**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**ІМЕНІ ТАРАСА ШЕВЧЕНКА**

**ФАКУЛЬТЕТ ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ**

**Кафедра прикладних інформаційних систем**

ЗВІТ

Про виконання лабораторної роботи №2

з дисципліни: «Data Science та Big Data»

Студентки групи ПП-42

Шкаленко Ганни

Київ – 2022

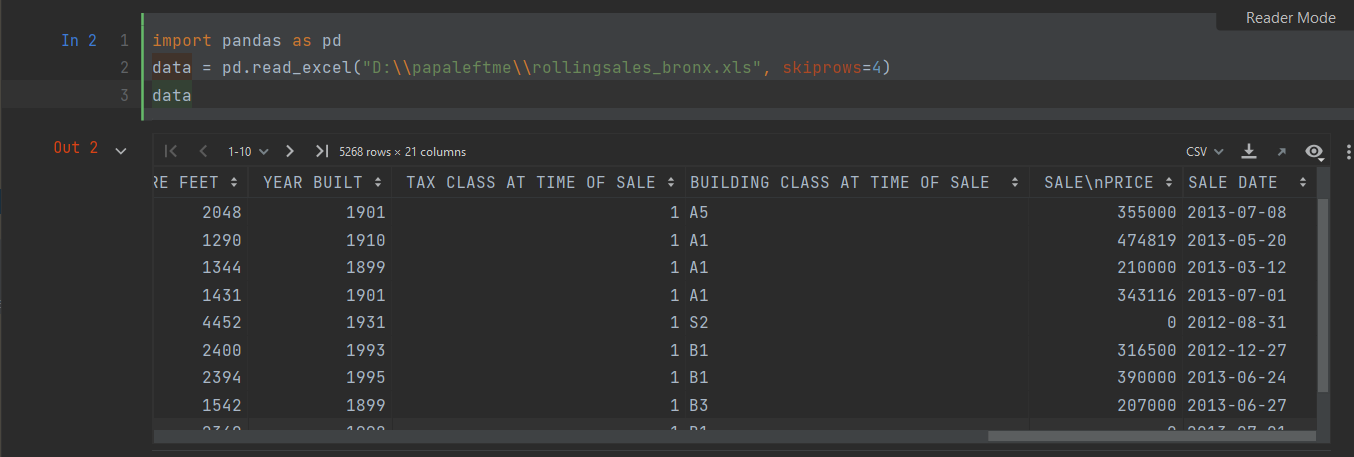
**Лабораторна робота 2.**

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

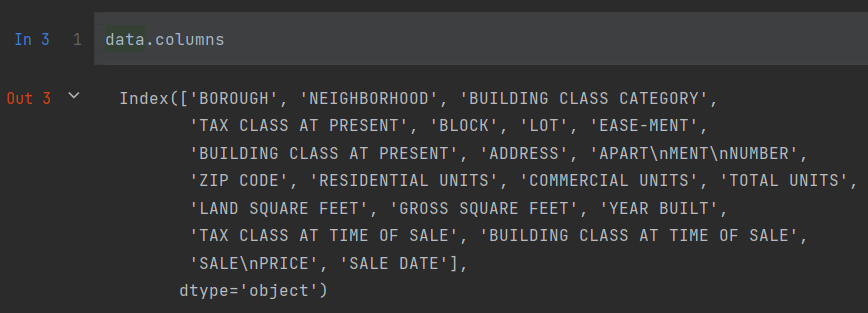
*Мета:* Метою лабораторної роботи є отримання практичних навичок виконання розвідувального аналізу даних, використовуючи пакети jupyter, pandas, seaborn. Ознайомлення з методологією складання аналітичного звіту для зовнішнього користувача інформаційного продукту.

Порядок виконання

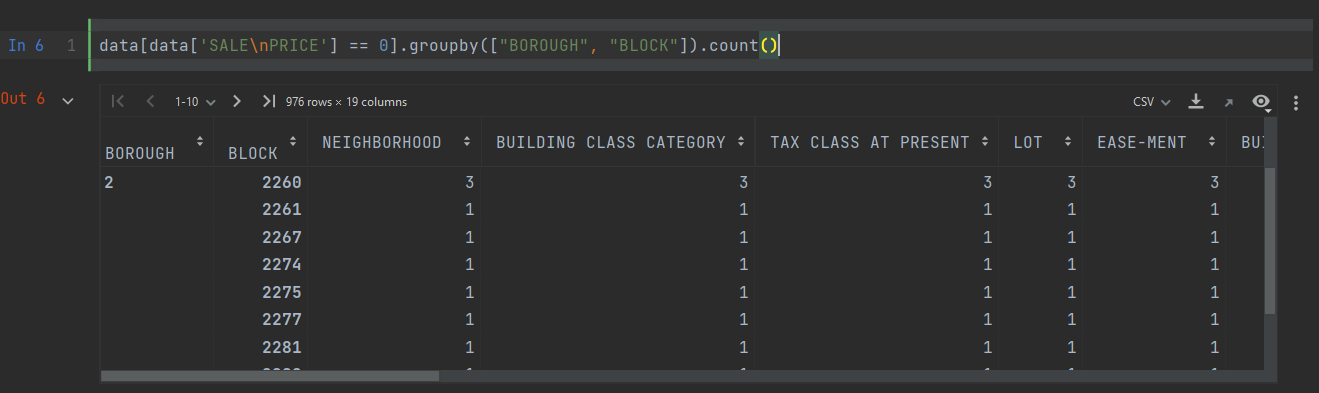
Для виконання роботи спочатку занесемо всі данні в змінну data.



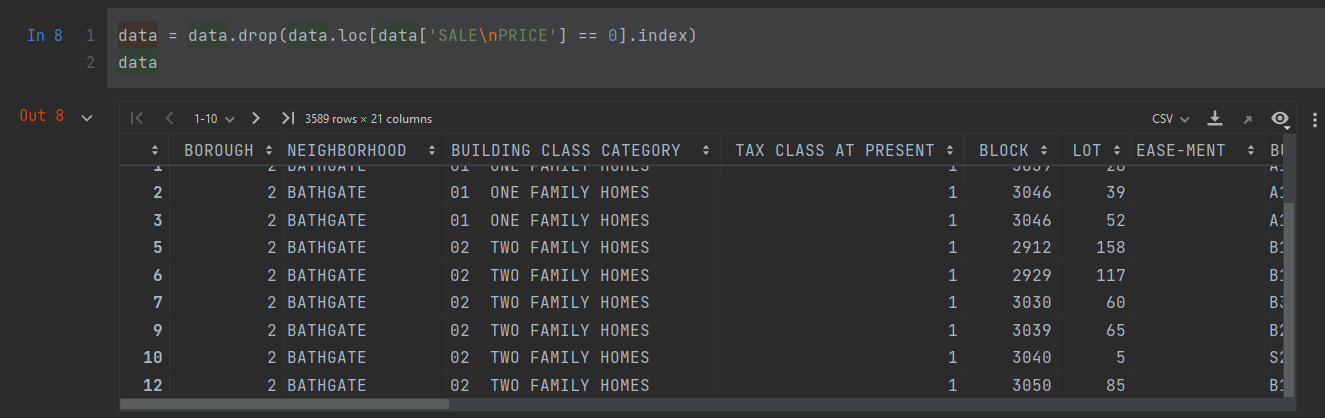
Далі виведемо всі доступні назви колонок – термінологія, що використовується для опису закономірностей ринку нерухомості.



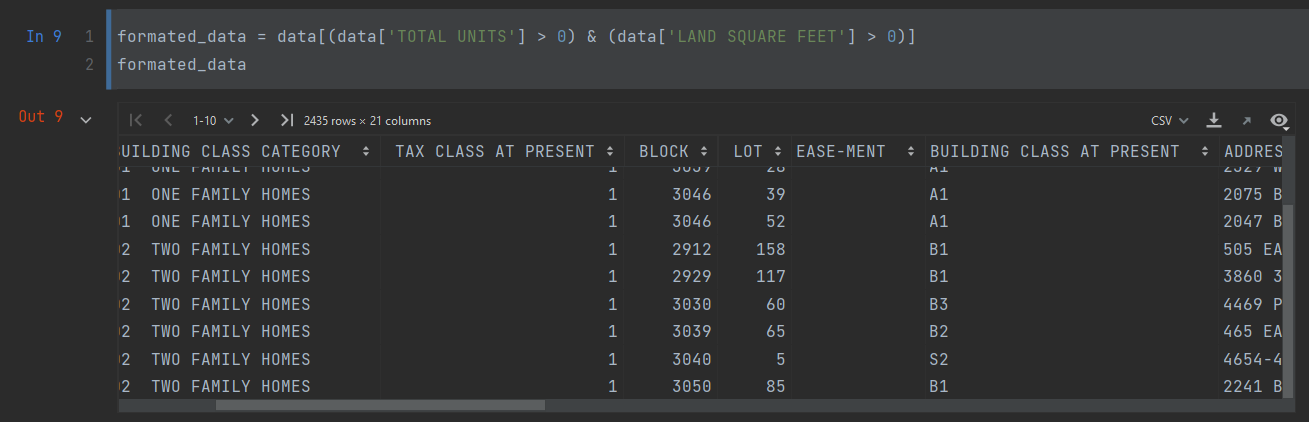
Далі виведемо інформацію про лоти, в яких ціна продажі нульова та обрахуємо їх кількість. При цьому будемо використовувати групування за місцем розташування та номером блоку.



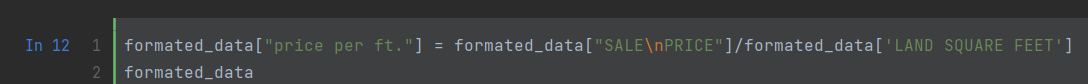
Після цього відредагуємо дату так, щоб виключити лоти з нульовою ціною.

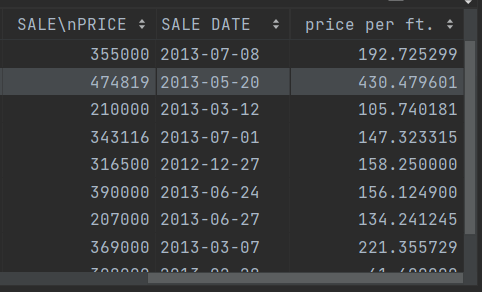


Таким же чином виключаємо всі квартири з нульовим розміром.



Наступним кроком визначимо ціну за квадратний метр та додаємо ці данні в таблицю.

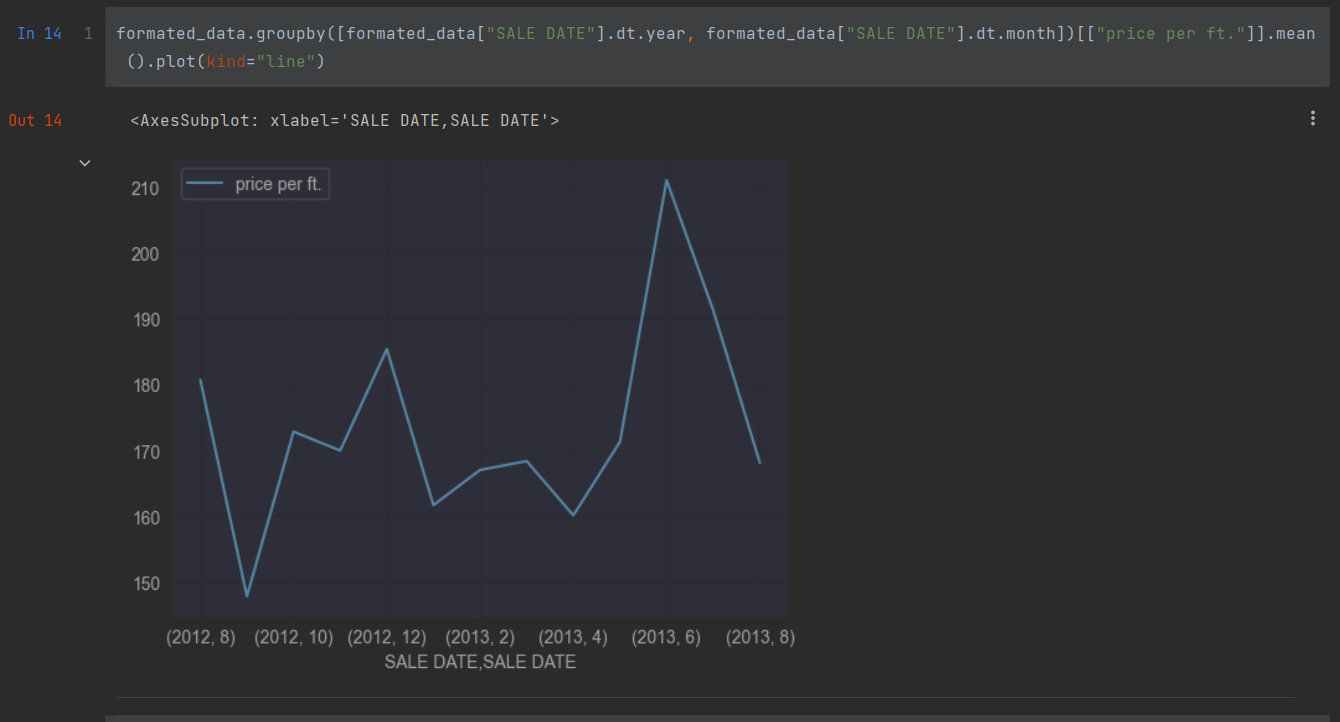




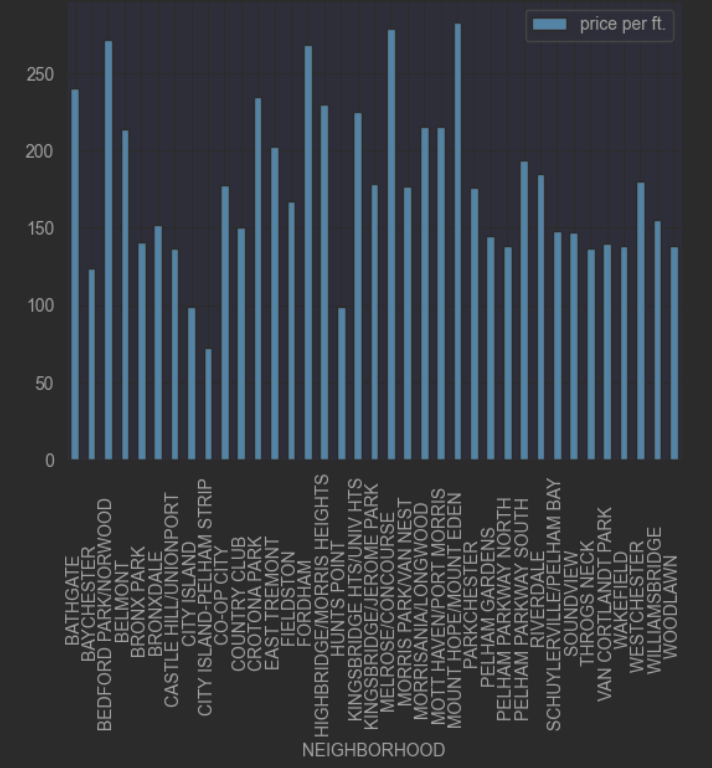
На цьому етапі можна розглянути деякі графіки. На першому з них зображене відношення між періодом, за який були продані квартири, та виручена ціна за кожен з місяців. Таким чином можна побачити, що найбільша сума була виручена в грудні 2012 року, а найменша – в наступному місяці, січні 2013 року.



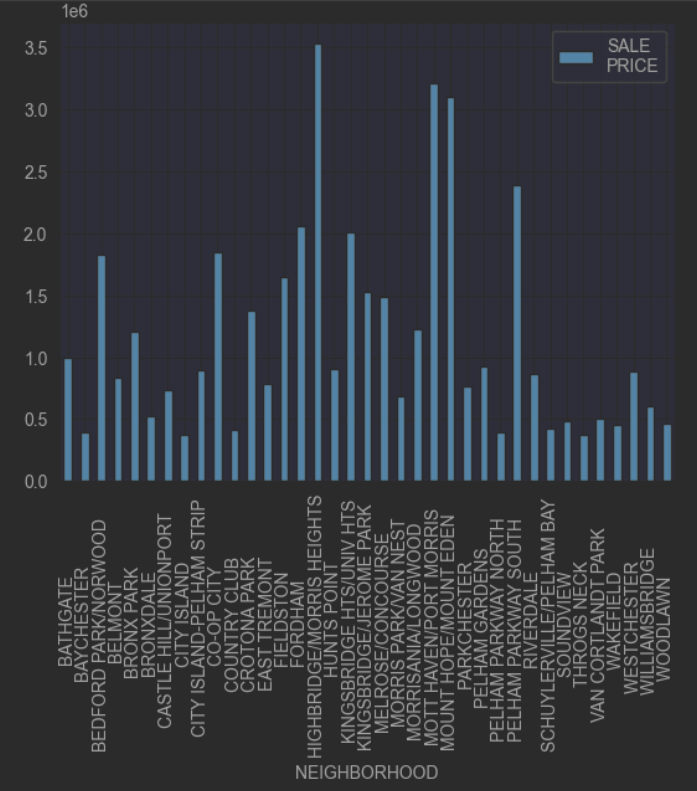
Далі виведемо графік зміни ціни за метр проданих квартир за даний нам рік. Можна побачити, що пік був у червні 2013, а найвигідніше квартиру можна було придбати у липні 2012.



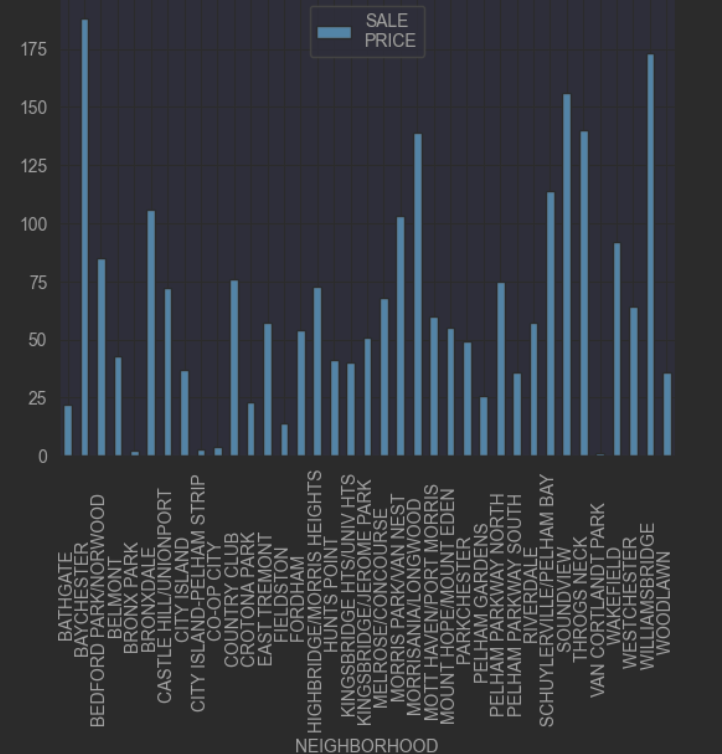
Наступний графік відображає залежність ціни за квадратний метр квартир у розрізі інфраструктури та ландшафту. Можна дійти висновку, що найменш рентабельні квартири знаходяться біля City Island-pelham strip. Але при цьому немає явного лідера або середини серед районів.



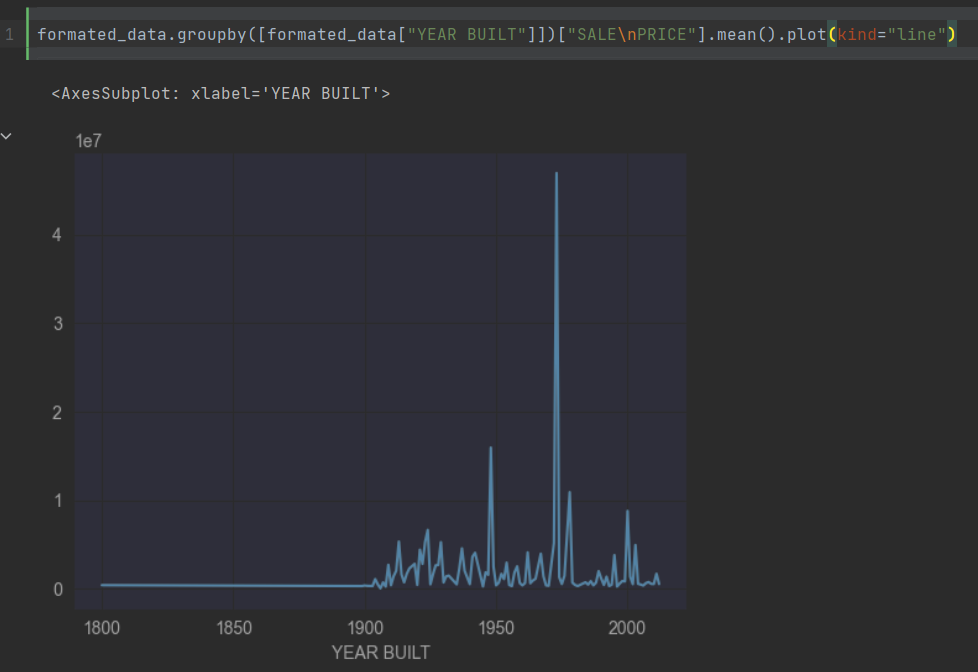
Наступний графік відображає залежність загальної ціни квартир у розрізі інфраструктури та ландшафту. Можна дійти висновку, що найбільш рентабельні квартири знаходяться біля Highbridge, Mott Haven, Mount Hope. При цьому всі інші райони, згідно цього графіку, виглядають більш менш посередньо.



На даному графіку представлено кількість проданих квартир в кожному районі за рік. Згідно з графіком, найпопулярнішими є район Baychester. На другому місці знаходиться WilliamsBridge.



Наступний графік відображає ціну квартири в залежності від року побудови будівлі. Квартири побудовані в 1975 році коштують найдорожче.



Наступний графік відображає кількість проданих квартир, дорожчих ніж півмільйона, що були продані за рік. В грудні 2012 їх було майже 90, а в серпні 2013 всього декілька.

