|  |  |  |
| --- | --- | --- |
| ДонГТУ | Методы анализа данных | СКС-23 |
| Кафедра СКС | Лабораторная работа №2 | Кукарин А.А. |

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Выполнение | | | | Защита | | | |
| Подпись  студента | Фамилия  преподавателя | Дата  выполн. | Подпись  препод. | Фамилия  преподавателя | Оценка | Дата  защиты | Подпись  препод. |
|  | Самойлов Д.В. |  |  | Самойлов Д.В. |  |  |  |

**Тема работы:** Формулировка и проверка статистических гипотез

**Цель работы:** научиться формулировать и проверять статистические гипотезы, научиться составлять программы их использованием языка Python.

Задание к лабораторной работе:

1.Загрузите набор данных по ссылке

https://dpds.weatheronline.co.uk/historical\_data/weather\_stations\_download

2. Загрузите набор данных в pandas при помощи функции read\_csv().Данные должны быть извлечены корректно (указана исходная кодировка, разобраны даты, все поля разделены). Установите индекс на один из столбцов.

3. Выведите набор на экран.

4. Выведите на экран одно из полей

5. Постройте график набора/поля при помощи функции plot().

Исходный код:

import pandas as pd

import matplotlib.pyplot as plt

# Загружаем данные, пропуская первые 9 строк

df = pd.read\_csv("dataset для лб2.csv", encoding="utf-8", skiprows=9, names=["timestamp", "temperature"])

# Преобразуем строковые даты в datetime

df["timestamp"] = pd.to\_datetime(df["timestamp"], errors="coerce", infer\_datetime\_format=True)

print(df["timestamp"].head(10))

# Преобразуем температуру в float

df["temperature"] = pd.to\_numeric(df["temperature"], errors="coerce")

# Установим индекс

df.set\_index("timestamp", inplace=True)

# Выведем первые строки

print(df.head())

# Построим график температуры

df["temperature"].plot(title="График температуры", figsize=(12, 5))

plt.xlabel("Дата")

plt.ylabel("Температура (°C)")

plt.grid()

plt.show()

Результат:



