

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Національний аерокосмічний університет ім. М.Є. Жуковського
“Харківський авіаційний інститут”

Кафедра комп’ютерних систем, мереж і кібербезпеки

Лабораторна робота № 3

з дисципліни “Обробка даних засобами Python”

Дослідження можливостей бібліотеки Pandas

Виконав студент гр. 555iM Литвинов.О.А

(підпис, дата)

Перевірів к.т.н зав.кафедри каф. 301

(науковий ступінь, вчене звання, посада)

(підпис, дата)

Дергачов.К.Ю
(П.І.Б.)

2023

Мета: сформувати у здобувачів компетенції використання засобів бібліотеки Pandas для вирішення прикладних завдань

Завдання на лабораторну роботу

В рамках проведення лабораторної роботи необхідно обробити набір даних вимірювань 2-антенного GPS-приймача, побудувати графічні залежності вимірювальних параметрів та оцінити статистичні характеристики вимірювань.

1. За даними отриманого файлу формату Novatel DualGPS необхідно побудувати графічні залежності кутів курсу(head(t) (стовпчик даних № 13)) та куту крену (pitch(t) стовпчик № 14) від часу вимірювань, час знаходиться у 7 стовпчику та вимірюється у секундах, а кути у градусах.

2. Необхідно додати графічні залежності для усереднених даних вказаних кутів із шагом 10 с, та 20 с від часу вимірювань

3. Отримати статистичні характеристики (середнє значення, середньо квадратичне відхилення та максимальне відхилення від середнього значення)

Варіант 6:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

# Считывание данных с файла
data = pd.read_csv('file.csv')
time = data['time']
heading = data['head']
pitch = data['pitch']

# Графические зависимости для задания 1, 2
def figures():
    # Графики без шага
    print_figure(time, heading, 'Залежність кутів курсу від часу')
    print_figure(time, pitch, 'Залежність кутів крену від часу')

    # усреднённые графики с шагом 10 секунд
    step = 10
    averaged_time_10_step, averaged_heading_10_step, averaged_pitch_10_step = averaged_with_step(step)
    print_figure(averaged_time_10_step, averaged_heading_10_step,
                  f'Залежність усереднених кутів курсу від часу (шаг {step} секунд)')
    print_figure(averaged_time_10_step, averaged_pitch_10_step,
                  f'Залежність усереднених кутів крену від часу (шаг {step} секунд)')

    # усреднённые графики с шагом 20 секунд
    step = 20
    averaged_time_20_step, averaged_heading_20_step, averaged_pitch_20_step = averaged_with_step(step)
```

```

print_figure(averaged_time_20_step, averaged_heading_20_step,
              f'Залежність усереднених кутів курсу від часу (шаг {step} секунд)')
print_figure(averaged_time_20_step, averaged_pitch_20_step,
              f'Залежність усереднених кутів крену від часу (шаг {step} секунд)')

# Подсчёт статистических характеристик
def statistical_characteristics():
    # Средние значения
    mean_heading = np.mean(heading)
    mean_pitch = np.mean(pitch)

    # Подсчёт среднеквадратического отклонения
    std_heading = np.std(heading)
    std_pitch = np.std(pitch)

    # Подсчёт максимальных отклонений от среднего значения
    max_dev_heading = np.max(np.abs(heading - mean_heading))
    max_dev_pitch = np.max(np.abs(pitch - mean_pitch))

    # Вывод результата в терминал
    print(f'Середнє значення куту курсу: {mean_heading} градуси')
    print(f'Середнє значення куту крену: {mean_pitch} градуси')
    print(f'Середньоквадратичне відхилення куту курсу: {std_heading} градуси')
    print(f'Середньоквадратичне відхилення куту крену: {std_pitch} градуси')
    print(f'Максимальне відхилення куту курсу: {max_dev_heading} градуси')
    print(f'Максимальне відхилення куту крену: {max_dev_pitch} градуси')

# Подсчёт усреднённых данных для указанного шага
def averaged_with_step(step):
    averaged_time_step = np.arange(min(time), max(time), step)
    averaged_heading_step = [np.mean(heading[(time >= t) & (time < t + step)]) for t in averaged_time_step]
    averaged_pitch_step = [np.mean(pitch[(time >= t) & (time < t + step)]) for t in averaged_time_step]
    return averaged_time_step, averaged_heading_step, averaged_pitch_step

# Создание графических фигур, с указанными данными, и названием
def print_figure(data_1, data_2, string):
    plt.figure(figsize=(12, 6))
    plt.plot(data_1, data_2)
    plt.xlabel('Час (с)')
    plt.ylabel('Кут (градуси)')
    plt.title(string)
    plt.show()

if __name__ == '__main__':
    figures()
    statistical_characteristics()

```

Середнє значення куту курсу: 97.11364867085796 градуси
Середнє значення куту крену: -0.12822110062457723 градуси
Середньоквадратичне відхилення куту курсу: 0.1232174735737523 градуси
Середньоквадратичне відхилення куту крену: 0.18688379992811197 градуси
Максимальне відхилення куту курсу: 0.42816804014204024 градуси
Максимальне відхилення куту крену: 0.6253494563754227 градуси

Рисунок 1 – Статистичні характеристики

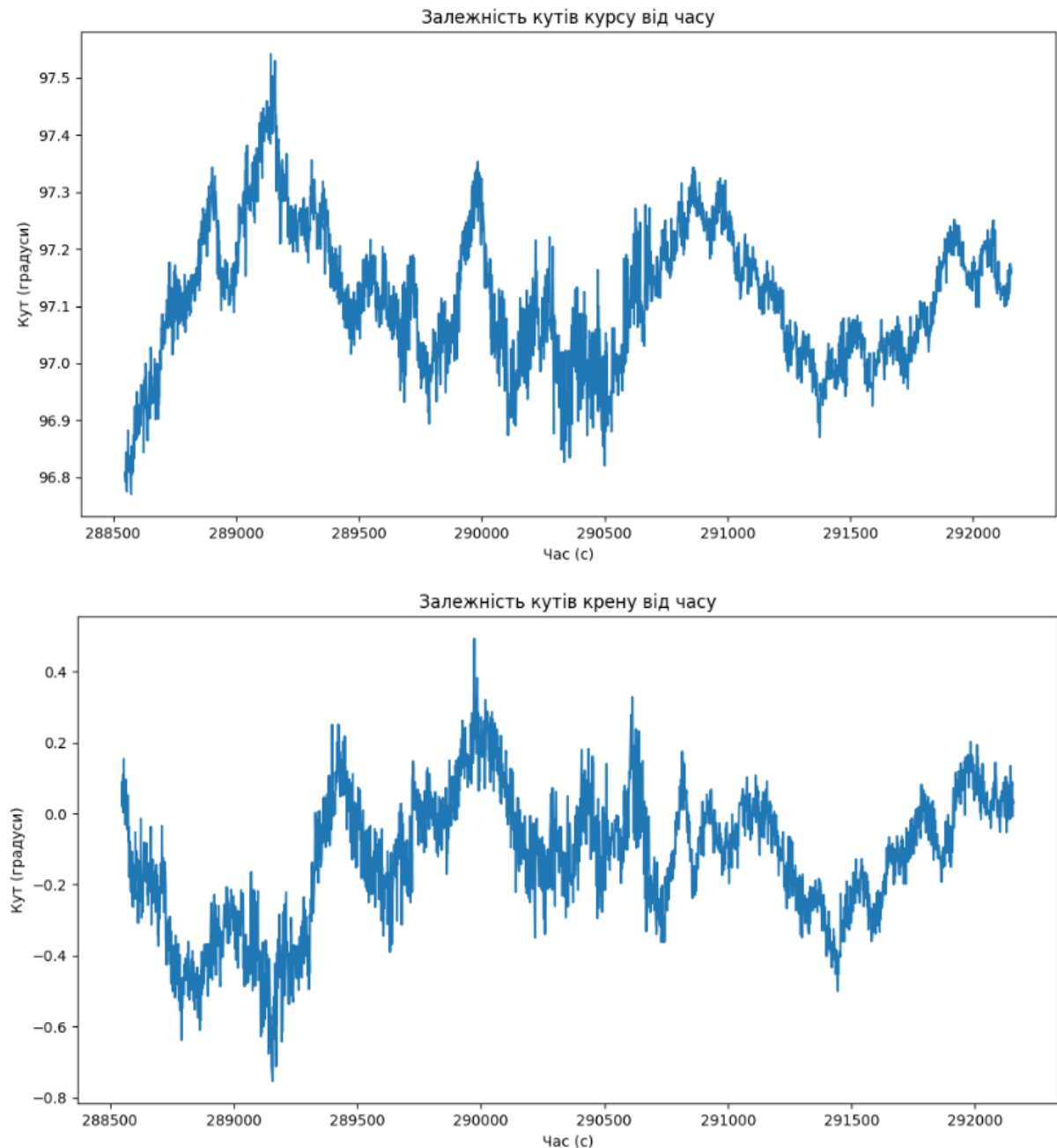


Рисунок 2 – Залежність кутів(яких саме, від чого саме, підписано над графіками)

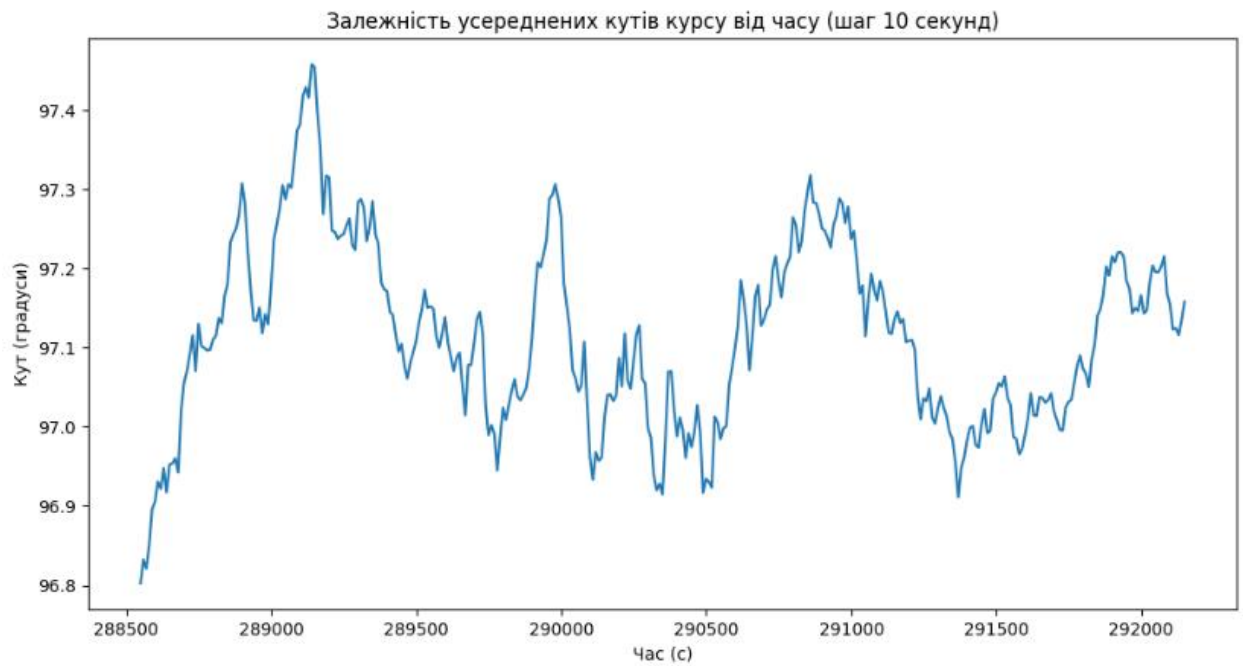


Рисунок 3 – Залежність кутів з кроком 10 секунд(яких саме, від чого саме, підписано над графіками)

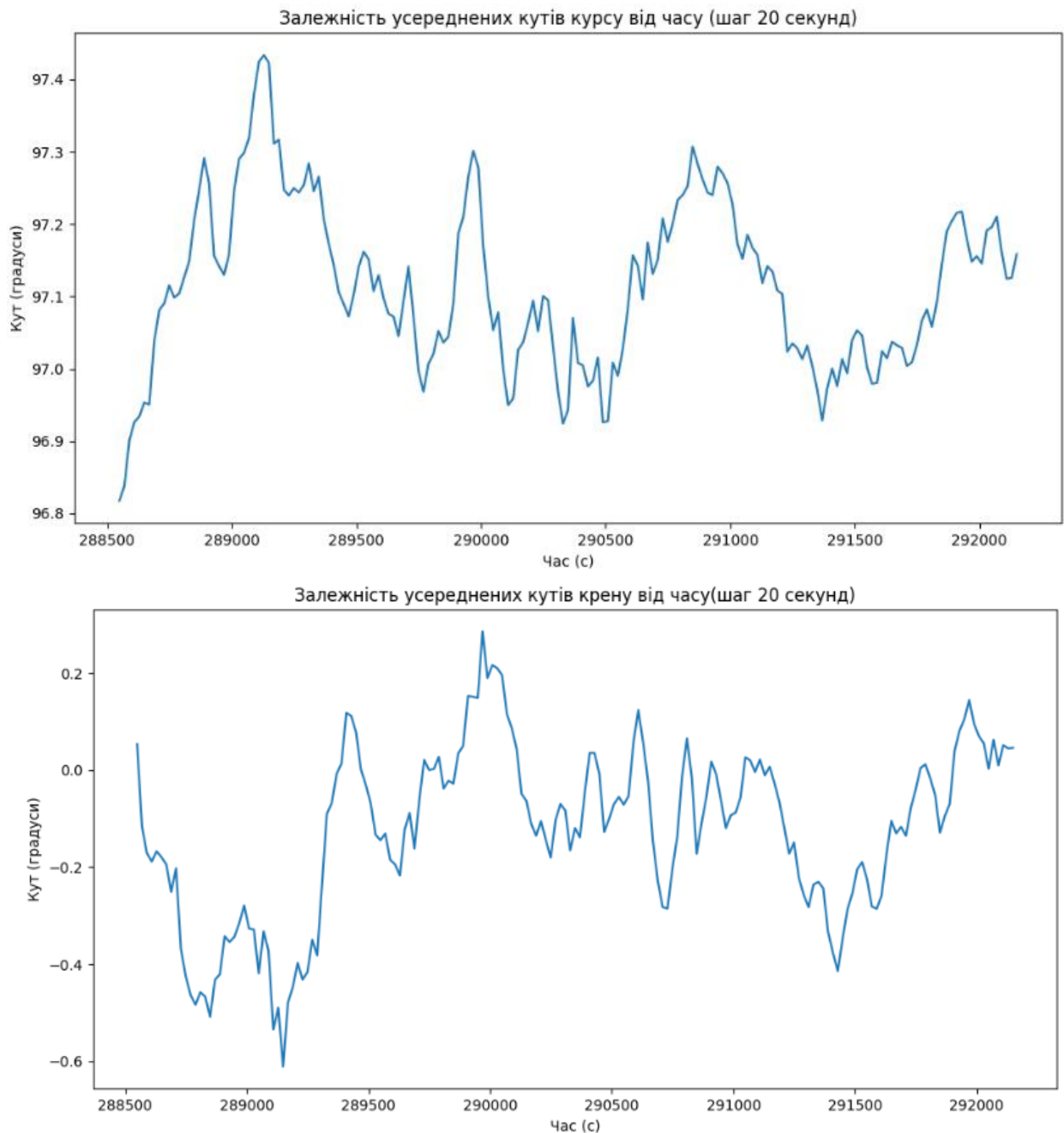


Рисунок 4 - Залежність кутів з кроком 20 секунд(яких саме, від чого саме, підписано над графіками)

Висновок:

Pandas, використовують для аналізу та маніпулювання даними. Для її роботи користувач надає DataFrame в випадку цієї роботи це NMEA. Для роботи було задіяно читання з файлу «pd.read_csv('file.csv')» та можливість витягування даних з конкретних стовбців `time = data['time']`.

Для підрахунків використовувалася numpy, а точніше:

- `np.mean` – використовується для обчислення середнього значення з набору числових даних;
- `np.std` – використовується для обчислення середньо-квадратичного відхилення з набору даних;
- `np.abs` – використовується для обчислення абсолютних відхилень між кожним значенням в масиві та середнім значенням;

- `pr.max` – використовується для знаходження максимального значення з набору даних.

Для виводу графіків `matplotlib.pyplot`.