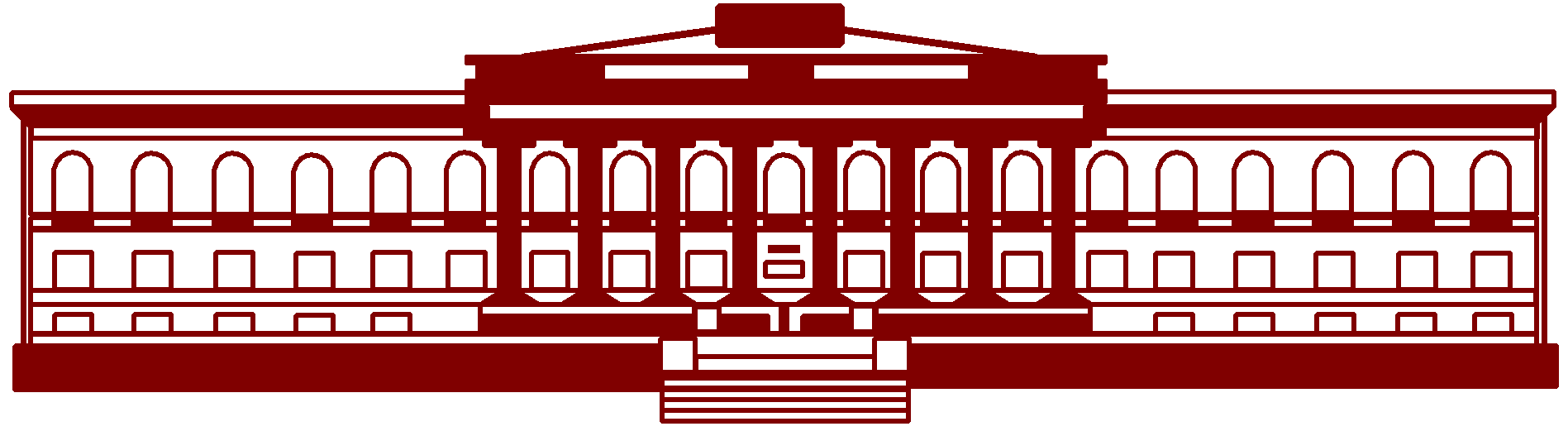
**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**імені ТАРАСА ШЕВЧЕНКА**

****

**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

**Кафедра прикладних інформаційних систем**

**Звіт до лабораторної роботи №2**

**з курсу**

**«Data Science та Big Data»**

*студентки 4 курсу*

*групи ПП-42*

*спеціальності 122 «Комп'ютерні науки»*

*ОП «Прикладне програмування»*

*Романович І.В*

*Викладач:*

*Шолохов О.В*

*Білий Р.О*

**Київ – 2023**

**Тема:** "**Розвідувальний аналіз даних (EDA). Складання аналітичного звіту**"

**Мета:** метою лабораторної роботи є отримання практичних навичок виконання розвідувального аналізу даних, використовуючи пакети jupyter, pandas, seaborn. Ознайомлення з методологією складання аналітичного звіту для зовнішнього користувача інформаційного продукту.

**Контекст:**

Ви – щойно нанятий data analyst у великій американській компанії, яка працює на ринку нерухомості США. На черговому засіданні ваш бос дав вам завдання зробити аналітичний звіт по цікавому йому сегменту ринку - Нью-Йорку

**Завдання:**

1. Виконайте дослідження domain experience стосовно американського ринку нерухомості. Ознайомтесь з декількома прикладами аналітичних продуктів від топових гравців на американському ринку, направлених на інвесторів. Питання, які потрібно опрацювати:
   1. Як топові компанії на ринку складають звіти по нерухомості?
   2. Які графіки використовуються для донесення інформації?
   3. Які співвідношення між якими даними по ринку є показовими для інвесторів / керівників агенцій нерухомості?
   4. Яка термінологія використовується для опису закономірностей на ринку нерухомості?
2. Завантажити файли з даними у папку проекту з посилання:

<https://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page>

1. Очистити дані.
2. Виконайте розвідувальний аналіз, щоб дізнатися, де є викиди або відсутні значення, вирішіть, як ви їх обробляти, переконайтеся, що дати відформатовані правильно, значення, які ви вважаєте числовими, розглядаються як такі і т.д.
3. Виконайте аналіз розвідувальних даних (отриманих результатів) для візуалізації та зіставлення за житловими масивами та за часом. Почніть шукати осмислені закономірності у цьому наборі.
4. Зберіть висновки до невеликий звіт для генерального директора (графіки, висновки з текстом у окремому файлі), який потребує належного оформлення висновків, структури тощо.

**Хід роботи:**

**Завдання 1**

a) Як топові компанії на ринку складають звіти по нерухомості?

За останні роки зростає важливість даних для прийняття обґрунтованих рішень у сфері нерухомості. Тому вивчення звітів провідних компаній стає ключовим етапом для інвесторів та фахівців галузі.

Великі компанії на ринку нерухомості використовують високоорганізований та методичний підхід до складання звітів, спрямованих на надання ключової інформації інвесторам та учасникам ринку. Цей процес включає обширний збір даних, їхній аналіз та прогнозування змін на ринку. Звіти представляються у вигляді чіткої та структурованої документації, яка дозволяє зацікавленим сторонам легко здобути необхідний уявлення про стан ринку та можливості для інвестицій. Такий комплексний та професійний підхід допомагає великим компаніям забезпечувати найвищий ступінь достовірності та корисності своїх звітів на ринку нерухомості.

Великі компанії на ринку нерухомості використовують обширний збір даних, їхній аналіз та прогнозування змін на ринку для створення звітів. Звіти представляються у вигляді чіткої та структурованої документації, яка дозволяє зацікавленим сторонам легко здобути необхідний уявлення про стан ринку та можливості для інвестицій.

b) Які графіки використовуються для донесення інформації?

У створенні звітів по нерухомості використовуються різні графічні елементи, що допомагають інвесторам та учасникам ринку краще розуміти та аналізувати дані. Зокрема, лінійні графіки дозволяють визначити тенденції цін з часом, стовпчасті графіки використовуються для порівняння різних районів за ціновою динамікою, а географічні карти вказують на розташування об'єктів у контексті інфраструктурних особливостей. Це візуальне підкріплення дозволяє учасникам ринку краще розуміти та швидше реагувати на зміни у сфері нерухомості [1].

Для створення таких графіків можна скористатися різноманітними інструментами та програмним забезпеченням. Наприклад, на сайті GraphicRiver доступно багато графічних елементів, дизайнів та шаблонів для звітів по нерухомості.

Такий підхід допомагає інвесторам та учасникам ринку швидше виявляти перспективні напрямки для інвестицій або ризики в конкретних регіонах. Крім того, використання графічних елементів робить дані більш доступними та легко зрозумілими.

с)Які співвідношення між якими даними по ринку є показовими для інвесторів / керівників агенцій нерухомості?

У світі нерухомості для визначення правильних стратегій і прийняття обґрунтованих рішень інвесторами та керівниками агенцій необхідно звернути увагу на певні ключові показники. Одним з таких показників є відношення ціни до доходу, яке вказує на доступність житла для місцевого населення в порівнянні з їхнім доходом. Високе відношення може свідчити про можливу переоціненість ринку.

Також важливим показником є відношення ціни до орендної плати, яке допомагає оцінити якість інвестицій в орендну нерухомість та визначити, чи є вигідним придбання чи оренда житла.

Крім того, відношення попиту до пропозиції є важливим показником, який вказує на взаємодію попиту та пропозиції на ринку нерухомості. Цей показник допомагає визначити життєздатність ринку та імовірність вигідних угод.

Крім того, вартість життя та рівень доходу в регіоні мають значення для визначення платоспроможності місцевого населення. Співвідношення між вартістю життя та рівнем доходу допомагає зрозуміти, наскільки люди можуть собі дозволити придбати нерухомість у даному регіоні.

Ці показники допомагають інвесторам та керівникам агенцій нерухомості зробити обґрунтовані рішення та визначити перспективні напрямки для інвестицій у сферу нерухомості.

d) Яка термінологія використовується для опису закономірностей на ринку нерухомості?

Цінова еластичність попиту в нерухомості вказує на те, наскільки змінюється кількість людей, які бажають придбати нерухомість, при зміні її ціни. Цей показник вимірює чутливість попиту до зміни ціни товару. Висока цінова еластичність попиту означає, що зміна ціни суттєво впливає на кількість попиту, тоді як низька еластичність попиту свідчить про меншу залежність кількості попиту від цінових змін.

Зважаючи на це, важливо враховувати цінову еластичність попиту при встановленні цін на нерухомості. Якщо ціна занадто висока і цінова еластичність попиту велика, це може призвести до зменшення кількості покупок та скорочення ринку. З іншого боку, якщо ціна надто низька і еластичність попиту низька, це може призвести до недоотримання максимальної вигоди від продажу нерухомості.

Цінова еластичність попиту є важливим фактором при аналізі ринку нерухомості і прийняття рішень щодо ціноутворення. Для більш точного визначення цих показників можуть використовуватися економетричні методи та статистичні моделі.

Щодо попиту та пропозиції, це два ключових поняття у сфері нерухомості. Попит вказує на кількість людей, які бажають придбати нерухомість, тоді як пропозиція означає наявну кількість нерухомості на ринку для продажу. Взаємодія між попитом та пропозицією визначає ринкову активність та можливості для угод.

Цикл нерухомості описує різні етапи активності на ринку. Включаються періоди зростання, спаду та стабільності. Цей цикл може бути впливований рядом факторів, таких як економічна кон'юнктура, зміни в політиці та інші зовнішні чинники.

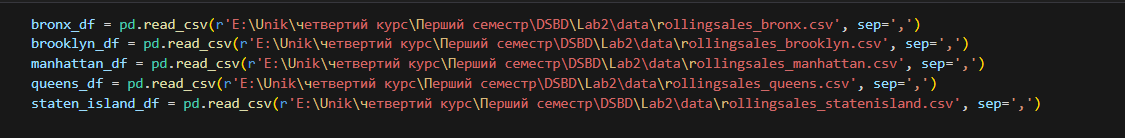
Вартість замовлення - це оцінка вартості будинку, яка проводиться перед його будівництвом або придбанням. Це оцінка потенційних витрат на будинок, включаючи матеріали, працю та інші фактори.

Ринкова вартість - це фактична ціна, за яку можна продати нерухомий об'єкт. Вона визначається попитом та пропозицією на ринку, а також іншими факторами, такими як стан нерухомості, розташування та економічна ситуація.

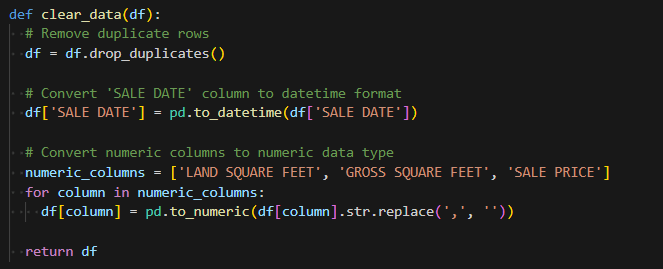
Іпотека та процентна ставка стосуються умов фінансування для придбання нерухомості. Іпотека - це вид кредиту, який надається для придбання житла, де саме житло служить гарантією повернення кредитних коштів.

**Завдання 2**

Об’єднаємо дані в один датафрейм

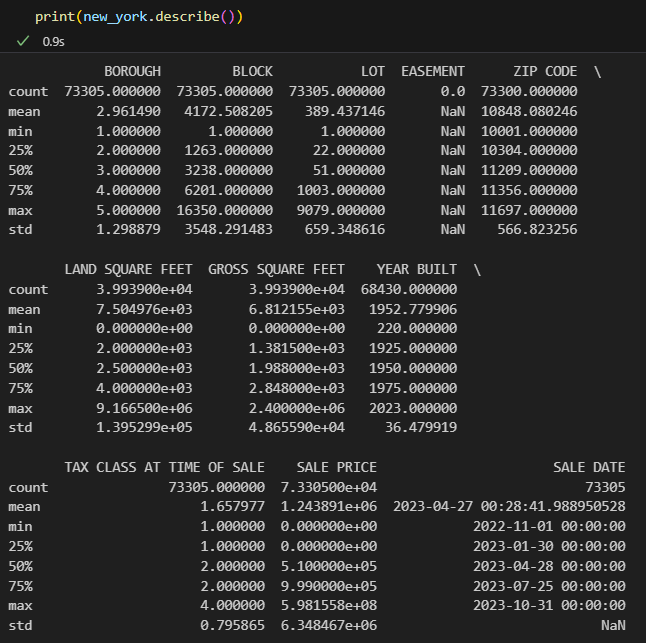


Очистимо дані. Для цього очистимо дані від дублікатів, приведемо дані до правильних типів даних, видалимо дані у яких відсутні необхідні значення (район,мікрорайон, ціна продажу, дата продажу). Для цього створимо функцію clear\_data:

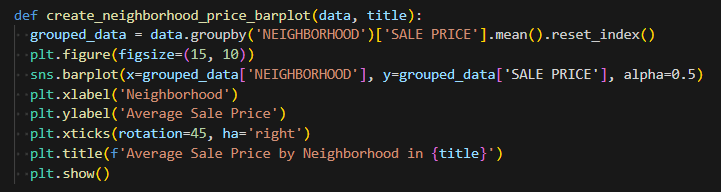


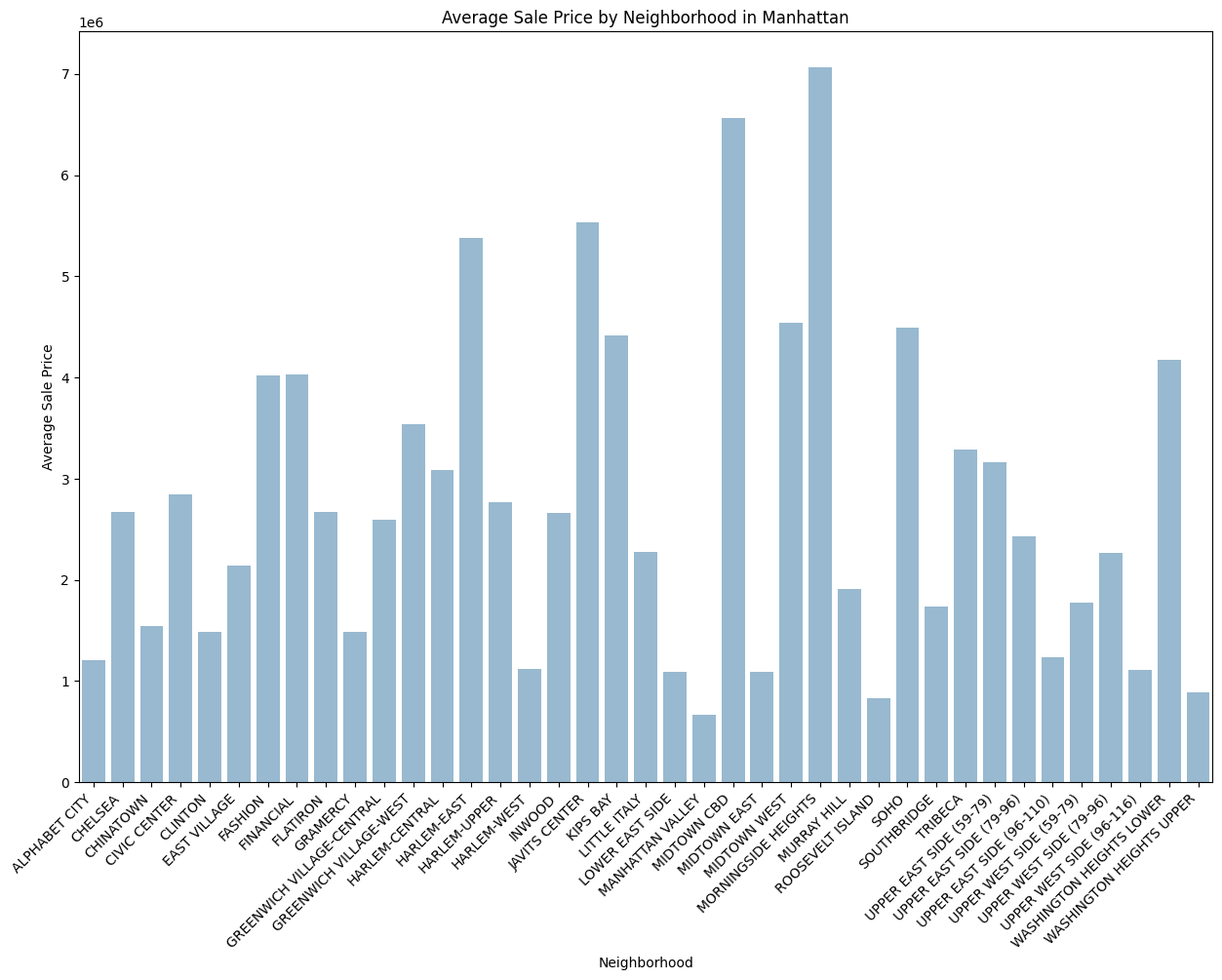
Проведемо розвідувальний аналіз даних

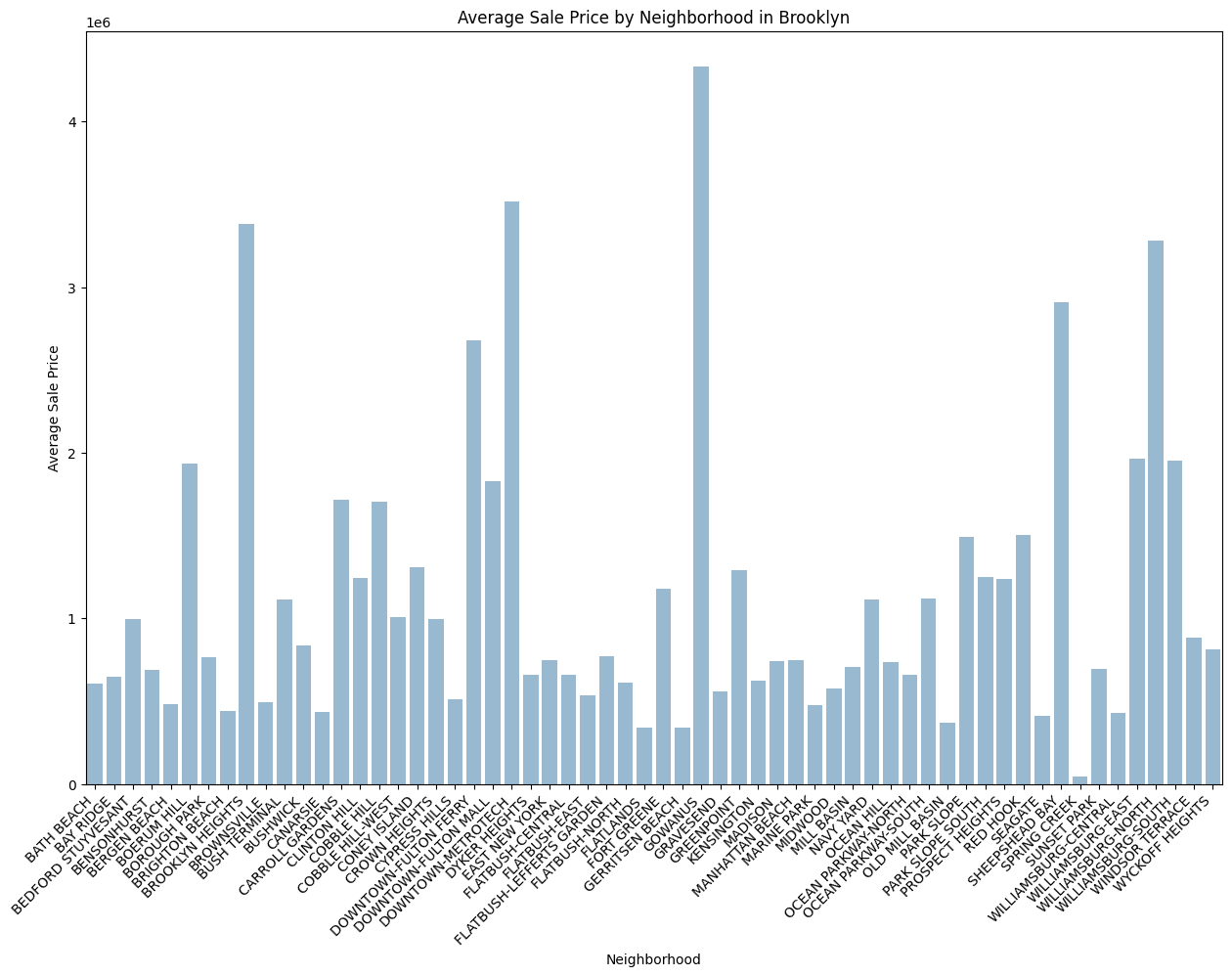
Отримаємо попередню оцінку даних (к-сть ненульових значень, максимум, мінімум, середнє значення, стандартне відхилення, процентиль 25%, 50%, 75% )

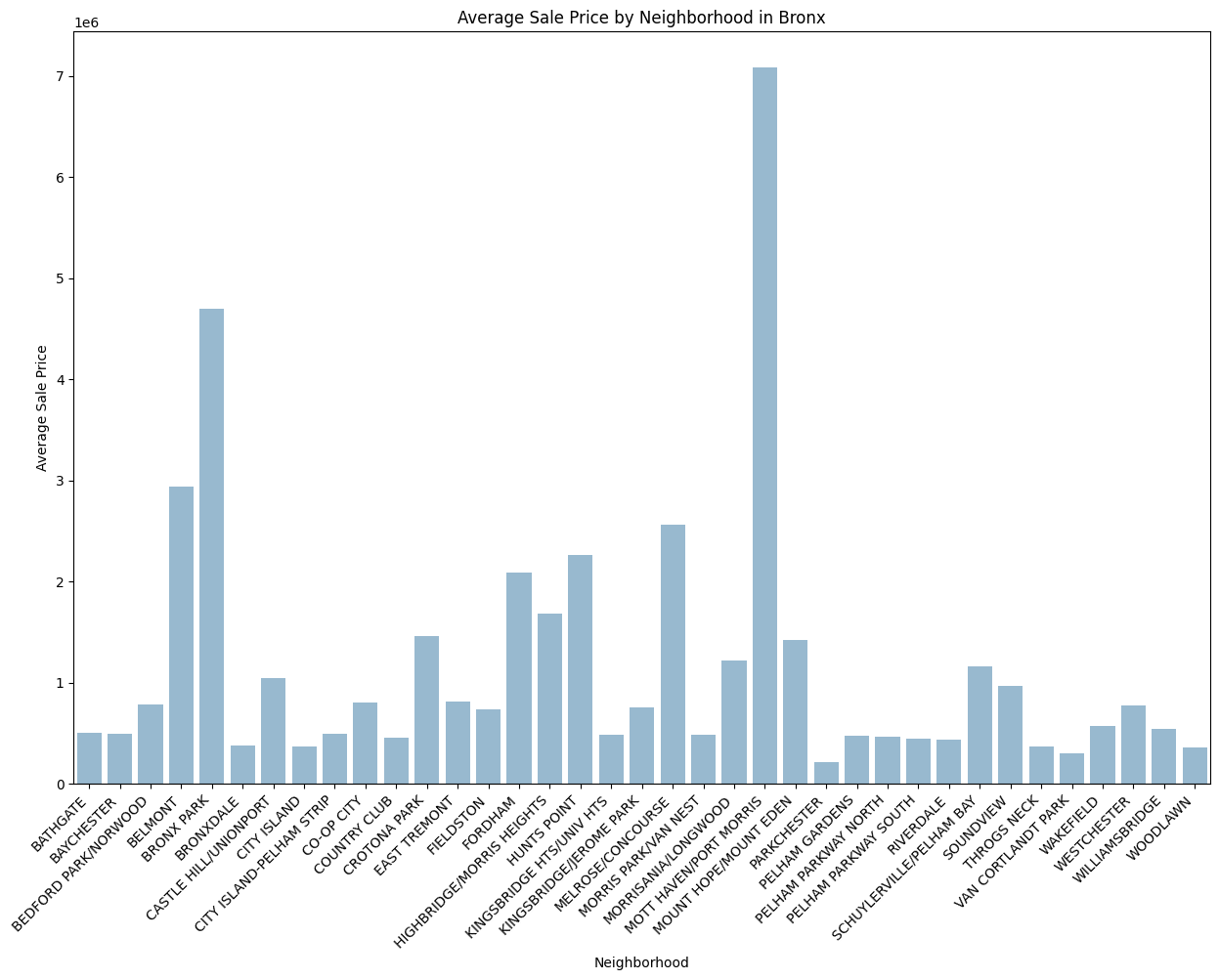


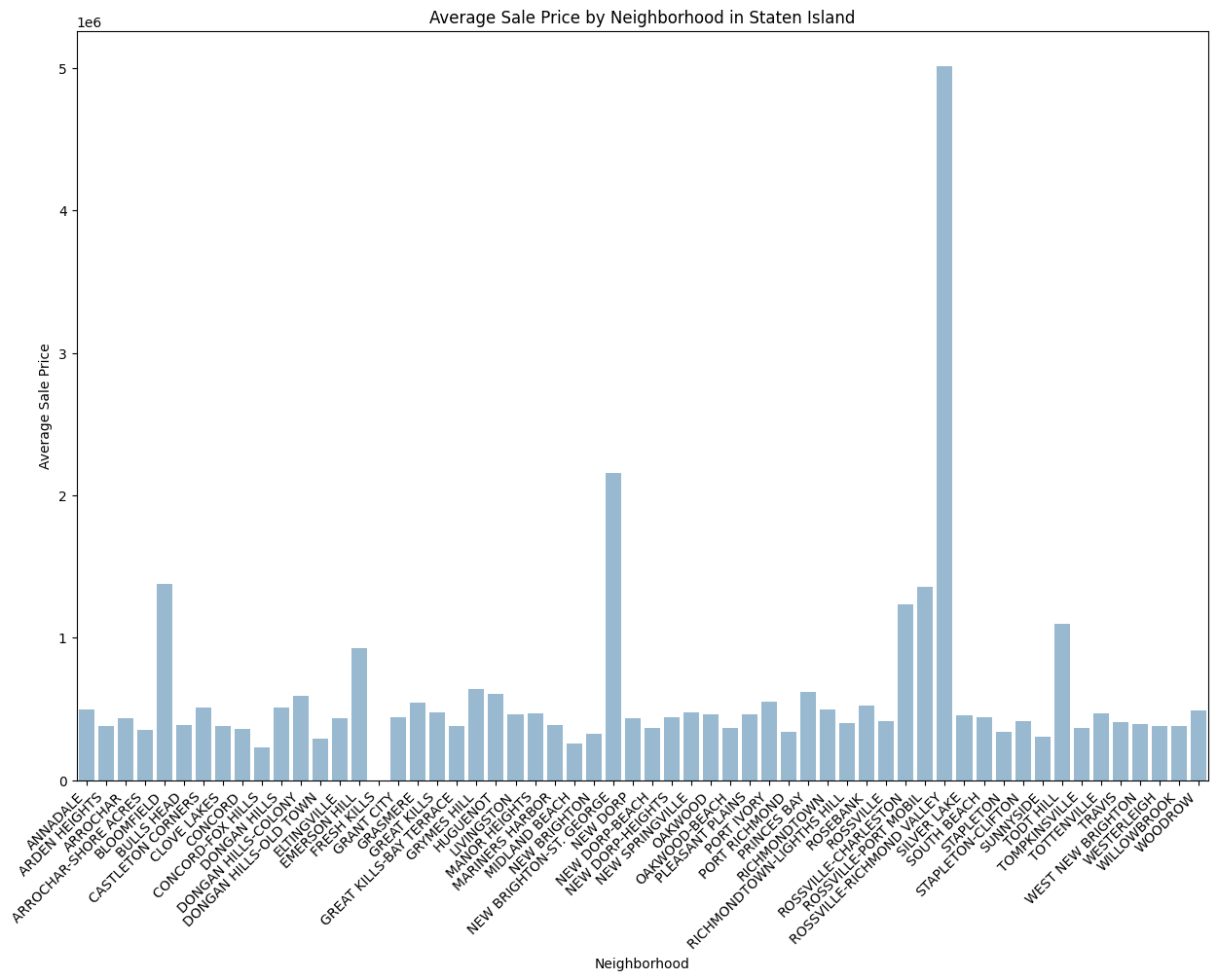
Далі дослідимо середню ціну для кожного мікрорайону района Нью-Йорка:





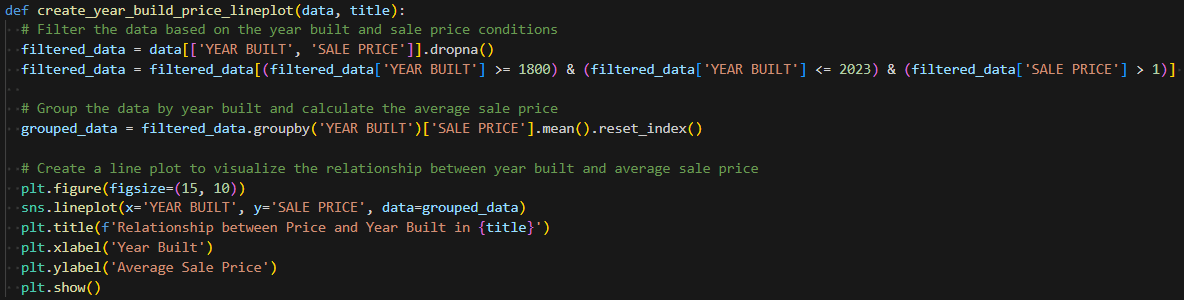
**

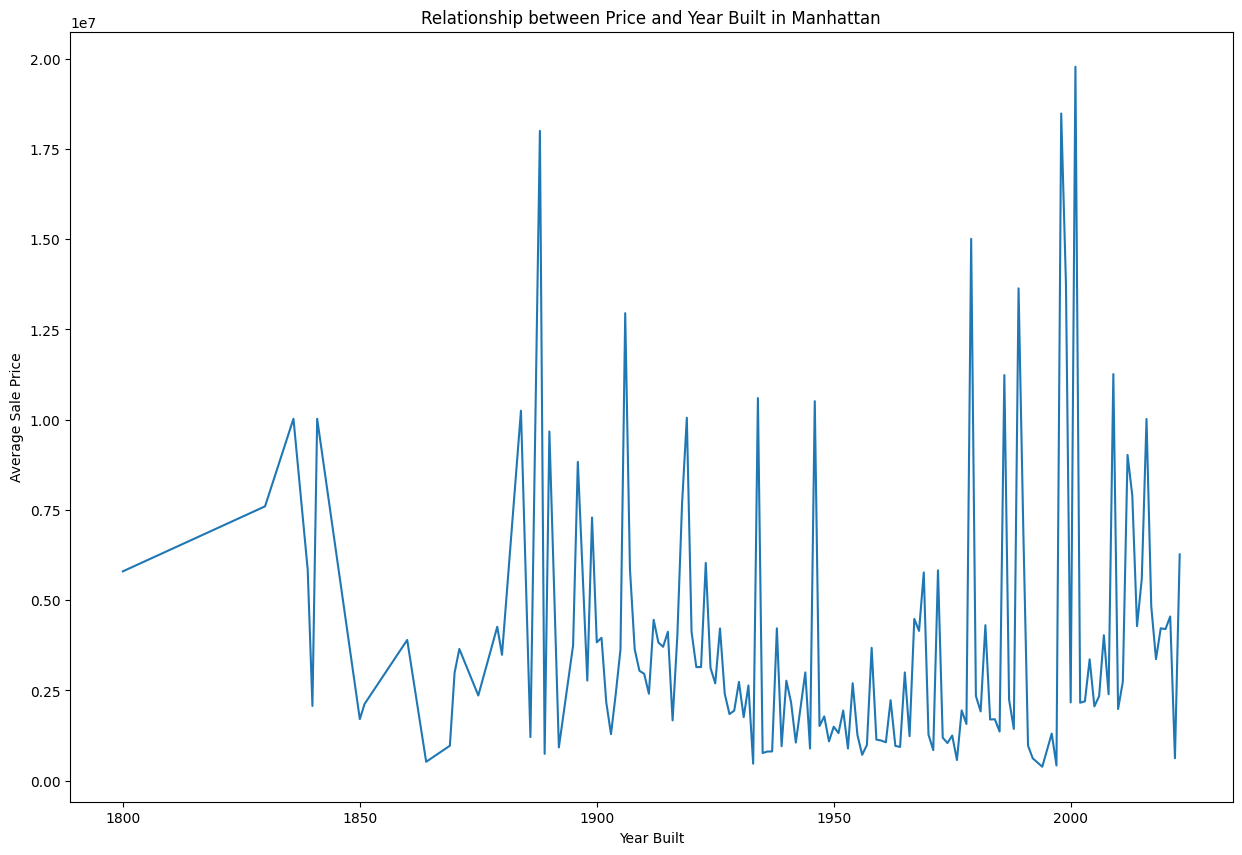
**

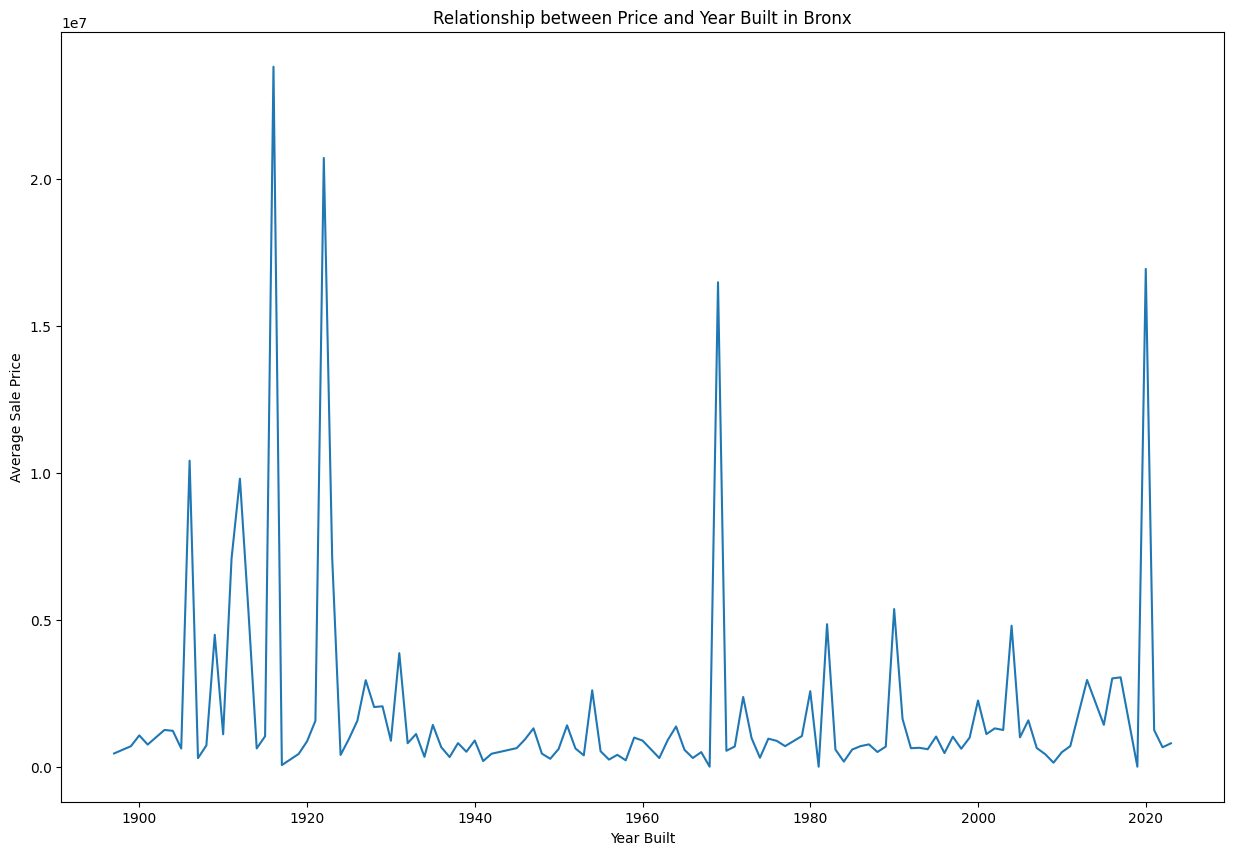
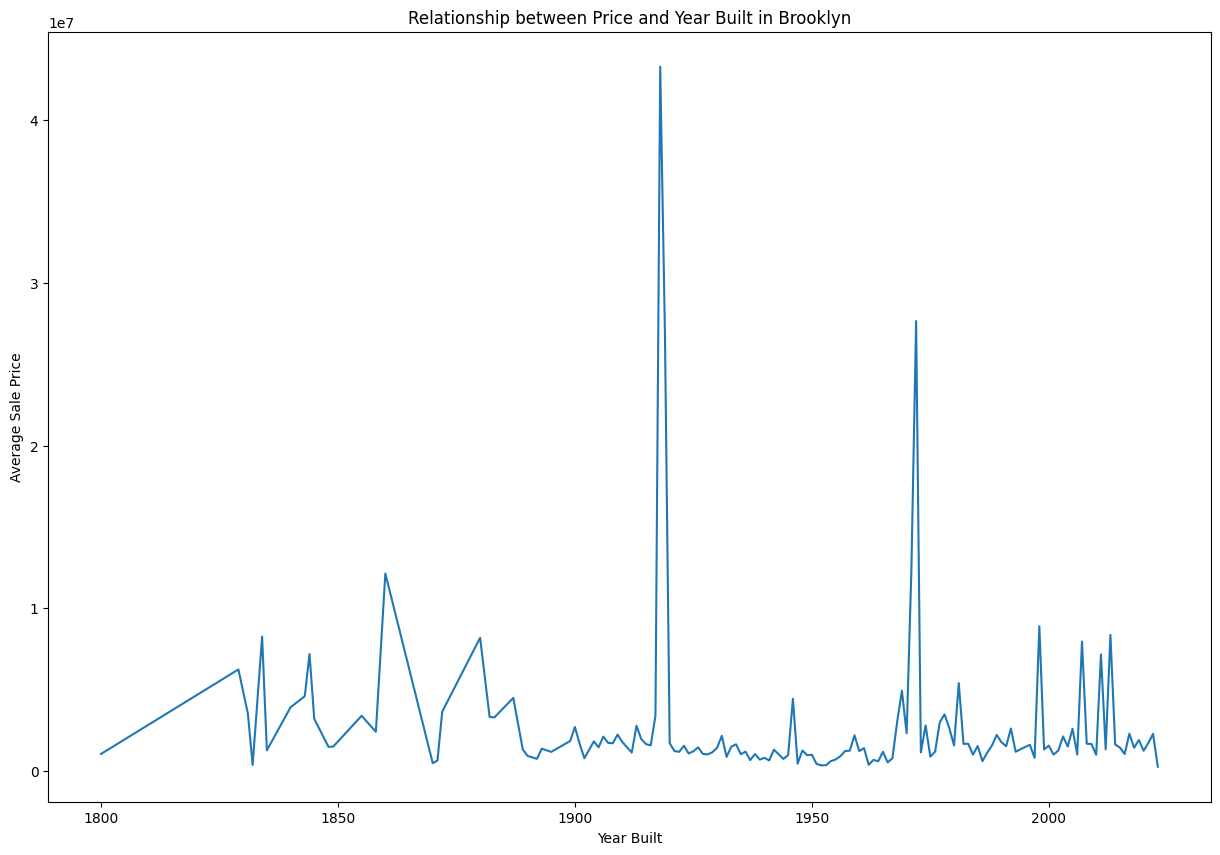
**

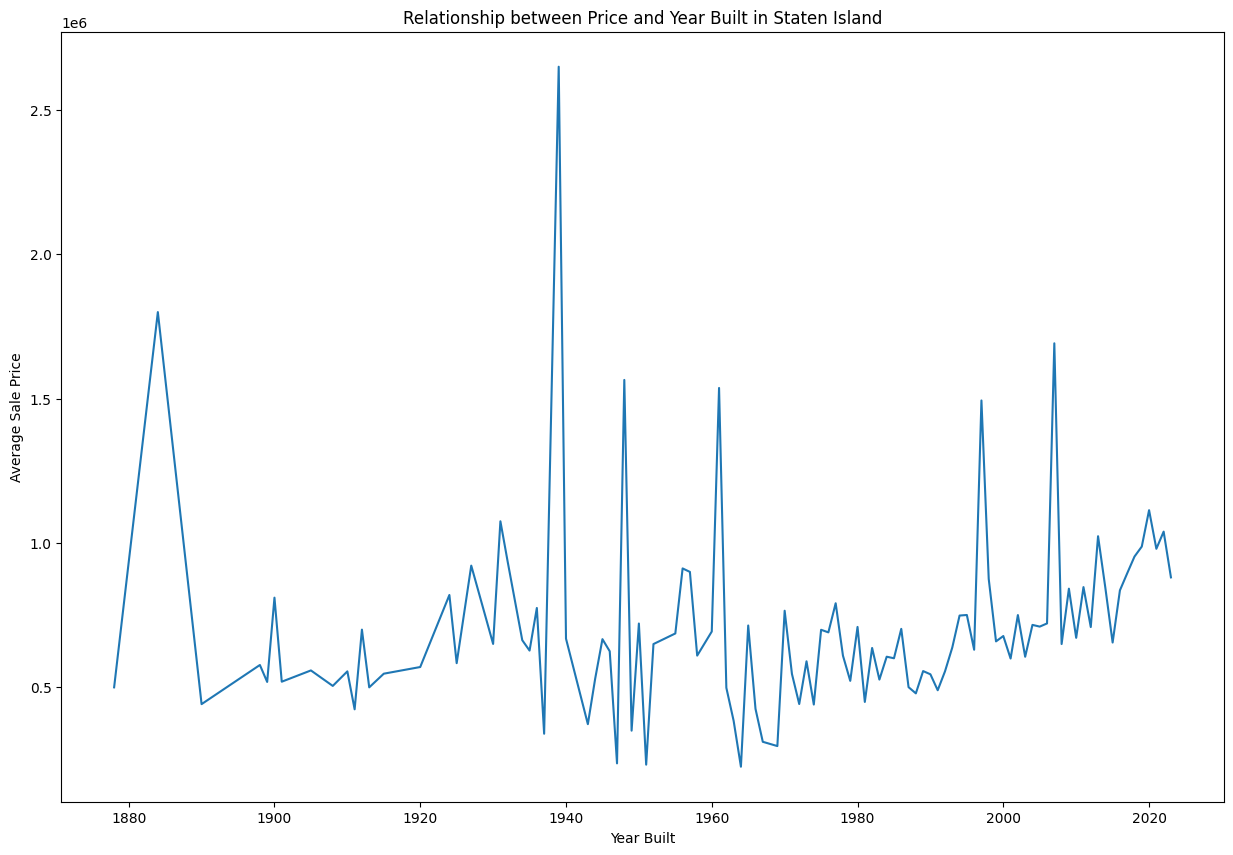
З візуалізації бачимо що найдорожчими мікрорайонами є Morningside Height у Манхетені, Govanus в Брукліні, Mott Haven/Port Morris в Бронксі та Rossville-Richmong Valley в Staten Island.

Дослідимо відношення ціни на будівлю до її віку



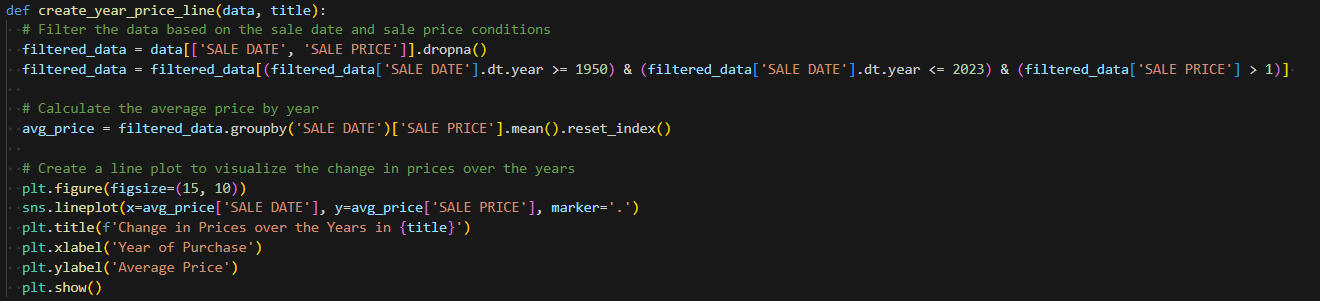
**

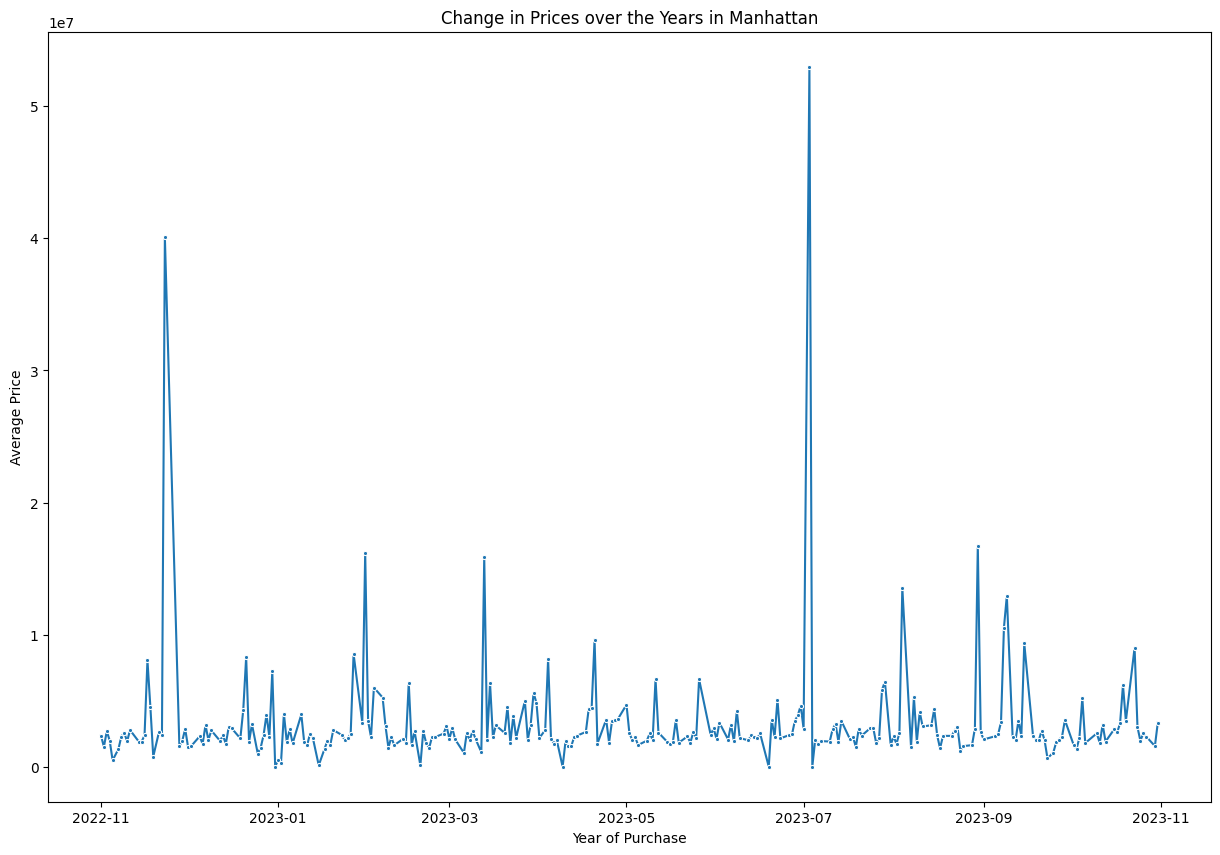
**

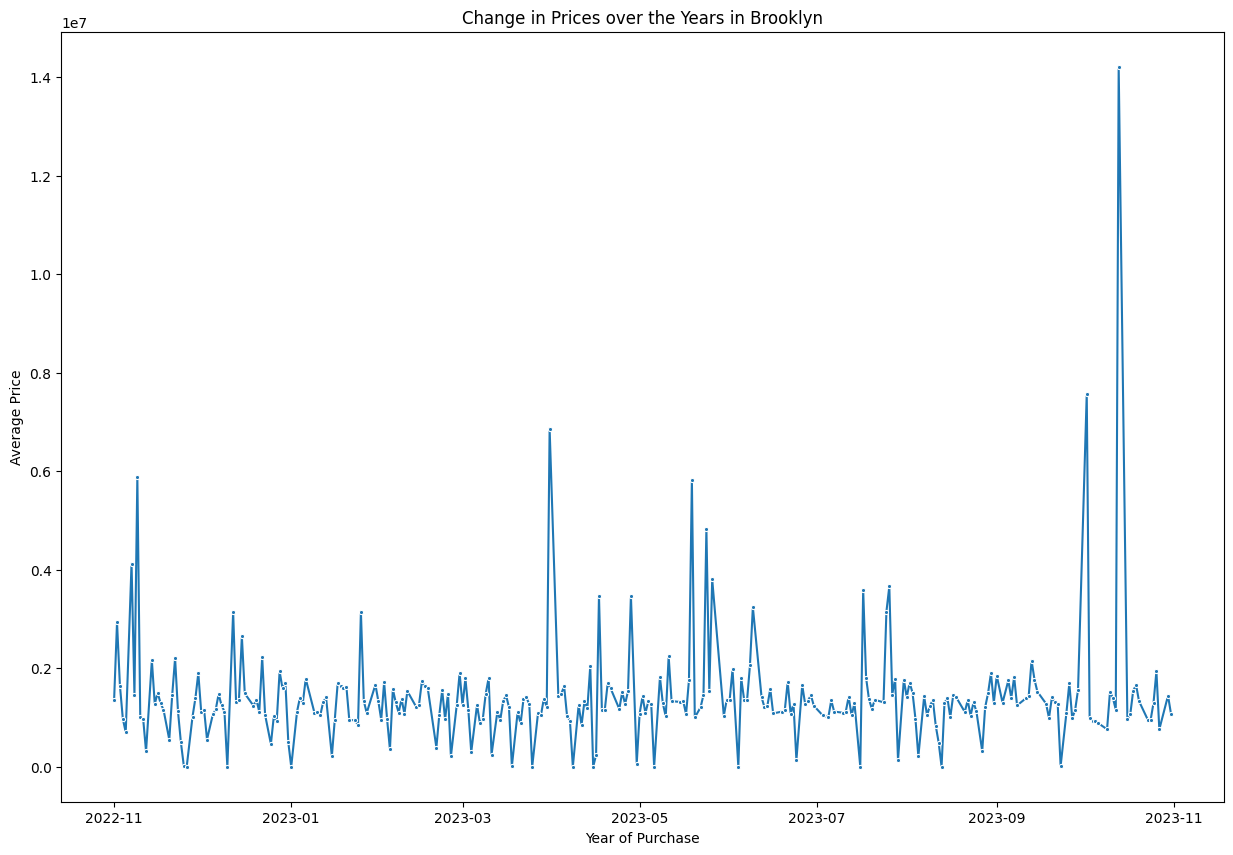
**

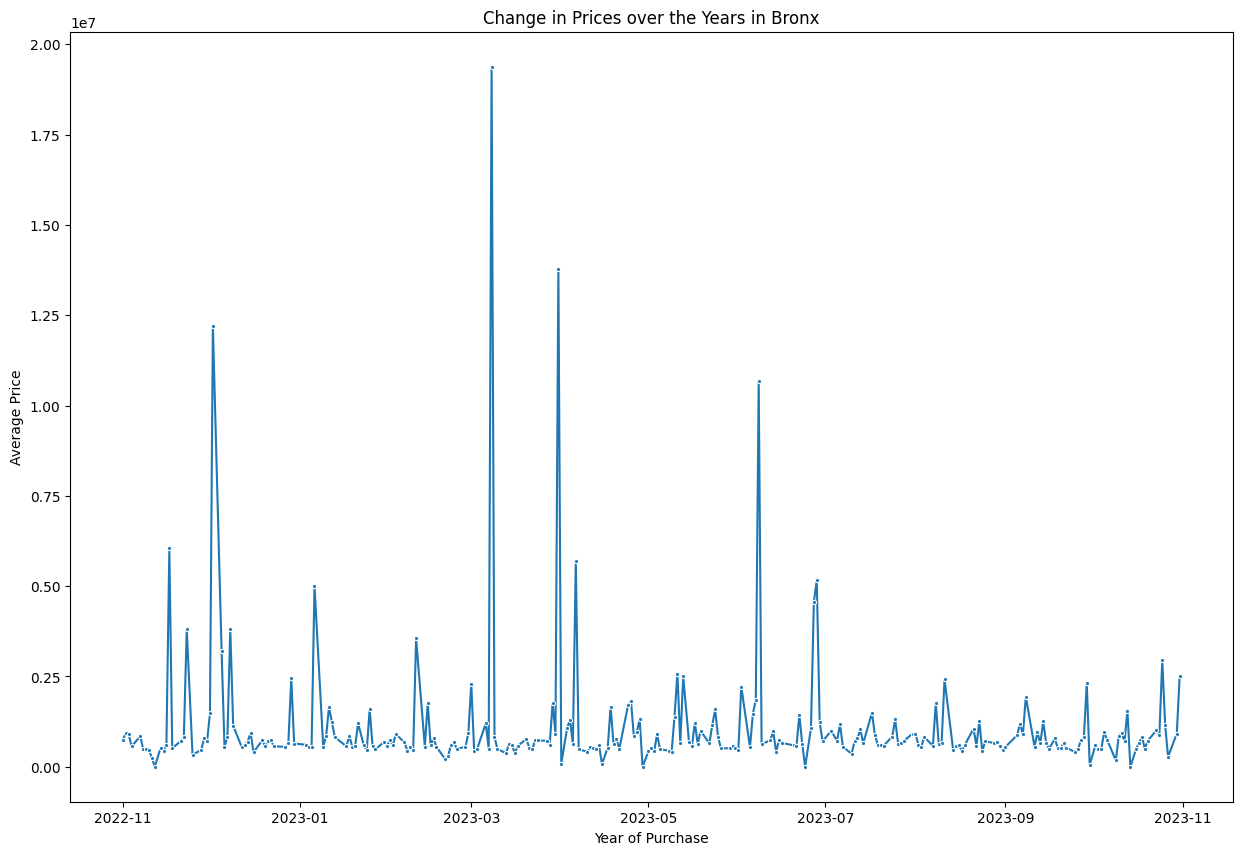
З графіків можна побачити що у Манхетені найбільш вартують будівлі кінця 19 століття та 1960х-70х років, у Брукліні – 1920х та 1970х років, у Бронксі в загальному найвищу ціну мають будівлі першох половини 19 століття, а у Стейтен Айленді – 1890 років, та будівлі періоду 1935-1965 років.

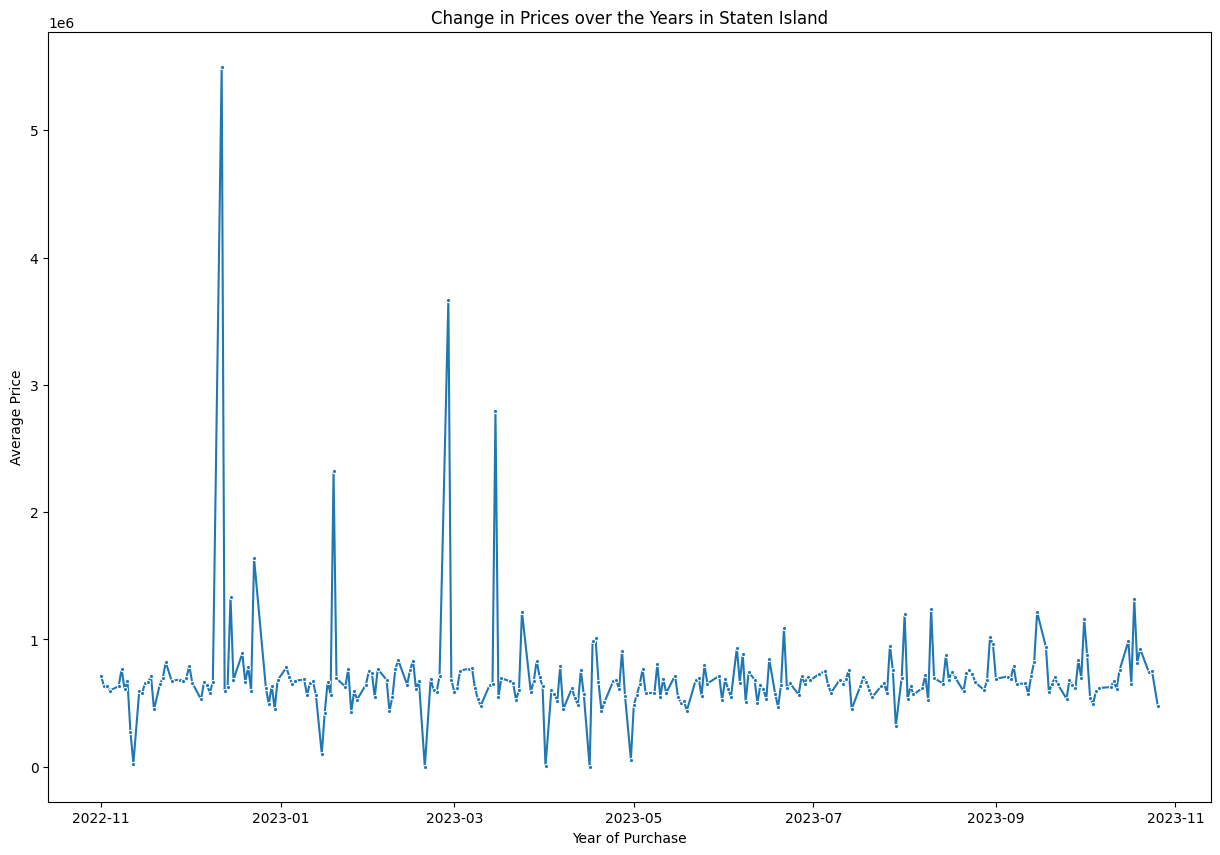
Далі дослідимо як мінялась середня вартість будівлі з часом





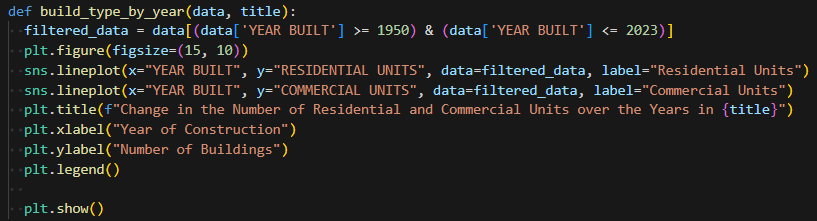


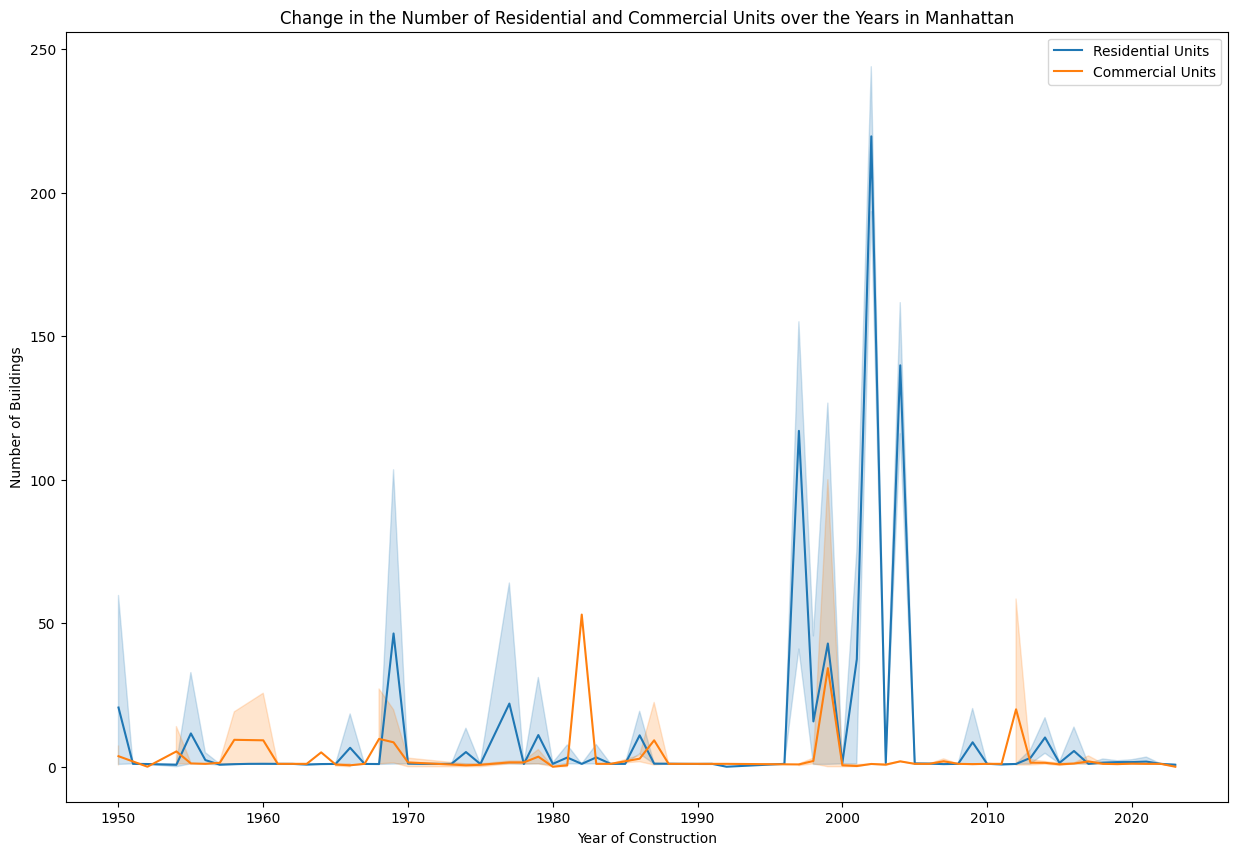


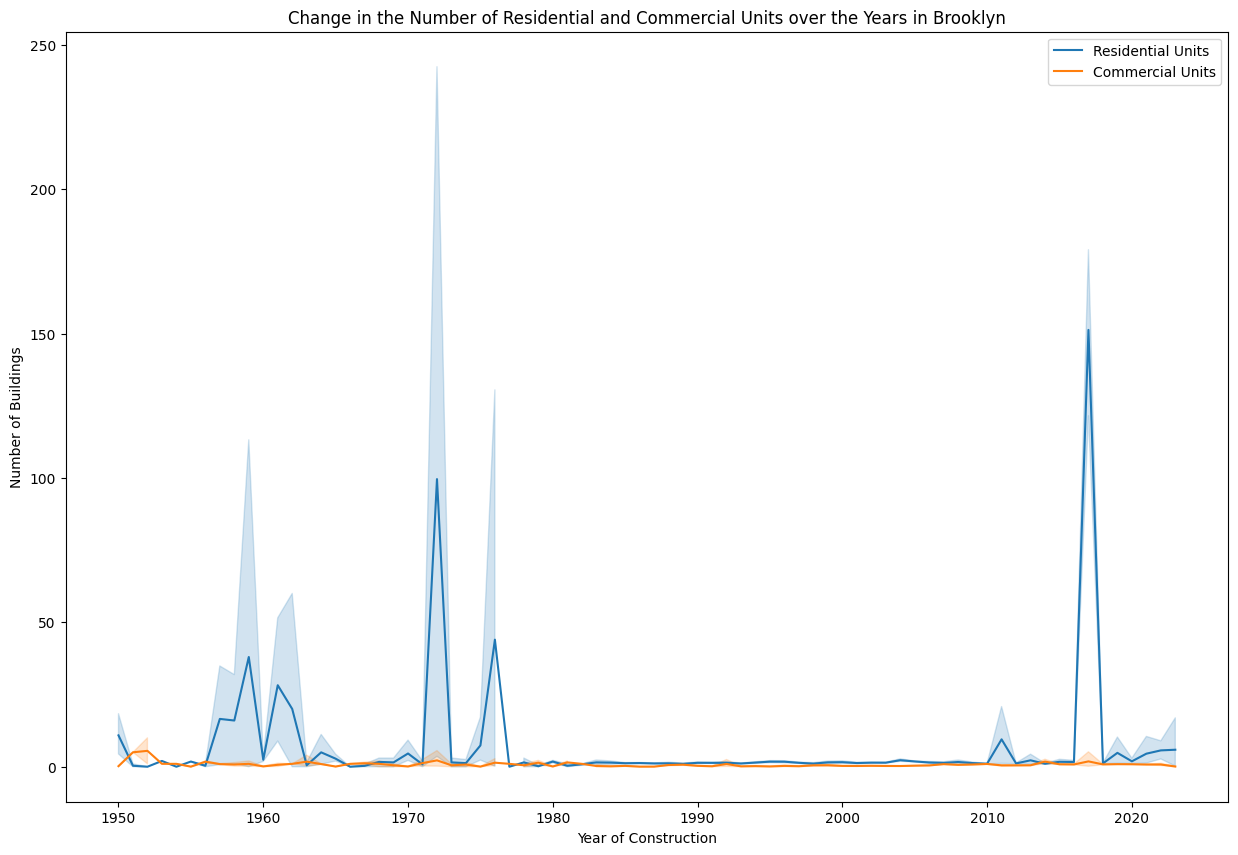


Згідно графіків, можна зробити висновок, що середня сума, яку люди віддавали за будівлі у Манхетені досягла максимума у листопаді 2022 та липні 2023, у Брукліні досягла свого максимума за осінь 2023 року, і зросла у 7 разів, у Бронксі – максимальна середня сума була отримана у період 2023-03 по 2023-05, а у Стейтен Айленді з 2022-11 по 2023-01.

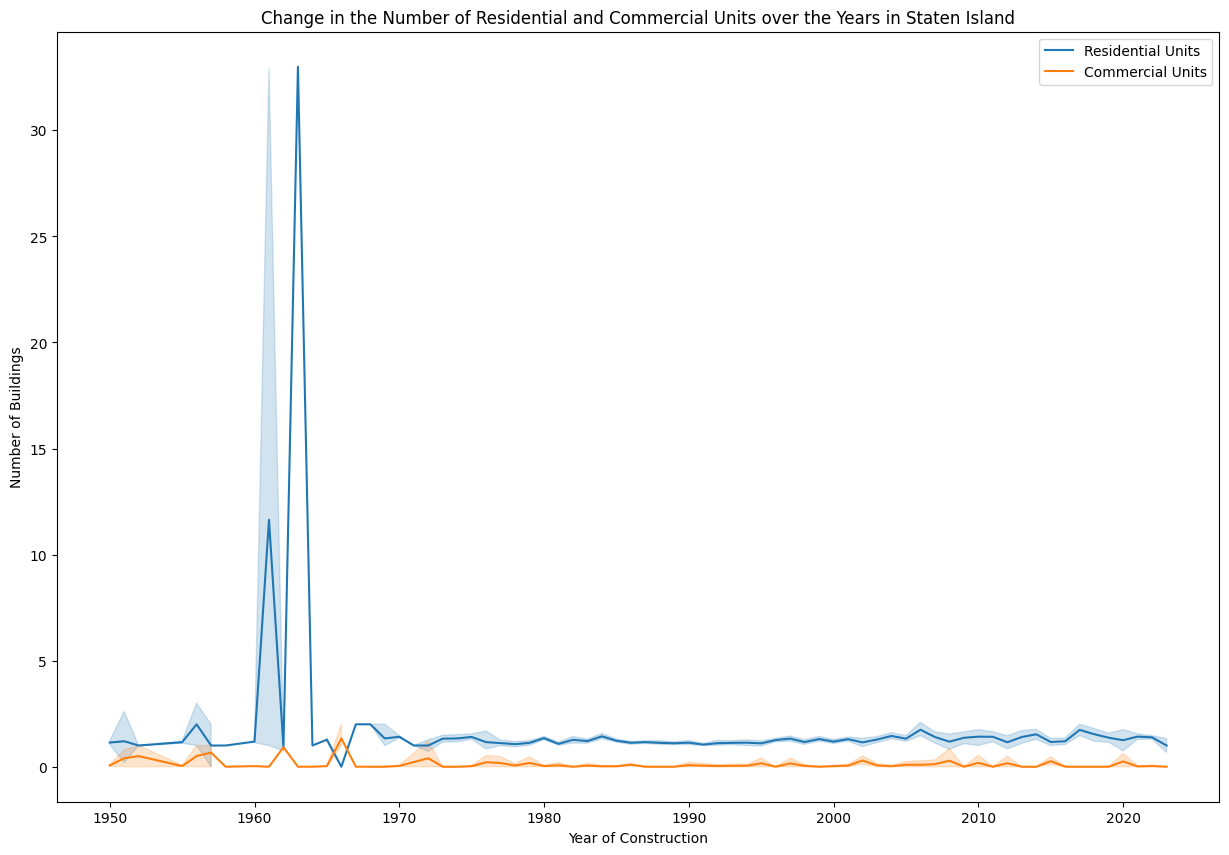
Далі дослідимо зміну кількості комерційних та житлових будівель з часом:



\

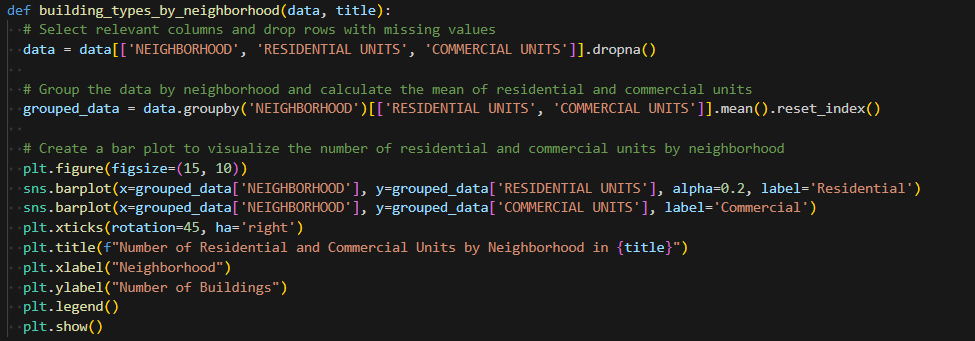


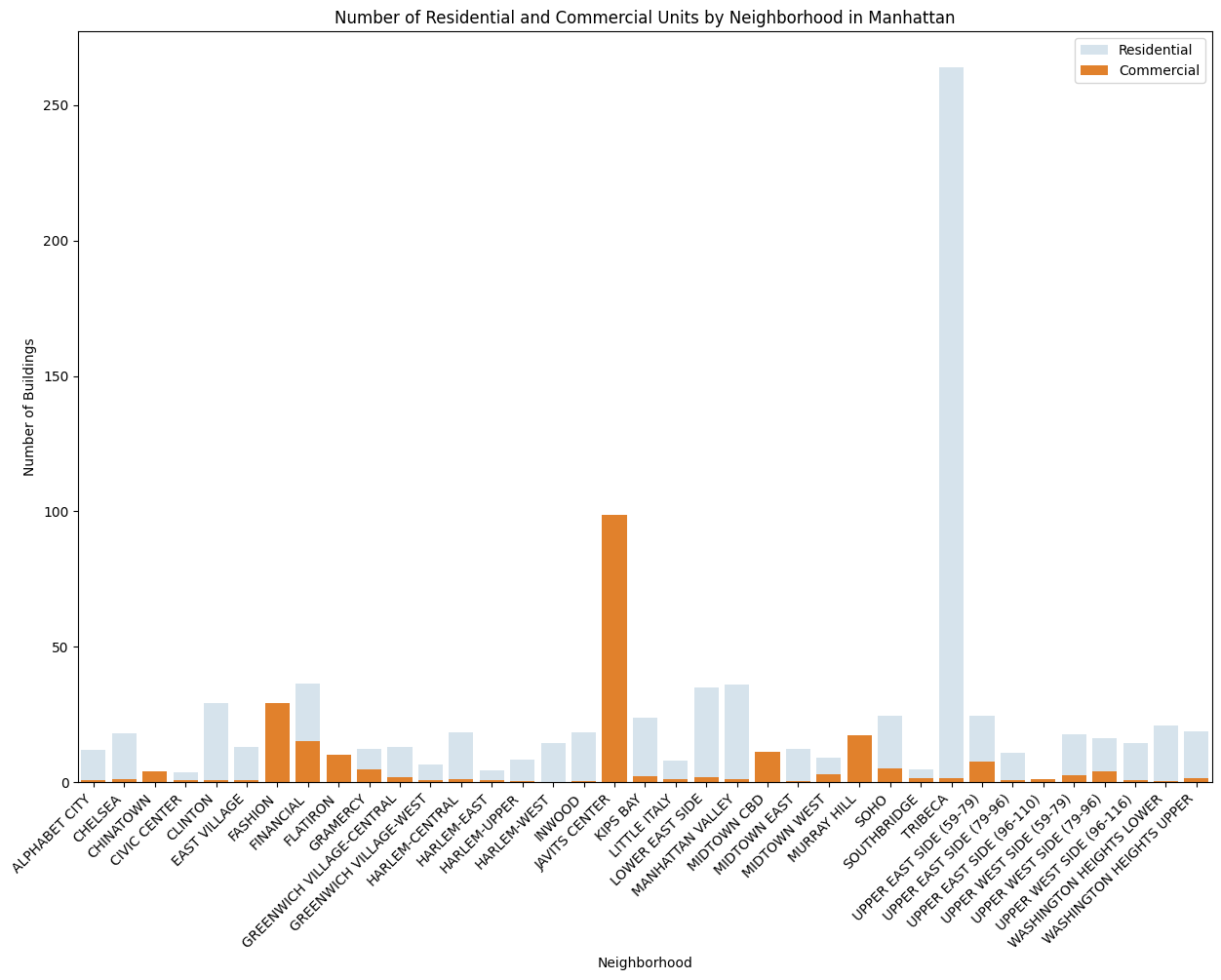


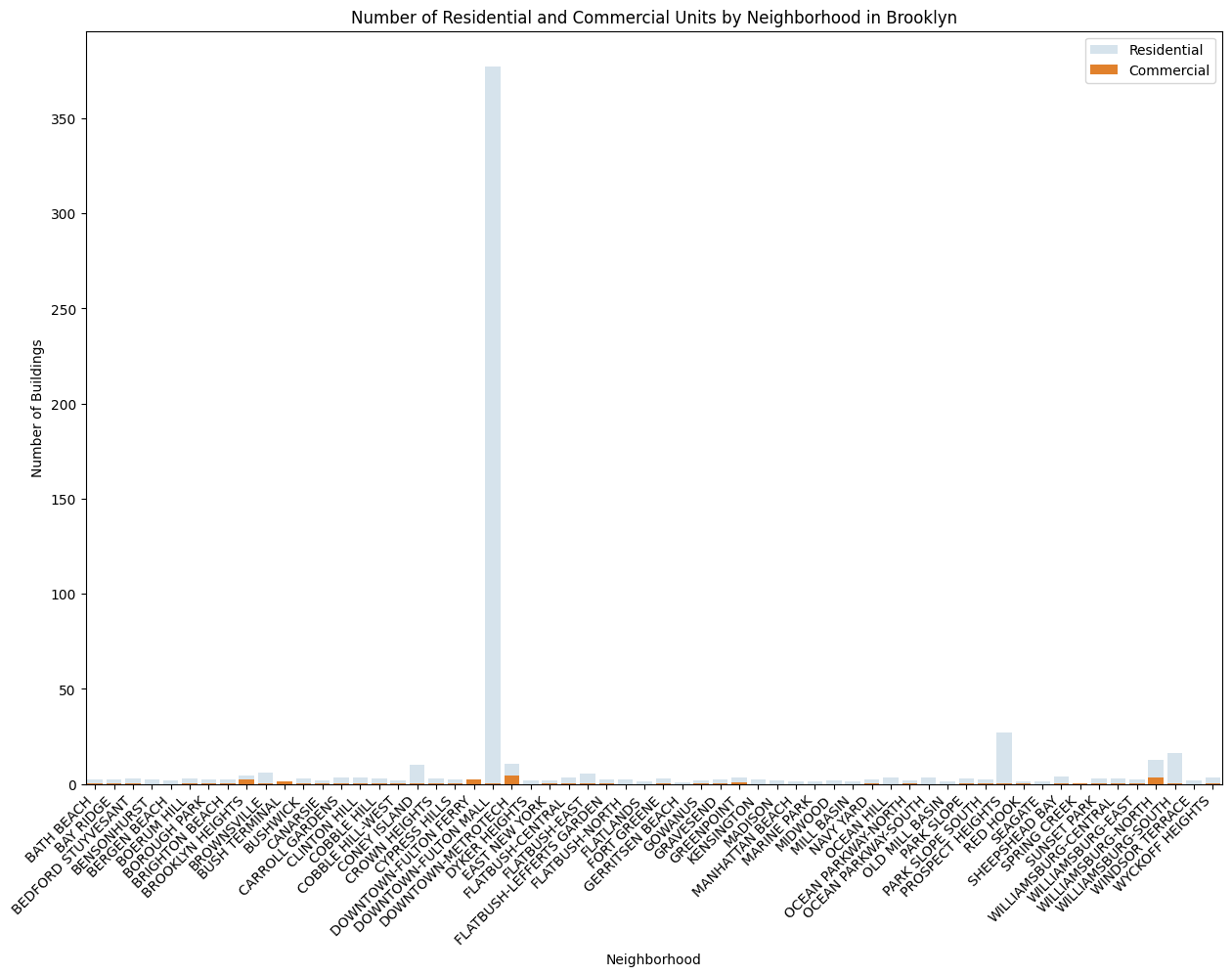


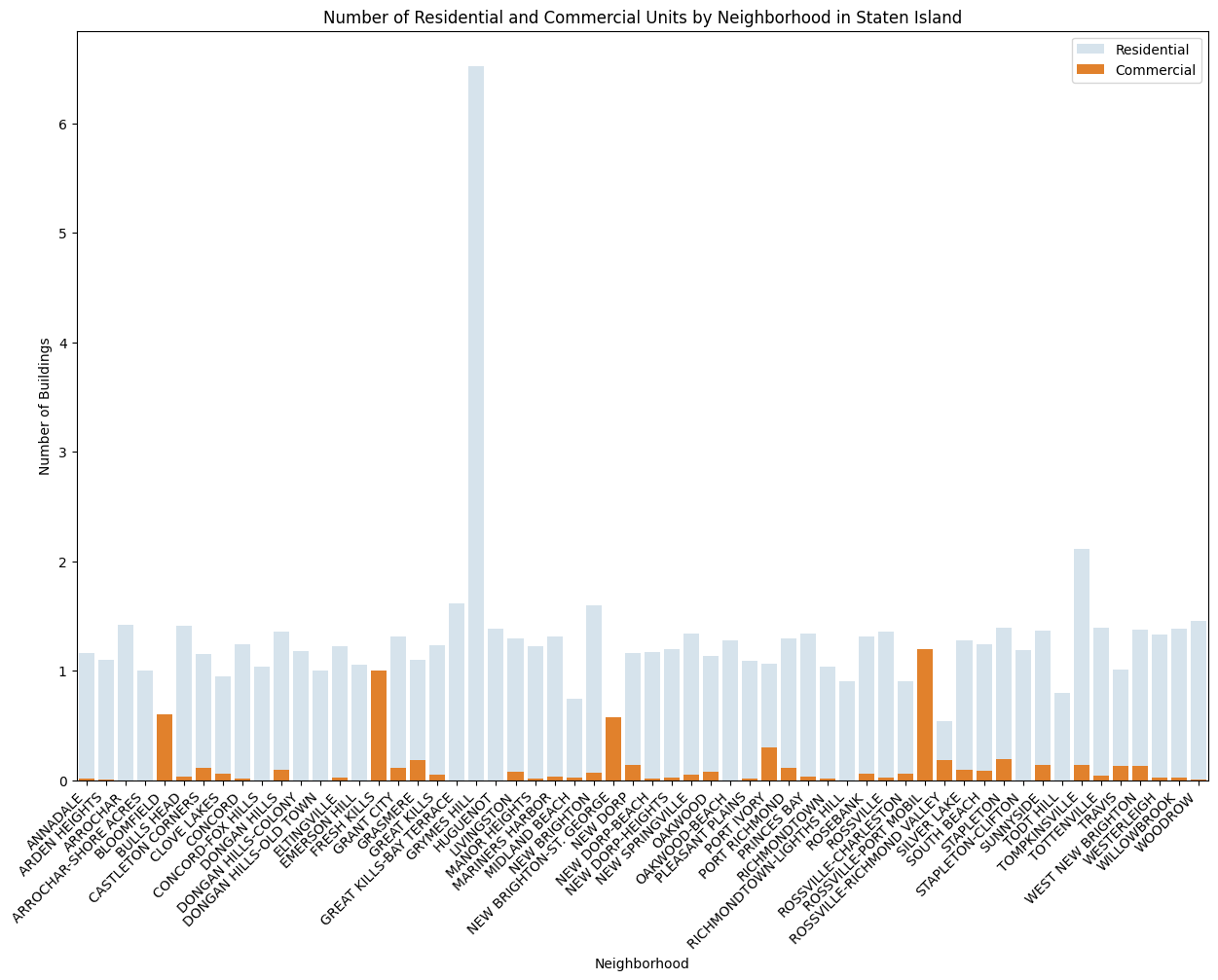
Згідно графіків, можемо сказати що у всіх районах, окрім Манхетена, переважає будування житла, а у Манхетені упродовж даного проміжку часу показник будування житла чергується з будування комерції, щоправда у період з 1995 по 2005 будування комерційних будівель значно випередило будівництво житла.

Дослідимо к-сть комерційних та житлових споруд по мікрорайонах



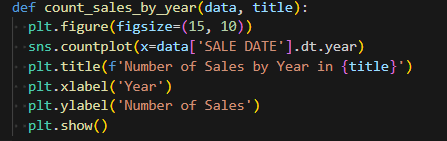


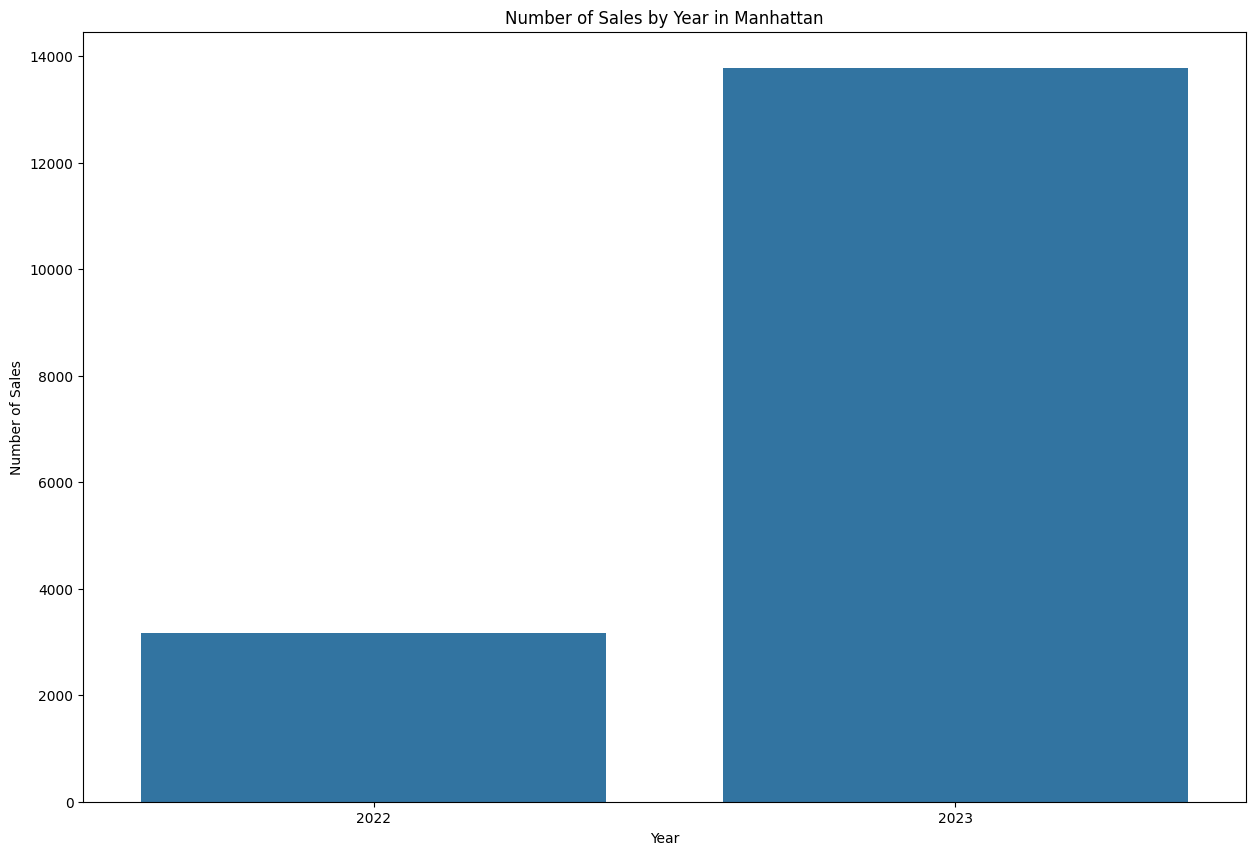


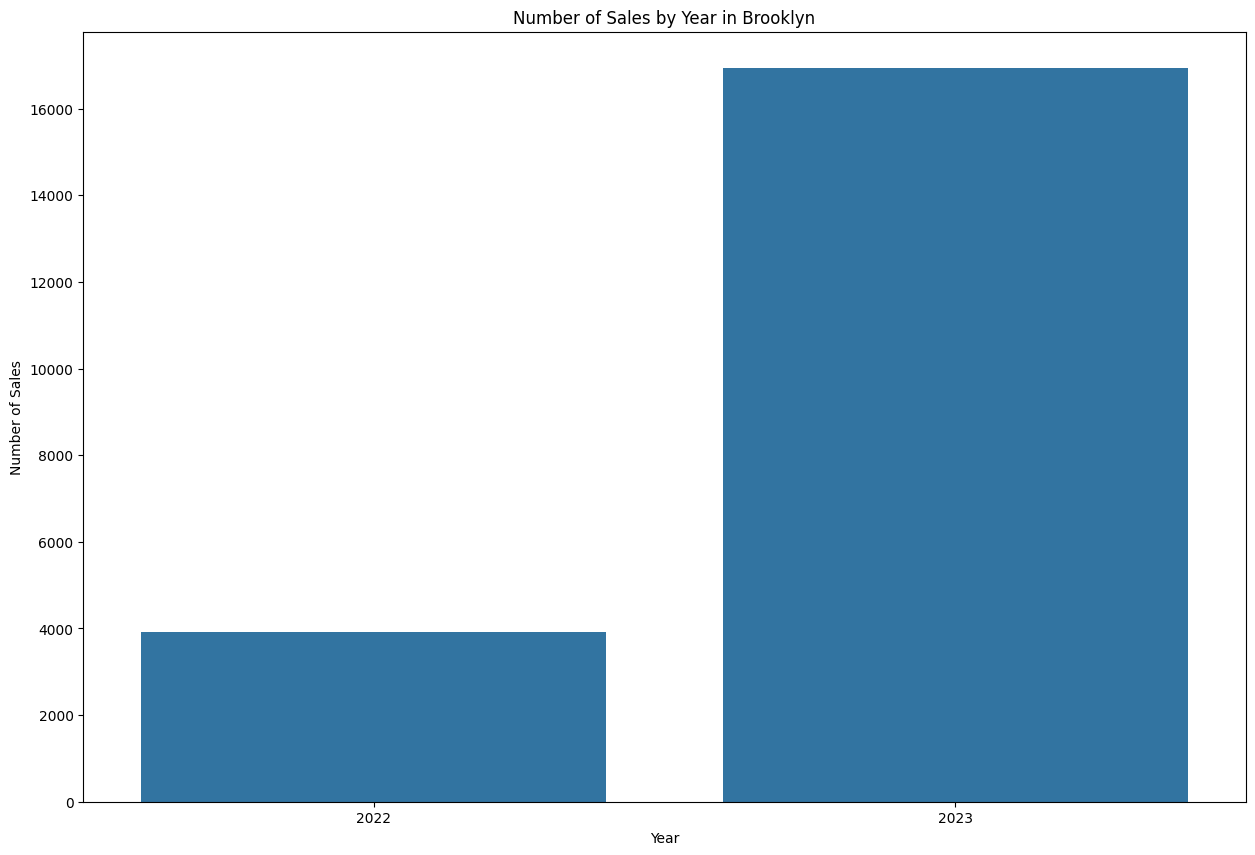


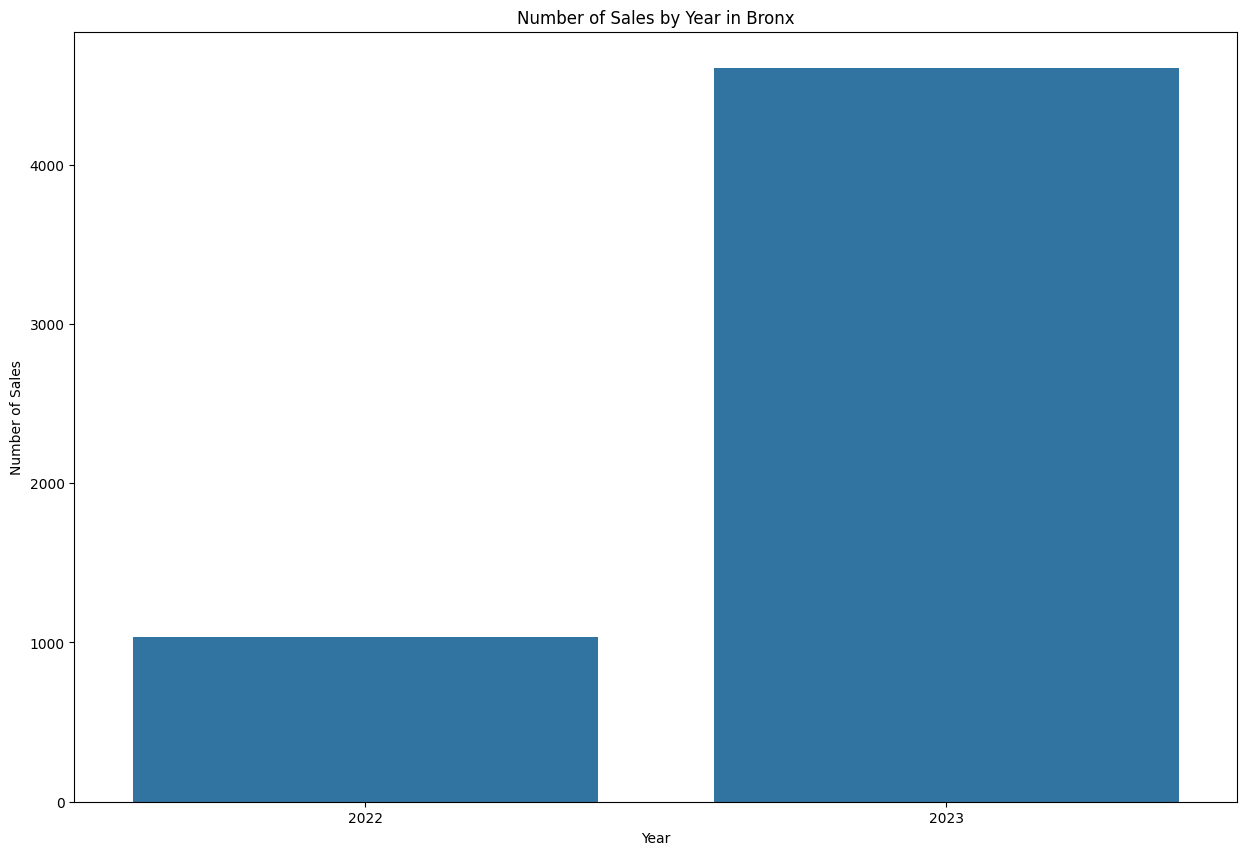
Бачимо, що у всіх районах, переважають житлові будівлі, проте у Манхетені виділяється мікро Javits Center, де переважають комерційні будівлі.

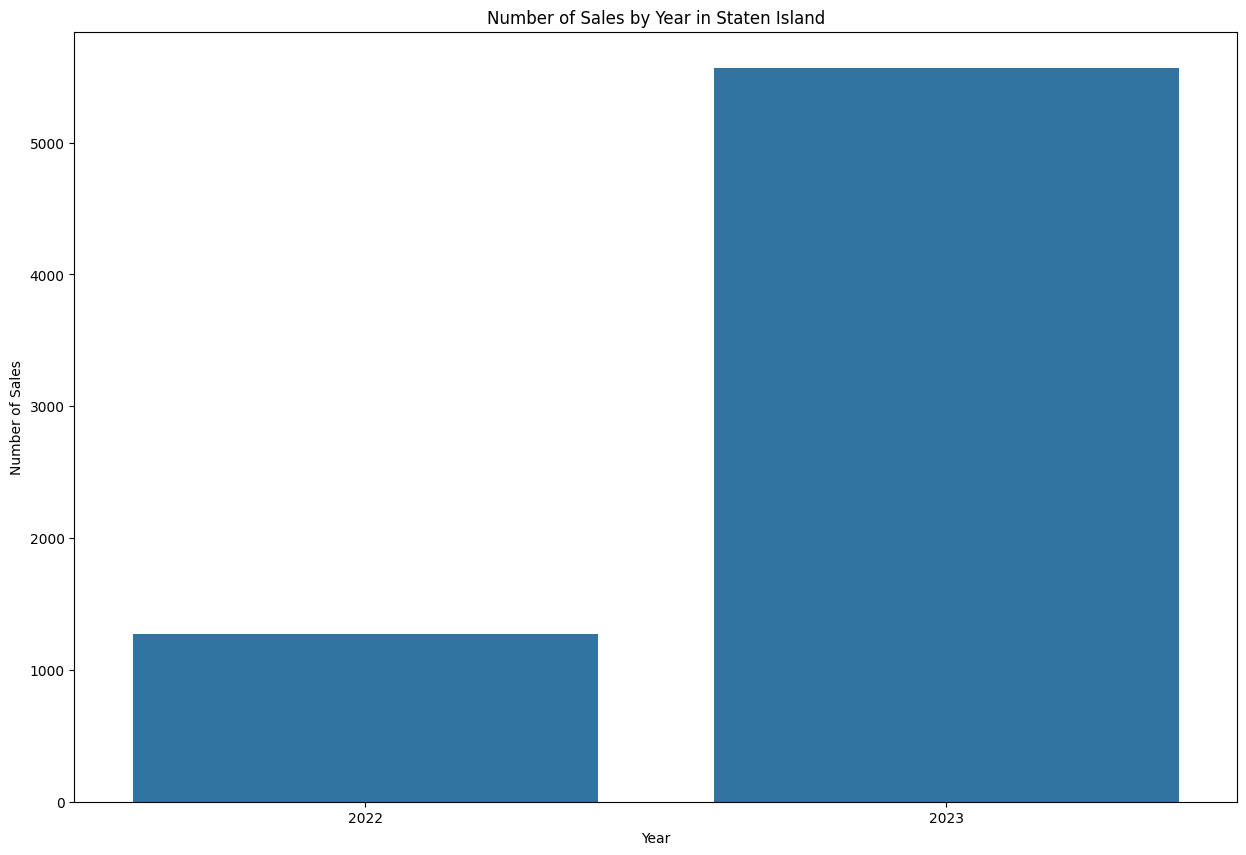
Відобразимо к-сть продажів будівель по мікрорайонам





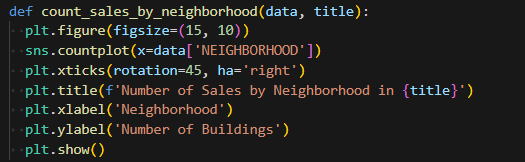


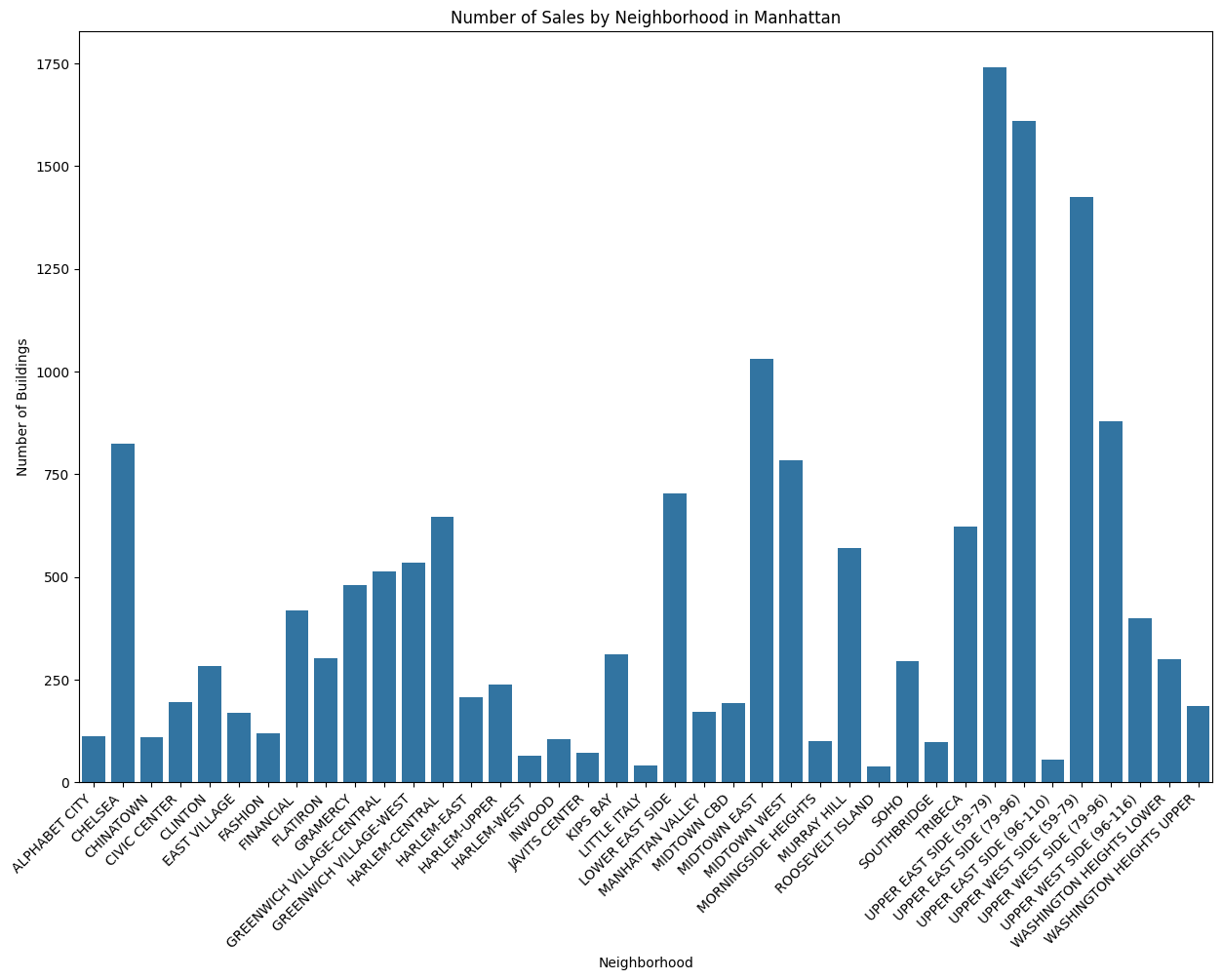


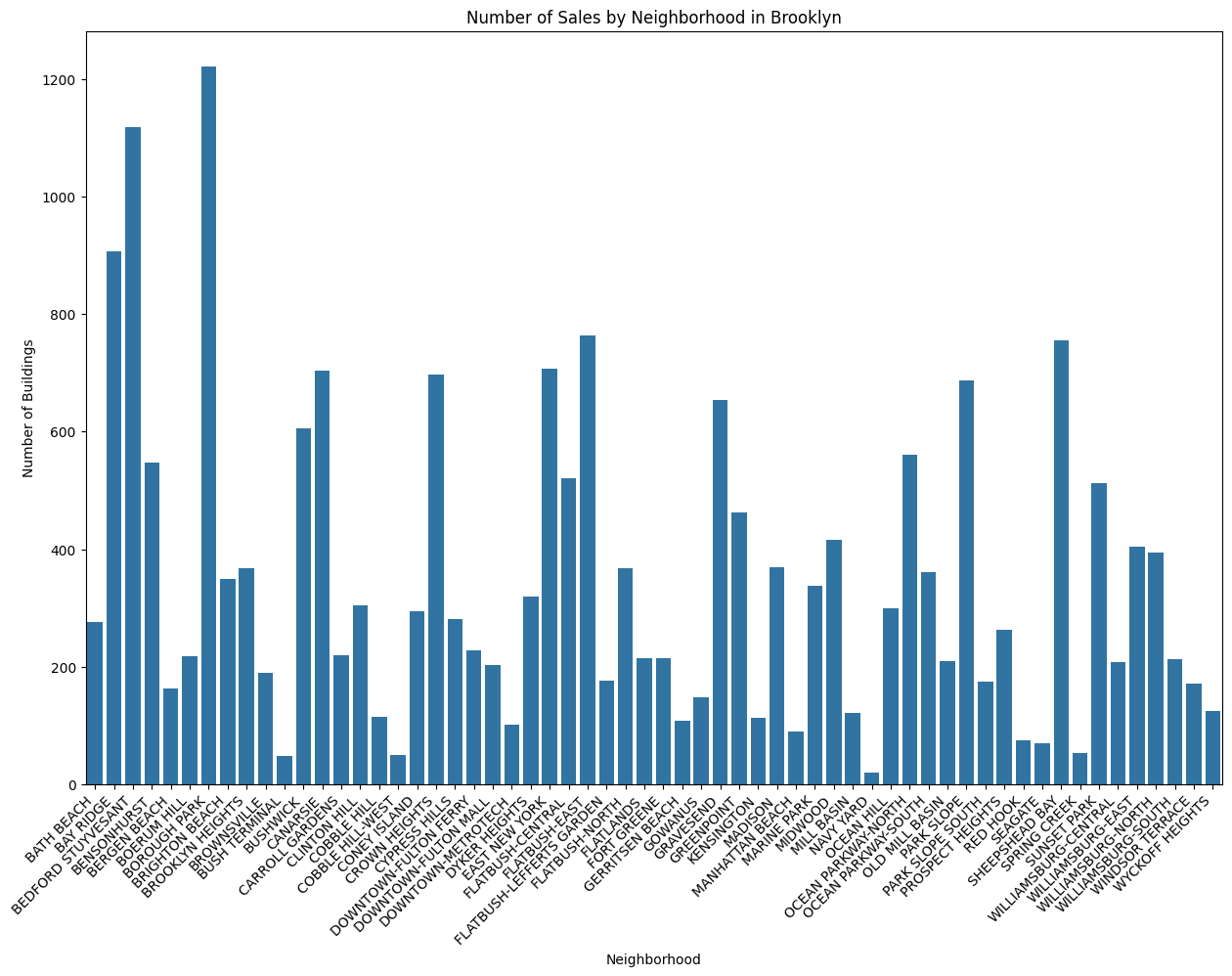


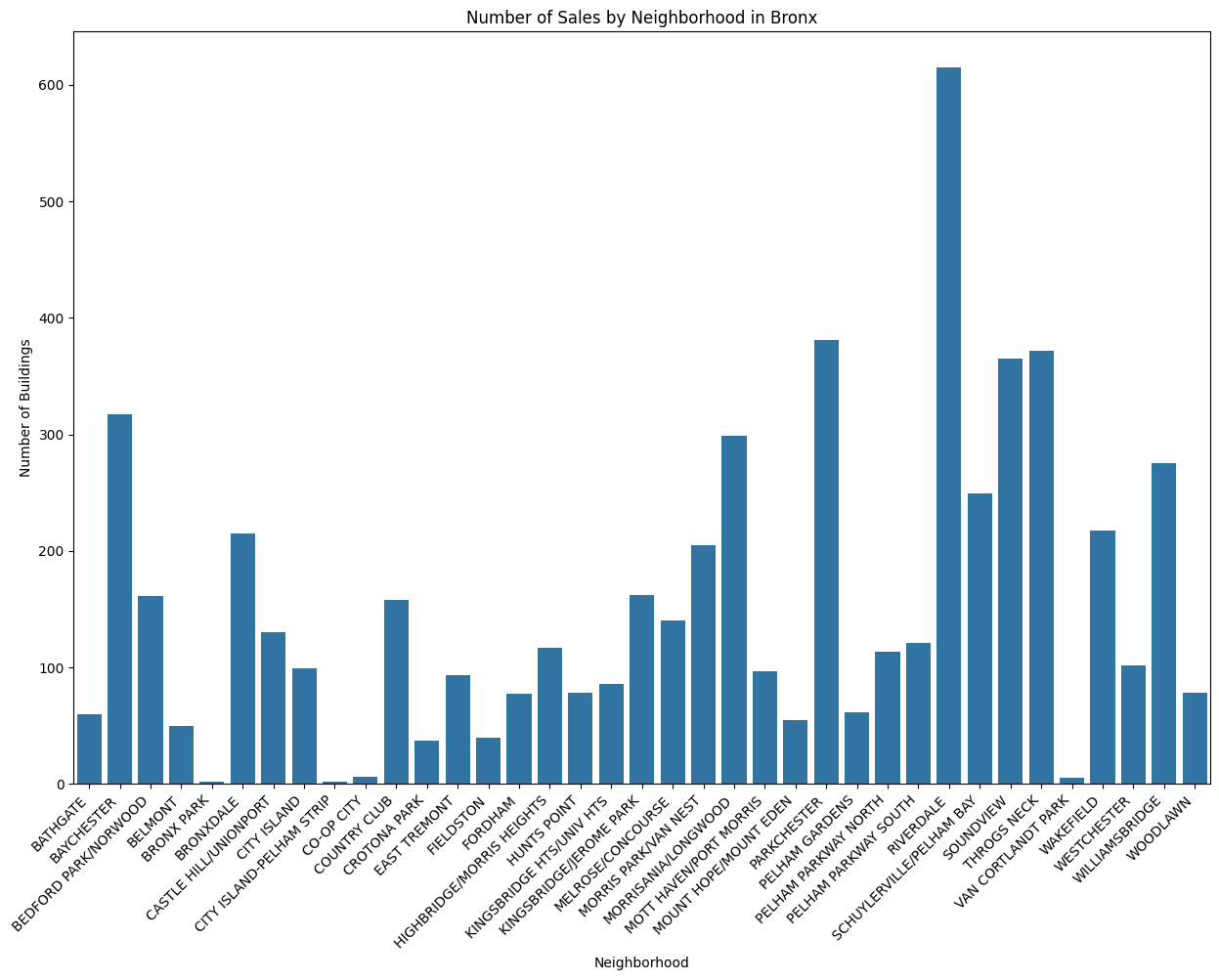
Бачимо, що продажі у 2023 році виросли у всіх районах, це пов’язано з реабіліацією світової економіки від кризи та з підвищення платоспроможності як компаній, так і людей.

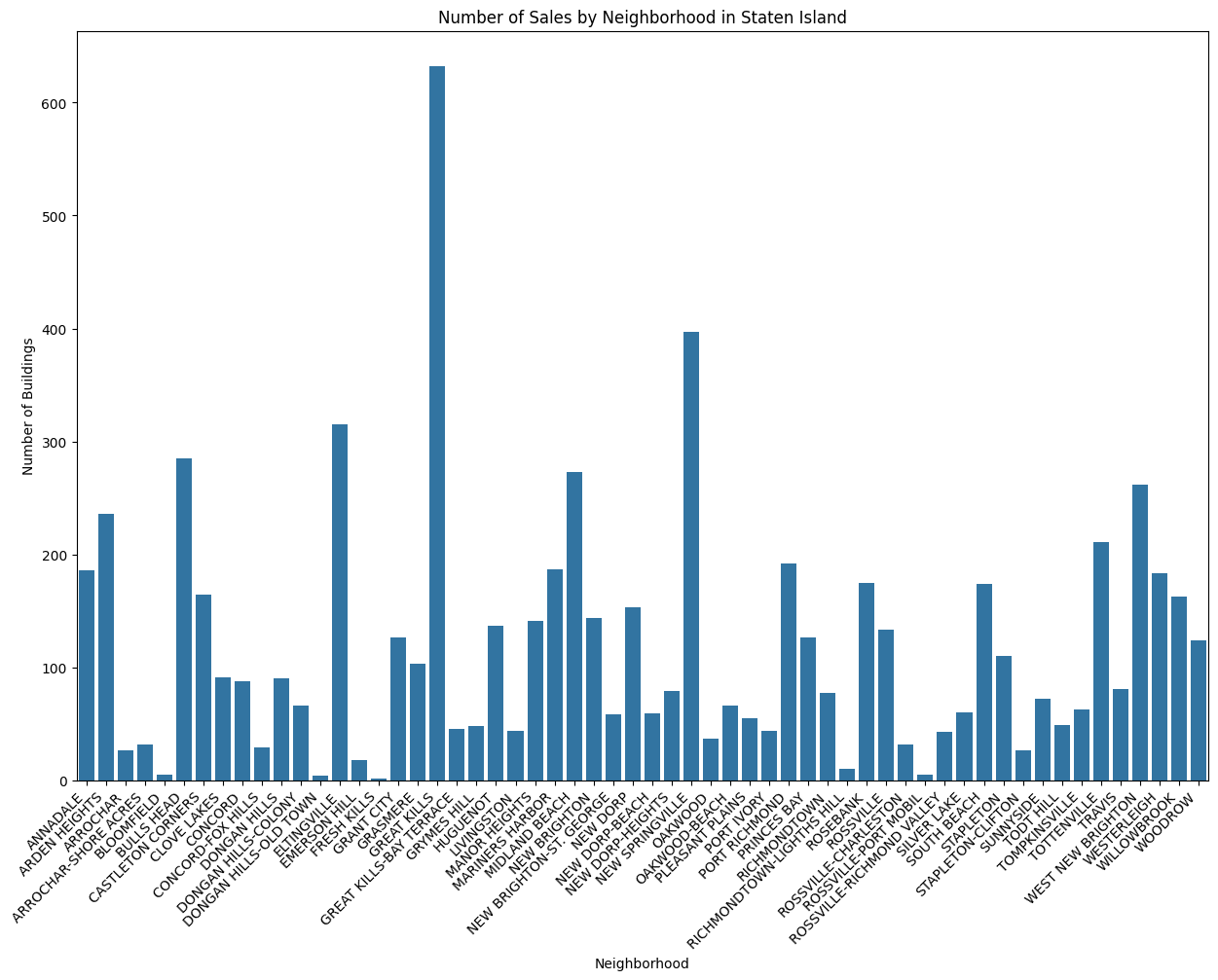
Дослідимо продажі будівель за мікрорайонами











По мікрорайонах за найбільшою к-стю проданих будівель виділяються Upper East Side(79-96), Upper East Side(96-110), Upper West Side(59-96), Upper West Side(79-96) в Манхетені; Borough Park, Bedford Stuyvesant, Bay Ridge – в Брукліні; Riverdale - в Бронксі; Great Kills, New Springvill – в Стейтен Айленді.

**Висновок:**

Географічний аспект та його вплив на ціни:

З результатів аналізу видно, що вартість нерухомості значно залежить від географічного розташування об'єктів. Конкретні мікрорайони, такі як Morningside Heights на Манхеттені, Gowanus в Брукліні, Mott Haven/Port Morris в Бронксі та Rossville-Richmond Valley на Staten Island, відзначаються цінами, що перевищують середні значення.

Важливість архітектурної спадщини:

Вивчення відношення ціни до року будівництва підкреслює, що історичні періоди суттєво впливають на вартість нерухомості. Об'єкти з архітектурною спадщиною дуже цінуються, і це важливо враховувати при розробці стратегій.

Динаміка будівництва та ринку:

Зміна кількості комерційних та житлових будівель з часом свідчить про гнучкість та адаптивність ринку до змін. Розуміння цієї динаміки є ключовим при формуванні інвестиційних та розвиткових стратегій.

Середня вартість будівлі та попит:

Аналіз зміни середньої вартості будівлі з роками вказує на періоди активності на ринку. З урахуванням цих піків, ми можемо адаптувати стратегії для досягнення оптимальних результатів.

Відновлення попиту в 2023 році:

Позитивний тренд у продажах у всіх районах у 2023 році свідчить про відновлення економіки та підвищення платоспроможності. Це відкриває нові можливості для наших інвестицій та проектів.

Отже, ринок нерухомості у Нью-Йорку виявляється різноманітним та динамічним. На основі отриманих даних рекомендую взяти до уваги врахування географічних, архітектурних та економічних факторів при розробці наших подальших стратегій. Це дозволить нам ефективно реагувати на зміни на ринку та максимізувати наш потенціал у цій конкурентній сфері.