Група ФБз-01

Міщенко О.М.

**Лабораторна робота №1**

**Наука про дані: підготовчий етап**

**Мета роботи**: ознайомитися з основними кроками по роботі з даними – workflow від постановки задачі до написання пояснювальної записки, зрозуміти постановку задачі та природу даних, над якими виконується аналітичні операції.

Постановка задачі ν Проаналізувати часові ряди глобальних продуктів по оцінці вегетаційного здоров’я VHI (vegetation health index) http://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/index.php, який надається Національною адміністрацією океанів та атмосфери США NOAA (http://www.noaa.gov/); ν Виявити особливості ходу індексу впродовж вегетаційного періоду (вересень року-попередника – липень поточного року) в розрізі областей України.

Семантика задачі **VHI** – вегетаційний індекс, який базується на відбитті видимого світла рослинним покривом, яке характеризує ступінь здоров’я рослинності. Цей індекс базується на поєднанні індексу VCI (Vegetation Condition Index), який описує ступінь пригніченості рослинного покриву та індексі 7 температурного режиму TCI (Temperature Condition Index), які були запропоновані в 1995 році Ф. Коганом.

VHI = 0.5\*VCI + 0.5\*TCI , (1.1)

Зміст VHI-індексу:

VHI 60 – сприятливі умови;

VHI < 15 – посуха, інтенсивність якої від середньої до надзвичайної;

VHI < 35 – посуха, інтенсивність якої від помірної до надзвичайної.

Хід виконання роботи:

1. Мова програмування: Python3.9. **province.csv** містить назви адміністративних одиниць та індекси, відповідні до індексів на сайті <http://www.star.nesdis.noaa.gov/smcd/emb/vci/VH/index.php> .Створено env в якому встановлені всі необхідні бібліотеки та налаштування для данної лабораторної роботи.

Встановлено : Pandas, Urllib.requests, datetime, numpy.

1. Наступні дії реалізовано в **Laba1.py:**

- для кожної із адміністративних одиниць України завантажити тестові структуровані файли, що містять значення VHI-індексу .Ця процедура автоматизована, параметром процедури є індекс (номер) області. При зберіганні файлу до його імені додається дата та час завантаження;

- зчитуються завантаженні текстові файли у фрейм;

- реалізована процедура, що змінює індекси областей, які використані на порталі NOAA відповідно до вимоги лабораторної роботи;

1. Наступні дії реалізовано в **Laba1\_1.py:**

* Реалізувано процедури для формування вибірок наступного виду (включаючи елементи аналізу): Ряд VHI для області за рік, пошук екстремумів (min та max);

1. Наступні дії реалізовано в **Laba1\_2.py:**

Ряд VHI за всі роки для області, виявити роки з екстремальними посухами, аналогічно для помірних посух.