**Дослідницький проект з курсу «Статистика»**

Дослідницький проект виконується у командах до 5 студентів.

**Опис завдання:**

1. Завантажити дані (таблицю із заданим набором даних) до середовища програмування Python.
2. Провести первинний аналіз структури і якості даних.
3. Здійснити загальний опис даних, що містяться у таблиці: яке явище або процес відображають дані; які є змінні у таблиці; кількість спостережень; наявність пропущених або потенційно некоректних спостережень; формат даних для кожної змінної.
4. У разі необхідності здійснити очищення й трансформацію вхідних даних, підготувавши їх до статистичного аналізу.
5. Описова статистика: представити основні параметри, що характеризують розподіл кожної змінної.
6. За потреби розрахувати допоміжні показники/змінні (на основі тих показників, що наявні у первинній таблиці).
7. Здійснити кореляційний і візуальний аналіз даних, побудувавши таблиці й графіки у різноманітних розрізах, які найкраще характеризують відповідну вибірку, на думку команди.
8. На основі проведеного аналізу коротко сформулювати ключові висновки щодо розподілу даних, основних тенденцій, закономірностей, аномалій, взаємозв’язків, які було виявлено.

**Формат презентації дослідницького проекту:** результати проведеного дослідження необхідно (i) підготувати у вигляді презентації та представити під час занять, а (ii) відповідний код (Python-скрипт) надіслати викладачу в текстовому форматі.

**Термін надсилання (і) презентації та (іі) Python-скрипту:** 5 грудня 2022 р.

**Оцінювання:**

* Максимальна кількість балів за дослідницьке завдання – 20 б.
* Оцінка залежить передусім від результатів командної роботи, представлених у презентації та файлі з Python-скриптом.
* На оцінку також впливають індивідуальні відповіді кожного з членів команди під час презентації дослідницького проекту.