

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

“КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ

імені ІГОРЯ СІКОРСЬКОГО”

Факультет прикладної математики

Кафедра програмного забезпечення комп’ютерних систем

**Лабораторна робота №** **1**

з дисципліни “Технології оброблення великих даних”

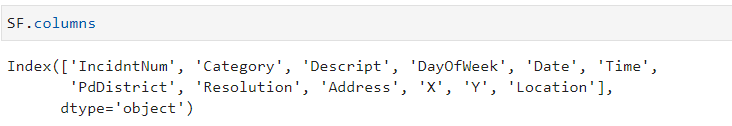
на тему: «Аналіз та візуалізація даних у Python»

|  |  |  |  |
| --- | --- | --- | --- |
| Виконав  студент III курсу  групи КП-93  Інюшев А.В. |  |  | Зарахована  “\_\_\_\_” “\_\_\_\_\_\_\_\_\_\_\_\_” 20\_\_\_ р.  викладачем  Олещенко Л. М. |

Київ 2022

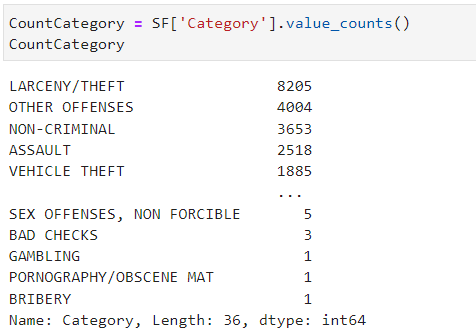
**Відповіді на питання:**

**1. Скільки змінних міститься у фреймі даних SF?**

****

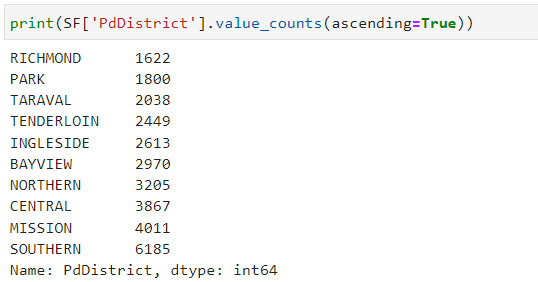
У нашому датасеті міститься 12 змінних ('IncidntNum', 'Category', 'Descript', 'DayOfWeek', 'Date', 'Time', 'PdDistrict', 'Resolution', 'Address', 'X', 'Y', 'Location')

**2. Якого виду злочину було скоєно найбільше?**

****

Найбільше було скоєно крадіжок(LARCENY/THEFT), а саме 8205 випадки подібних злочинів було зафіксовано в Сан-Франциско

**3. В якому PdDistrict було найбільше випадків зареєстрованих злочинів? Надайте команди Python, які використовуються для підтримки вашої відповіді.**

****

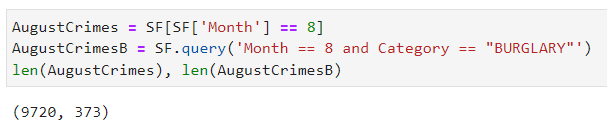
Найбільше випадків злочинів було виявлено у районі SOUTHERN, а саме 6185 випадків.

Можна отримати район з найбільшою кількістю злочинів можна декількома способами.

print(max(SF['PdDistrict'].value\_counts(ascending=True))) – за допомогою функції max()

print(SF['PdDistrict'].value\_counts()[0]) – вибрати значення по спаданню та взяти перше значення.

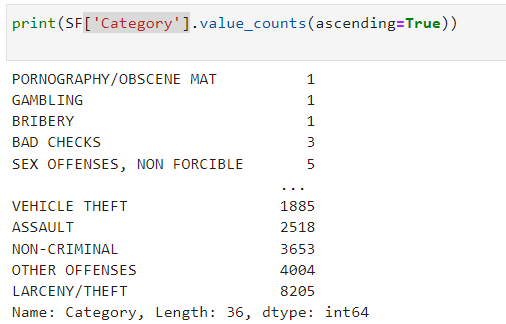
**4. Скільки випадків злочинів було за серпень? Скільки квартирних крадіжок було зареєстровано у серпні?**



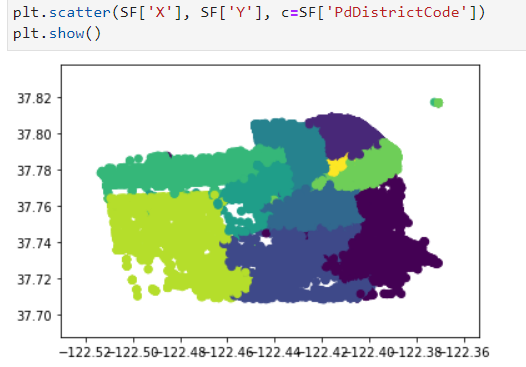
За серпень було скоєно 9720 злочинів, з них 373 квартирних крадіжок

**Інші скріншоти:**

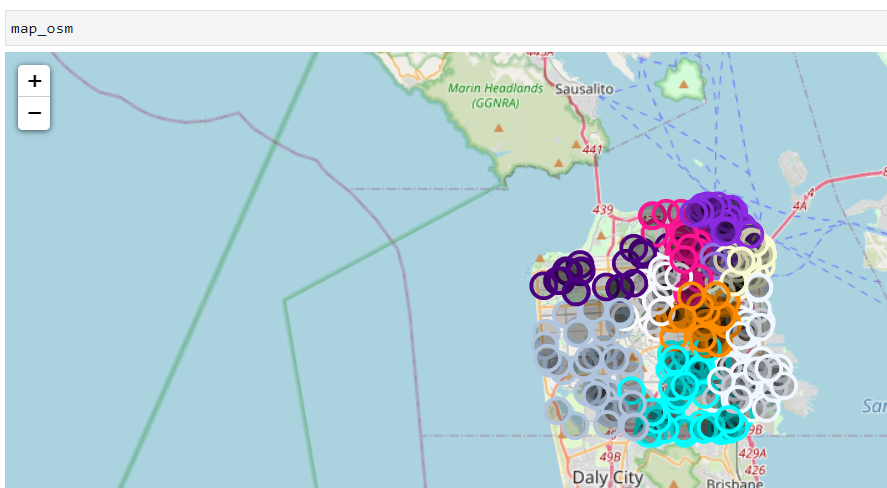
**Злочини по категоріях**

****

**Злочини по районам та координатам**

****

**Карта злочинів**

****