

Поймай баланс

если сможешь

Команда № 22

PIECE of DATA

Кондакова Капитолина

Бузмакова Кира

Васина Екатерина

Дьяков Иван

Толмачев Никита

Миров Иван

Структура базы данных

“Как быть успешным в учёбе”

Качественные:

- degree
- sex
- stress
- coping_x
- yesterday_typical
- enough_x

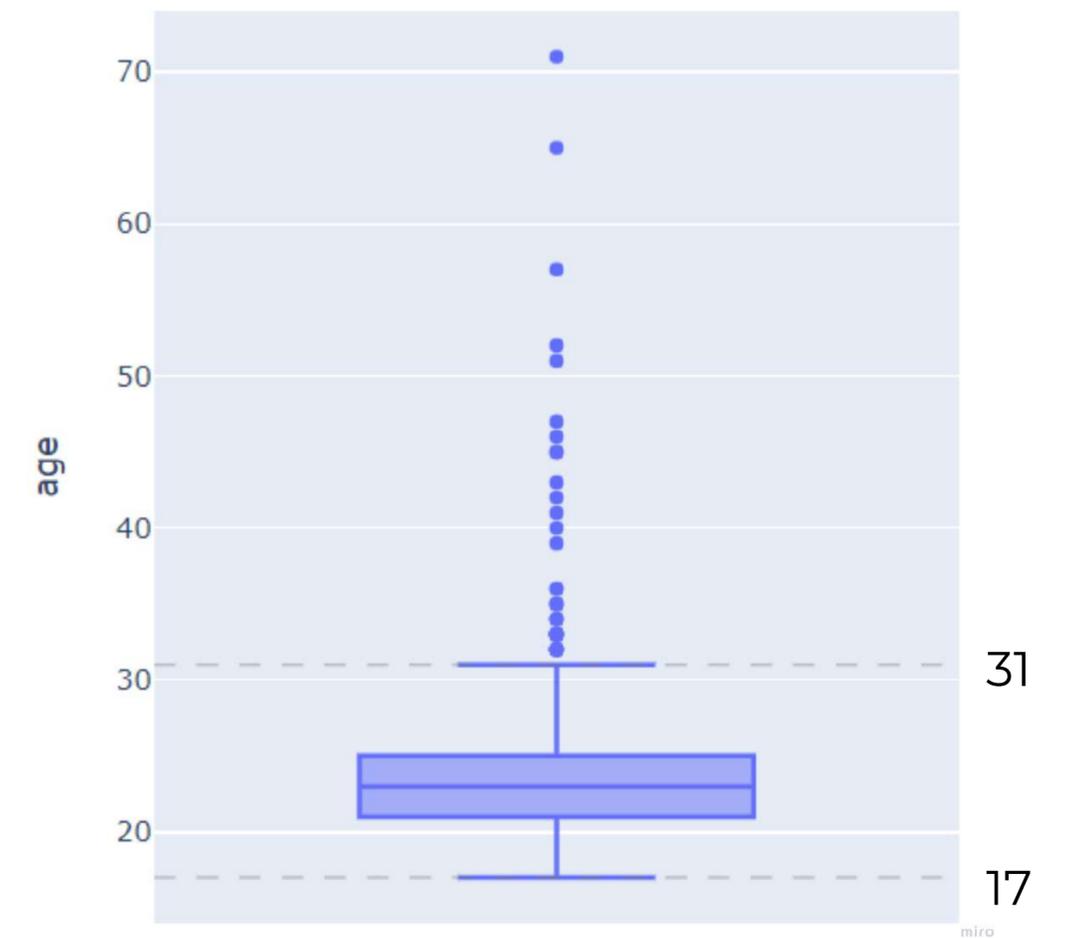
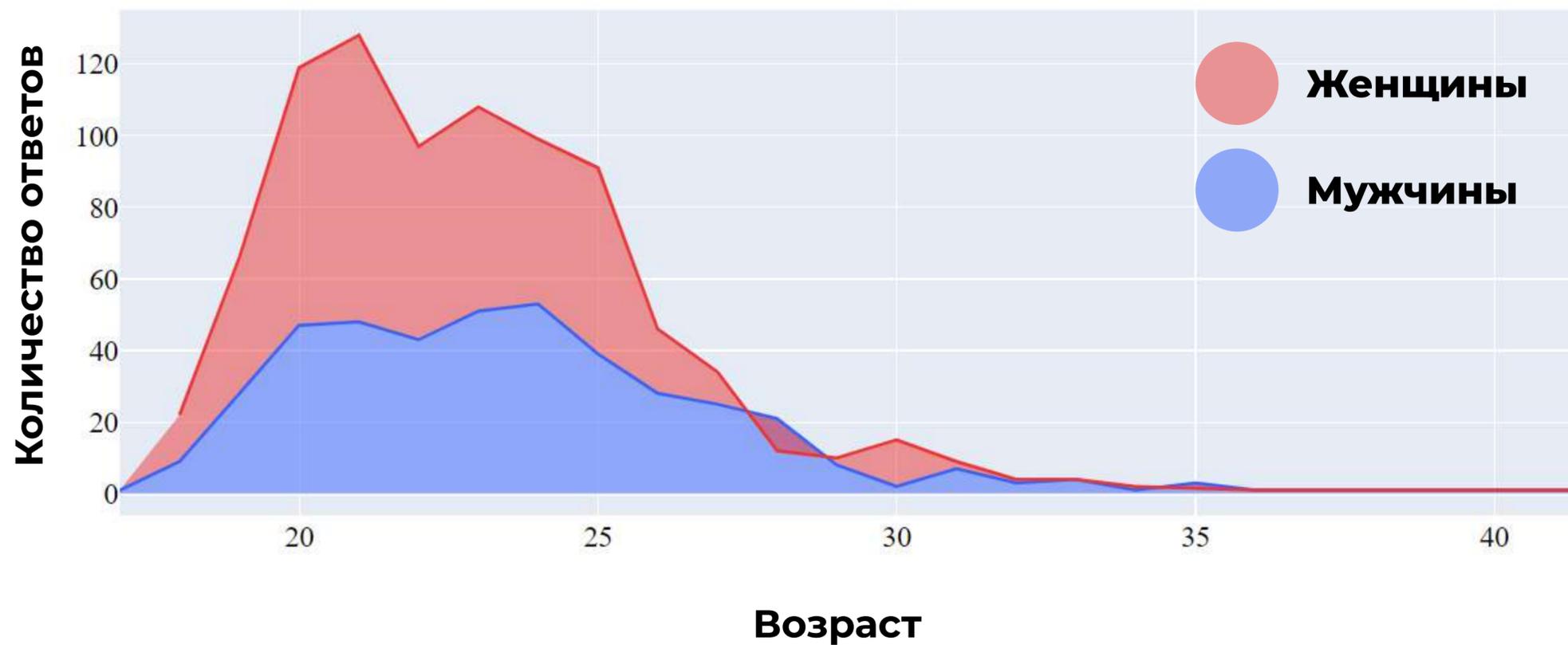
Количественные:

- gpa
- semester
- age
- survey_time
- source_x
- week_x
- yesterday_x

- 1559 опрашиваемых
- 42 критерия
- NAN присутствует

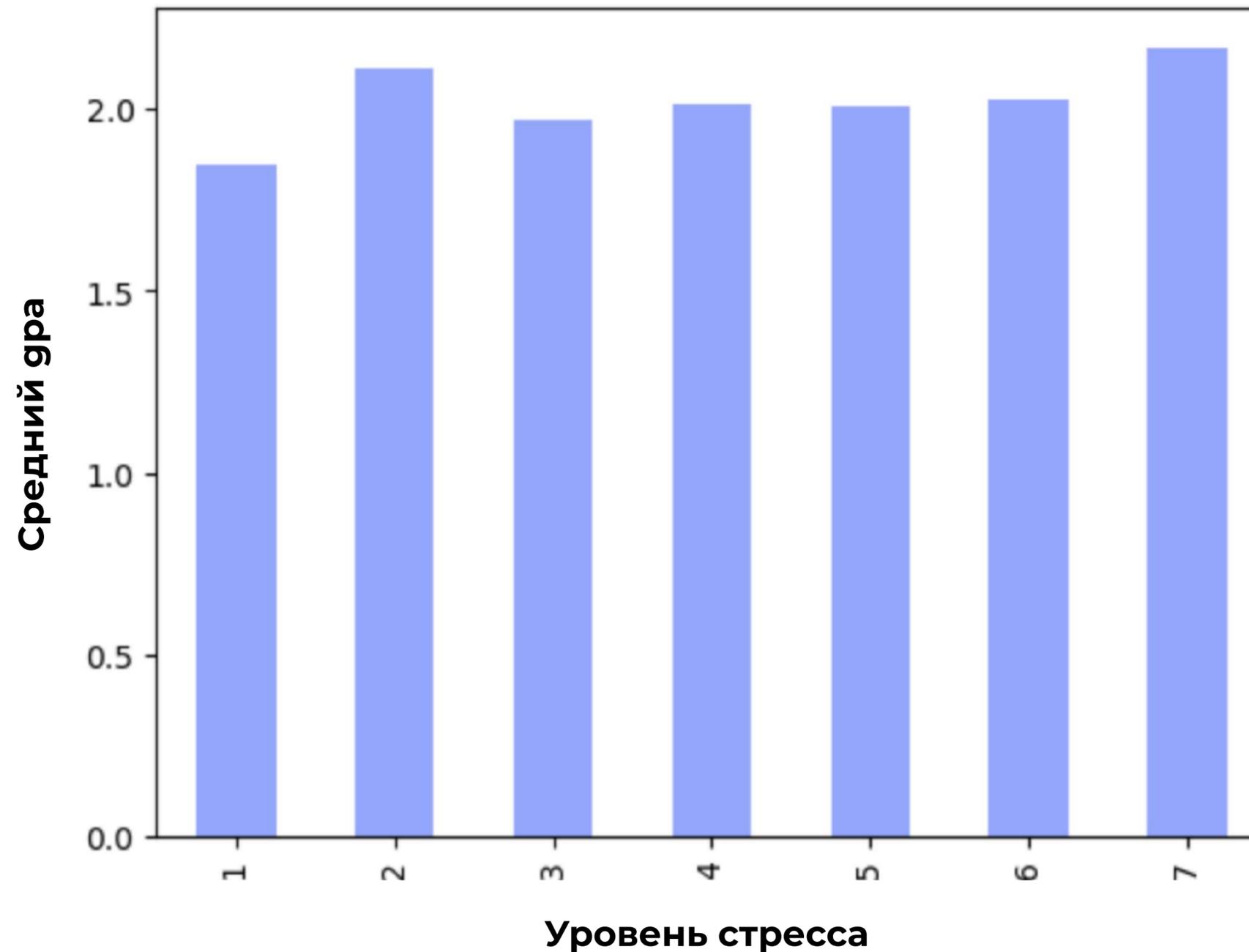
Предварительный анализ

Распределение по возрасту и полу

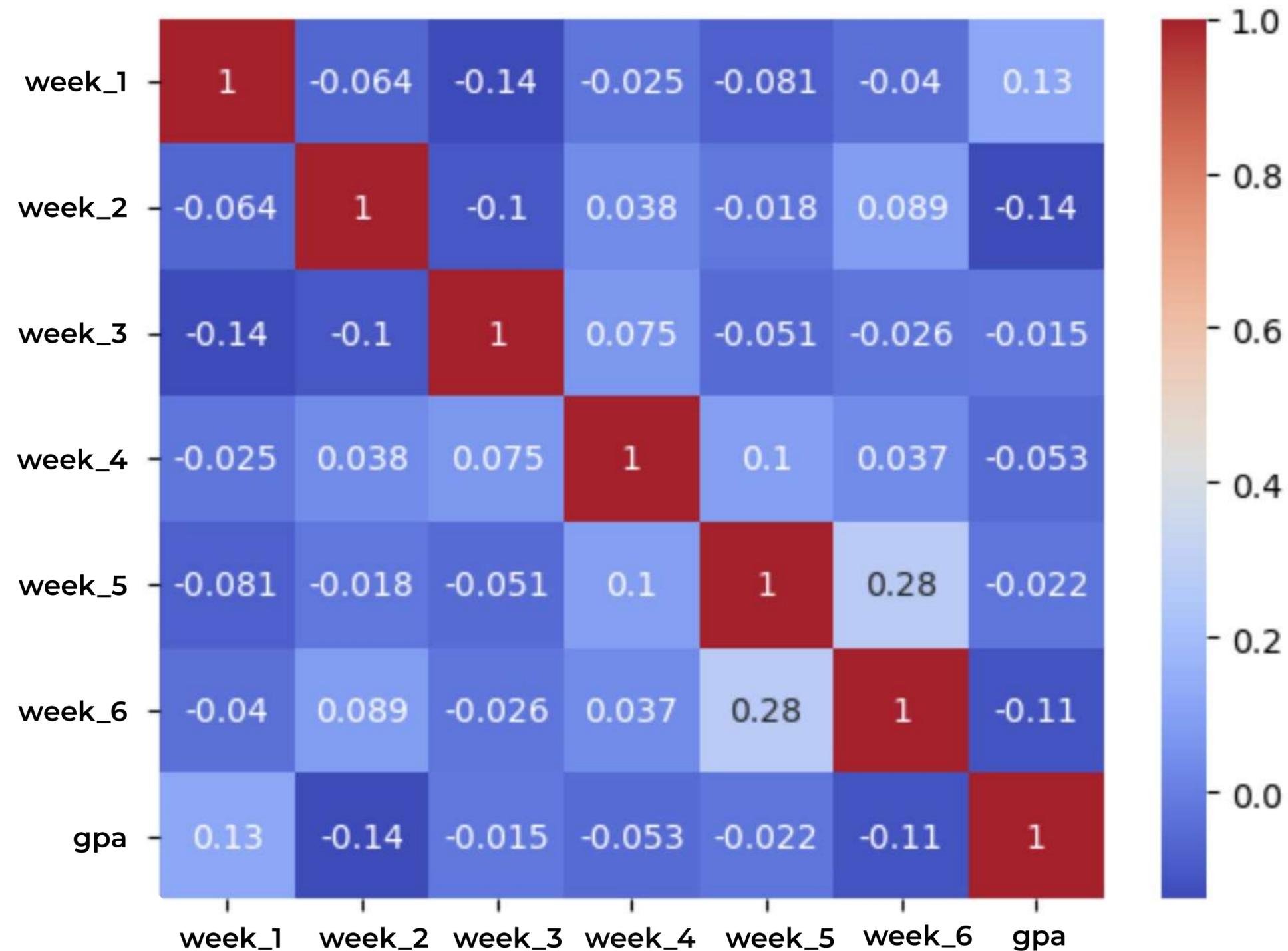


Предварительный анализ

Средний гра по уровню стресса



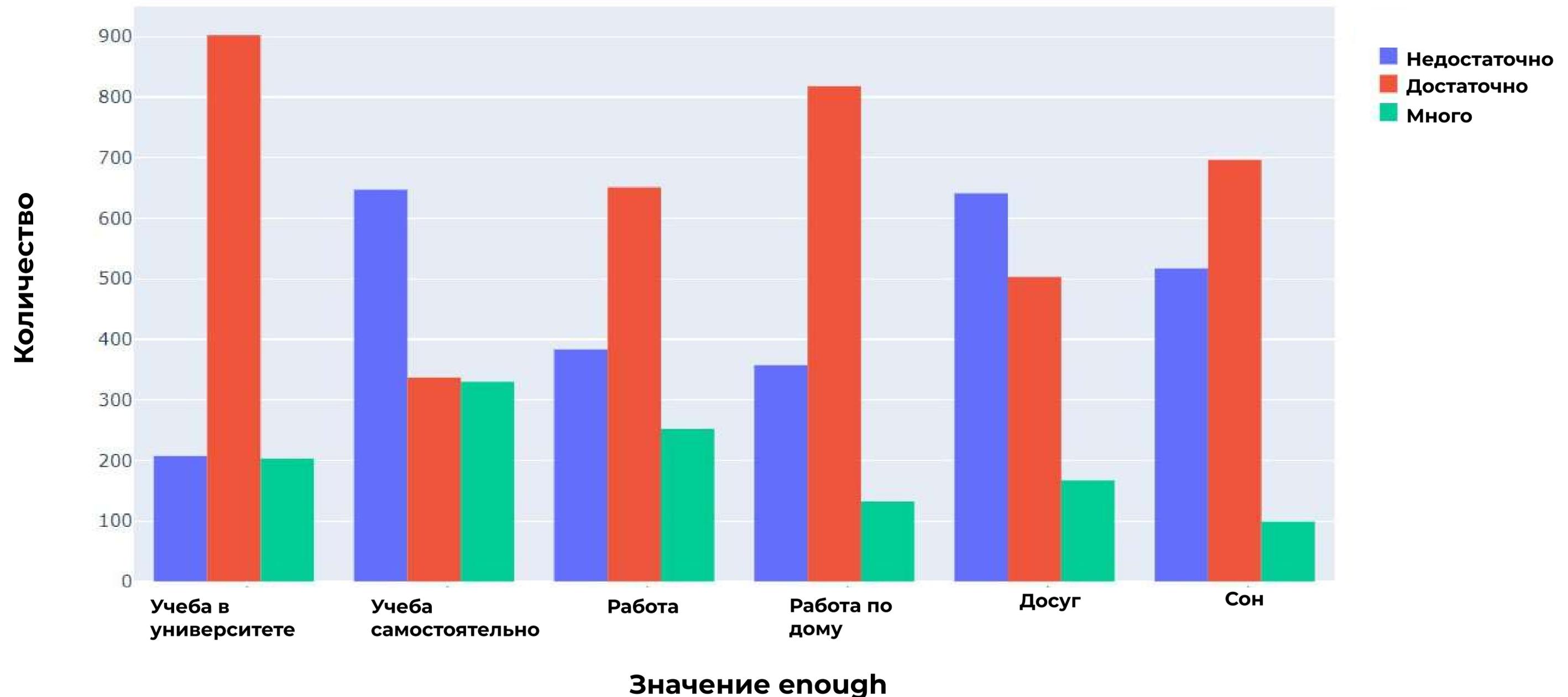
Предварительный анализ



Предварительный анализ

Считает ли студент, что уделяет достаточно времени следующим видам деятельности?

Распределение переменных enough



Подготовка базы данных

Объединение/добавление переменных:

week_x; yesterday_x

1. Объединение параметров времени по учёбе за неделю и за вчерашний день

yesterday_typical

2. Бинарное представление типичности вчерашнего дня

enough_x

3. Объединение схожих видов деятельности по данному параметру

week_day_x

4. Среднее время, потраченное за день, на различные виды деятельности : среднее время в неделю, деленное на 7

Подготовка базы данных

Чистка данных:

- Студенты, у которых не указана информация ни по одной из переменных
- Неизвестен средний балл
- Отбор по возрасту: от 17 до 31

Размер выборки: 1559



Размер выборки: 653

Исследовательский вопрос и гипотеза

Исследовательский вопрос

Что делает студента успешным в учебе?

Гипотеза

Если **распределение времени** на различные виды деятельности у студента **стабильно**, то это делает его **более успешным** в учебе

Механизм

Студент **стабильно** и **равномерно** распределяет время на различные виды деятельности

Студент тратит примерно **одинаковое количество** часов на подготовку к занятиям

+

Студент соблюдает **study-life balance**

Эффективность учебной деятельности повышается

Средний балл студента выше

Исследуемые показатели

Количество часов, которое студент тратит на каждый из видов деятельности **в неделю**

Количество часов, которое студент потратил на каждый из видов деятельности **вчера**

Типичность вчерашнего дня

Достаточно ли время, студент тратит на различные виды деятельности

Анализ

0.05Принимаемый
уровень значимости

Как измерять стабильность?

Сравниваем распределение переменных `week_day*` и `yesterday**` при условии, что студент ответил, что вчера — типичный день. Для каждого студента рассчитывается U-критерий Манна-Уитни.

H0

Нет различий между средним в день распределением видов деятельности и распределением в предыдущий день

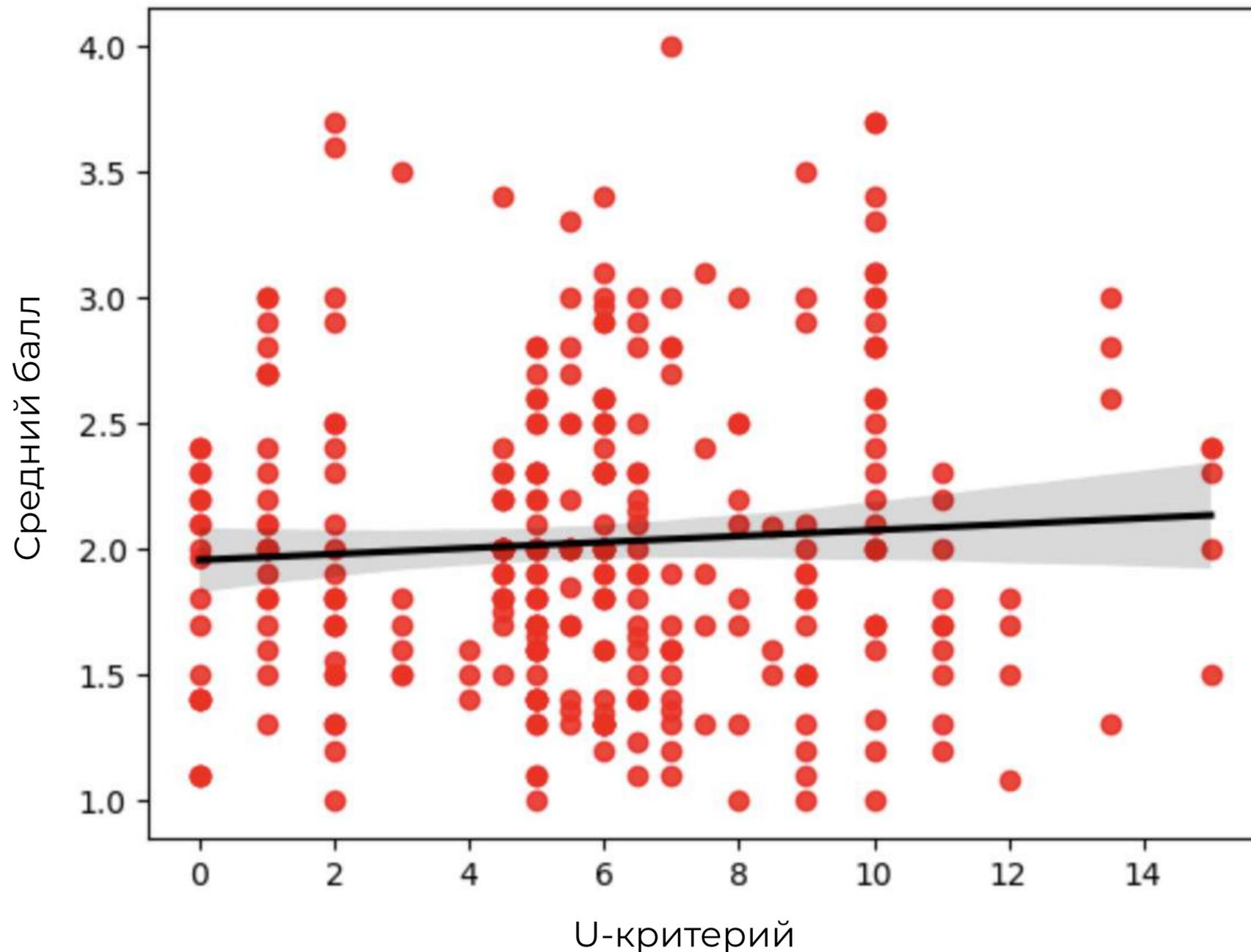
H1

Есть различия между средним в день распределением видов деятельности и распределением в предыдущий день

*`Week_day` - часов в среднем в день на различные виды деятельности

**`Yesterday` - часов в среднем вчера на различные виды деятельности

Математическое подтверждение



Объясняющая переменная -
У-критерий Манна-Уитни

Объясняемая переменная -
средний балл

КОРРЕЛЯЦИЯ

0.07

P-VALUE

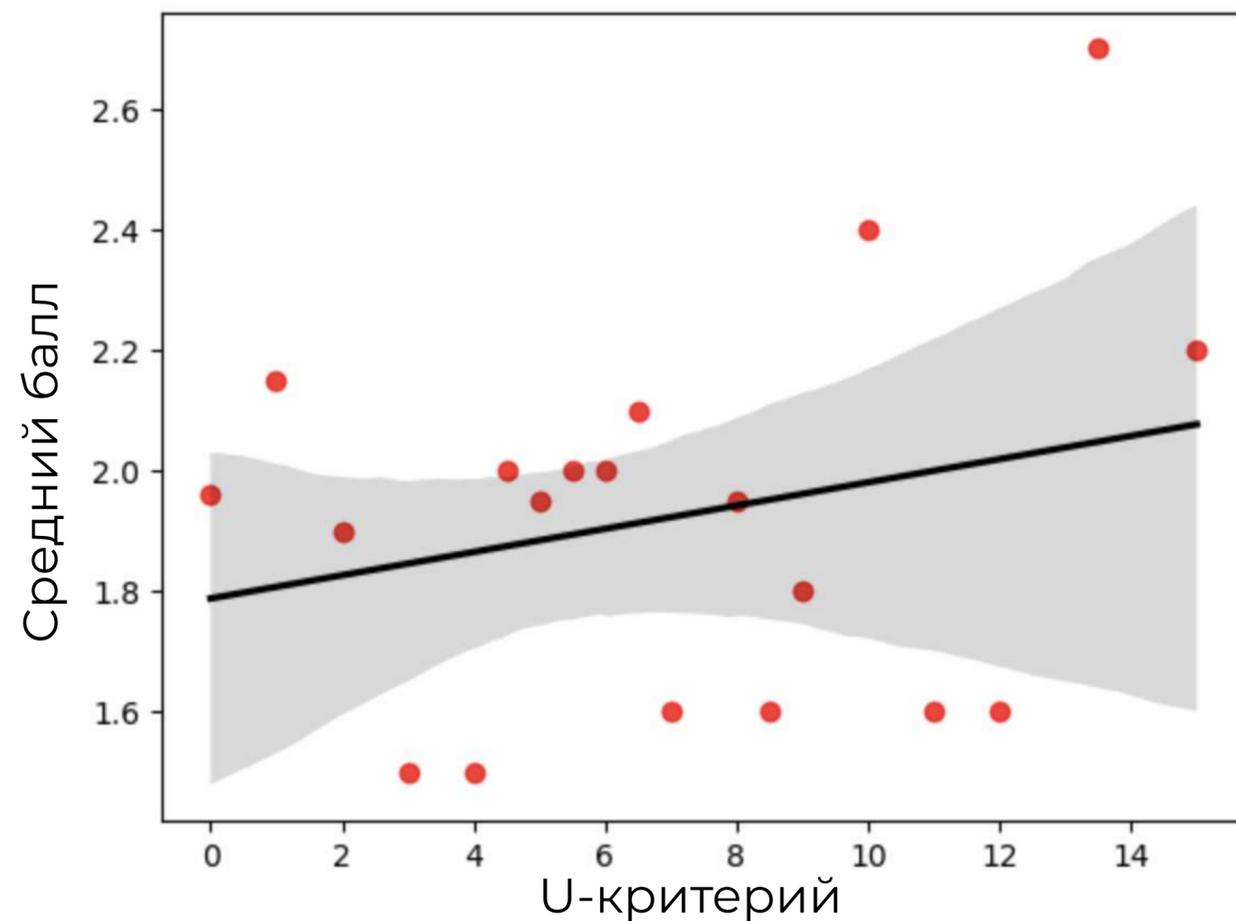
0.24

Математическое подтверждение.

Подвыборки. Отдых

Как вы считаете, достаточно ли вы тратите времени на отдых, досуг и сон?

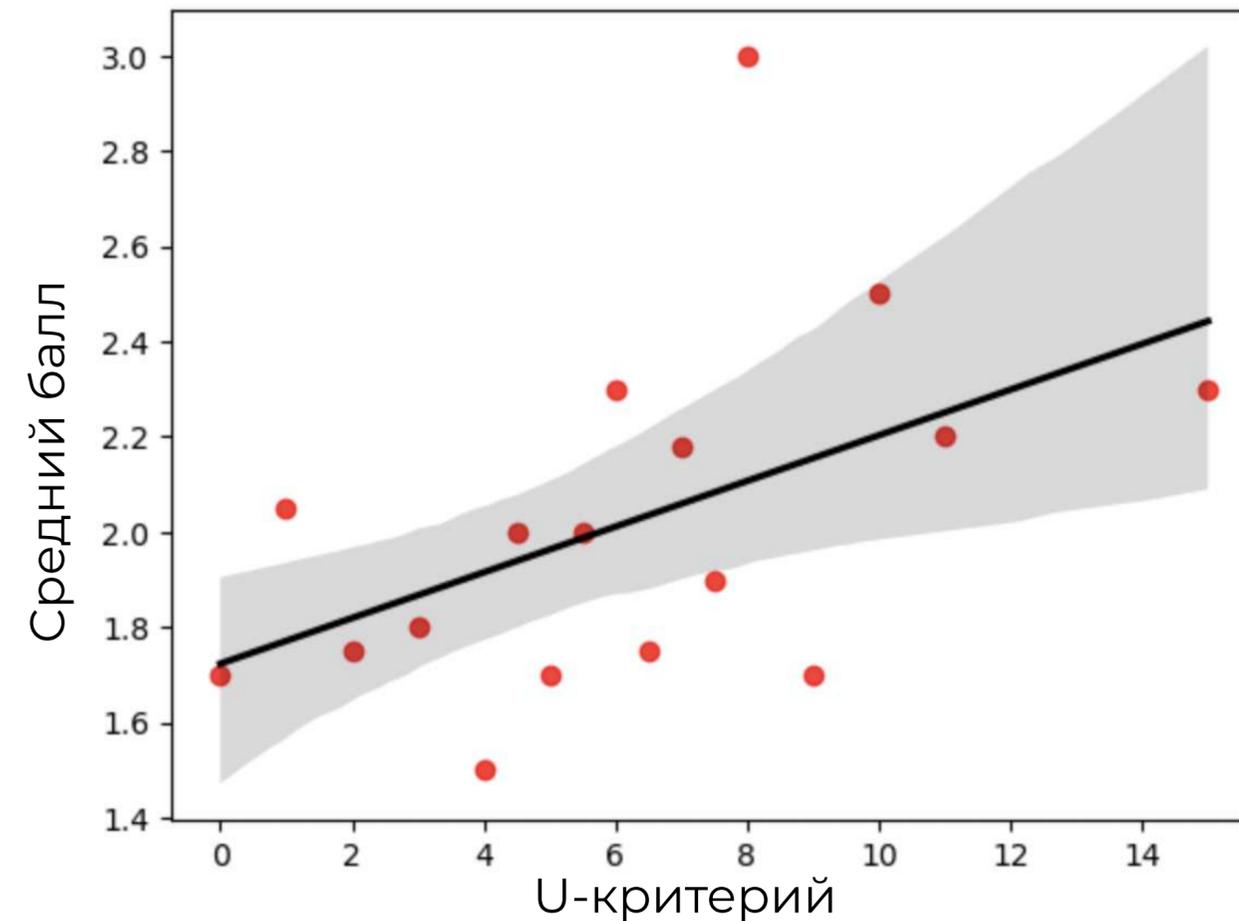
Недостаточно



Cor: 0.25

P-value: 0.3

Достаточно



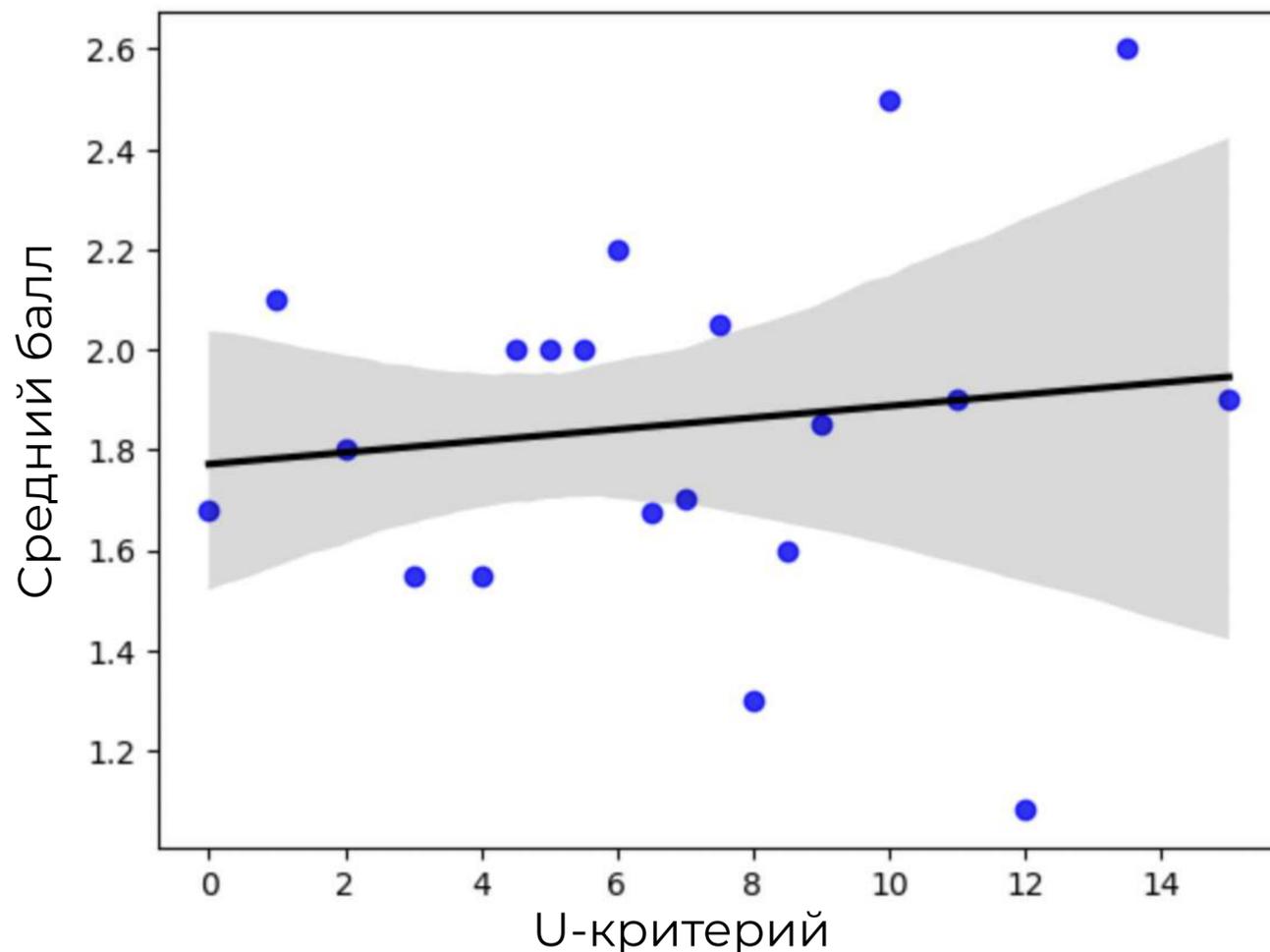
Cor: 0.49

P-value: 0.04

Математическое подтверждение. Подвыборки. Учеба

Как вы считаете, достаточно ли вы тратите времени на учебу?

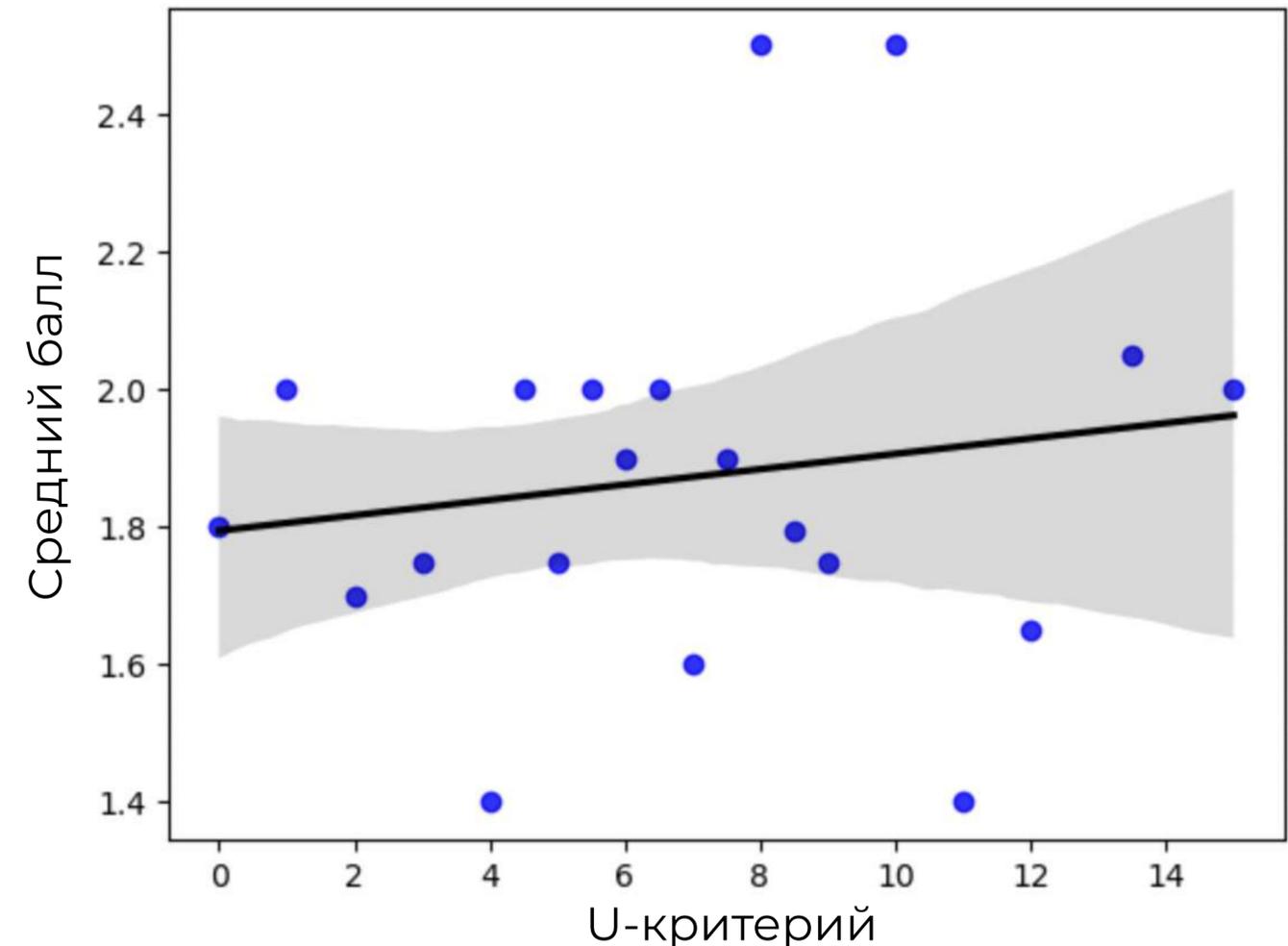
Недостаточно



Cor: 0.13

P-value: 0.59

Достаточно



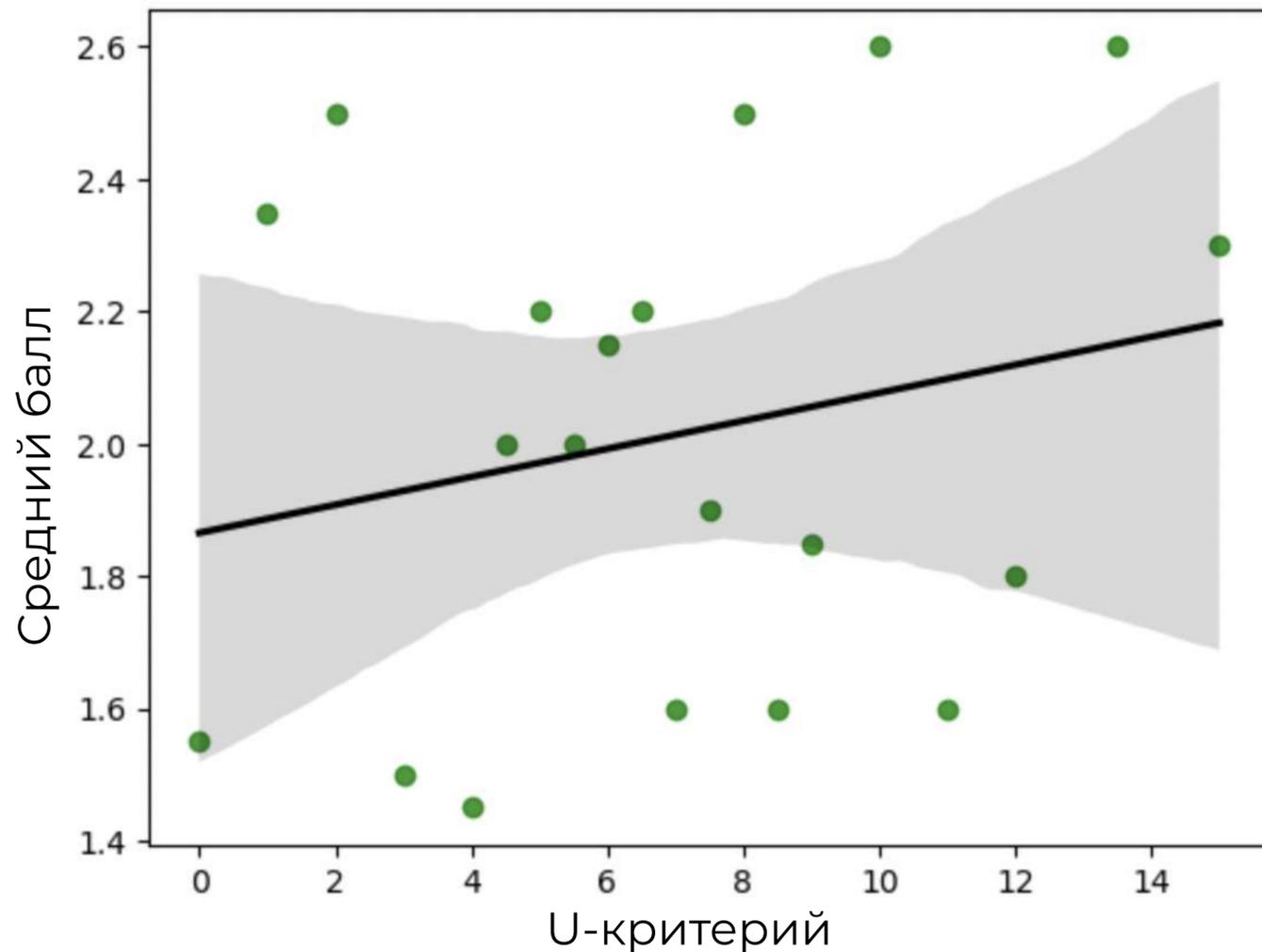
Cor: 0.16

P-value: 0.5

Математическое подтверждение. Подвыборки. Работа

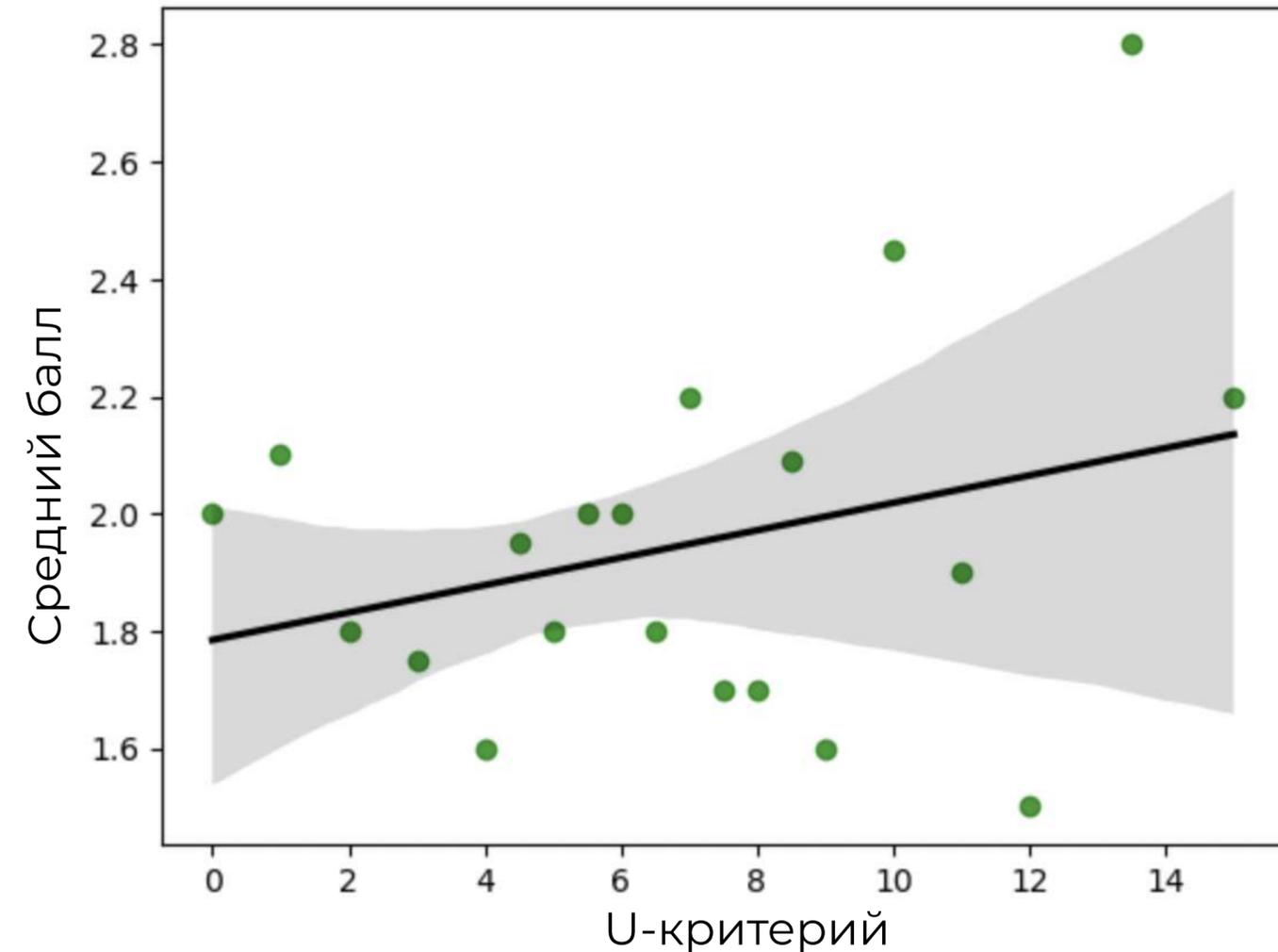
Как вы считаете, достаточно ли вы тратите времени на работу?

Недостаточно



Cor: 0.22 P-value: 0.35

Достаточно



Cor: 0.31 P-value: 0.19

Интерпретация



Гипотеза не подтвердилась на общем уровне, но подтвердилась на одной из подвыборок

Студенты, считающие, что достаточно времени тратят на отдых и сон и распределяющие свое время более стабильно, учатся успешнее

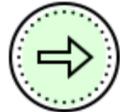
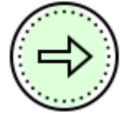
Ограничения

-  Маленький объем выборки, основанная только на студентах одного вуза
-  Субъективность ответов
-  Особенности 2020 года
-  Нет информации о дне прохождения опроса
-  Лишь один фактор для определения успешности студента(гра)

Перспективы

-  Собрать данные о студентах других вузов/других стран

Улучшить качество опроса:

-  Подтвердить достоверность ответов студентов
-  Провести подобные исследования несколько раз в течение года/ в разные времена года

Практическая польза

- 1 **Стабильность** в распределении времени для студентов
- 2 Понимание **значения** отдыха для студентов

Рекомендации

Адаптация принципов распределения времени под разные группы* студентов

*на основе удовлетворенности количеством времени, которое студент тратит на различные виды деятельности

Популяризация важности отдыха и study-life balance для повышения успеха студентов в учебе



Piece of Data



Дьяков Иван
визуализатор



Толмачев Никита
программист



Миров Иван
исследователь ресурсов



Кондакова Капитолина
мат. модель



Бузмакова Кира
data-аналитик



Васина Екатерина
бизнес-аналитик

Приложение

degree distribution

