. Similar to our home country Kenya, locations such as Runda, Karen, Muthangari, Riverside etc. all have disproportionately high prices in comparison to locations such as South C, Umoja, Kawangware etc. even though these neighborhoods have greater population densities.

Why that matters: price per square footage bundles together scarcity, prestige, access to jobs/transit, and local amenities. It’s a compact way to see where the city’s market concentrates value. This explains why living a short subway stop farther out can change rent dramatically.

**4) Median Price by Number of Bedrooms — the ladder of household choices**

* This chart reads like a life map. Moving from 1 → 2 → 3+ bedrooms is not just a bigger place; it typically moves you into a new price bracket. The median prices jump at these thresholds because buyer needs change — more bedrooms mean different household types (couples, families) and therefore different willingness to pay. The tier with the fastest growth often shows where demand is strongest right now (could be young professionals sharing 2-beds, or families seeking 3-beds).
* Why it matters: Let’s say, for example, that you’re a student thinking of roommates vs solo living, this shows the financial trade-offs clearly: a roommate arrangement that moves you from 1 bed to split a 2-bed could cut per-person cost dramatically — and conversely, moving up to a private 3-bedroom will cost you a lot more.

**5) Inventory vs Median Sale Price — the short-term weather of the market**

* This dual-axis story is the “weather report” for housing. When inventory falls — when there are fewer houses actively for sale — prices tend to rise soon after because buyers compete for scarce options. Conversely, when inventory builds, prices often stabilize or fall. This chart often predicts which way the market will tilt in the next season.
* Why this matters: for buyers, this is the closest thing to a short-term signal you can use: months-of-supply or inventory drops mean landlords and sellers will have more bargaining power and rents/sales prices may spike. If inventory is up, you might have more leverage to negotiate.
* If you’re hunting for an apartment and you see inventory dropping in your target area, speed up your search and be prepared with documents — that’s when landlords fill units quickly. If inventory is high, you can slow down and negotiate.

**6) Median Sales Price vs Median List Price — who’s deciding the deal?**

* This chart tells you about leverage: when list prices are consistently above sale prices, sellers are asking more than the market will pay and buyers have negotiating power; when sale prices meet or exceed list prices, buyers are competing, and the market is hot. For anyone that rents or hopes to buy one day, this is the clearest gauge of how aggressively the real estate market operates.
* Why: list prices are expectations; sale prices are outcomes. Gaps reveal whether sellers overestimate demand or whether bidding wars are pushing prices through the roof.
* This calls people to watch this gap as you plan timing. If list > sale persistently, offer politely below asking and try to negotiate concessions; if sale ≥ list, assume competition and act quickly. In the long run, persistent sale ≥ list indicates a market where investing in real estate yields stronger returns — but it also suggests affordability pressures for residents.

**Conclusion**

We have set out to “explain what makes NYC’s housing market strong and resilient.” The data show resilience stems from a large, deep market where location and functional housing features keep demand strong; supply constraints and market fluidity create price surges. New York’s housing market looks huge on the headlines but is mostly driven by location and short-term supply swings — a few ultra-expensive sales skew averages while median buyers live a very different reality. When supply tightens, prices spike; when incomes don’t keep up with list prices, affordability drops — and that’s the story today’s data analysis has told.

**What can we conclude having looked at one of the most robust and influential real estate markets in the world?**

* For those hunting housing today: prioritize usable space and track the inventory within your local market — those two things will help you find better deals fast. If you rent, you’re on the front line — use inventory and neighborhood price-per-sqft to time moves and negotiate.
* For future buyers: Buying is about location + function, not facades. Focus on neighborhoods with stable price per sqft and consider functional features (houses with more bedrooms will have greater base resale values, but trade-offs include higher upfront prices) more than premiums such as large yard space, extremely optimal locations.
* For everyone thinking about policy or investing: Median matters more than mean — talk about typical affordability, not headline prices that skew data. NYC shows that even in an above average market, the median’s tell the most about the value of houses in the area. Policies that address supply where incomes lag behind prices will move a market towards more affordability.

**Data Methodology**

New York’s housing market is famous — and complicated. To tell a clear, trustworthy story from messy public data, I cleaned and prepared several datasets (listing-level sales, yearly market snapshots, and historical income & minimum-wage series). The cleaned data powers the visuals and insights on this site so you can understand what really drives price and affordability in the city. Browse the visual story, download the cleaned data, or read the full technical write-up (link).

**What I started with:**

* A detailed listing dataset of New York property sales (prices, bedrooms, bathrooms, square footage, broker, address).
* Yearly market summaries (median list/sale price, inventory, months-of-supply, homes sold).
* Historical income series (NY and U.S.) and minimum wage series (NY and U.S.) spanning decades.

**Core cleaning steps:**

1. **Standardized names and columns**. I made every column consistent (no weird spacing or punctuation) so analyses are repeatable.
2. **Fixed price units**. Some prices were written as “315” (meaning $315k). I detected those patterns and converted values to consistent dollars.
3. **Coerced text to numbers**. Square footage, bedroom counts and price fields were converted to numeric types, with bad values turned into blanks rather than broken results.
4. **Address de-duplication**. Duplicate listings were removed using address + key attributes so each row represents one unique record.
5. **Outlier handling for visuals**. For readable charts, I capped extreme values (top 1%) or used log scales — this preserves the story without letting a few headline-making sales dominate every plot.
6. **Exported final files**. Clean CSVs and interactive charts were generated and saved for visualization and presentation.

**What you’ll find on this site:**

* Interactive charts showing price distribution, price vs. size, and supply vs. price trends.
* Neighborhood snapshots that highlight where prices and incomes are diverging.
* A clear set of recommendations for buyers, investors, and policymakers based on the cleaned data.