МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ

«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ ІМЕНІ ІГОРЯ СІКОРСЬКОГО»

**WEB-Аналітика**

ЛАБОРАТОРНА РОБОТА №2

|  |  |  |
| --- | --- | --- |
|  |  | Виконав  студент групи ФІ-42мн  Бєш Радомир Андрійович |

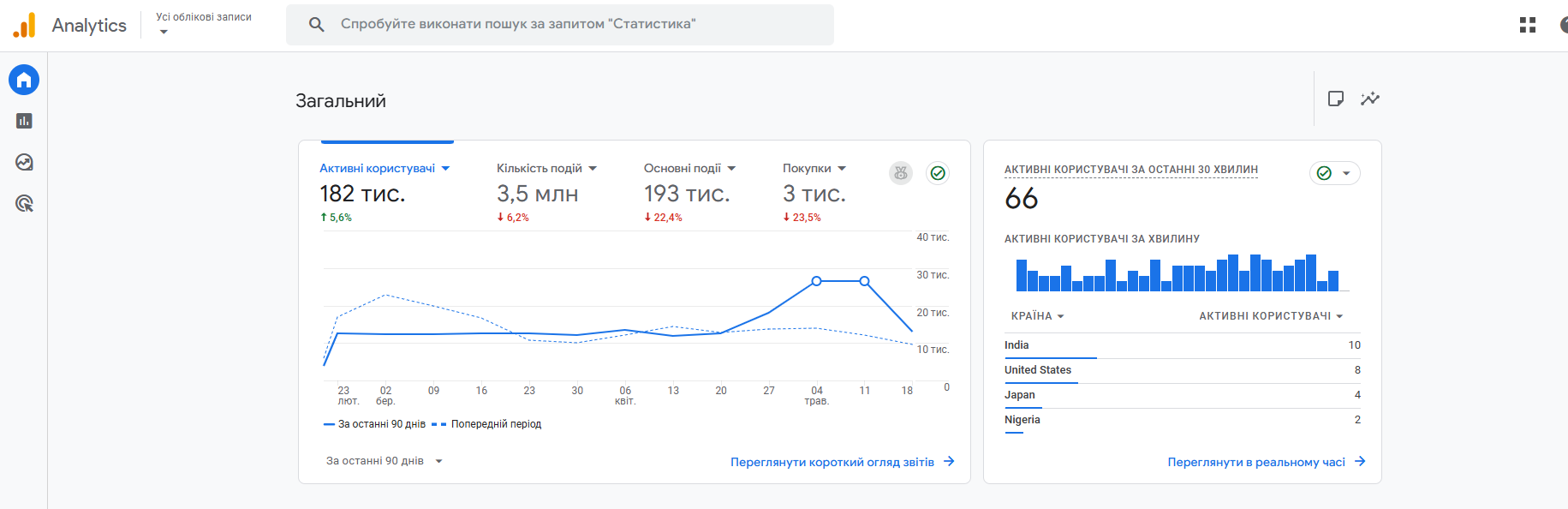
Київ 2025

**Завдання:**

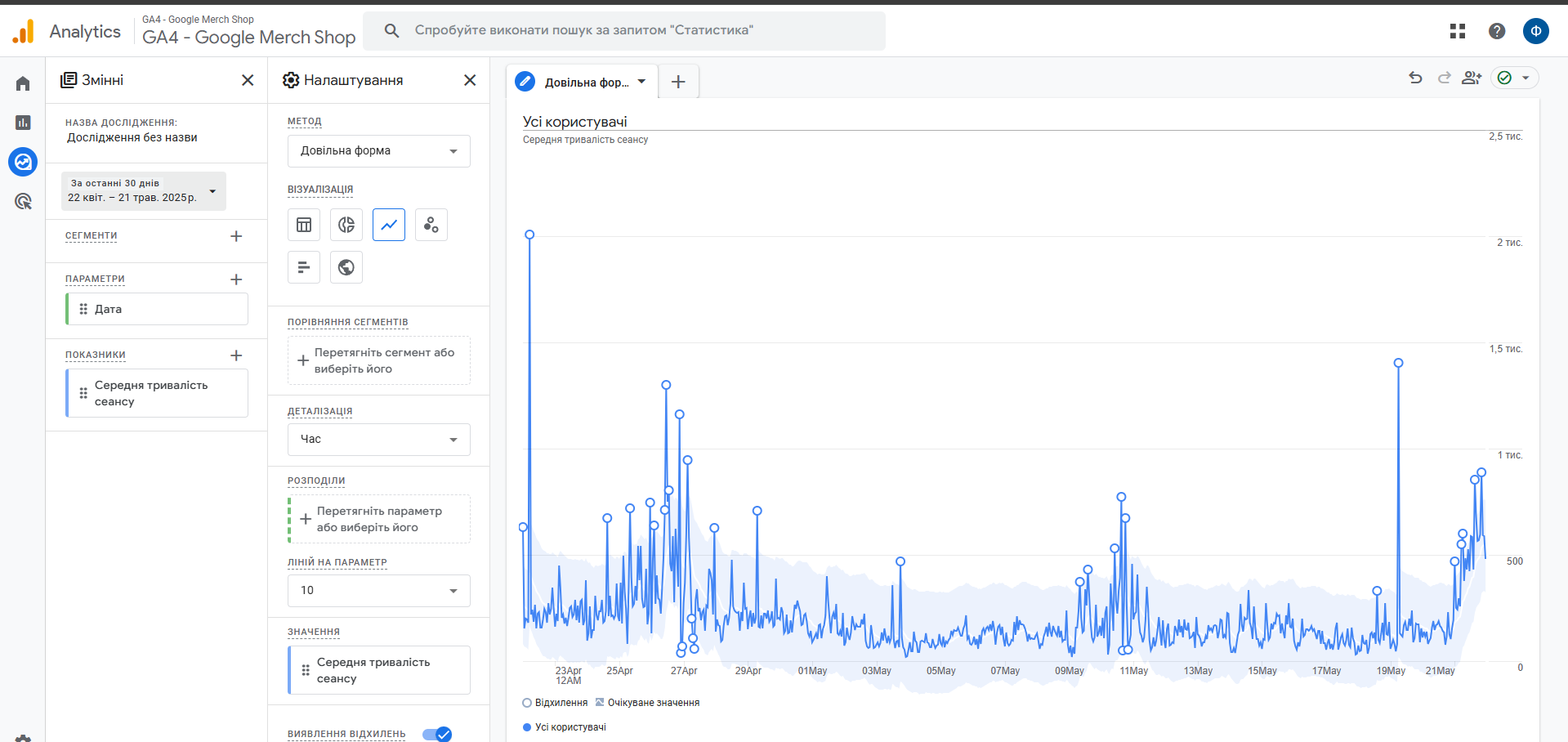
На основі даних з Google Analytics Demo account вибрати ТРИ різні часові ряди і на їх основі шляхом застосування методів визначення аномалій визначити аномалії в поведінці системи.

**Хід роботи**

1. Першим етапом лабораторної роботи було сформувати та завантажити датасет з демо акаунту в Google Analytics, але на превеликий жаль, сервіс в мене не працював.



Я не міг обрати демо акаунт. Навіть коли я створював нове дослідження я просто не міг завантажити дані в форматі .csv, та взагалі не міг завантажити будь-що.

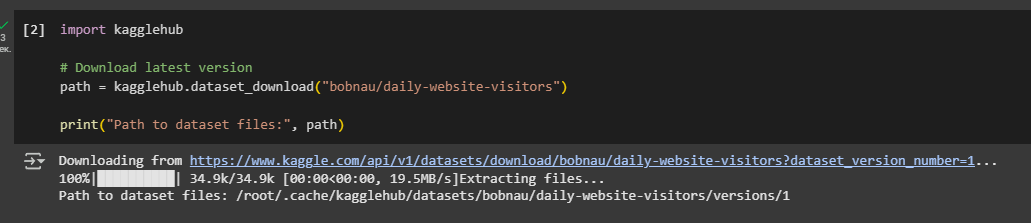


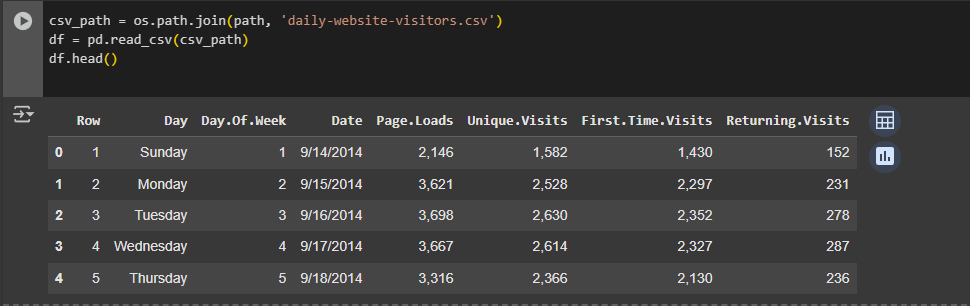
Тому було прийнято рішення знайти вже сформований датасет на Kaggle.

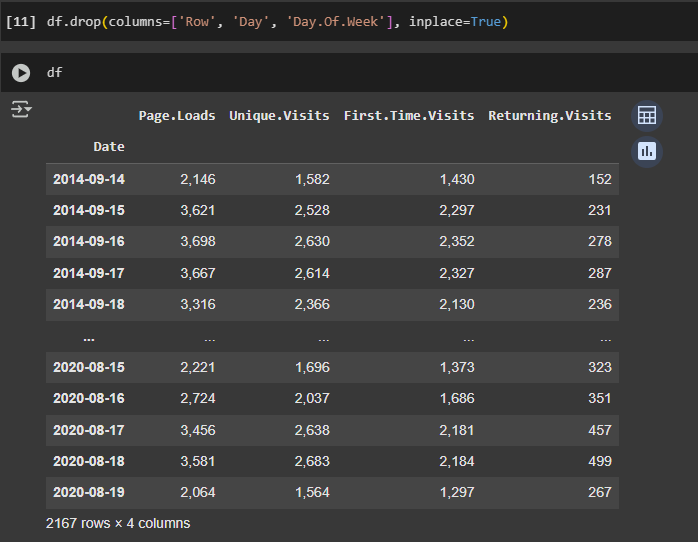
1. Обраний датасет:

Посилання: [https://www.kaggle.com/datasets/bobnau/daily-website-visitors](https://www.kaggle.com/datasets/bobnau/daily-website-visitors%20)

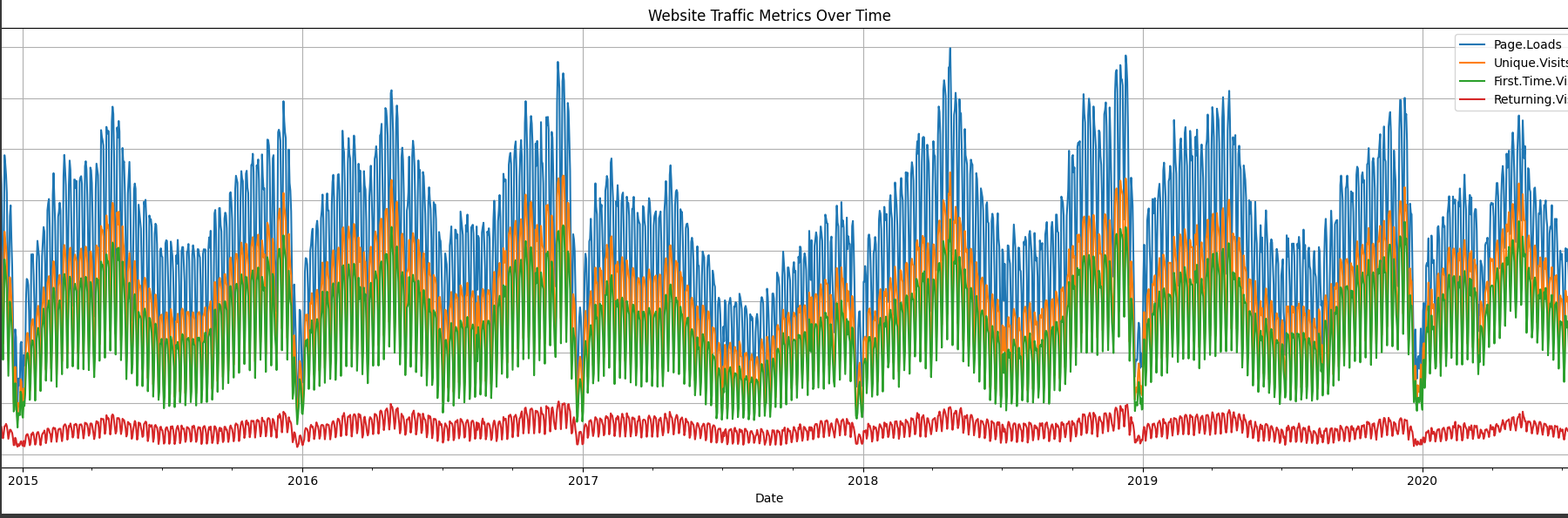
Початок роботи з датасетом, завантаження , читання , первинна обробка.



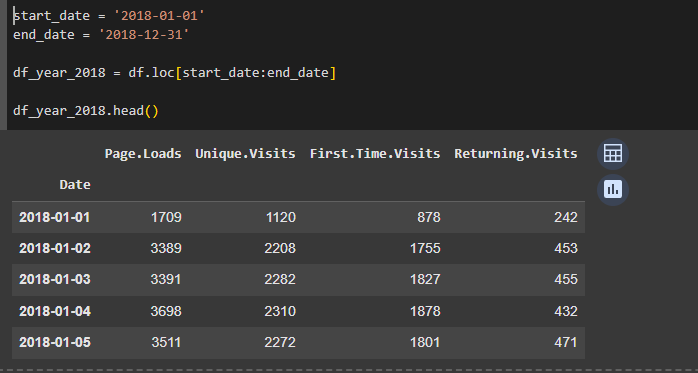




Датасет містить дані за період з 2015 по 2020 роки.

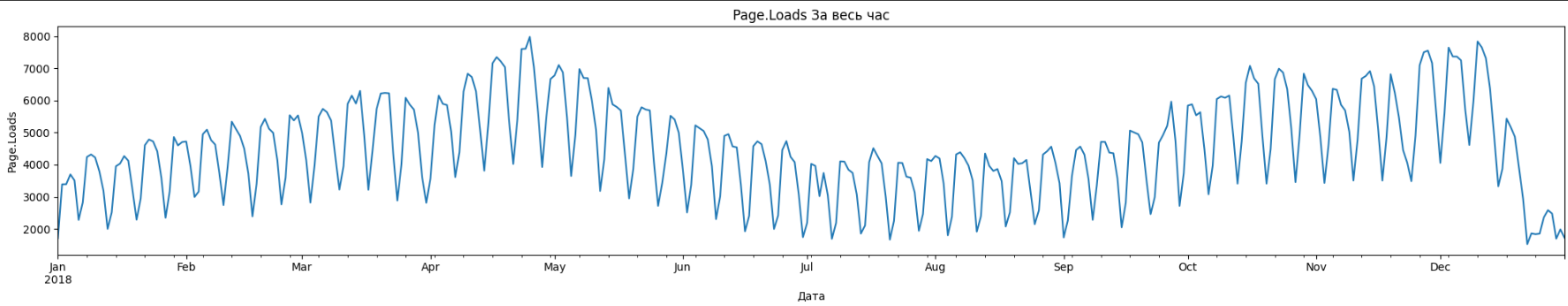


Але для нашого дослідження оберемо часовий проміжок за 2018 рік.

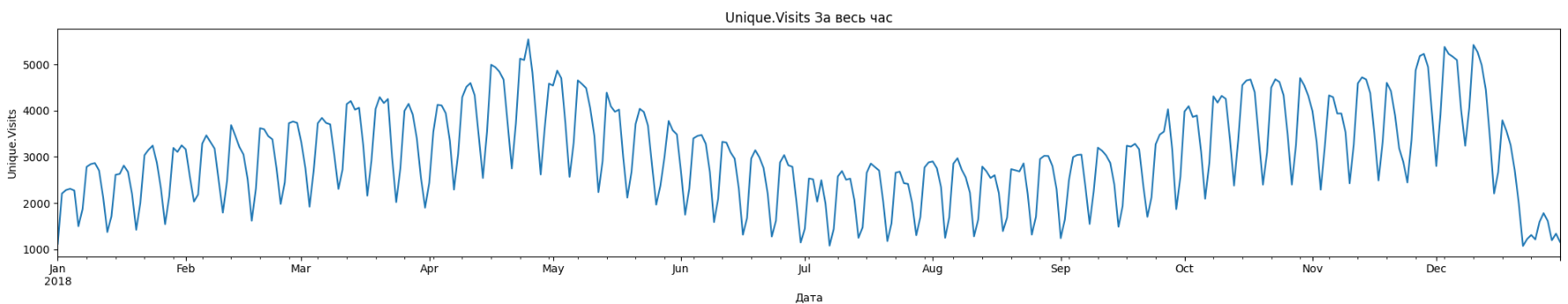


1. Будуємо графіки трьох часових рядів

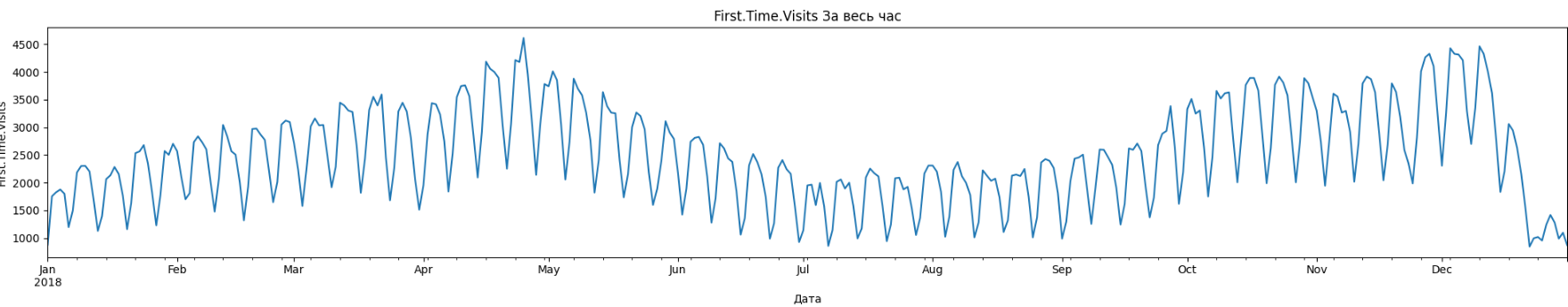
Завантаження сторінки



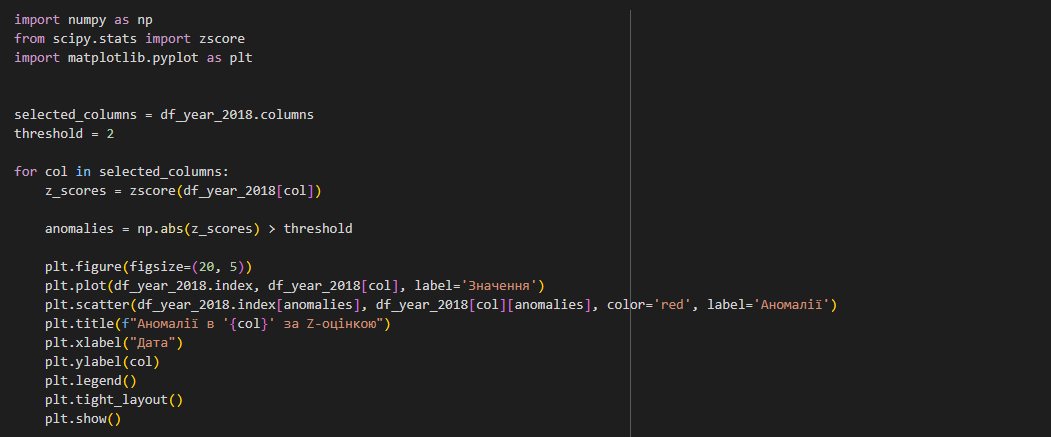
Унікальні користувачі



Перший візит на сторінку

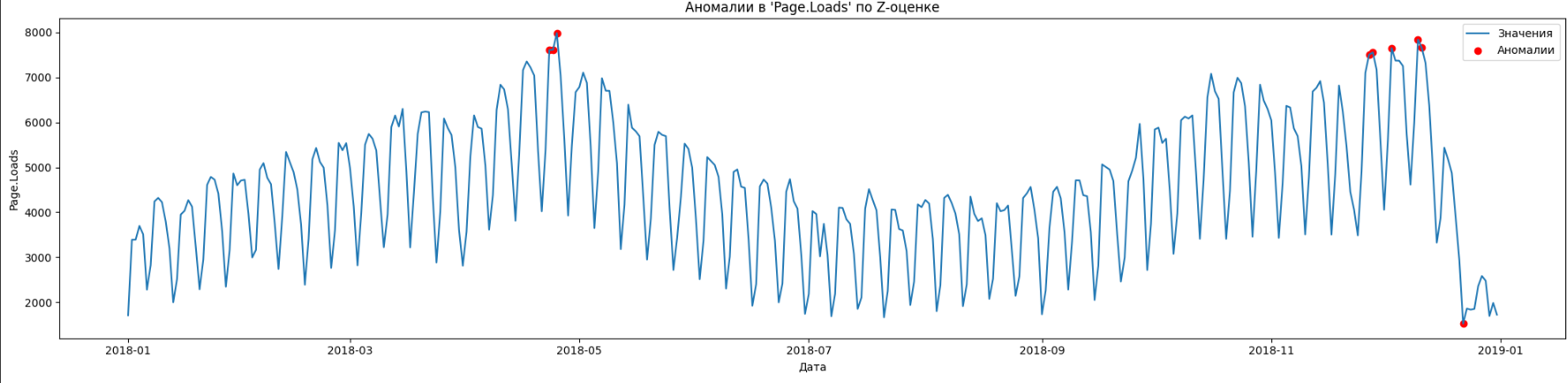


1. Z-Score метод для побудови аномалій

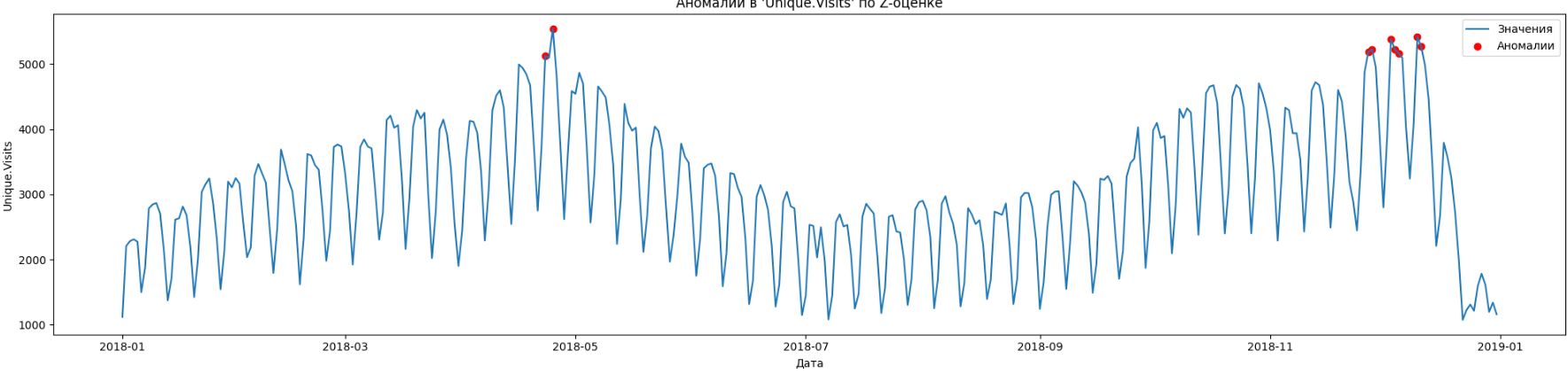


Результати:

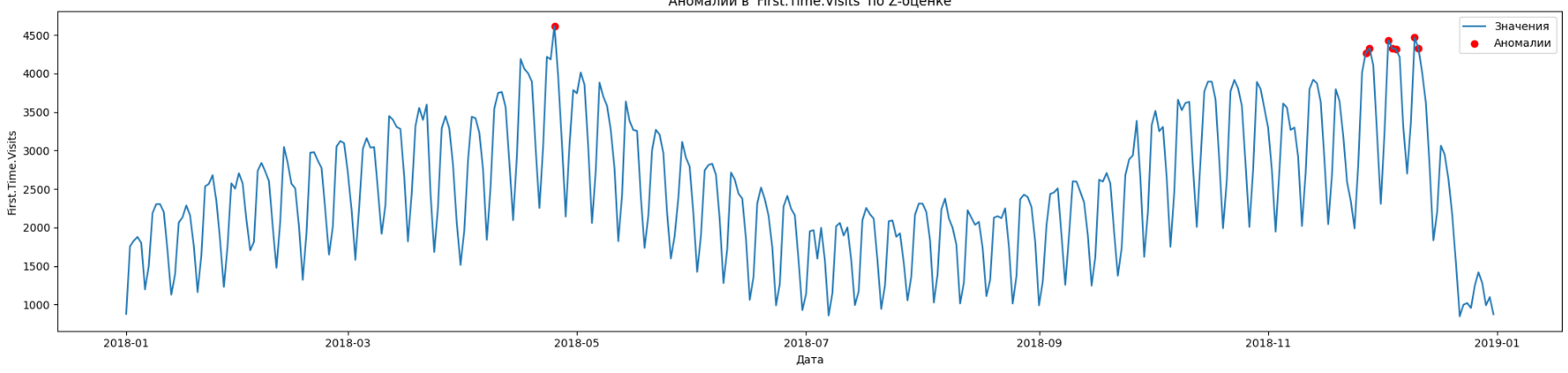
Завантаження сторінки



Унікальні користувачі

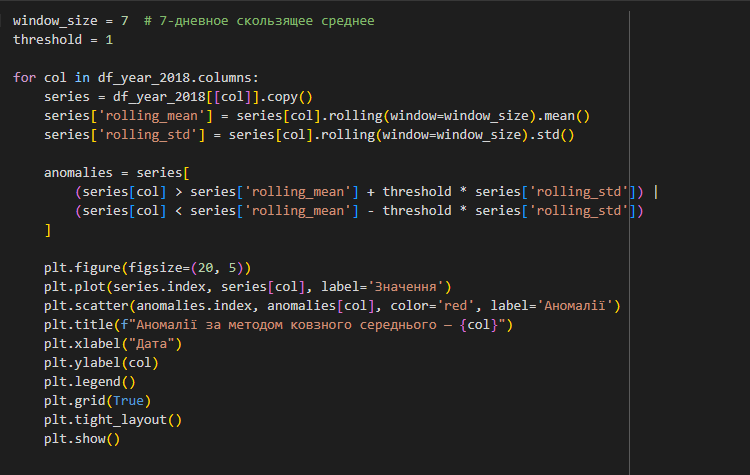


Перший візит на сторінку



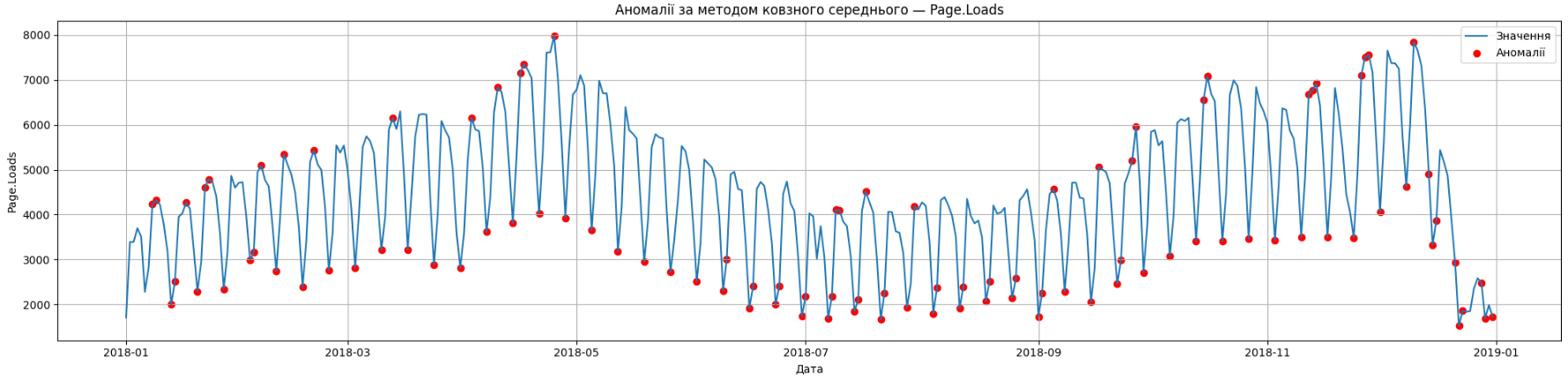
1. Метод Moving Avarage або Ковзне середнє

Обрали вікно в тиждень

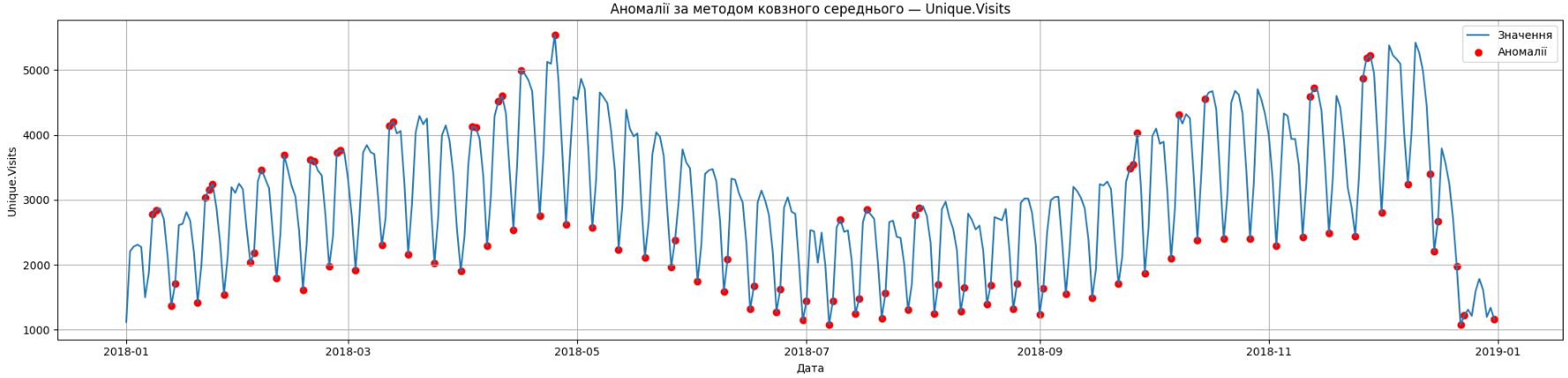


Результати:

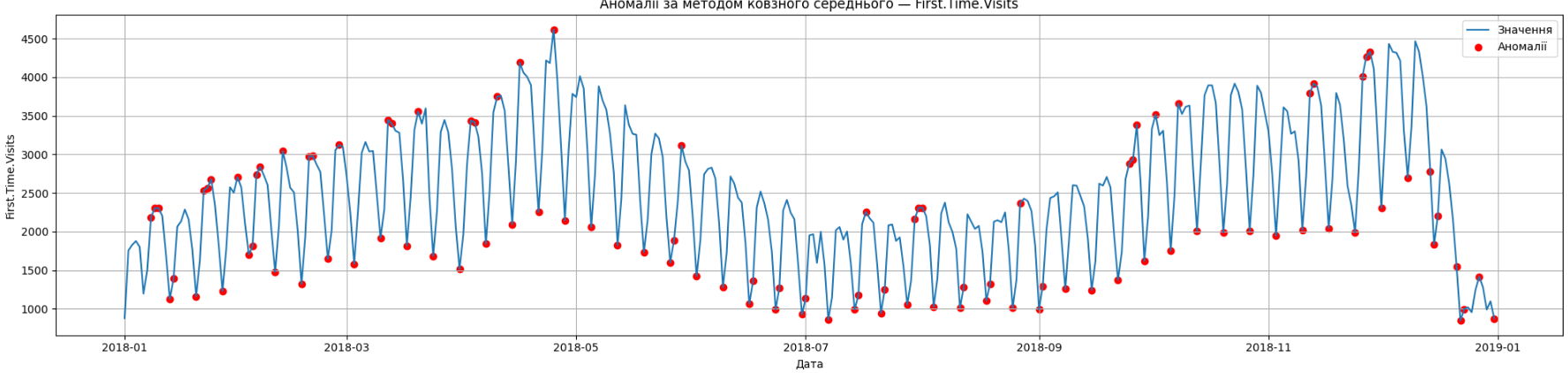
Завантаження сторінки



Унікальні користувачі



Перший візит на сторінку

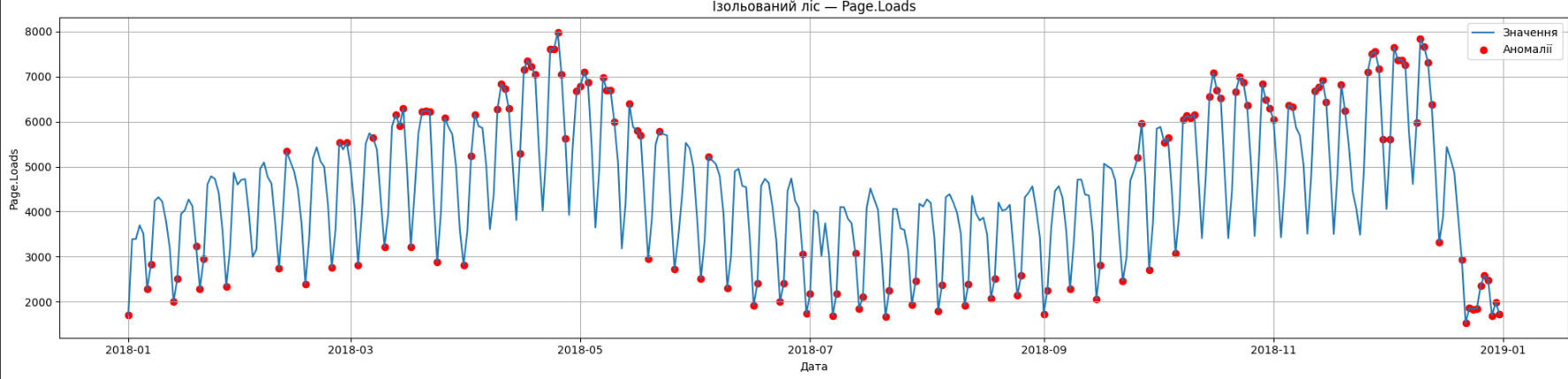


1. Метод ізольованого лісу

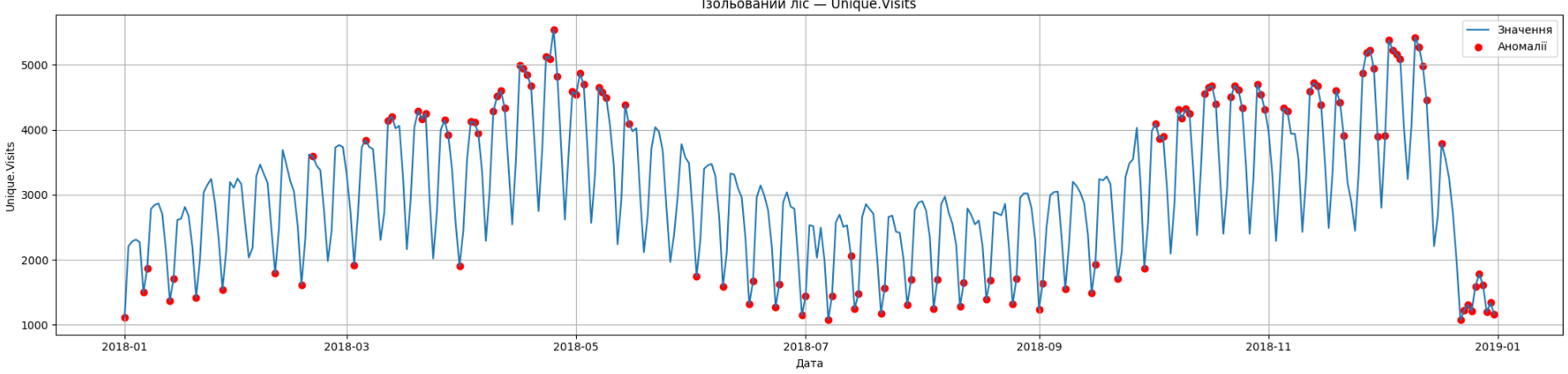


Результати:

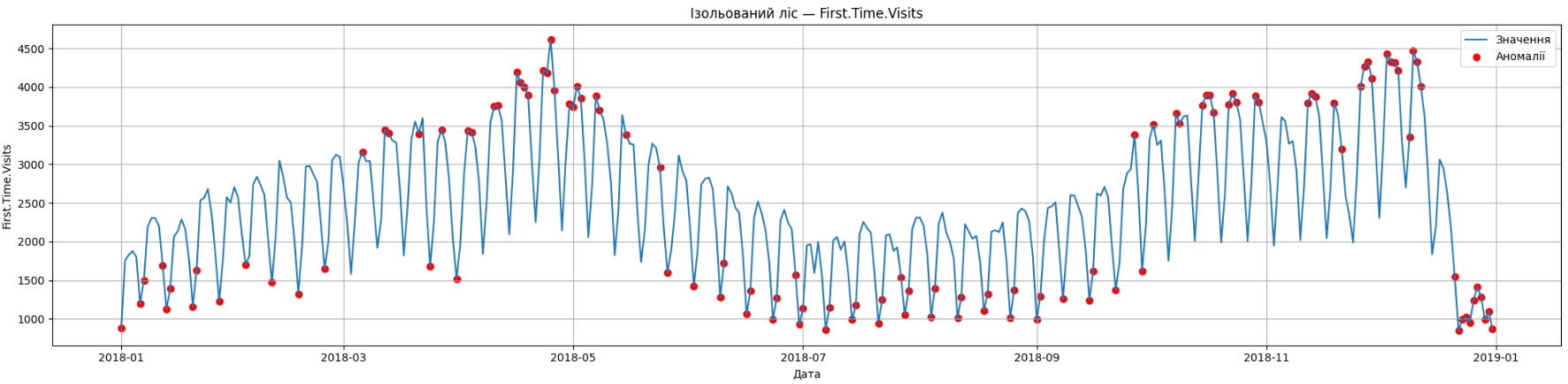
Завантаження сторінки



Унікальні користувачі



Перший візит на сторінку



**Висновки:** у ході лабораторної роботи було проведено аналіз часових рядів на основі даних із Google Analytics. Було обрано три часові ряди за 2018 рік: завантаження сторінки, унікальні користувачі та перші візити. Для виявлення аномалій застосовувалися три різні методи: Z-Score, ковзне середнє (Moving Average) та ізольований ліс. Метод Z-Score показав себе як чутливий до різких піків і провалів у даних, дозволяючи швидко виявити одиничні аномальні дні. Ковзне середнє із вікном у сім днів забезпечило згладжування рядів і дало можливість виявляти відхилення від локальних трендів. Метод ізольованого лісу, який базується на машинному навчанні, виявив нестандартні структури та зміни в поведінці користувачів, зокрема тривалі спади чи підйоми активності.

Загалом, усі три методи підтвердили наявність аномалій у поведінці користувачів на вебсайті протягом 2018 року. Різні підходи дозволили виявити різні типи аномалій, тому їх комбінація є найбільш ефективною для повного розуміння поведінкових змін. Отримані результати можуть бути корисними для глибшого аналізу користувацької активності, виявлення технічних проблем, впливу маркетингових кампаній або нетипової поведінки, пов’язаної з ботами чи зовнішніми факторами.