Отчет по лабораторной работе №1.

Дисциплина: «Структуры данных и парадигмы программирования».

Отчет выполнила Ефремова Вера

группа НБИ-01-22, №студ. билета: 1032225672

Задание:

5. Data.gov - [https://www.data.gov/](https://www.data.gov/)

- Портал открытых данных правительства США.

№ 1. Чтение данных из CSV файла import pandas as pd df = pd.read\_csv('file.csv') print(df.head())

№ 2. Запись DataFrame в CSV файл [df.to](https://df.to/)\_csv('new\_file.csv', index=False)

№ 3. Выбор столбца из DataFrame column = df['column\_name'] print(column)

№ 4. Выбор строки по индексу row = df.iloc[0] print(row)

№ 6. Группировка данных по столбцу grouped\_data = df.groupby('column\_name').sum() print(grouped\_data)

№ 7. Объединение двух DataFrame df1 = pd.DataFrame(data1) df2 = pd.DataFrame(data2) merged\_df = pd.merge(df1, df2, on='common\_column') print(merged\_df)

№ 8. Расчет основных статистических показателей print(df.describe())

№ 9. Сортировка данных sorted\_df = df.sort\_values(by='column\_name', ascending=True) print(sorted\_df)

№ 10. Построение простого графика df['column\_name'].plot(kind='bar')

№ 11. Сохранение DataFrame в Excel файл [df.to](https://df.to/)\_excel('output.xlsx', index=False)

Код программы:

import pandas as pd

import matplotlib.pyplot as plt

# 1. Чтение данных из CSV файла

df = pd.read\_csv('https://www.ncei.noaa.gov/pub/data/cdo/samples/LCD\_sample\_csv.csv')

print("Первые пять строк данных:")

print(df.head())

# 2. Запись DataFrame в CSV файл

df.to\_csv('new\_file.csv', index=False)

print("\nДанные записаны в файл 'new\_file.csv'")

# 3. Выбор столбца из DataFrame

column = df['HourlyDryBulbTemperature']

print("\nСтолбец 'HourlyDryBulbTemperature':")

print(column)

# 4. Выбор строки по индексу

row = df.iloc[0]

print("\nПервая строка данных:")

print(row)

# 5. Фильтрация данных по условию

filtered\_data = df[df['HourlyDryBulbTemperature'] > 50]

print("\nДанные с температурой выше 50:")

print(filtered\_data)

# 6. Группировка данных по столбцу

grouped\_data = df.groupby('StationName').sum()

print("\nСгруппированные данные по станциям:")

print(grouped\_data)

# 7. Объединение двух DataFrame

data1 = {'StationName': ['A', 'B'], 'Value1': [1, 2]}

data2 = {'StationName': ['A', 'B'], 'Value2': [3, 4]}

df1 = pd.DataFrame(data1)

df2 = pd.DataFrame(data2)

merged\_df = pd.merge(df1, df2, on='StationName')

print("\nОбъединенные данные:")

print(merged\_df)

# 8. Расчет основных статистических показателей

print("\nОсновные статистические показатели:")

print(df.describe())

# 9. Сортировка данных

sorted\_df = df.sort\_values(by='HourlyDryBulbTemperature', ascending=True)

print("\nОтсортированные данные по температуре:")

print(sorted\_df)

# 10. Построение простого графика

df['HourlyDryBulbTemperature'].plot(kind='bar')

plt.title('Hourly Dry Bulb Temperature')

plt.xlabel('Index')

plt.ylabel('Temperature')

plt.show()

# 11. Сохранение DataFrame в Excel файл

df.to\_excel('output.xlsx', index=False)

print("\nДанные сохранены в файл 'output.xlsx'")