

Возьми миллион – отдай два



**ITMO**



**Tinkoff**



# Датасет

- ✓ Кредитование в Tinkoff Bank
- ✓ В датасете представлены данные о заявках на кредитование в период с марта 2021 года по ноябрь 2022 года.
- ✓ Всего представлено 226к записей, что соответствует 211к заявок от клиентов

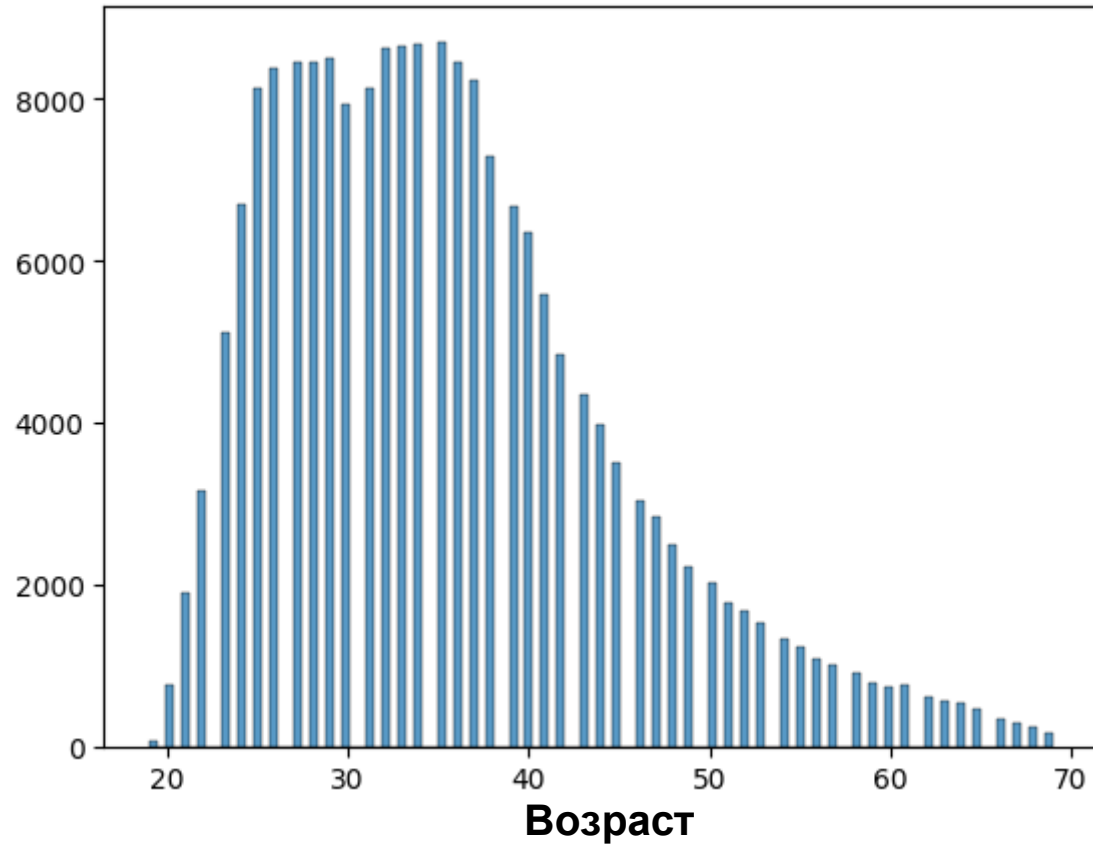


# Основные показатели, их статистики

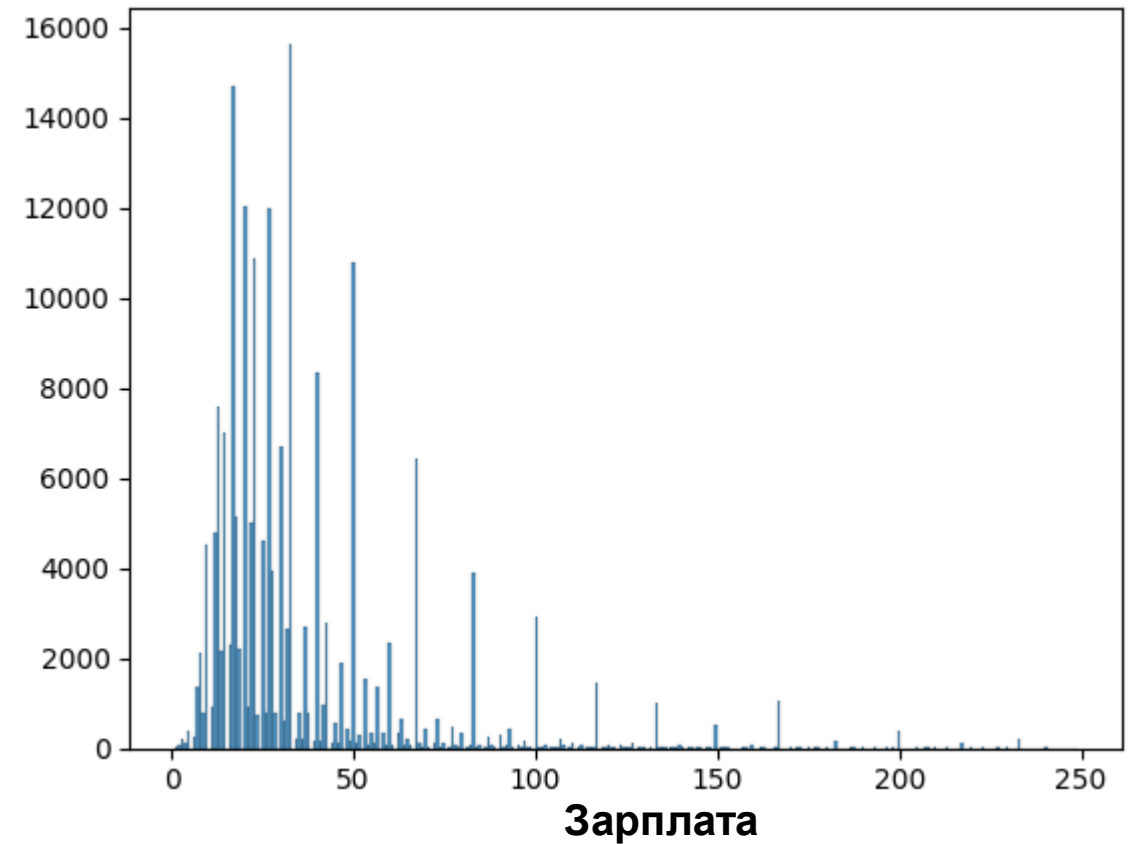
|                       | <b>Возраст, лет</b> | <b>Зарплата,<br/>условные единицы</b> | <b>Запрошенная сумма<br/>кредитования,<br/>Условные единицы</b> |
|-----------------------|---------------------|---------------------------------------|---|
| Количество значений   | 198233              | 198233                                | 198019  |
| Медианное значение    | 34                  | 27                                    | 100   |
| Среднее значение      | 35,6                | 37,9                                  | 188,9   |
| Минимальное значение  | 19                  | 1                                     | 7   |
| Максимальное значение | 69                  | 1852                                  | 8000  |

# Распределения основных показателей

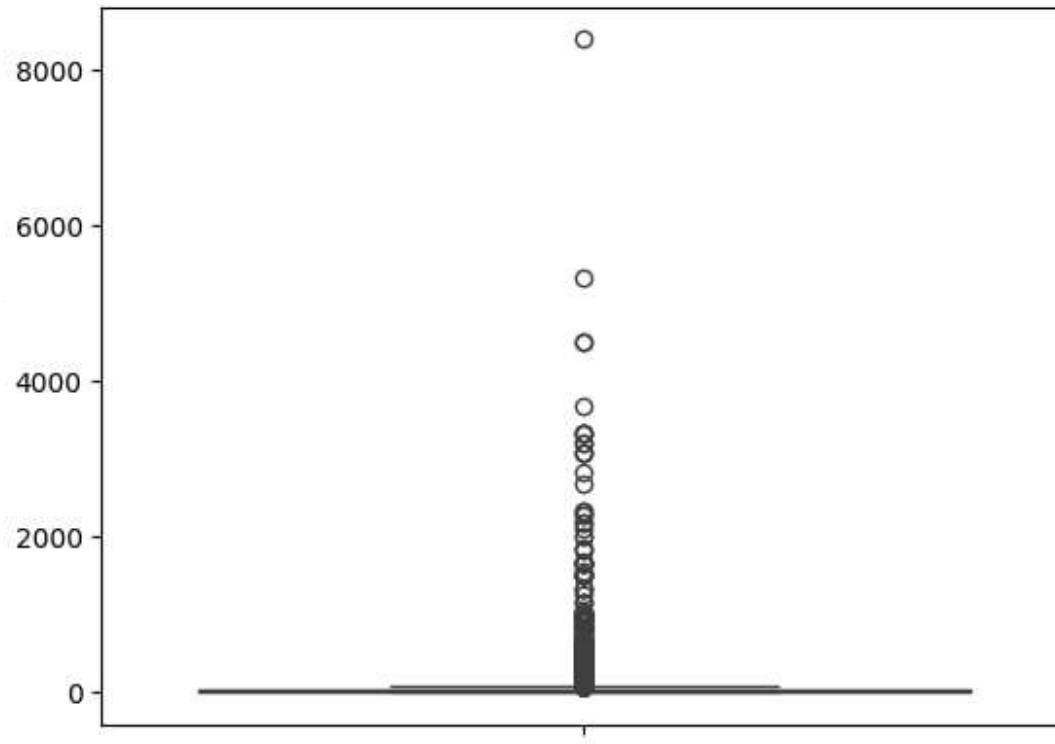
Количество



Количество



# Удаление выбросов

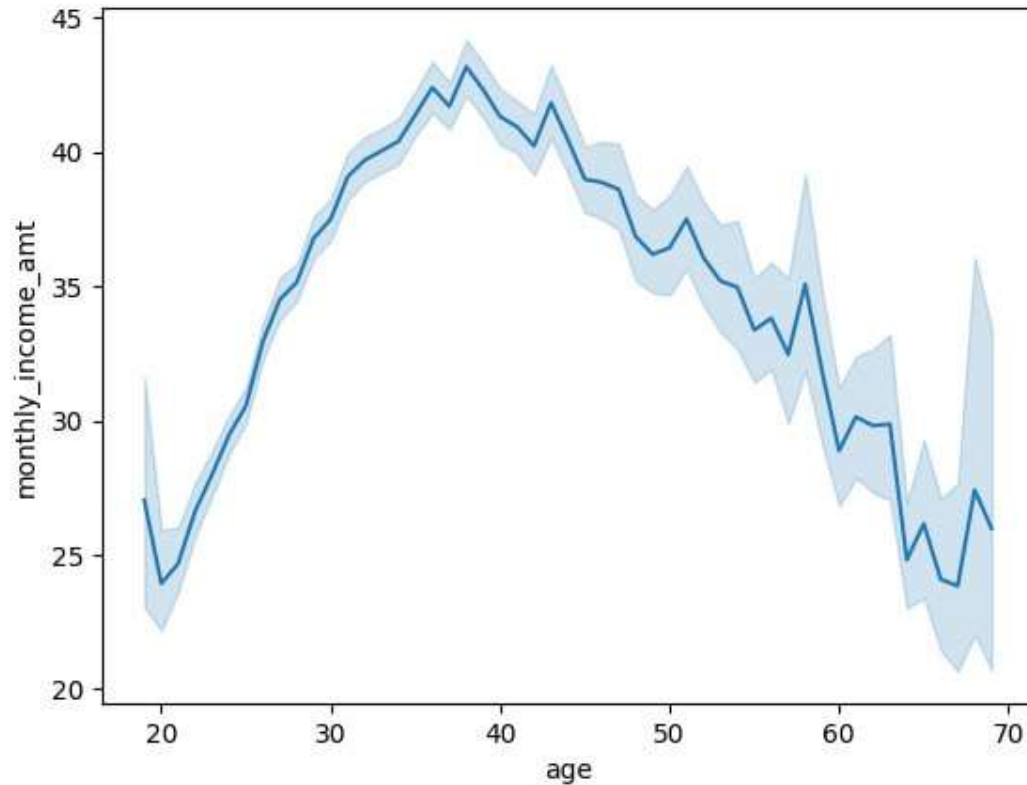


Распределение по зарплате

Было удалено  
27721 строк

- Вне 97,5 перцентиля зарплаты
- Вне диапазона 18-69 лет
- Дубликаты по id операции

# Зависимость среднего заработка от возраста



Пик заработка – около 35 лет

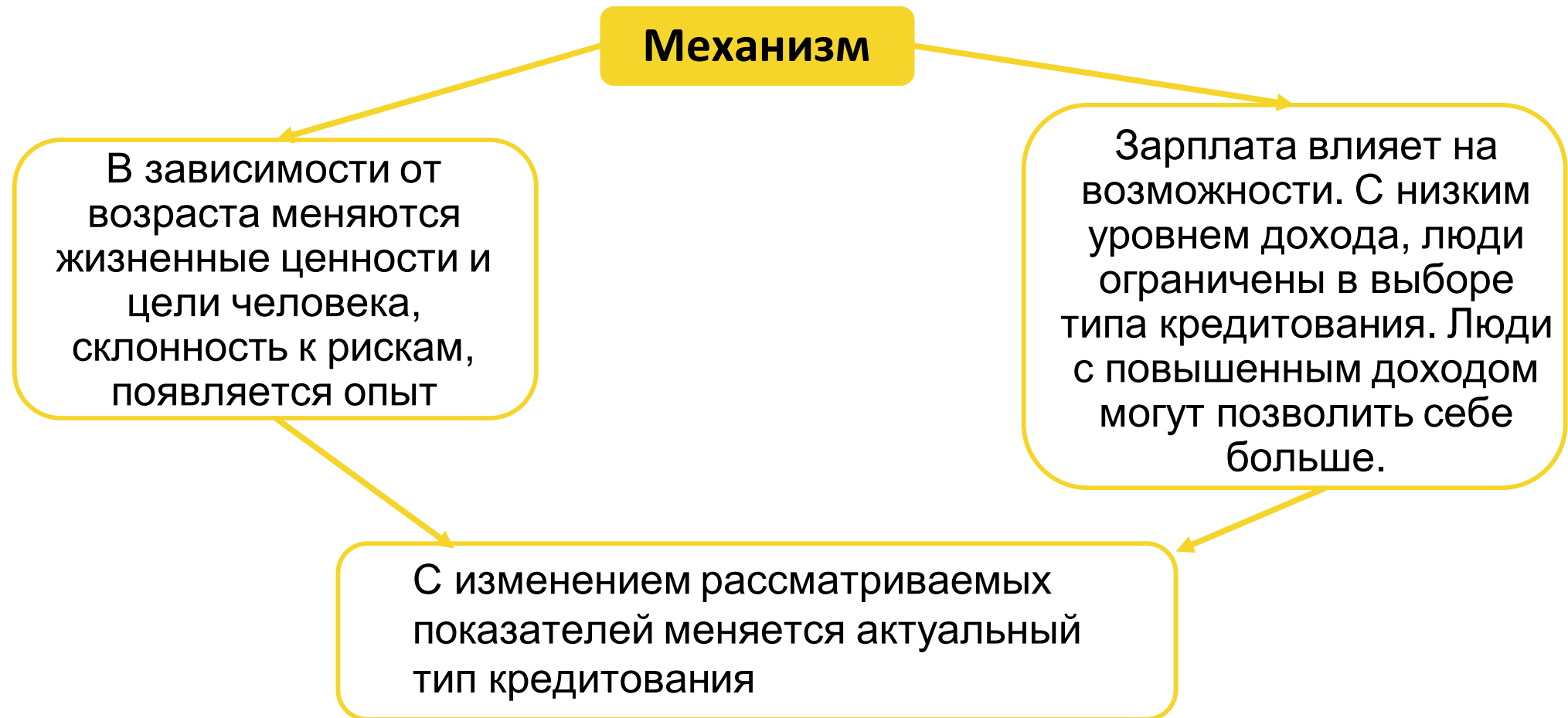
# Исследовательский вопрос

Какие характеристики клиента определяют, какой вид кредитования для него наиболее актуален?



# Гипотеза

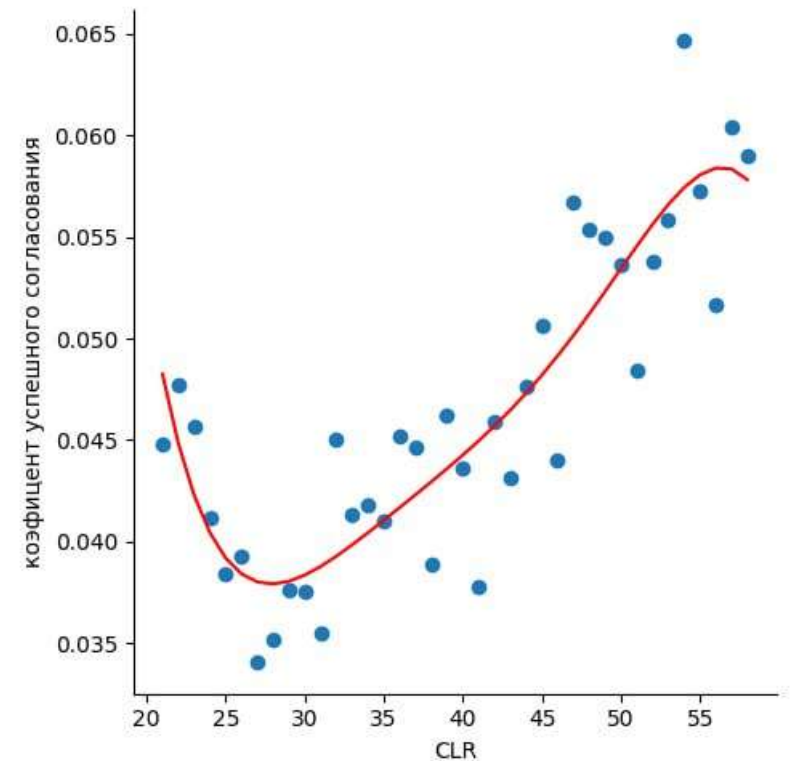
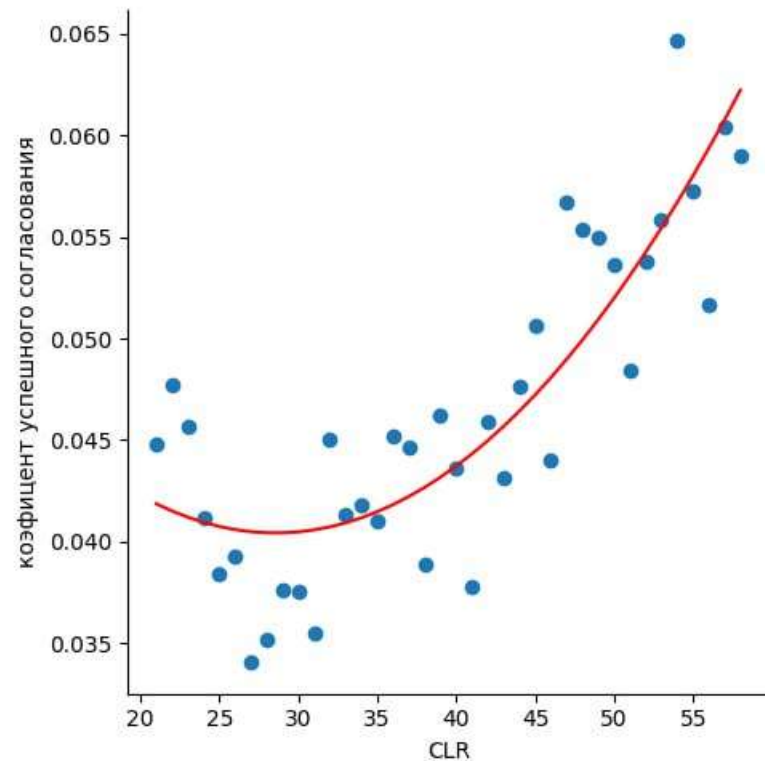
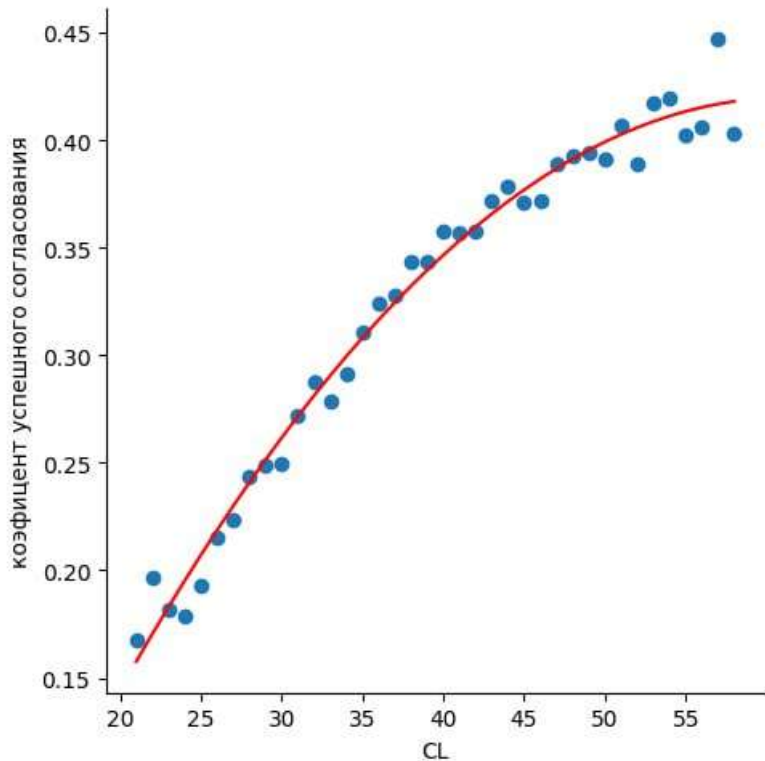
**Возраст и зарплата клиента значительно влияют на актуальный для него тип кредитования**





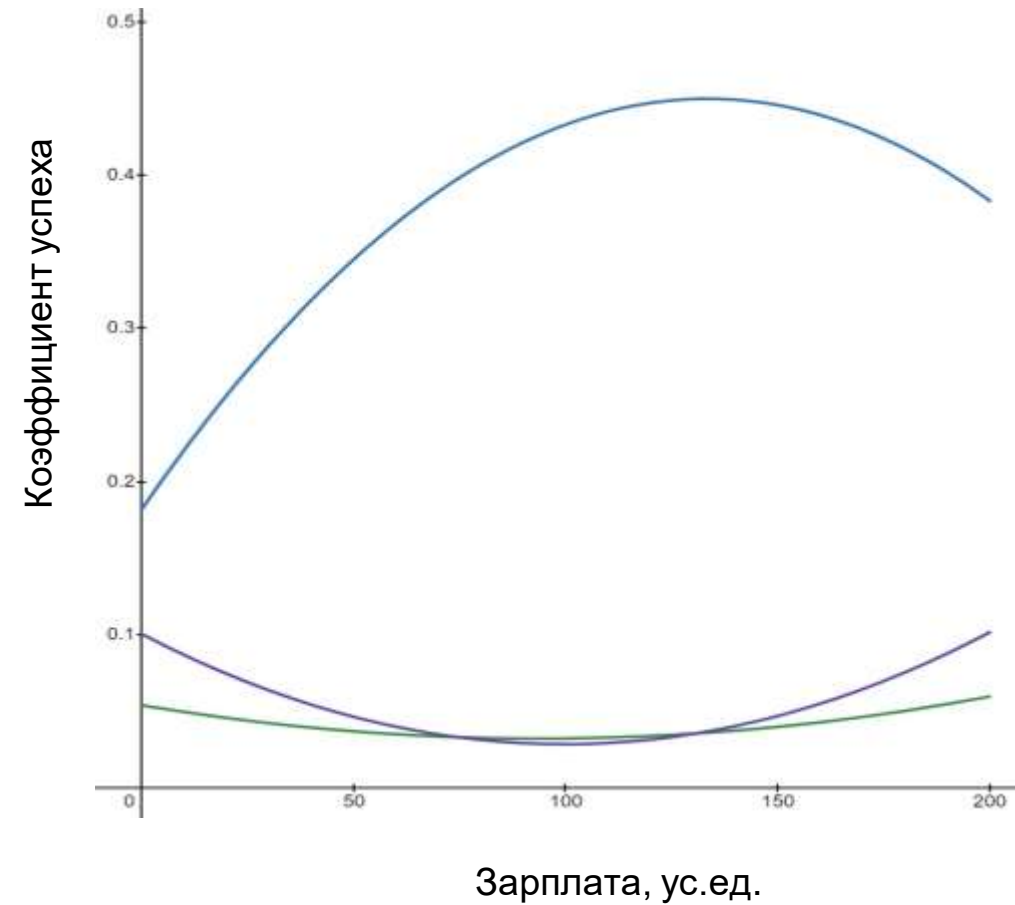
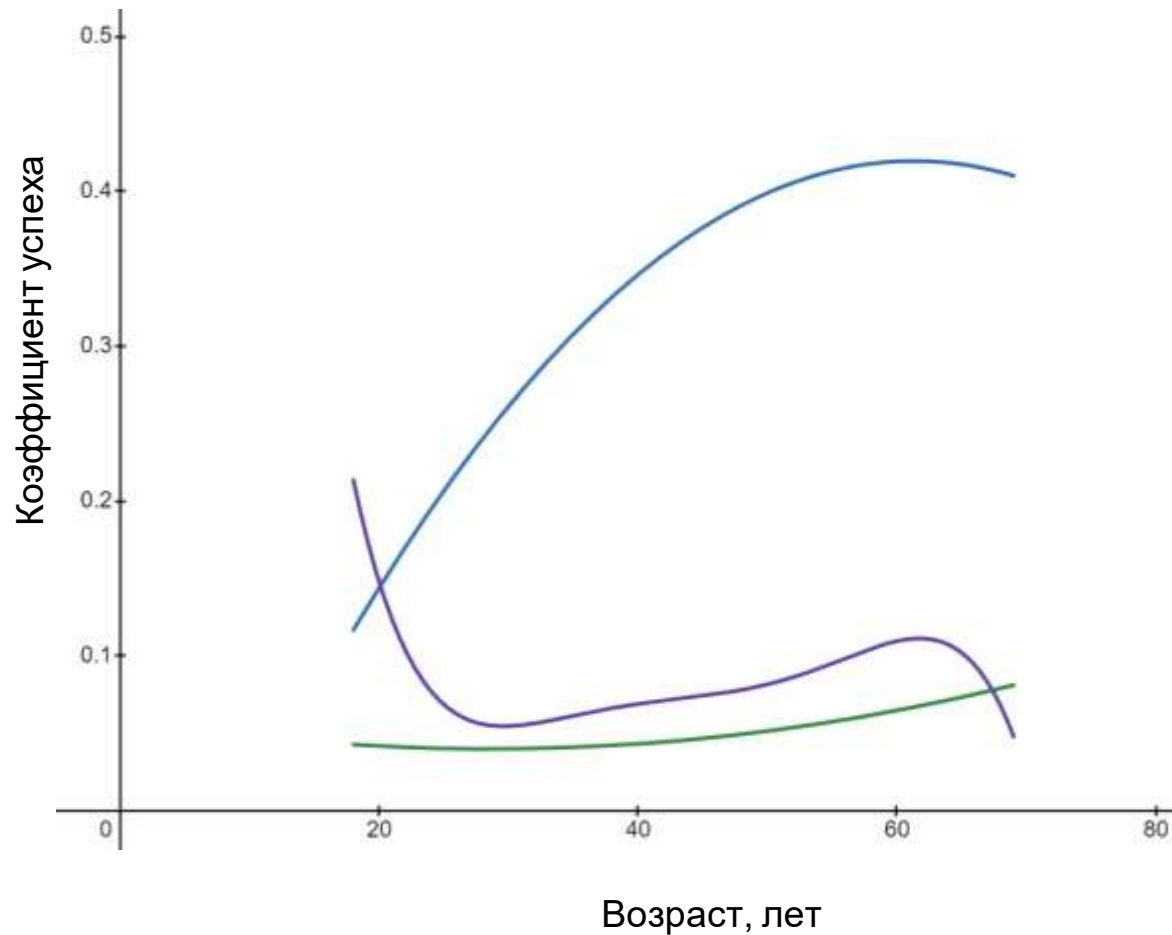
# Мат. модель

Коэффициент успеха – доля успешных согласований условий, среди исследуемой группы людей, которым предложили определенный тип кредита



# Мат. модель

Синий график – CL, фиолетовый график – CC, зеленый график - CLR



# Мат. модель

## U-тест Манна-Уитни

| Исследуемая характеристика и тип кредита | Исследуемые группы   | Полученные значения | Табличные значения |
|--|----------------------|---------------------|--------------------|
| Возраст, CL                              | До 25; после 60      | 64                  | 7                  |
| Возраст, CC                              | До 25; после 60      | 33                  | 2                  |
| Возраст, CLR                             | До 25; после 60      | 81                  | 11                 |
| Зарплата, CL                             | До 30; от 100 до 140 | 100                 | 16                 |
| Зарплата, CC                             | До 30; от 100 до 140 | 49                  | 14                 |
| Зарплата, CLR                            | До 30; от 100 до 140 | 24                  | 2                  |

# Интерпретация

Гипотеза подтвердилась, действительно для разных зарплатных и возрастных групп отличается эффективность рекомендаций различных типов кредитования

С возрастом растет популярность наличного кредитования и кредитования под залог недвижимости. Тогда как кредитные карты эффективнее предлагать молодым людям

Наличное кредитование пользуется популярностью у людей среднего заработка. Тогда как кредитование под залог недвижимости и кредитные карты у людей других категорий дохода.

# Policy implication

Полученные данные могут быть использованы для улучшения системы рекомендаций типа кредитования. Это увеличит прибыль банка за счет увеличения коэффициента успеха.

## Ограничения

- Отсутствие данных о страховке
- Отсутствие кредитной истории
- Информация об автомобиле

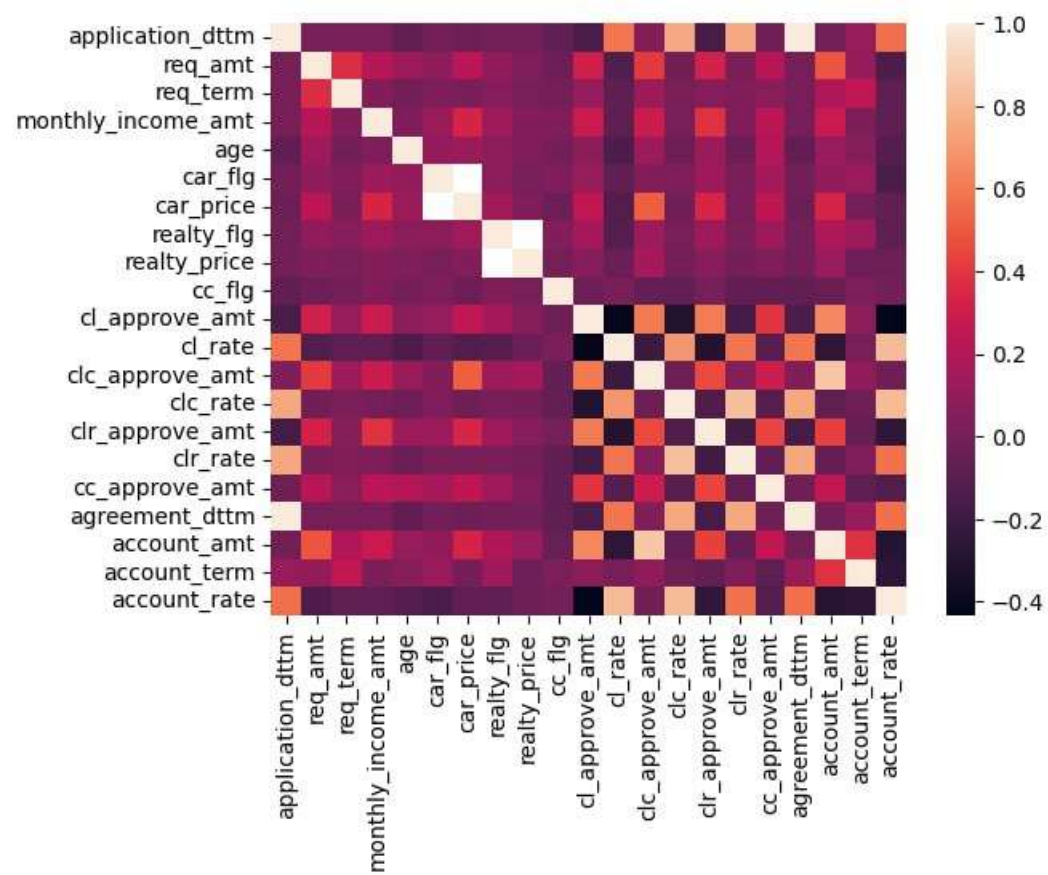
## Перспективы

- Увеличение количества данных по кредиту под залог машины
- Расширение по временному и географическому отрезку

# Stradigi в людях

**Пронин Михаил Васильевич  
Козлов Дмитрий Максимович  
Пономарев Андрей Михайлович  
Трошкин Антон Владимирович  
Горшков Кирилл Юрьевич  
Трескова Снежана Николаевна**

# Подполье



# Подполье

