**КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ**

**імені ТАРАСА ШЕВЧЕНКА**

**Факультет інформаційних технологій**

**Кафедра прикладних інформаційних систем**

****

**Звіт**

до виконання лабораторної роботи № 2

з дисципліни **«**Data Science та Big Data**»**

на тему:

«Розвідувальний аналіз даних (EDA). Складання аналітичного звіту»

**Виконано:**

студ. групи ПП-42, підгрупа 1

Дикий Максим

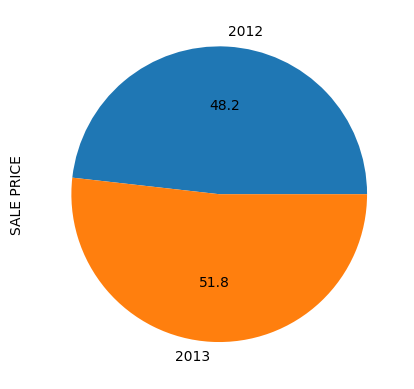
**Київ – 2022**

**Лабораторна робота №2.**

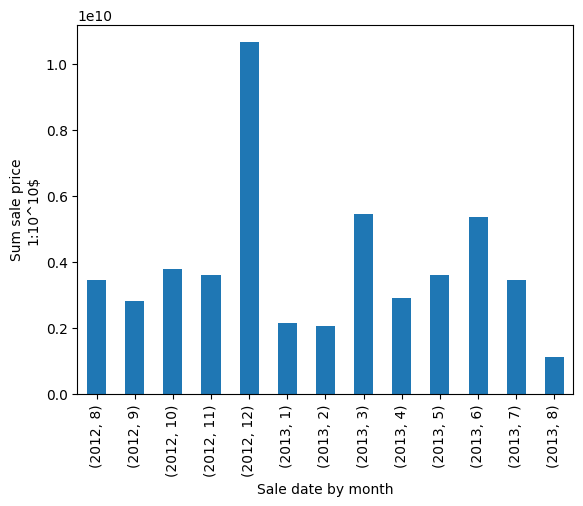
**Завдання для лабораторної роботи**

* 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:
* Як топові компанії на ринку складають звіти по нерухомості?
* Які графіки використовуються для донесення інформації?
* Які співвідношення між якими даними по ринку є показовими для інвесторів / керівників агенцій нерухомості?
* Яка термінологія використовується для опису закономірностей на ринку нерухомості?
* Завантажити файли з даними у папку проекту з посилання:
* https://www1.nyc.gov/site/finance/taxes/property-rolling-sales-data.page
* Очистити дані.
* Виконайте розвідувальний аналіз, щоб дізнатися, де є викиди або відсутні значення, вирішіть, як ви їх будете обробляти, переконайтеся, що дати відформатовані правильно, значення, які ви вважаєте числовими, розглядаються як такі і т.д.
* Виконайте аналіз розвідувальних даних (отриманих результатів) для візуалізації та зіставлення за житловими масивами та за часом. Почніть шукати осмислені закономірності у цьому наборі.
* Зберіть висновки у невеликий звіт для генерального директора (графіки, висновки з текстом у окремому файлі), який потребує належного оформлення висновків, структури тощо.
* Завантажити звіт та файл ipynb з виконаними завданнями на git в окрему папку з відповідною назвою лабораторної роботи

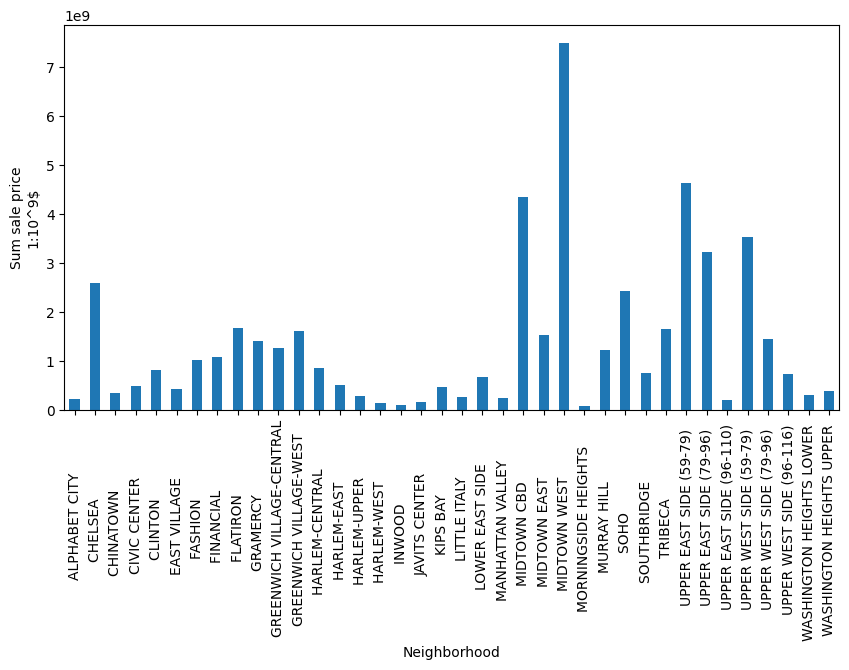
**Хід роботи:**



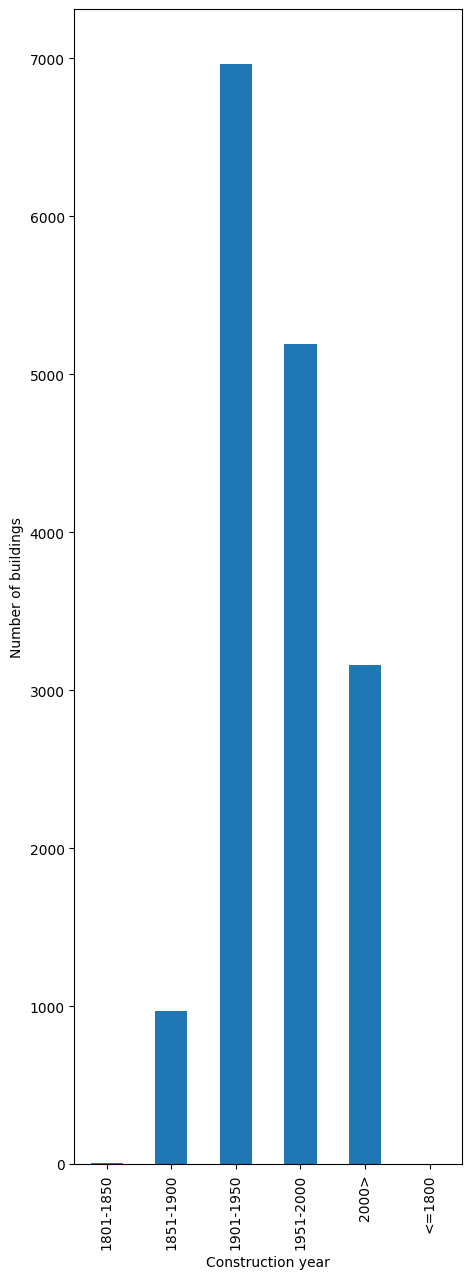
Відношення загальної суми продажів за 2012 та 2013 роки



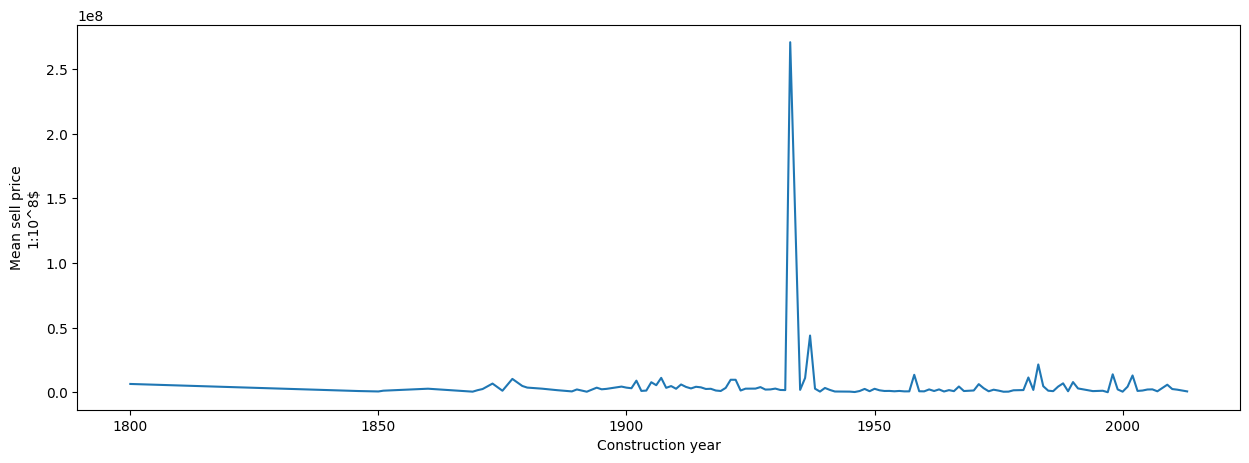
Продажі за кожен місяць кожного року, наданого в документі



Сума продажів за районами



Кількість будівель побудованих у певний проміжок часу



Середня ціна будівель об’єднаних за роками

**Висновки:**

Відносно 2012 року, дохід у 2013 році зріс, однак не дуже сильно. Найбільше продажів було у грудні 2012 року. Загалом, найбільш популярними районами є – Midtown West, Upper East Side (59-79) та Midtown CBD. Більше всього будівель було сконструйовано між 1901 та 1950, спостерігається поступовий спад кількості нових будівель. Також, можна побачити, що найбільш коштовна нерухомість була збудована у проміжку між 1930-1950 років.