МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ

РОССИЙСКОЙ ФЕДЕРАЦИИ

Федеральное государственное автономное образовательное учреждение

высшего образования

«ТЮМЕНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ»

Практическое занятие № 21\_22.

**“АНАЛИЗ ВЗАИМОСВЯЗИ КАЧЕСТВА СНА И ОБРАЗА ЖИЗНИ С ПОКАЗАТЕЛЯМИ ЗДОРОВЬЯ”**

**Выполнили** студенты:

Альзинский Егор Дмитриевич

Ибраев Дамир Беркутбаевич

Тюмень-2025

**1. Вступление.**

***1.1 Актуальность***

Нарушения сна, такие как бессонница и апноэ во сне, являются распространёнными проблемами, негативно влияющими на физическое и психическое здоровье. Недостаточная продолжительность и низкое качество сна связаны с повышенным уровнем стресса, снижением физической активности, риском сердечно-сосудистых заболеваний и ухудшением общего качества жизни.

Современный образ жизни, включая высокий уровень стресса, малоподвижность и несбалансированное питание, усугубляет эти проблемы. Анализ взаимосвязи между сном, физической активностью, стрессом и другими факторами поможет выявить ключевые закономерности и разработать эффективные рекомендации для улучшения здоровья.

***1.2 Проблема***

Из-за отсутствия инструментов для визуализации и анализа взаимосвязи между повседневной рутиной и академической успеваемостью невозможно выстроить действительно эффективные стратегии обучения и помочь студентам скорректировать свой образ жизни.

***1.3 Цель***

Проанализировать и предобработать dataset, описывающий повседневные привычки студентов и их академическую успеваемость, а также разработать программу на языке Python для визуального представления данных и выявления значимых взаимосвязей.

***1.4 Задачи***

**1)**Изучить и установить библиотеки pandas, matplotlib, tkinter и textwrap, plotly.express, seaborn, numpy.

**2)** Загрузить dataset и очистить его от некорректных значений.

**3)** Провести анализ структуры данных: диапазон значений, типы переменных, нормализовать данные для выявления средних значений.

**4)** Для наглядного представления работы кода реализовать блок-схему программы.

**4)**Реализовать графики для визуализации зависимости уровня сна рассматриваемых людей от следующих факторов: уровень стресса, гендера, кол-во шагов, пройденных за сутки в созависимости от времени сна. Также представить нормализованные средние значения всех разбираемых признаков, для прослеживания закономерностей, где 100 = максимум.

**5)** Реализовать отображение исходного датасета в виде таблицы.

**6)** Реализовать меню для управления готовой программой (выбор между показом датасета и аналитикой).

**7)** Оценить, как каждый из рассмотренных факторов влияет на уровень сна человека, сделать вывод.

**2. Основная часть.**

***2.1 Описание датасета***

Dataset Student Habits vs Academic Performance, который был использован в проекте, является публичным и взят с сайта <https://www.kaggle.com/>. Данный смоделированный, но реалистичный, набор данных состоит из 400 строк и 13 столбцов, охватывающих широкий спектр переменных, связанных со сном и ежедневными привычками. Он включает такие данные, как пол, возраст, род занятий, продолжительность сна, качество сна, уровень физической активности, уровень стресса, категория ИМТ, артериальное давление, частота сердечных сокращений, ежедневные шаги и наличие или отсутствие нарушений сна. Этот набор данных, предназначенный для машинного обучения, регрессии, классификации, интеллектуального анализа данных и визуализации данных, идеально подходит для образовательного анализа и экспериментов в проектах по анализу данных.

***2.2 Описание колонок датасета***

Person ID: идентификатор для каждого человека.

Gender: пол человека (мужчина/женщина).

Age: возраст человека(год).

Occupation: профессия или род занятий человека.

Sleep Duration (hours): количество часов, которые человек спит в день.

Quality of Sleep (scale: 1-10): субъективная оценка качества сна в диапазоне от 1 до 10.

Physical Activity Level (minutes/day): количество минут, в течение которых человек занимается физической активностью ежедневно.

Stress Level (scale: 1-10): субъективная оценка уровня стресса, испытываемого человеком, в диапазоне от 1 до 10.

BMI Category:  категория ИМТ человека (например, недостаточный вес, нормальный вес, избыточный вес)

Blood Pressure (systolic/diastolic): измерение артериального давления человека, указанное как систолическое давление над диастолическим давлением..

Heart Rate (bpm): частота сердечных сокращений человека в состоянии покоя в ударах в минуту.

Daily Steps: количество шагов, которые человек делает за день.

Sleep Disorder: наличие или отсутствие нарушения сна у человека (отсутствует, бессонница, апноэ во сне).

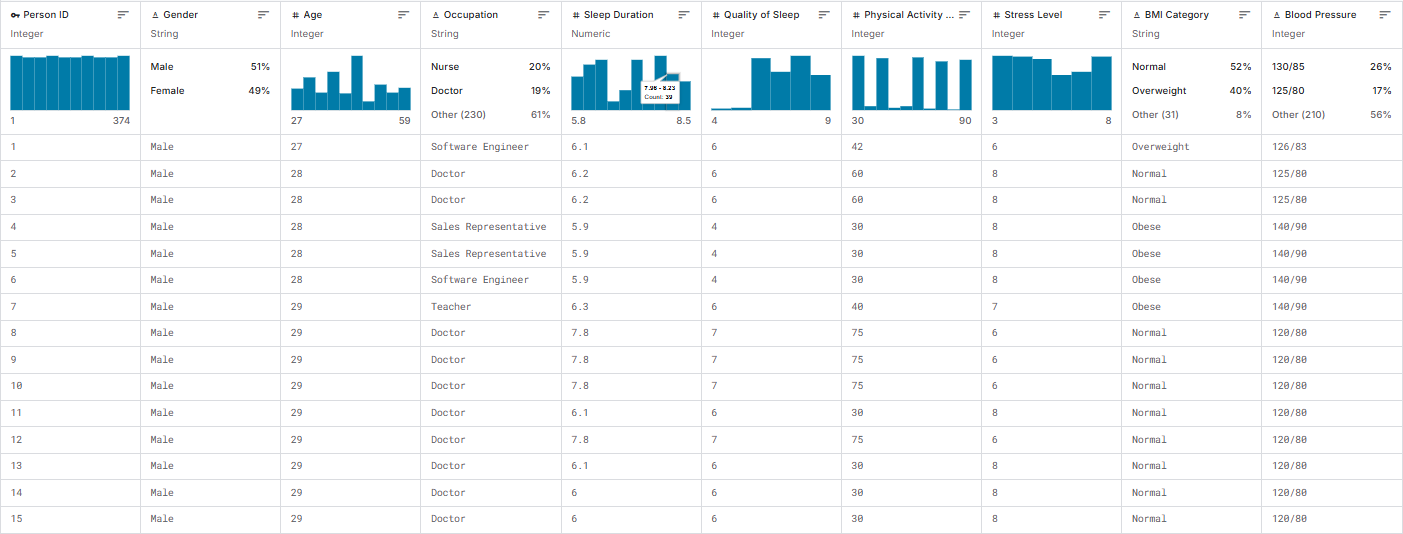


Рис. 1 – файл Sleep\_health\_and\_lifestyle\_dataset.csv

***2.3 Структура программы***

**Язык программирования**: Python

**Использованные библиотеки:**

**1)** pandas - обработка и анализ данных.

**2)** matplotlib – визуализация данных в виде графика.

**3)** tkinter - отображение исходного датасета в виде таблицы.

**4)** textwrap – форматирование текста.

**5)** numpy – работа с многомерными массивами и выполнения математических операций над ними.

**6)** seaborn – визуализации данных, построенная на основе Matplotlib.

**7)** plotly.express – создания интерактивных графиков и диаграмм.

***2.4 Принцип работы кода***

**1)** Подключение необходимых библиотек для анализа данных (pandas), построения графиков (matplotlib.pyplot, seaborn, plotly.express), создания графического интерфейса (tkinter) и форматирования текста (textwrap). (Рис. 3), реализация блок-схемы для наиболее наглядного представления работы программы. (Рис. 2)

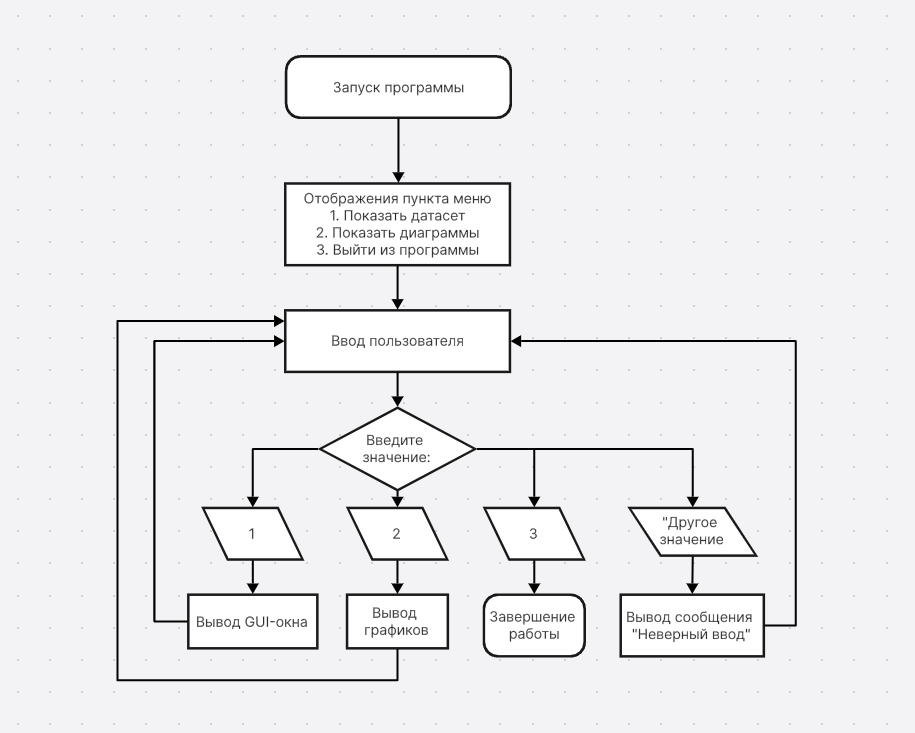


Рис. 2 – Блок-схема программы.

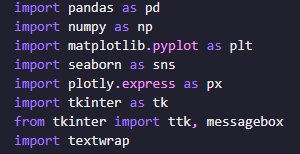


Рис. 3 – Импорт библиотек.

**2)** Попытка загрузки CSV-файла с последующей предобработкой. При ошибке загрузки выводится соответствующее сообщение. (Рис. 4)

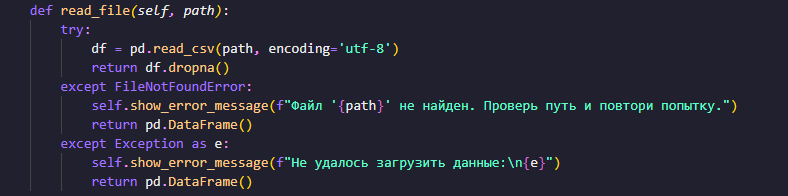


Рис. 4 – Загрузка данных.

**3)** Визуализация данных. Каждый блок визуализации создаёт график, который отражает зависимость определённого фактора на уровень сна.

**3.1)** Зависимость кол-во шагов, пройденных за сутки в созависимости от времени сна, на качество сна. Строится график роя в зависимости от рассматриваемых данных (Рис. 5)

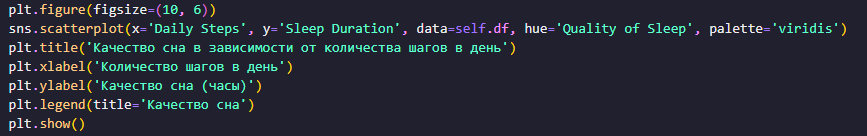


Рис. 5 – Продолжительность сна в зависимости от шагов в день.

**3.2)** Зависимость уровня стресса на уровень качество сна. Строится столбчатая диаграмма зависимости уровня стресса (0-10) на уровень качества сна. (Рис. 6)

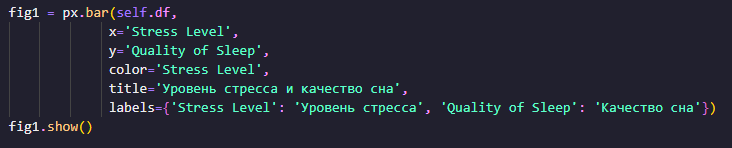


Рис. 6 – Зависимость уровня стресса на уровень качество сна.

**3.3)** Оценка влияния гендера на качество сна. Строится скрипичная диаграмма зависимости гендера на уровень качество сна. (Рис. 7)

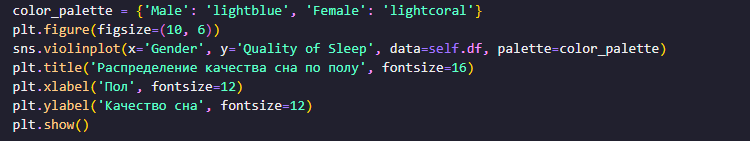


Рис. 7 – Оценка влияния гендера на качество сна.

**3.4)** Для построения следующей диаграммы необходима нормализация данных по общему значению (100). После нормализации строится круговая диаграмма, где каждый сектор отвечает за отдельный рассматриваемый параметр. (Рис. 8)



Рис. 8 – Влияние учёбы на оценку за экзамен.

**4)** Реализация меню программы. (Рис. 9)

Опции меню:

1. Показать датасет

2. Показать диаграммы

3. Выйти из программы

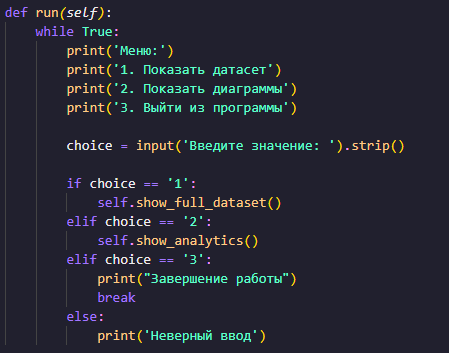


Рис. 9 – Влияние учёбы на оценку за экзамен.

**7)** Запуск программы. (Рис. 10)

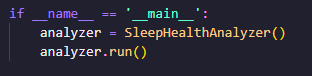


Рис. 10

***2.5 Результаты работы кода***

При запуске программа выводит меню с тремя различными действиями.

(Рис. 11)

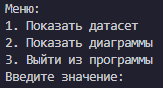


Рис. 11 – Интерфейс меню.

При выборе 1 пункта меню (Показать датасет) программа выводит датасет в виде таблицы. (Рис. 12)

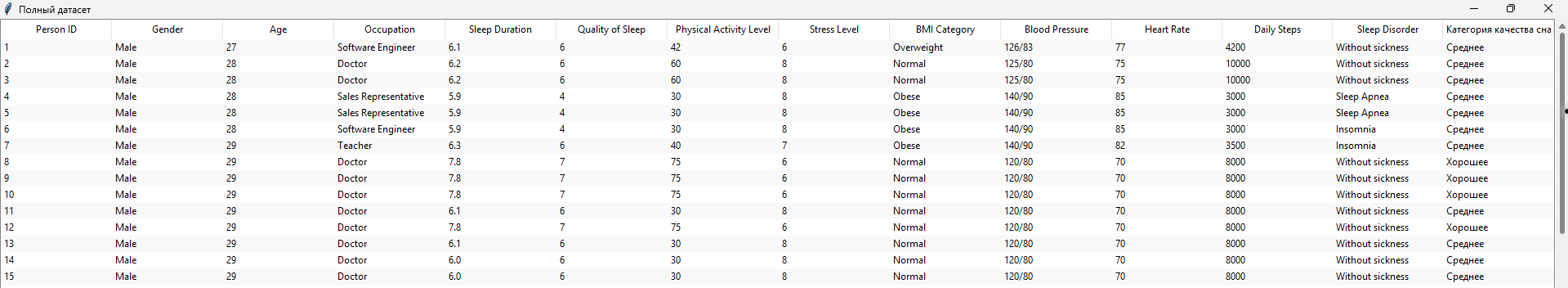


Рис. 12 – Датасет

При выборе 2 пункта меню (Показать диаграммы) программа выводит 4 графика:

**1)** На графике (Рис. 13) представлена зависимость между уровнем стресса и качеством сна. Чётко прослеживается отрицательная корреляция: чем выше уровень стресса, тем ниже качество сна. У респондентов с минимальным стрессом (1-3 балла) качество сна достигает 8-9 баллов, тогда как при высоком стрессе (7-10 баллов) этот показатель падает до 4-5 баллов. Это подтверждает важность управления стрессом для улучшения сна.

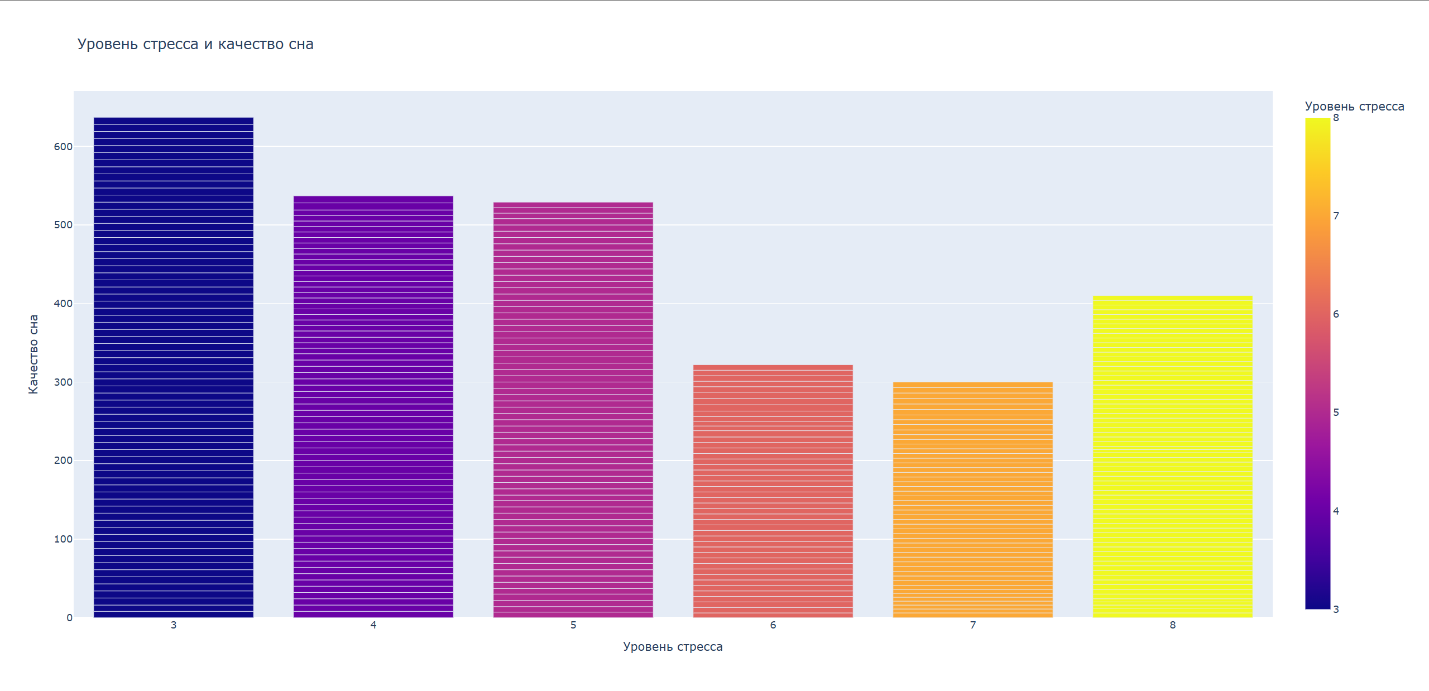


Рис. 13 – Влияние уровня стресса на качество сна.

**2)** График (Рис. 14) демонстрирует связь между ежедневной физической активностью (количество шагов) и качеством сна. Наблюдается положительная тенденция: у людей, проходящих более 8000 шагов в день, качество сна в среднем выше (7-8 баллов), чем у тех, кто ведёт малоподвижный образ жизни (менее 5000 шагов – 5-6 баллов). Это подчёркивает роль физической активности в регуляции сна.

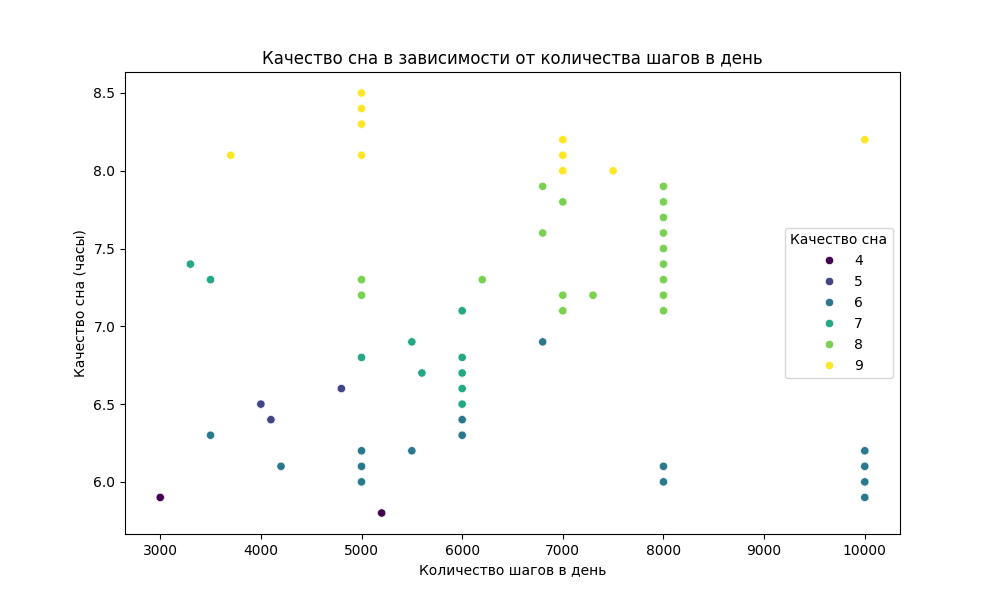


Рис. 14 – Влияние сна на оценку.

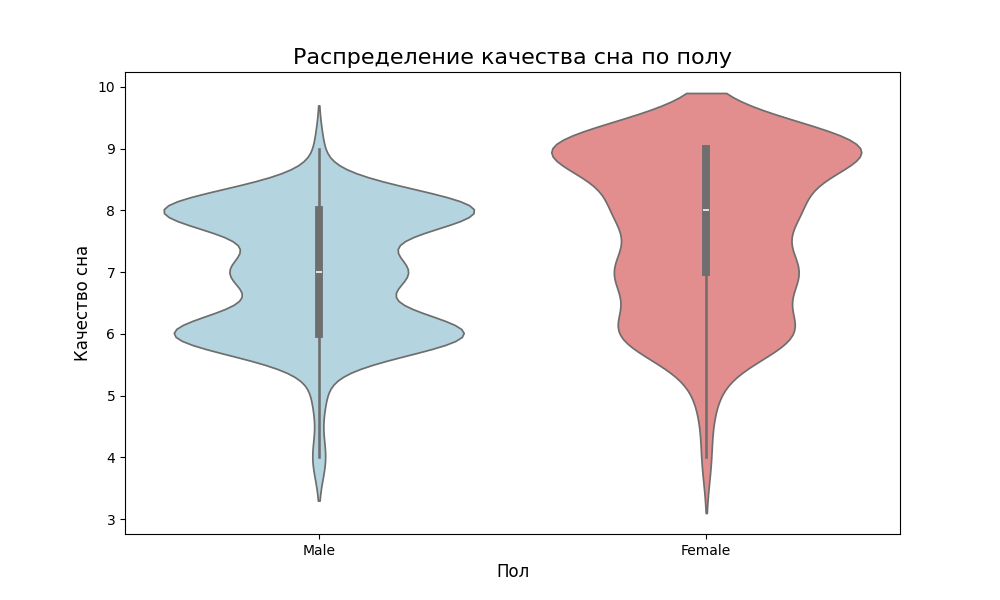
**3)** На графике (Рис. 15) сравнивается качество сна у мужчин (Male) и женщин (Female). Средние значения у женщин несколько выше (6.8 балла против 6.3 у мужчин), что может быть связано с биологическими или поведенческими различиями. Однако разница незначительна, что указывает на необходимость дальнейшего изучения факторов, влияющих на сон у разных полов.

Рис. 15 – Влияние соцсетей на оценку.

**4)** График (Рис. 16) отображает нормализованные средние значения ключевых признаков:

* **Качество сна (66.3)** – один из самых высоких показателей, что подчёркивает его важность.
* **Физическая активность (48.6)** и **уровень стресса (47.7)** находятся на среднем уровне, указывая на потенциал для улучшения.
* **Частота сердечных сокращений (24.6)** – самый низкий показатель, что требует дополнительного анализа.



Рис. 16 – Влияние соцсетей на оценку.

**3. Заключение.**

Проведенный анализ данных о сне и образе жизни позволил выявить ключевые закономерности, влияющие на качество сна и общее состояние здоровья. Разработанный программный код обеспечил комплексную обработку данных и визуализацию полученных результатов.

1) Основные выводы исследования свидетельствуют, что физическая активность и уровень стресса являются определяющими факторами качества сна. Установлена четкая зависимость: увеличение ежедневной активности до 8000 шагов и более способствует улучшению показателей сна, тогда как высокий уровень стресса закономерно приводит к его ухудшению. Эти результаты подтверждают важность сбалансированного образа жизни для поддержания здорового сна.

2) Нормализованные данные продемонстрировали, что среди всех рассматриваемых параметров именно качество сна имеет наивысший показатель значимости. При этом такие факторы как физическая активность и уровень стресса, несмотря на их средние значения, остаются важными переменными, влияющими на общее состояние здоровья.

3) Полученные результаты создают основу для дальнейших исследований в области изучения факторов, влияющих на качество сна и здоровье человека. Выявленные закономерности подчеркивают комплексный характер взаимодействия различных аспектов образа жизни и их влияние на физиологическое состояние организма.

**4. Распределение задач в проекте.**

Таблица 1. Распределение обязанностей.

|  |  |  |
| --- | --- | --- |
| **№ п.п.** | **Вид работ** | **Ответственный** |
| 1. | Поиск и выбор датасета | Альзинский Е. Д. Ибраев Д. Б. |
| 2. | Изучение библиотек pandas и textwrap, numpy | Альзинский Е. Д. |
| 3. | Изучение библиотек matplotlib и tkinter, seaborn, plotly.express | Ибраев Д. Б. |
| 4. | Предобработка данных датасета | Альзинский Е. Д. |
| 5. | Визуализация работы программы в виде блок-схемы | Альзинский Е. Д. |
| 6. | Разработка вывода датасета в виде таблицы через tkinter | Ибраев Д. Б. |
| 7. | Визуализация графиков для аналитики | Ибраев Д. Б. |
| 8. | Оформление отчёта | Альзинский Е. Д. |

**5. Список литературы.**

5.1 Sleep Health and Lifestyle Dataset [Электронный ресурс] набор данных / Kaggle. – URL: https://www.kaggle.com/datasets/uom190346a/sleep-health-and-lifestyle-dataset/data (дата обращения: 24.05.2025).

5.2 textwrap: все, что можно сделать с помощью модуля textwrap в Python. – https://uproger.com/vse-chto-mozhno-sdelat-s-pomoshhyu-modulya-textwrap-v-python/ (дата обращения: 31.05.2025).

5.3 pandas: официальная документация. – URL: https://pandas.pydata.org/docs/ (дата обращения: 31.05.2025).

5.4 Matplotlib: официальная документация. – URL: https://matplotlib.org/stable/contents.html (дата обращения: 31.05.2025).

5.5 tkinter: справочник по библиотеке / Python Software Foundation. – URL: https://docs.python.org/3/library/tkinter.html (дата обращения: 31.05.2025).

5.6 seaborn: как строить красивые графики на Python с Seaborn. – URL: https://habr.com/ru/companies/otus/articles/540526/ (дата обращения: 31.05.2025).

5.7 csv: модуль для работы с CSV-файлами. – URL: https://docs.python.org/3/library/csv.html (дата обращения: 31.05.2025).

5.8 plotly.express: офицальный сайт библиотеки. – URL: https://plotly.com/python/plotly-express/ (дата обращения: 31.05.2025).

**6. Приложение.**

**Код программы:**

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

import plotly.express as px

import tkinter as tk

from tkinter import ttk, messagebox

import textwrap

class SleepHealthAnalyzer:

def \_\_init\_\_(self, path="Sleep\_health\_and\_lifestyle\_dataset.csv"):

self.df = self.read\_file(path)

self.setup\_quality\_categories()

def read\_file(self, path):

try:

df = pd.read\_csv(path, encoding='utf-8')

return df.dropna()

except FileNotFoundError:

self.show\_error\_message(f"Файл '{path}' не найден. Проверь путь и повтори попытку.")

return pd.DataFrame()

except Exception as e:

self.show\_error\_message(f"Не удалось загрузить данные:\n{e}")

return pd.DataFrame()

def setup\_quality\_categories(self):

if not self.df.empty:

sleep\_quality\_bins = [0, 3, 6, 10]

sleep\_quality\_labels = ['Плохое', 'Среднее', 'Хорошее']

self.df['Категория качества сна'] = pd.cut(

self.df['Quality of Sleep'],

bins=sleep\_quality\_bins,

labels=sleep\_quality\_labels,

include\_lowest=True

)

def show\_error\_message(self, message):

messagebox.showerror("Ошибка", textwrap.fill(message, width=60))

def show\_full\_dataset(self):

if self.df.empty:

self.show\_error\_message("Нет данных для отображения")

return

root = tk.Tk()

root.title('Полный датасет')

root.geometry('1000x600')

tree = ttk.Treeview(root, columns=list(self.df.columns), show='headings')

vsb = ttk.Scrollbar(root, orient='vertical', command=tree.yview)

hsb = ttk.Scrollbar(root, orient='horizontal', command=tree.xview)

tree.configure(yscrollcommand=vsb.set, xscrollcommand=hsb.set)

for col in self.df.columns:

tree.heading(col, text=col)

tree.column(col, width=120, anchor='w')

for idx, row in enumerate(self.df.head(200).itertuples(index=False), start=1):

tree.insert('', 'end', values=row, tags=('evenrow' if idx % 2 == 0 else 'oddrow',))

tree.tag\_configure('oddrow', background='#f9f9f9')

tree.tag\_configure('evenrow', background='#ffffff')

vsb.pack(side='right', fill='y')

hsb.pack(side='bottom', fill='x')

tree.pack(fill='both', expand=True)

root.mainloop()

def show\_analytics(self):

if self.df.empty:

messagebox.showerror("Ошибка", "Нет данных для анализа")

return

fig1 = px.bar(self.df,

x='Stress Level',

y='Quality of Sleep',

color='Stress Level',

title='Уровень стресса и качество сна',

labels={'Stress Level': 'Уровень стресса', 'Quality of Sleep': 'Качество сна'})

fig1.show()

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

sns.scatterplot(x='Daily Steps', y='Sleep Duration', data=self.df, hue='Quality of Sleep', palette='viridis')

plt.title('Качество сна в зависимости от количества шагов в день')

plt.xlabel('Количество шагов в день')

plt.ylabel('Качество сна (часы)')

plt.legend(title='Качество сна')

plt.show()

color\_palette = {'Male': 'lightblue', 'Female': 'lightcoral'}

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

sns.violinplot(x='Gender', y='Quality of Sleep', data=self.df, palette=color\_palette)

plt.title('Распределение качества сна по полу', fontsize=16)

plt.xlabel('Пол', fontsize=12)

plt.ylabel('Качество сна', fontsize=12)

plt.show()

categories = ['Sleep Duration', 'Quality of Sleep', 'Physical Activity Level', 'Stress Level', 'Heart Rate', 'Daily Steps']

colors = ["#59a6dd", "#ffa75b", "#80f580", "#f76666", "#c792f8", "#fd9d89"]

normalized\_values = []

for category in categories:

max\_val = self.df[category].max()

min\_val = self.df[category].min()

norm\_val = ((self.df[category].mean() - min\_val) / (max\_val - min\_val)) \* 100

normalized\_values.append(norm\_val)

width = 2 \* np.pi / len(categories)

fig, ax = plt.subplots(figsize=(8, 8), subplot\_kw=dict(polar=True))

angles = np.arange(len(categories)) \* width

bars = ax.bar(angles, normalized\_values, width=width, color=colors, alpha=1, edgecolor='black', linewidth=1.5)

circle\_angles = np.linspace(0, 2 \* np.pi, 100)

circle\_radius = np.full\_like(circle\_angles, 100)

ax.plot(circle\_angles, circle\_radius, linestyle='--', color='grey', linewidth=1)

ax.text(np.pi/2, 105, '100 — максимум', ha='center', va='bottom', fontsize=9, color='grey')

for i, (label, angle) in enumerate(zip(categories, angles)):

x = angle

y = normalized\_values[i] + 5

ax.text(x, y, f"{label}\n{normalized\_values[i]:.1f}", ha='center', va='center', fontsize=8, color='black')

ax.set\_theta\_offset(np.pi / 2)

ax.set\_theta\_direction(-1)

ax.set\_ylim(0, 110)

ax.set\_yticks([])

ax.set\_xticks([])

ax.spines['polar'].set\_visible(False)

plt.title('Нормализованные средние значения признаков (100 = максимум)')

plt.tight\_layout()

plt.show()

def run(self):

while True:

print('Меню:')

print('1. Показать датасет')

print('2. Показать диаграммы')

print('3. Выйти из программы')

choice = input('Введите значение: ').strip()

if choice == '1':

self.show\_full\_dataset()

elif choice == '2':

self.show\_analytics()

elif choice == '3':

print("Завершение работы")

break

else:

print('Неверный ввод')

if \_\_name\_\_ == '\_\_main\_\_':

analyzer = SleepHealthAnalyzer()

analyzer.run()