**Homework 4**

**Знайдіть набір даних для фінального проекту**

**Лебяк Маркіян**

1. **Посилання на набір даних**

“Інформація про створені електронні медичні записи про вакцинацію населення від Корона вірусної хвороби”:

<https://data.gov.ua/en/dataset/3fcbfe9e-7cec-4f69-b9ca-be49daae2369>

1. **Короткий опис того, чому ви вважаєте цей набір даних цікавим, і що саме ви плануєте у ньому дослідити / візуалізувати**

Даний набір даних складається з 4 файлів, а саме:

* immunization\_covid19\_2021\_16.11.2021.csv (саме дані про вакцинованих людей; .csv)
* immunization\_legal\_entities\_info.csv (інформацію про заклади, що вносили вакцинацію; .csv)
* immunization\_vaccine\_codes.csv (інформація про вакцини; .csv)
* opis-atributiv-naboru-danikh.xlsx (таблиця .xlsx, яка описує вміст усіх трьох .csv файлів)

Мене цікавлять два перших .csv файли які містять такі дані:

Table

Description automatically generated

Table

Description automatically generated

Дані на мою думку є досить актуальними для аналізу, що підтверджує вибір даного датасету. Перший файл містить більш ніж 20млн записів, що при візуалізації дасть досить повну картину.

Серед можливих візуалізацій:

* розподіл вакцинації по віковій категорії
* розподіл вакцинації по регіонам України (карта)
* розподіл по етапам вакцинації
* статистика проведення вакцинації по дням (зацікавленість людей у вакцинації)