МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

СУМСЬКИЙ ДЕРЖАВНИЙ УНІВЕРСИТЕТ

КАФЕДРА ІНФОРМАЦІЙНИХ ТЕХНОЛОГІЙ

**ЗВІТ З ЛАБОРАТОРНОЇ РОБОТИ №15**

з дисципліни «Програмування мовою Python»

на тему «Аналіз даних за допомогою Python бібліотек»

Виконав студент групи КН-33

Глумний Тимур

Перевірила Парфененко Ю.В.

Варіант 3

Суми 2024

**Завдання 1**

1. **Постановка задачі**

Перетворіть словник, створений у лабораторній роботі №6, на датафрейм, у разі необхідності доповніть словник даними, виведіть його на екран. Виконайте агрегацію та групування даних із заданої предметної області.

1. Код програми

import pandas as pd

# Дані про автомобілі

data = {

"Автомобіль": ["Автомобіль 4", "Автомобіль 9", "Автомобіль 5", "Автомобіль 10",

"Автомобіль 6", "Автомобіль 7", "Автомобіль 2", "Автомобіль 3"],

"Потужність (к.с.)": [200.0, 170.0, 150.0, 140.0, 130.0, 95.0, 80.0, 50.0],

"Вартість (грн)": [160000.0, 140000.0, 120000.0, 120000.0, 110000.0, 75000.0, 60000.0, 30000.0]

}

# Створення DataFrame

df = pd.DataFrame(data)

# Виведення DataFrame на екран

print("DataFrame з даними про автомобілі:\n", df)

# Обчислення середньої потужності

average\_horsepower = df["Потужність (к.с.)"].mean()

print(f"\nСередня потужність автомобілів = {average\_horsepower:.2f} к.с.")

# Обчислення загальної вартості

total\_cost = df["Вартість (грн)"].sum()

print(f"Загальна вартість автомобілів = {total\_cost:.2f} грн")

# Мінімальна і максимальна вартість автомобілів

min\_cost = df["Вартість (грн)"].min()

max\_cost = df["Вартість (грн)"].max()

print(f"Мінімальна вартість автомобіля = {min\_cost:.2f} грн")

print(f"Максимальна вартість автомобіля = {max\_cost:.2f} грн")

1. **Результат виконання**

Зображення, що містить текст, знімок екрана, Шрифт, меню

Автоматично згенерований опис

**Завдання 2**

1. **Постановка задачі**

Створіть датафрейм з даними використання велодоріжок за рік, заданий варіантом. CSV файл завантажте на сторінці http://donnees.ville.montreal.qc.ca/dataset/velos-comptage (рис.1).   
2018 рік. Визначте, який місяць найбільш популярний у велосипедистів.

1. Код програми

import pandas as pd

import matplotlib.pyplot as plt

import matplotlib.ticker as ticker

# Завантажуємо дані з CSV

df = pd.read\_csv('lab 15.2.csv')

# Перетворюємо колонку з датами на формат datetime, враховуючи, що день вказаний першим

df['Date'] = pd.to\_datetime(df['Date'], format='%d/%m/%Y', dayfirst=True)

# Фільтруємо дані за 2018 рік

df = df[df['Date'].dt.year == 2018]

# Додаємо колонку з місяцем для подальшого групування

df['Month'] = df['Date'].dt.month

# Групуємо за місяцем та сумуємо кількість велосипедистів для кожного місяця

monthly\_counts = df.groupby('Month').sum(numeric\_only=True).sum(axis=1)

# Знаходимо місяць з найбільшою кількістю велосипедистів

most\_popular\_month = monthly\_counts.idxmax()

most\_popular\_count = monthly\_counts.max()

# Список назв місяців

month\_names = [

'Січень', 'Лютий', 'Березень', 'Квітень', 'Травень', 'Червень',

'Липень', 'Серпень', 'Вересень', 'Жовтень', 'Листопад', 'Грудень'

]

# Виводимо результати

print("Місяць з найбільшою кількістю велосипедистів:", month\_names[most\_popular\_month - 1])

print("Кількість велосипедистів у цьому місяці:", most\_popular\_count)

# Побудова графіка

plt.figure(figsize=(10, 6))

monthly\_counts.plot(kind='bar', color='skyblue')

plt.title('Використання велодоріжок по місяцях за 2018 рік')

plt.xlabel('Місяць')

plt.ylabel('Кількість велосипедистів')

# Форматуємо вісь Y, щоб відображати тисячі, мільйони чи мільярди

def format\_ticks(x, pos):

if x >= 1e9:

return f'{x\*1e-9:.1f} млрд'

elif x >= 1e6:

return f'{x\*1e-6:.1f} млн'

elif x >= 1e3:

return f'{x\*1e-3:.1f} тис'

else:

return f'{x:.0f}'

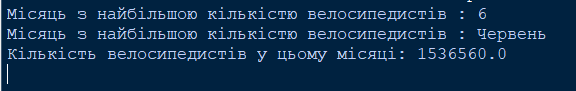
plt.gca().yaxis.set\_major\_formatter(ticker.FuncFormatter(format\_ticks))

# Відображаємо графік

plt.xticks(rotation=0)

plt.show()

1. Результат роботи програми



**https://github.com/doomedd/python-laba-15**