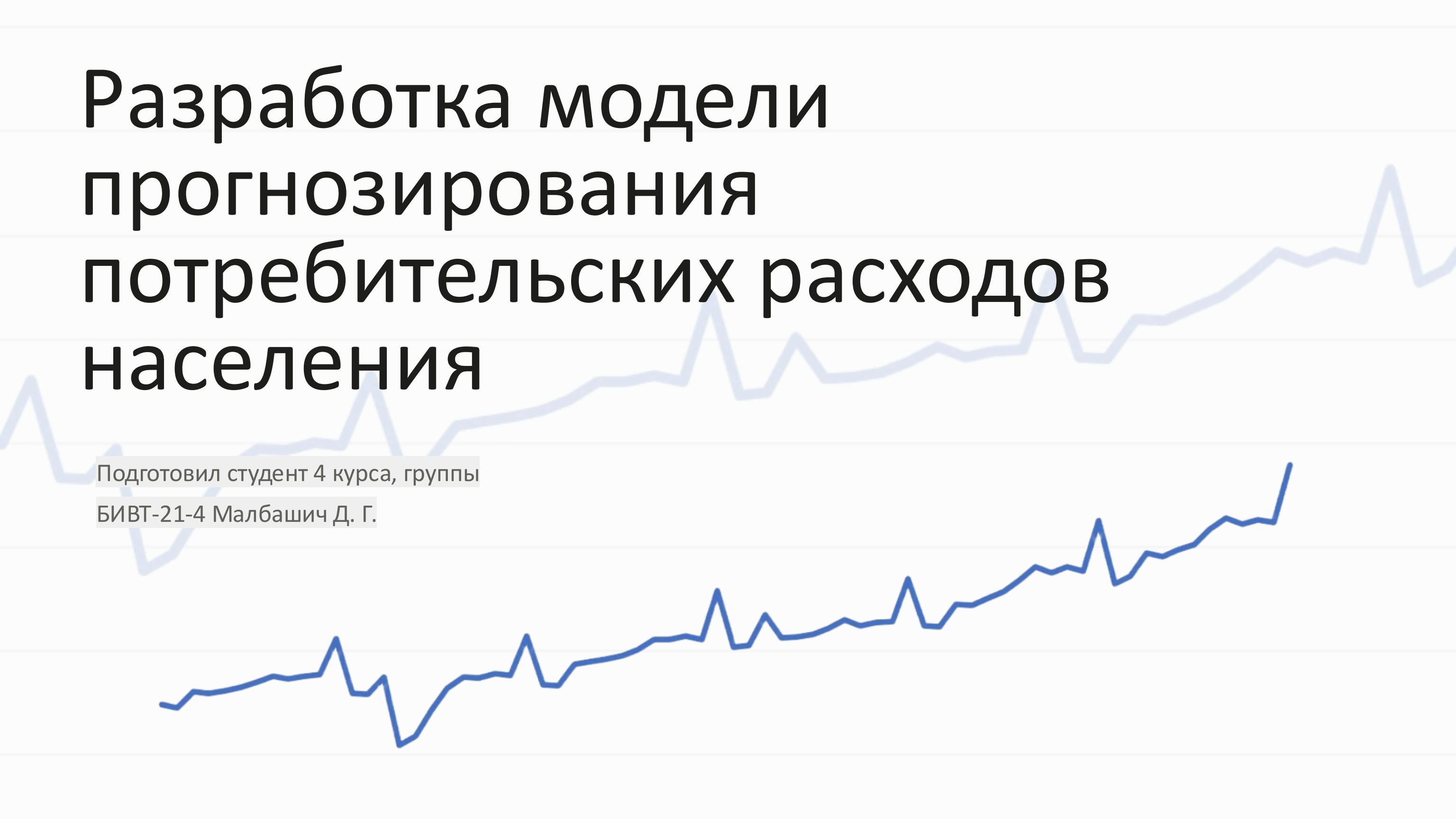


Разработка модели прогнозирования потребительских расходов населения



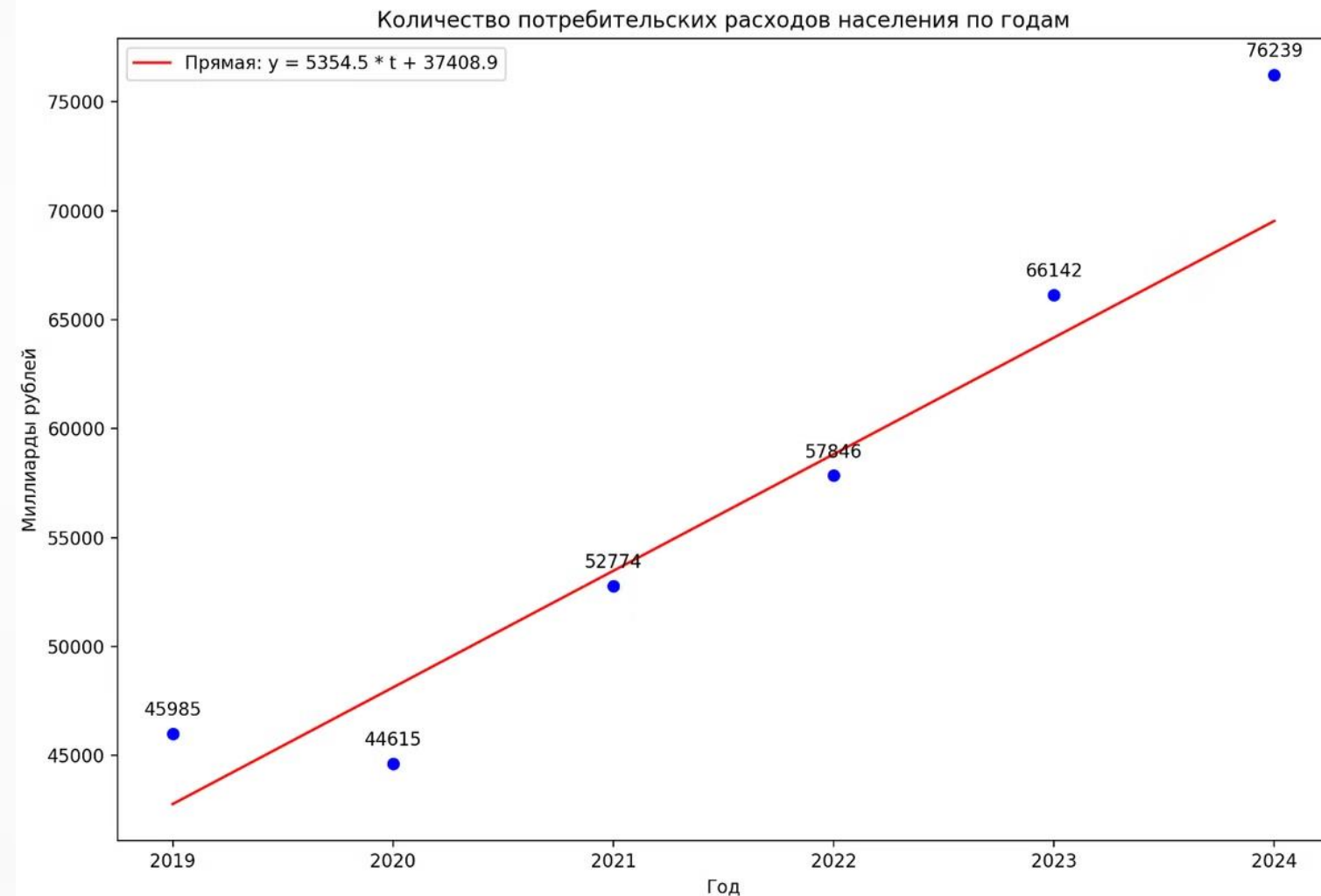
Подготовил студент 4 курса, группы
БИВТ-21-4 Малбашич Д. Г.

Моделируемый объект

- Объектом исследования являются **потребительские расходы населения.**
- Цель состоит в создании регрессионной модели, которая позволит предсказать будущие расходы на основе статистических данных.



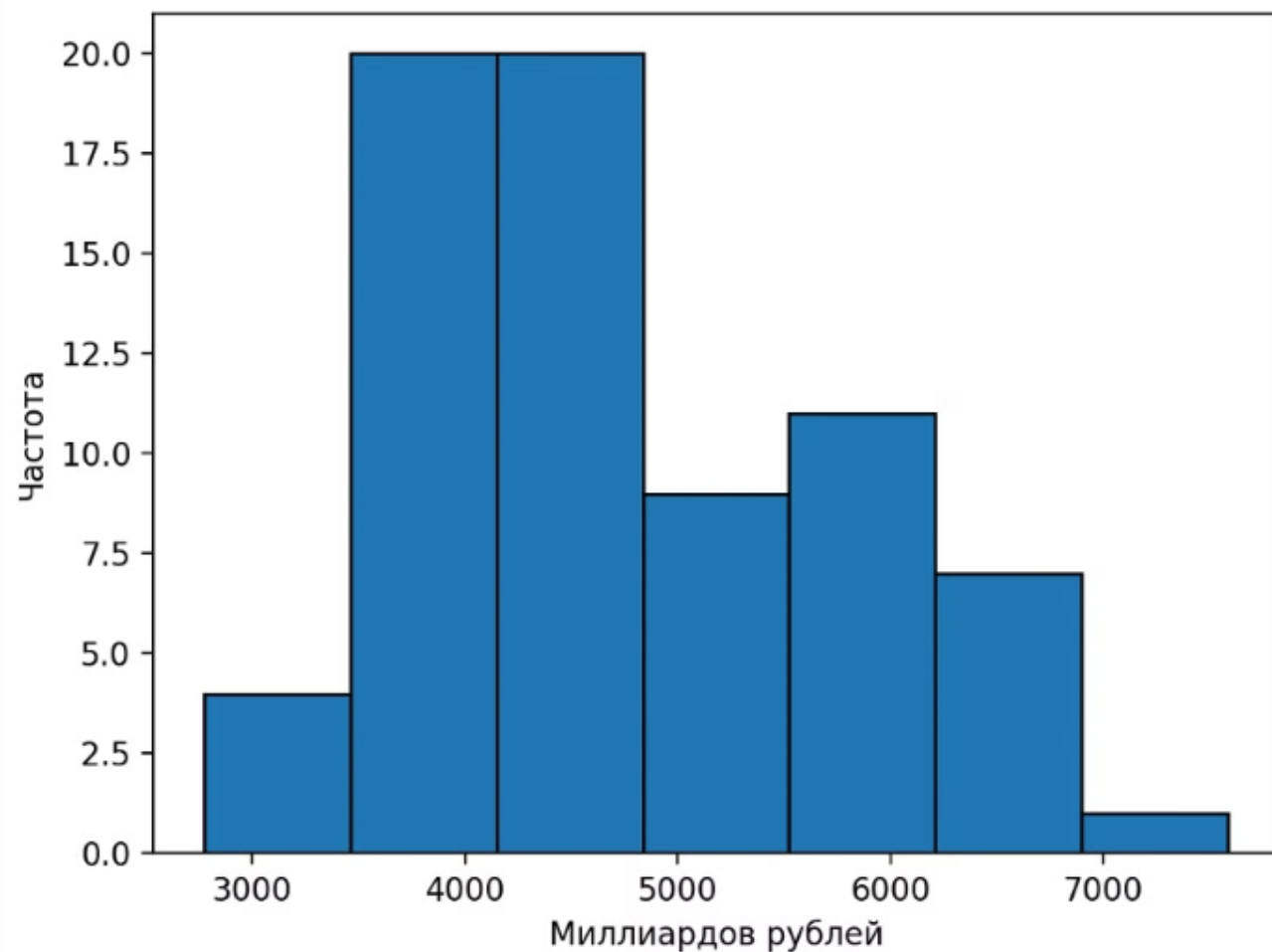
Статистические показатели



- Средний абсолютный прирост: 5040 (млрд. руб.)
 - Средний темп прироста: 9.5%
-
- Прогноз по среднему абсолютному приросту: 71182 (млрд. руб.)
 - Прогноз по среднему темпу роста: 72435 (млрд. руб.)
 - Предсказание методом аналитического выравнивания: 69536 (млрд. руб.)
-
- Относительная погрешность по среднему абсолют. приросту: 6.63%
 - Относительная погрешность по среднему темпу роста: 4.99%
 - Относительная погрешность по МНК: 8.79%

Распределение потребительских расходов

Гистограмма потребительских расходов



Проверка на нормальное распределение

Нормальное распределение по теореме 3-х сигм

Процент вхождений в интервал 1 сигмы: **68.06**

Процент вхождений в интервал 2 сигм: **98.61**

Процент вхождений в интервал 3 сигм: **100**



Распределение близко нормальному

Нормальное распределение по критерию Пирсона

Хи-квадрат наблюдаемое **11.374**

Хи-квадрат критическое **9.488**

Рассчитанное значение **1.199**

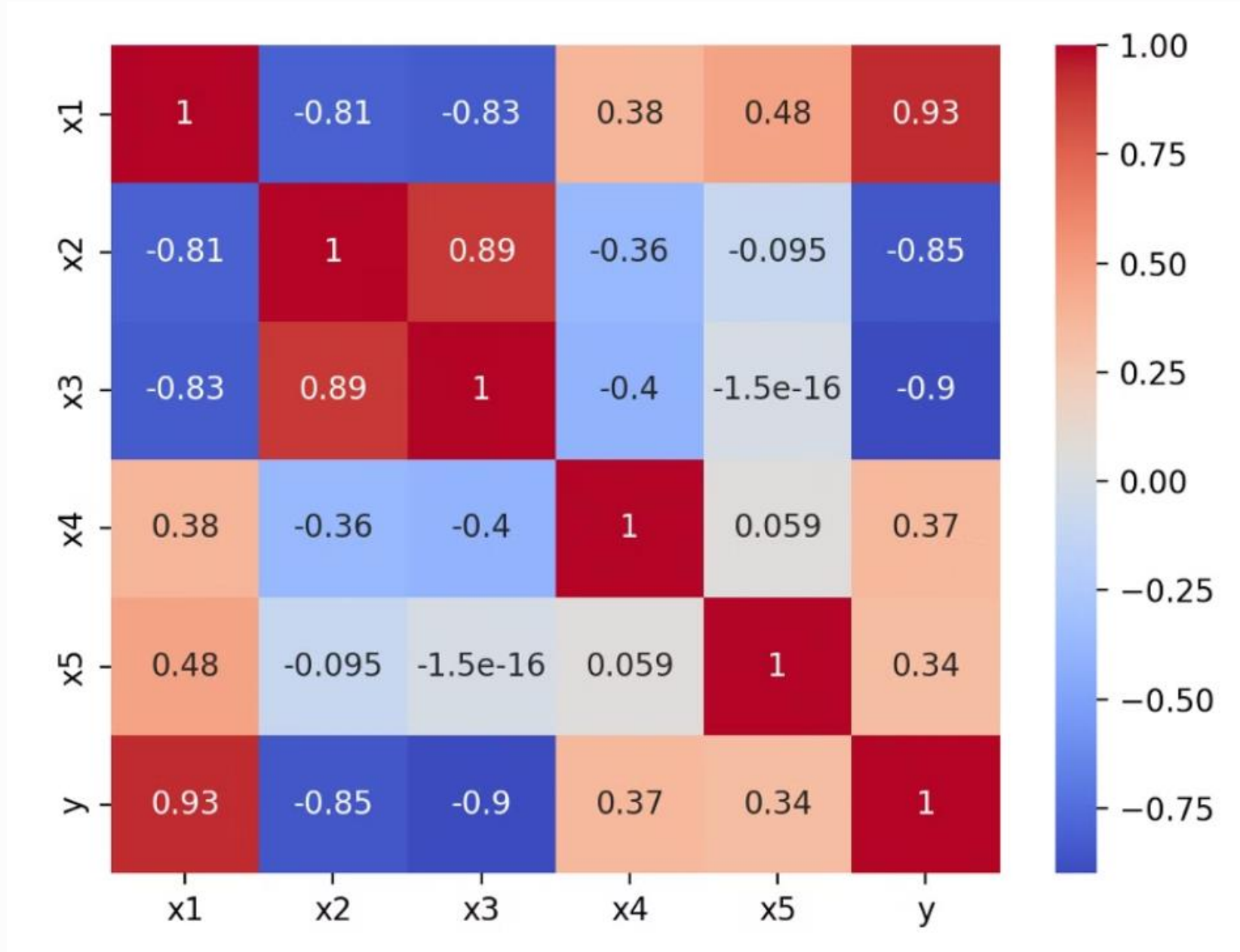


Отвергаем H_0 , распределение не является нормальным

Формализация и классификация переменных

- | | |
|--|---|
| → x1 - Среднедушевые доходы населения, млрд. руб. | → x2 - Численность безработных, тыс. чел |
| → x3 - Численность населения, млн. чел. | → x4 - Инфляция, % |
| → x5 - Номер месяца | × y . Сумма потребительских расходов, млрд. руб. |

Корреляционный анализ



Матрица корреляций

Определитель матрицы парных корреляций **0.00096**



Уравнение модели

Уравнение множественной линейной регрессии

$$Y = 142000 + 0,0266 \cdot x_1 - 0,0841 \cdot x_2 - 940,7479 \cdot x_3 - 3,7077 \cdot x_4 + 51,8012 \cdot x_5$$

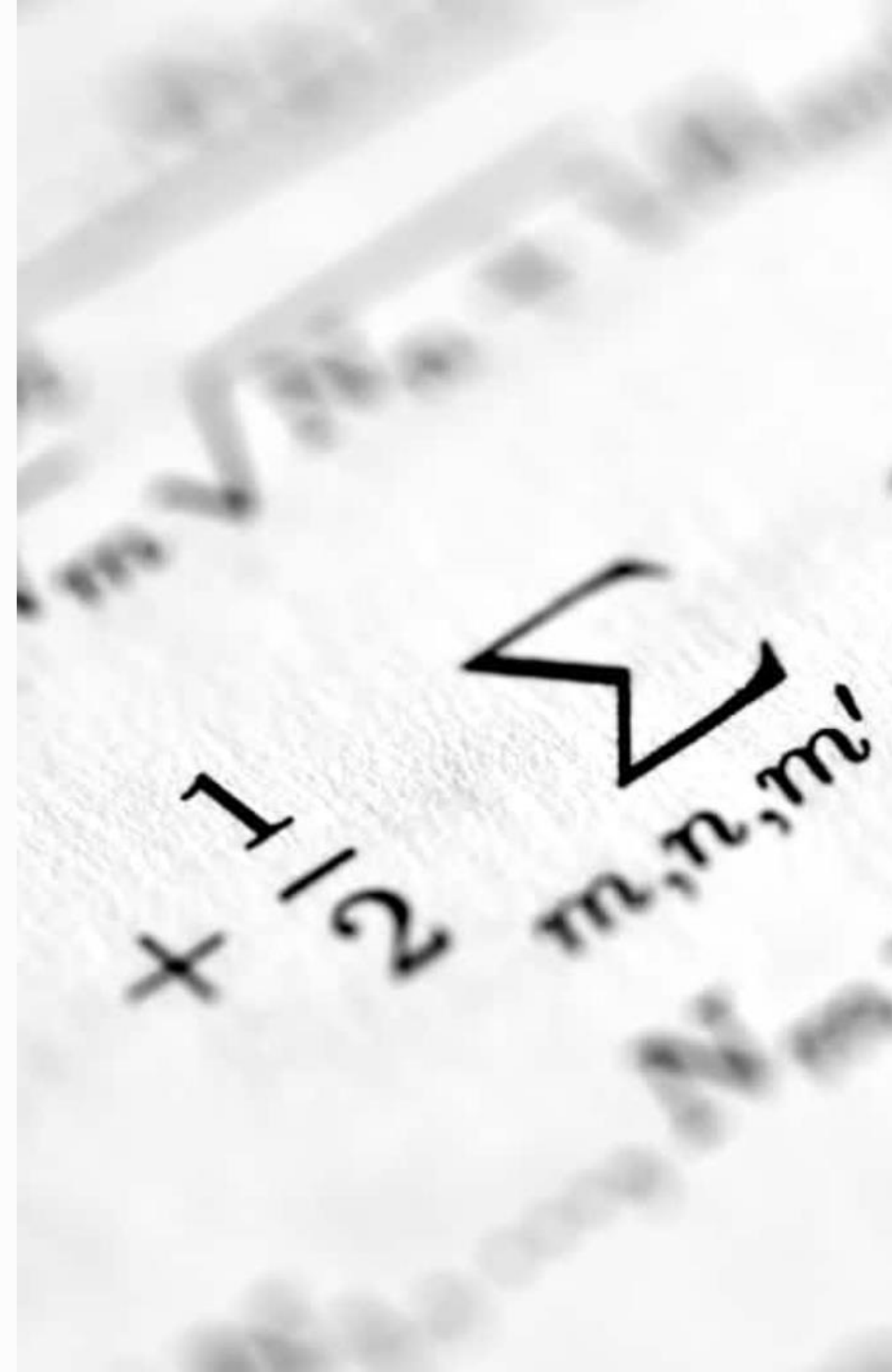
Исследование и улучшение модели

Шаговый регрессионный анализ

$$Y = 150400 + 0,028 * x_1 - 1000.68 * x_3 + 52,57 * x_5$$

Адаптивный коэффициент детерминации

Увеличился с **0,916** до **0,924**, после исключения статистически незначимых признаков.



Оценка модели

66234

MSE

Среднеквадратичная ошибка на тестовых данных

195

MAE

Средняя абсолютная ошибка на тестовых данных

0.935

R^2

Коэффициент детерминации

0.924

Adj R^2

Адаптивный коэффициент детерминации

Выводы

Были произведены

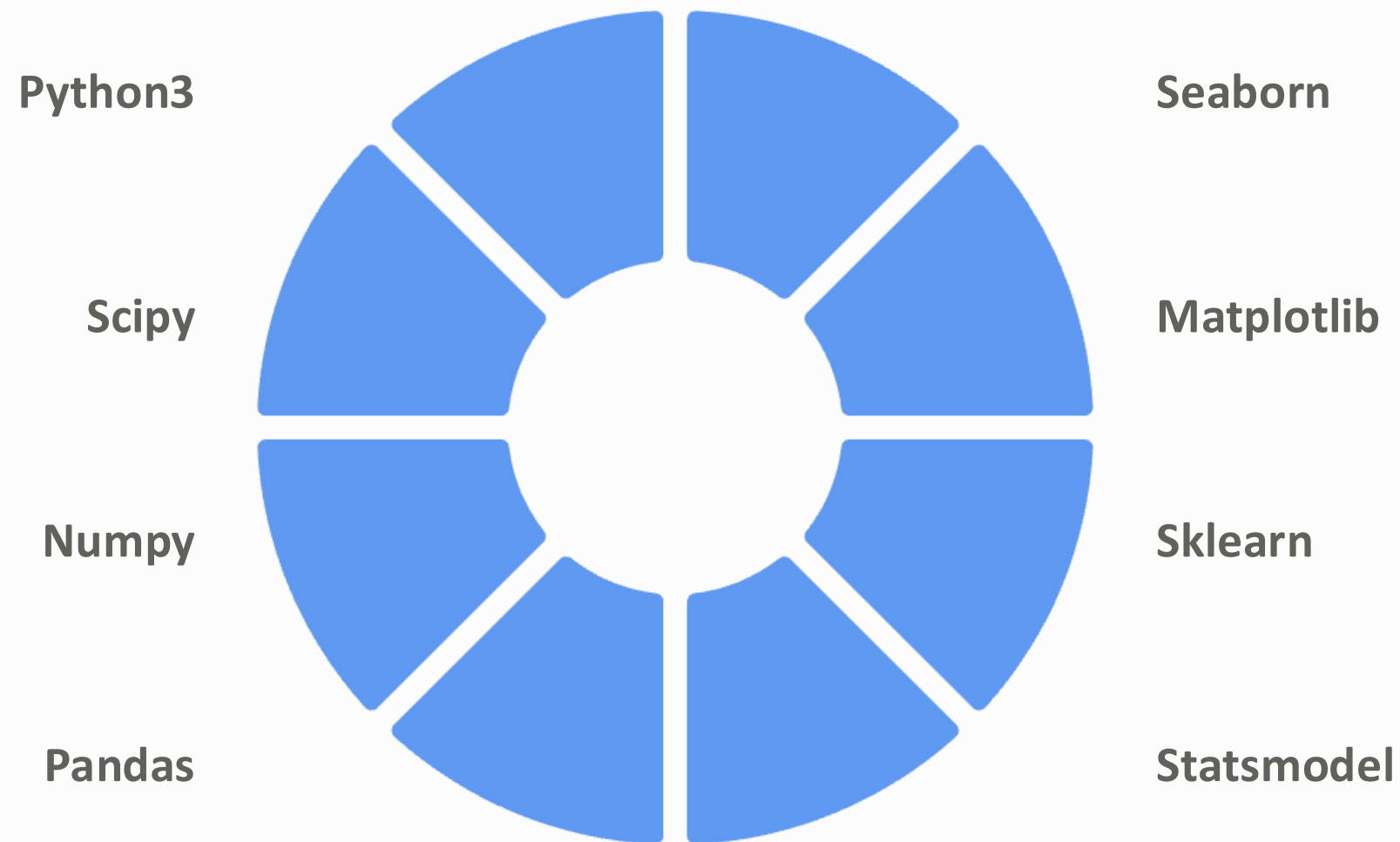
- Статистический анализ выходной величины
- Формулировка постановки задачи
- Анализ и обработка признаков
- Корреляционный анализ
- Уравнение множественной линейной линейной регрессии

В результате получена

Модель прогнозирования, пригодная для
пригодная для прикладного использования
использования



Программная реализация



Репозиторий: <https://github.com/>