NYC Crime & Rent Analysis (FRE6191)

Author: Sunny Chang

1. Executive Summary

This project explores the relationship between crime patterns and housing rental prices in New York City using data from 2015 to 2024. Using exploratory data analysis techniques in Python, we created both static and interactive visualizations to uncover trends across time, geography, and socioeconomic variables. Our analysis highlights important relationships such as rising rent in high-crime areas, precinct-level danger zones, and the potential mismatch between cost and safety. This report complements a working interactive dashboard built in Dash and a video walkthrough presentation.

2. Dataset Overview

I utilized and integrated the following three datasets:

- **StreetEasy Data Dashboard**: Provided historical rent data for NYC neighborhoods.
- **NYC Zipcodes (Kaggle)**: Used for spatial mapping and GeoJSON choropleths.
- **NYPD Complaint Data Historic**: Offered detailed historical crime complaint data by date, type, and geography.

After merging and processing the datasets, I created a unified dataframe `merged_df` with the following **key variables**:

- `date`, `median_rent`, `count` (crime count), `danger_ratio`, `precinct_area`, `Borough, Neighborhood`, `ZIP Codes`

Data Preparation

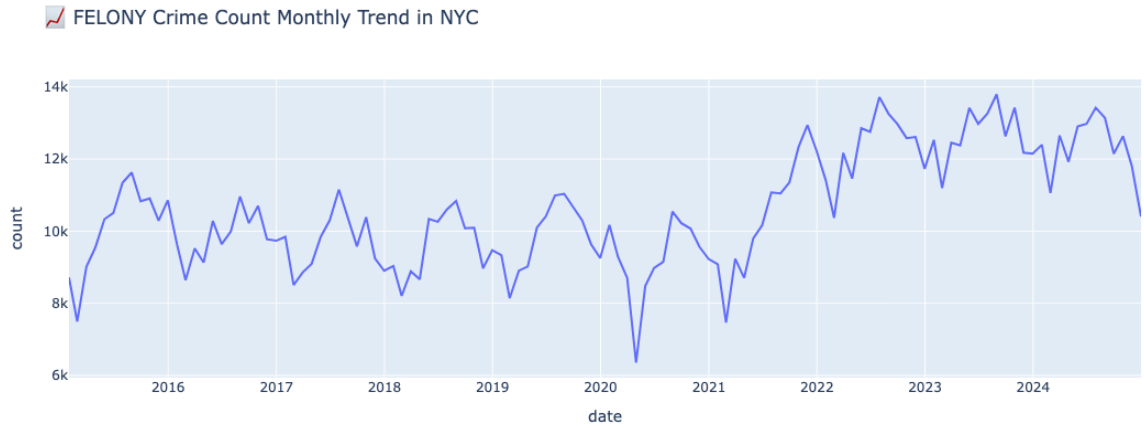
- Reshaped data from wide to long format where needed.
- Constructed a **mapping dictionary** to align neighborhood labels across datasets.
- Converted `date` columns to datetime format.
- Flattened lists of ZIP Codes into individual rows.
- Removed rows with null or zero values.
- Aggregated metrics by time (monthly) and space (precinct area, ZIP code).

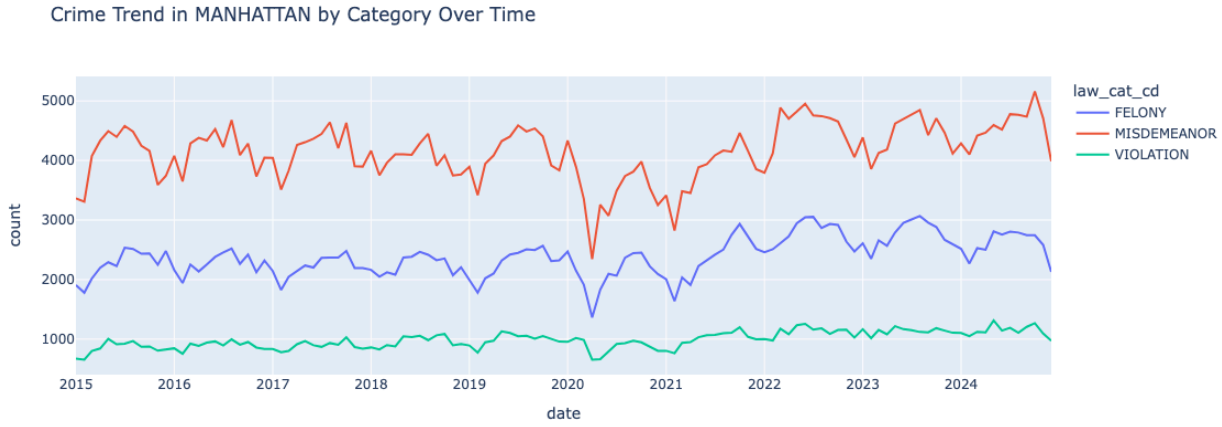
3. Analytical Questions

1. How have crime trends changed over time across NYC?
2. Is there a correlation between rent levels and crime frequency?
3. Which precincts have the highest danger ratios, and how do they evolve over time?
4. What boroughs offer the best balance between affordability and safety?
5. Is your rent truly worth it? A comparison of affordability, safety, and convenience in your neighborhood.

4. Visual Analysis & Findings

4.1 Monthly Trend of FELONY Crime Count

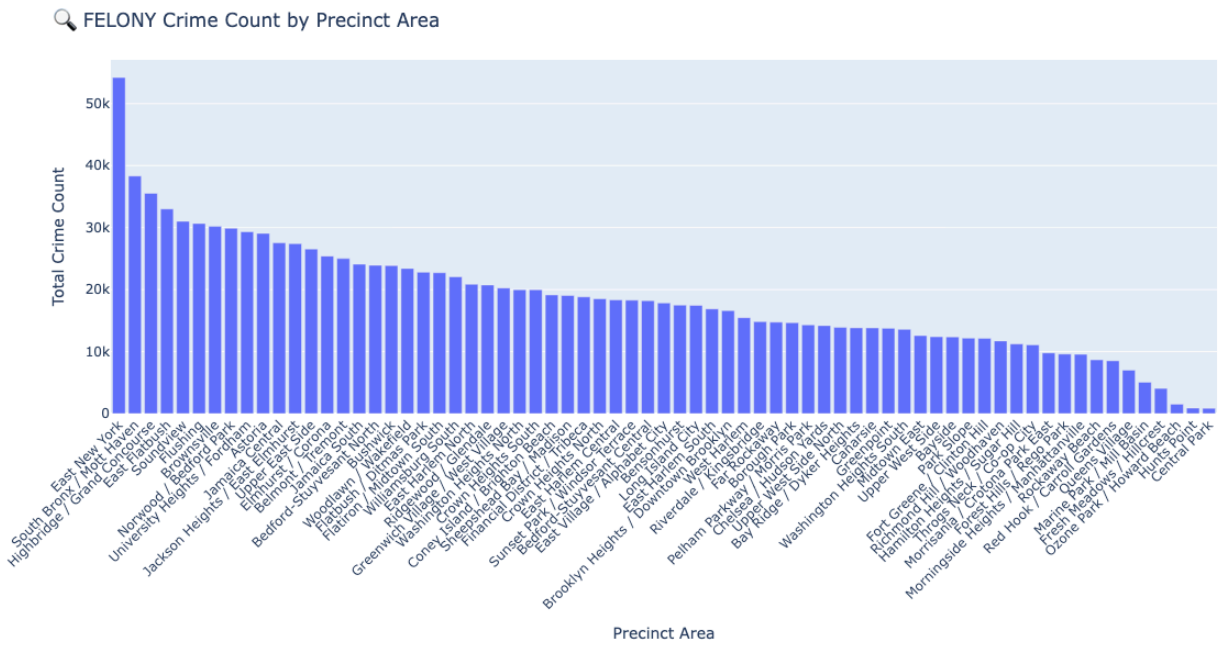




- **Insight**: Crime count dropped during the pandemic but rose post-2021. Seasonal spikes are visible.
- The overall downward trend from 2015 to 2019 was sharply reversed post-pandemic, potentially due to shifting law enforcement patterns or urban migration.

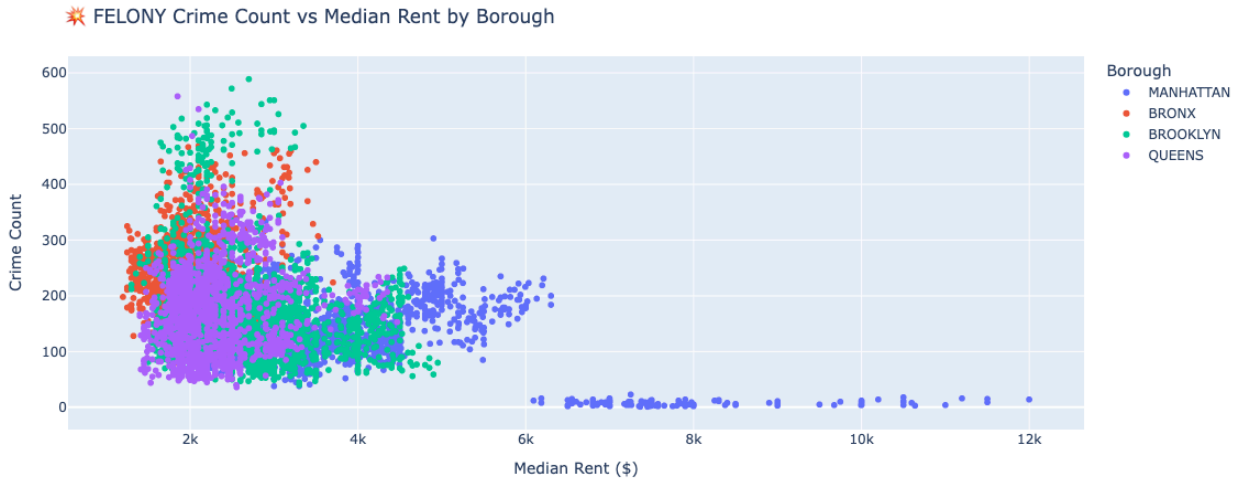
- This plot helps establish a baseline for how FELONY activity fluctuates with larger socio-political events.

4.2 Top 10 Precinct Areas by Crime Count



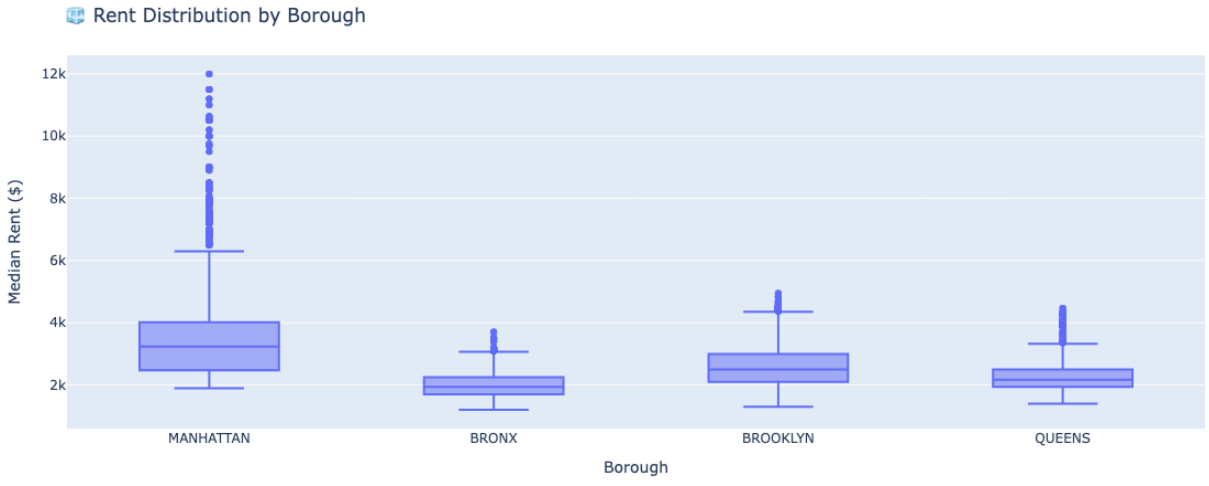
- **Insight**: Certain precincts consistently show higher crime volumes, requiring focused attention.
- These hotspots align with areas known for higher population density and commercial activity.
- Strategic resource allocation in policing and infrastructure could target these zones for intervention.

4.3 Crime Count vs Median Rent by Borough



- **Insight**: Boroughs like the Bronx have high crime counts with relatively lower rent, indicating affordability-risk tradeoff.
- Conversely, Manhattan and parts of Brooklyn cluster toward higher rent but exhibit mixed crime levels.
- This visualization emphasizes the need for context-aware urban policy: rent alone is not a measure of livability.

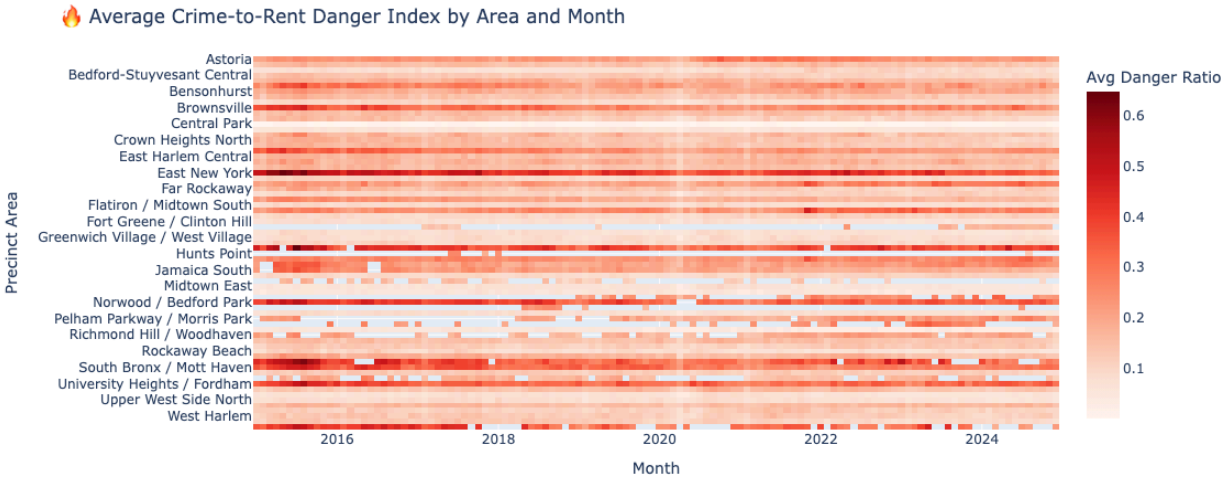
4.4 Rent Distribution by Borough



- **Insight**: Manhattan has the highest rent spread.
- The boxplot reveals a significant number of high-rent outliers in Manhattan, indicating the presence of ultra-premium properties.
- Rent distribution in the Bronx is more compact and skewed lower, making it more affordable but potentially associated with other tradeoffs such as safety or access.

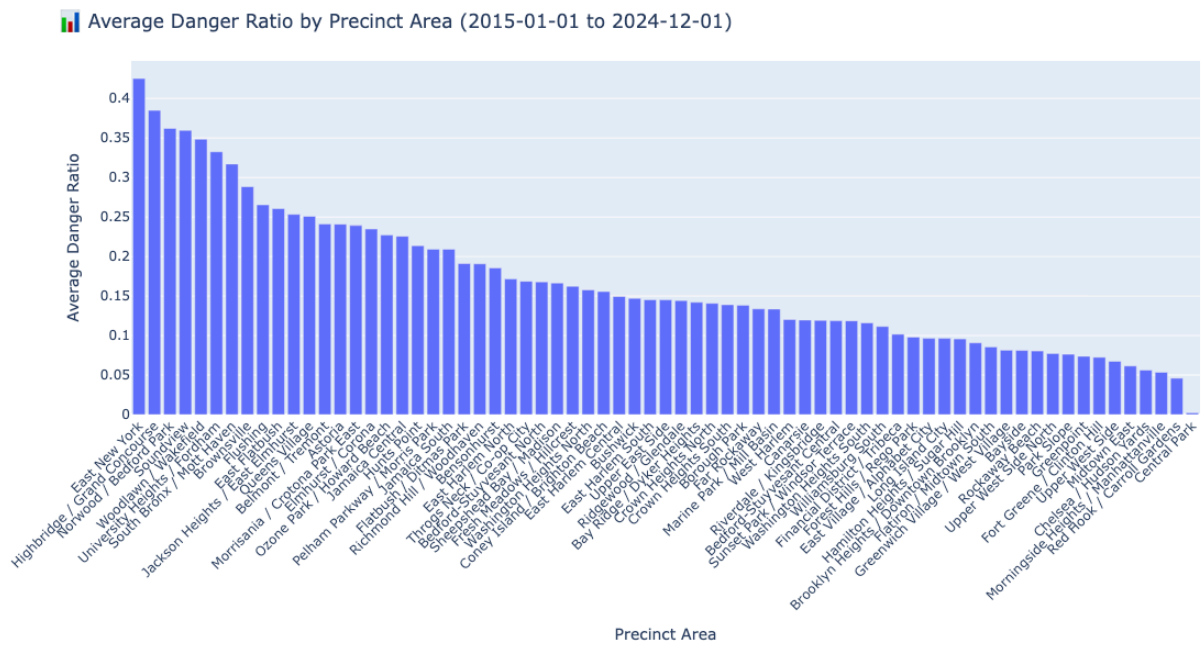
- Brooklyn shows a wide interquartile range, reflecting its mix of gentrified areas and more affordable neighborhoods.
- This visualization is useful for comparing borough-level affordability and rental volatility.

4.5 Heatmap: Avg Danger Ratio Over Time by Precinct



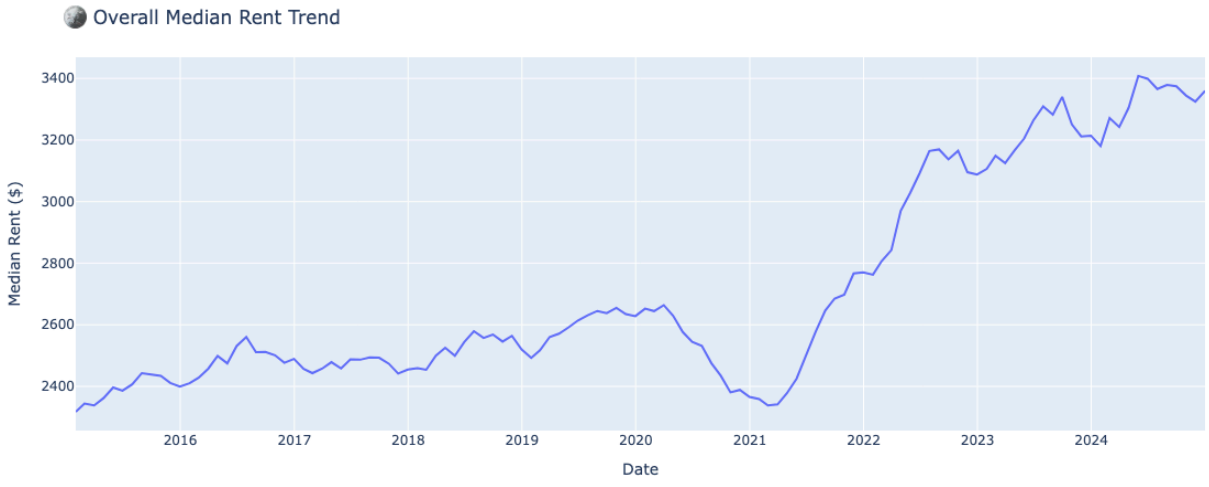
- **Insight**: Visualizes which areas are persistently dangerous and whether trends improve or worsen.
- Some precincts have consistently high danger ratios across years, while others show seasonal or policy-driven fluctuations.
- This heatmap is a strong candidate for identifying long-term policy impact or structural issues in neighborhood safety.

4.6 Avg Danger Ratio by Precinct



- **Insight**: Danger ratio isolates areas where crime outweighs cost.
- This metric captures the imbalance between housing affordability and neighborhood safety.
- High danger ratios in lower-rent areas suggest a "hidden cost" to cheap housing.
- The plot reveals that some mid-priced areas have unusually high ratios, challenging assumptions that only low-rent neighborhoods are high-risk.

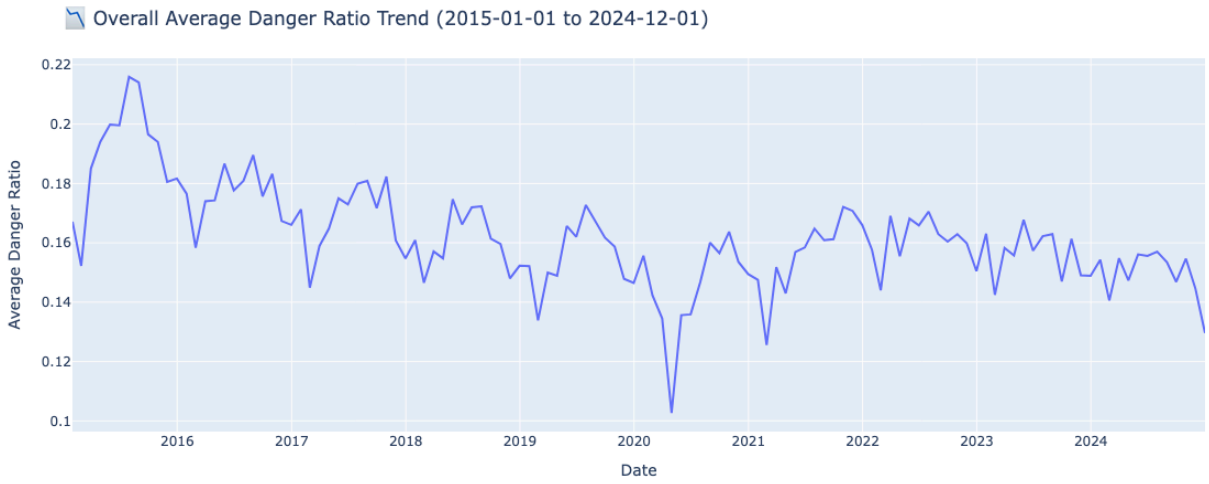
4.7 Overall Median Rent Trend



- **Insight**: Steady increase in rent post-2020, with slight dip during COVID onset.

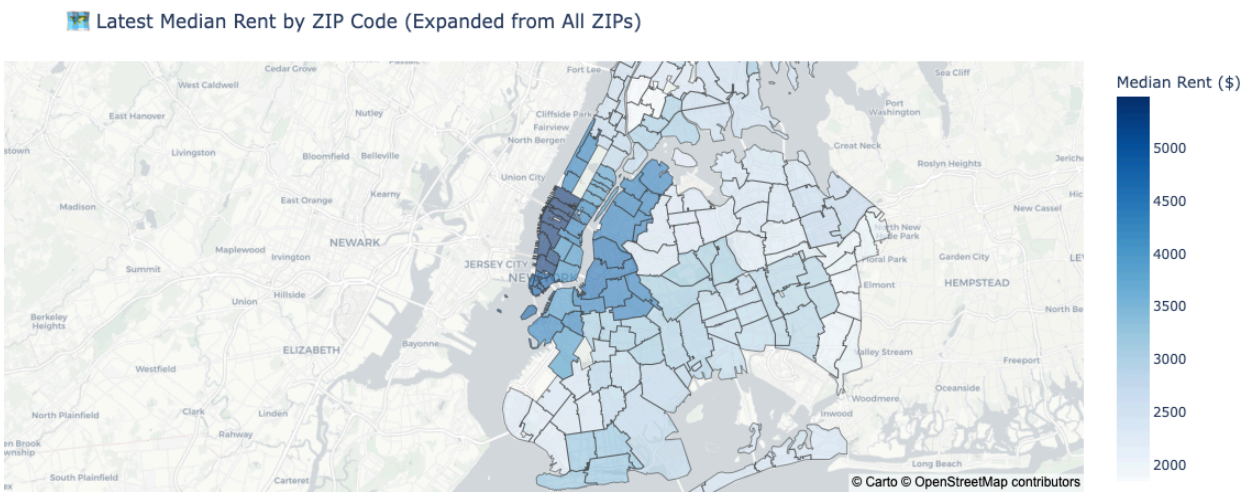
- This trend aligns with broader economic recovery and population return to NYC.
- Rent growth appears sharper in the last two years, reflecting housing supply pressures and shifting urban demand.
- Visualizing this trend helps contextualize affordability alongside rising crime or safety concerns.

4.8 Overall Danger Ratio Trend



- **Insight:** Overall decrease in danger ratio over time, with fluctuations tied to socio-economic events.
- Noticeable dip in 2020 likely linked to COVID-19 lockdowns and reduced urban activity.
- Danger ratio rebounds post-2020, but never returns to early high levels.
- Latest drop in late 2024 could indicate rapid rent inflation or declining crime — warrants further investigation.
- Tracking this metric helps interpret how “safety per dollar” changes over time and complements rent trend analysis.

4.9 Choropleth: Median Rent by ZIP Code



- **Insight**: Geographic disparities in rent, with some ZIP codes significantly above average.
- Visualization reveals clear clusters of high rent zones, especially in central Manhattan and parts of Brooklyn.
- Useful for layering additional metrics like danger ratio or amenities in future spatial dashboards.

5. Dashboard Design & Interaction

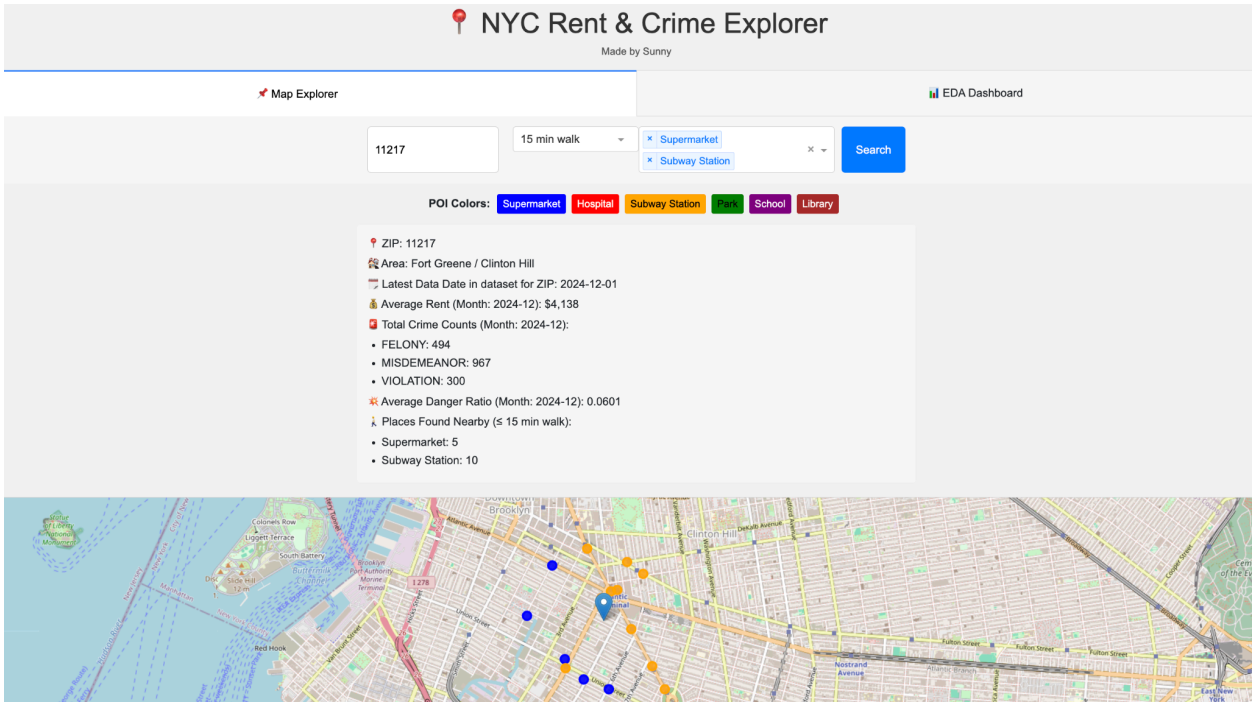
The dashboard is titled **"NYC Rent & Crime Explorer"** and was built using Plotly Dash. It provides an intuitive interface for users to explore neighborhood-level crime and rent data across New York City. The design focuses on interactivity, accessibility, and user empowerment through visual exploration.

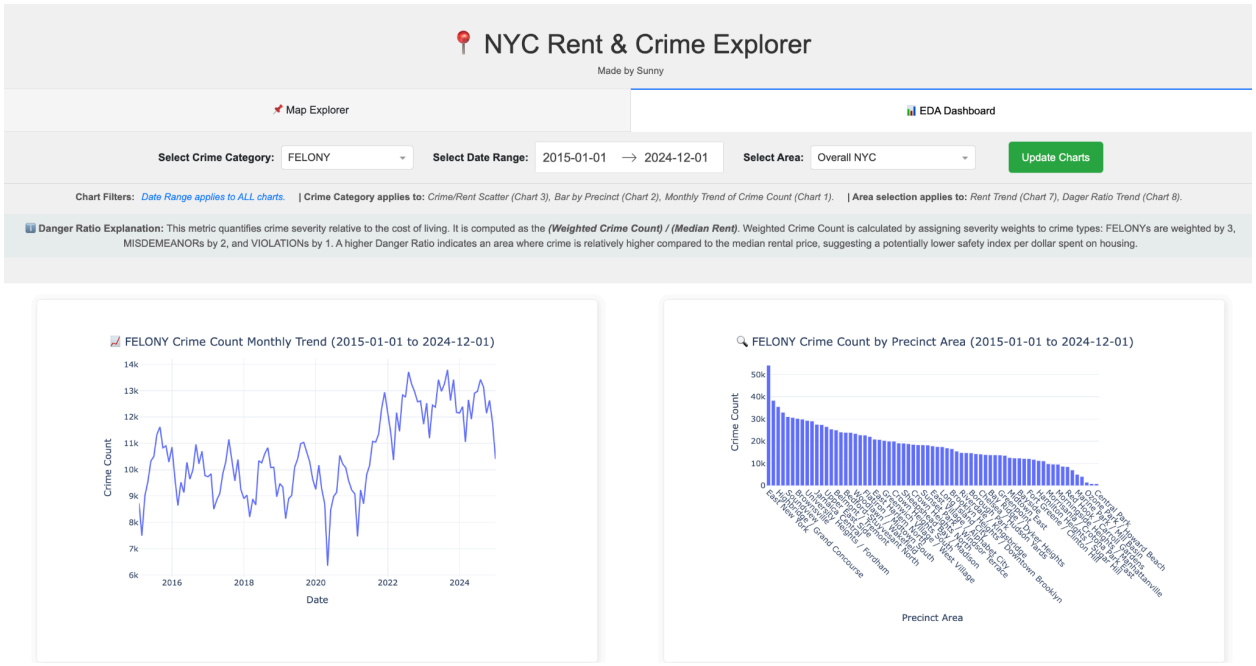
Purpose:

This dashboard bridges complex datasets with human-centered questions:

- Is this ZIP code worth living in?
- How does rent compare to crime?
- Are enough amenities within walking distance?

It serves as a decision-making tool for renters, urban planners, or real estate investors looking to weigh cost, safety, and accessibility in one unified view.





Tab 1: Map Explorer (Local ZIP Lookup)

This tab is powered by the **Google Places API**, which allows users to retrieve accurate Points of Interest (POIs) around any ZIP code input. Once a user enters a ZIP code and selects a walk

radius, the app fetches and visualizes nearby POIs — such as supermarkets, subway stations, and hospitals — directly on the map.

This provides not only a view of crime and rent data, but also neighborhood **convenience and accessibility**.

Users can then analyze hyper-local conditions using:

- **ZIP Code search**: Enter any NYC ZIP code and immediately retrieve stats for that area.
- **Walkability radius**: Choose a buffer (5, 10, or 15 minutes on foot) to assess proximity of POIs (Points of Interest).
- **POI category filters**: Visualize nearby amenities like supermarkets, hospitals, subway stations, parks, schools, and libraries using color-coded pins.
- **Summary box** for selected ZIP:
    - Area name
    - Average rent (latest month)
    - Crime breakdown by FELONY, MISDEMEANOR, and VIOLATION
    - Calculated danger ratio (crime-to-rent index)
    - Count of POIs within the defined walkable radius

This interface is designed to answer the question: “Is this neighborhood livable and balanced in terms of cost, safety, and convenience?”

Tab 2: EDA Dashboard (Macro Trends & Filtering)

This tab provides a broader exploratory experience for understanding how trends shift over time or differ across boroughs:

- **Dropdown: Crime Category** — switch between FELONY, MISDEMEANOR, VIOLATION.
- **Date Range Picker** — apply temporal filters across all charts.
- **Area Selector** — filter plots by borough or examine all of NYC.
- **Dynamic charts**:
    - Crime & Rent trend lines
    - Bar charts by precinct
    - Scatter plot for Rent vs Crime
    - Time series of danger ratio
    - Area-level heatmap

6. Theory of Visualization

This project leverages a number of powerful features from Dash and Plotly to create a dynamic and insightful visual analysis experience.

Dash & Plotly Features Utilized:

- **Interactive Graphs**: Line, bar, scatter, and choropleth maps allow users to hover and zoom for better detail exploration.
- **Dropdowns and Sliders**: Enable filtering by date, area, and crime category in real time.
- **Tabs**: Separate complex dashboards into clean, digestible sections (Map Explorer and EDA Dashboard).
- **Live Updates**: Charts respond instantly to user interaction without reloading the page.
- **Geospatial Integration**: Choropleth maps and clickable POIs provide spatial awareness for rent and safety.
- **Dynamic Text**: Area names, metrics, and explanatory labels update based on user input, enhancing context.

Theory of Data Visualization Concepts Applied:

- **Chart Junk**: Avoided unnecessary visual clutter, focusing on essential elements only.
- **Multifunctional Elements**: Combined spatial, categorical, and time-based dimensions in single views (e.g., heatmaps, choropleths).
- **Data-Ink Ratio**: Prioritized high-utility marks like lines and bars over decorative elements.
- **Small Multiples**: Organized charts by borough or precinct for easy comparison.
- **Consistency**: Maintained unified scales, colors, and label conventions across all charts.
- **Annotation & Context**: Explanatory text (e.g., Danger Ratio formula) embedded within the UI for user education.

7. Conclusion

This project bridges the gap between raw urban data and real-world decision-making by offering a deep dive into the intersection of crime, housing cost, and neighborhood dynamics in New York City. Through both EDA and an interactive dashboard, I illuminated patterns that are often hidden in spreadsheets—revealing that affordability and safety don't always go hand in hand.

My visualizations made it clear that while rent has steadily increased across most boroughs, safety levels have not consistently improved in tandem. Tools like the danger ratio allow for nuanced comparisons across ZIP codes, empowering residents, investors, and policymakers with actionable insights.

The final dashboard offers an intuitive and exploratory platform that translates complex data into accessible intelligence. Whether you're a renter searching for your next neighborhood, a policymaker targeting crime hotspots, or a real estate analyst weighing value, this project provides a versatile toolset for informed decisions.

Future improvements could include incorporating real-time data feeds. Overall, this project sets a strong foundation for understanding urban livability through the lens of data.