|  |
| --- |
|  |
| МИНОБРНАУКИ РОССИИ  Федеральное государственное бюджетное образовательное учреждение высшего образования  **«МИРЭА**  **Российский технологический университет»**  **РТУ МИРЭА** |

**Институт информационных технологий (ИИТ) Кафедра прикладной математики (ПМ)**

## ОТЧЕТ ПО ПРАКТИЧЕСКОЙ РАБОТЕ №5

по дисциплине «Языки программирования для статистической обработки данных»

|  |  |
| --- | --- |
| Студент группы | *ИМБО-11-23 Журавлев Ф. А.* |
|  | (подпись) |
| Преподаватель | *Трушин С. М.* |
|  | (подпись) |

Москва 2025 г.

# 1 ЦЕЛЬ И ЗАДАЧИ

## Цель практической работы:

Освоить методы анализа корреляций и построения линейных регрессионных моделей с использованием Python, R.

## Задачи практической работы:

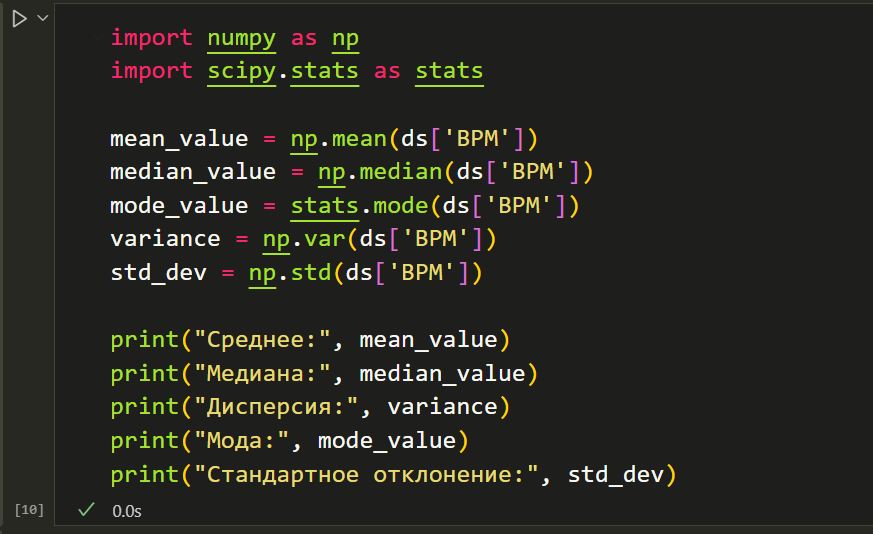
1. Рассчитать корреляцию между переменными:
   * Вычисление коэффициентов корреляции Пирсона и Спирмена.
   * Построение и визуализация корреляционной матрицы.
   * Python: pandas, seaborn.
   * R: cor(), ggcorrplot.
2. Построить линейную регрессионную модель:
   * Анализ зависимости одной переменной от другой.
   * Интерпретация результатов регрессии.
   * Python: statsmodels, sklearn.
   * R: Im().
3. Сравнить результаты между Python, R.

## РЕЗУЛЬТАТЫ ПРАКТИЧЕСКОЙ РАБОТЫ

## Расчет статистических показателей в Python

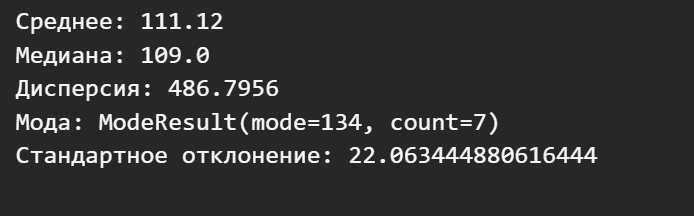
После загрузки данных в Python, напишем коды, которые помогут рас- считать различные статистические показатели.

*рисунок 2.1.1 – стастистических показателей*

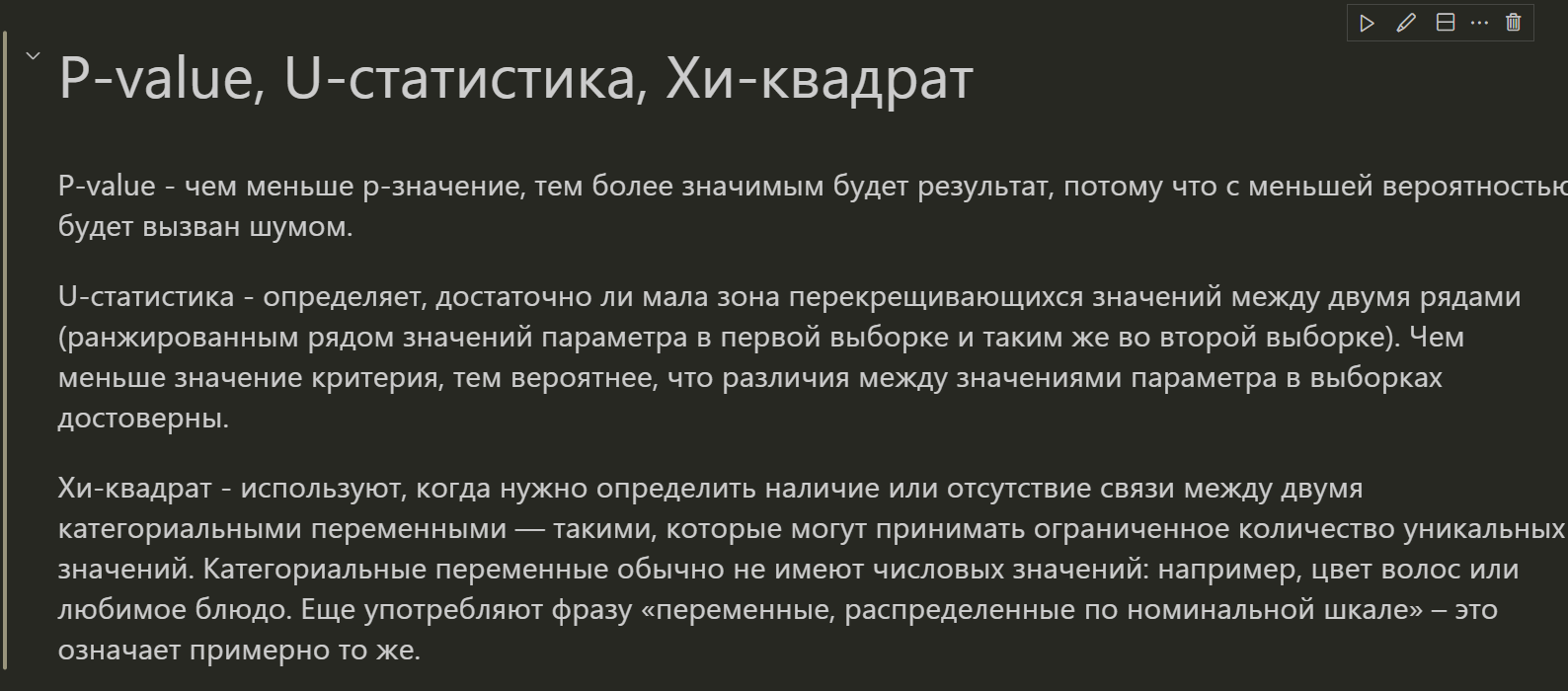
**

Далее посмотрим какие мы получили значения среднего, медианы, дисперсии, моды и стандартного отклонения.

*рисунок 2.1.2 – значения показателей.*

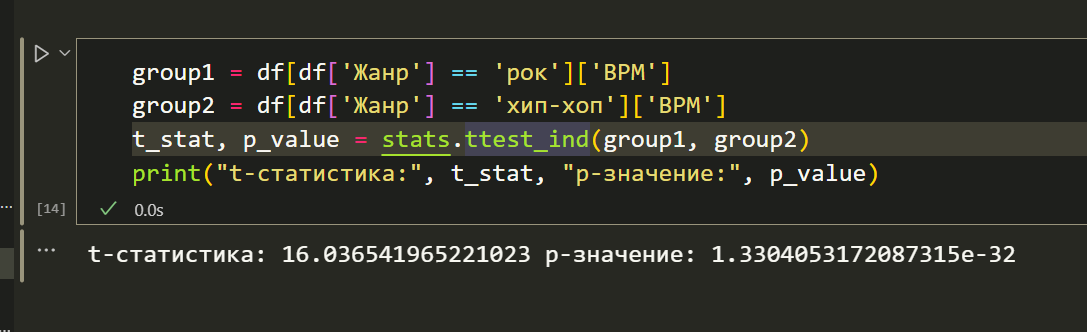
**

Далее рассмотрим код, благодаря которому сможем проанализировать показатели с помощью критерия Мана-Уитни и T-теста.

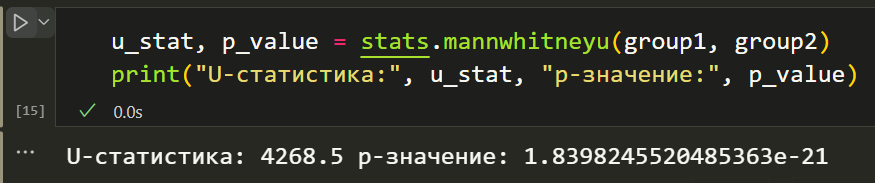




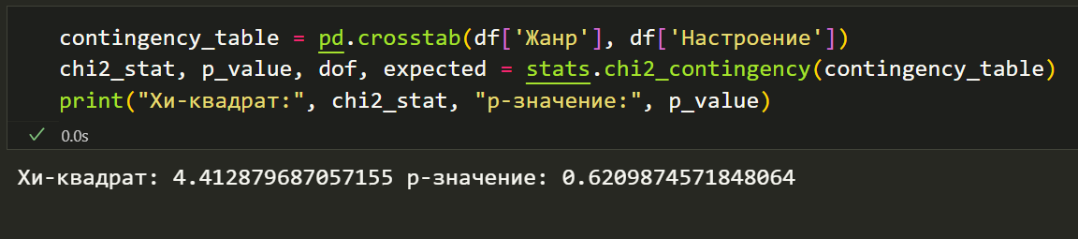
*Рисунок 2.1.3 – код T-test*

**

*Рисунок 2.1.4 – код Ман-Уитни.*

**

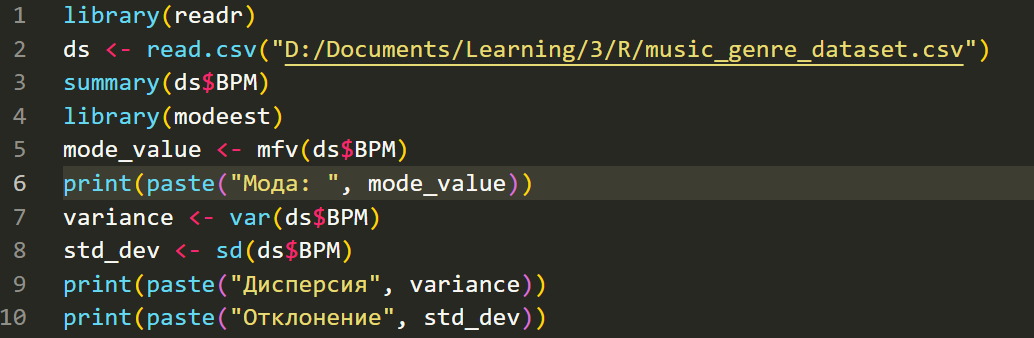
*Рисунок 2.1.5 – Хи-квадрат тест*.



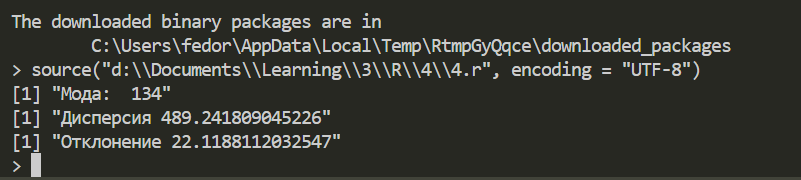
## Статистические данные в R.

Теперь рассмотрим реализацию подобного кода в R.

*Рисунок 2.2.1 – Статистические данные в R*

**

*Рисунок 2.2.2– Вывод статистических данных в R*

**

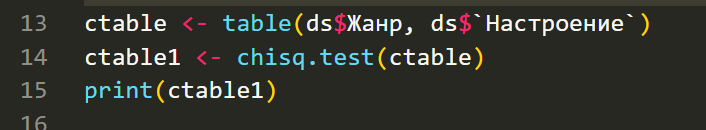
Теперь рассмотрим код, который показывает нам значения всех тестов, что мы использовали ранее. Статистические характеристики t-test, u-статистику нельзя выполнить на моих данных в R, так как у меня более 2 уникальных категориальных данных (3). Если я запущу код, то получу ошибку:

*Рисунок 2.2.4 – Ошибка*

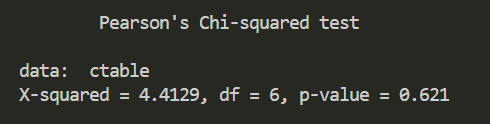
Изображение выглядит как текст, Шрифт, снимок экрана, линия

Автоматически созданное описание

*Рисунок 2.2.4 – Хи-квадрат тест.*

**

*Рисунок 2.2.5 — Хи-квадрат тест, вывод*

**

# ИТОГИ И ВЫВОДЫ:

В результате практической работы были проведены поиски статистических значений с помощью различных функций в обоих языках программирования, а также с помощью разных тестов таких, как t-test, Хи-квадрат тест и критерий Мана-Уитни. Работа со всеми этими характеристиками в разы удобнее на языке программирования Python. Так как в R нельзя анализировать любые данные произвольного размера для нахождений статистических характеристик (t-test, u-статистика).