

# Исследование зависимости стоимости тарифа страхования КАСКО от данных страхователя

Шарак Александр  
Стекачев Владимир  
Гайдин Михаил  
Волков Алексей

# Содержание

1. Введение
2. Данные
3. Методы
4. Результаты
5. Обсуждение

# Введение

Что может влиять на желание клиента застраховать авто?

Гипотеза: клиенты из центральных регионов России берут более дорогие тарифы страхования.



# Данные

Переменные:

- Сумма премии выпущенного полиса
- Цена авто
- Пробег авто
- Убытки за предыдущие периоды
- Год выпуска автомобиля
- Флаг кредитного авто
- Стаж
- Пол
- Возраст

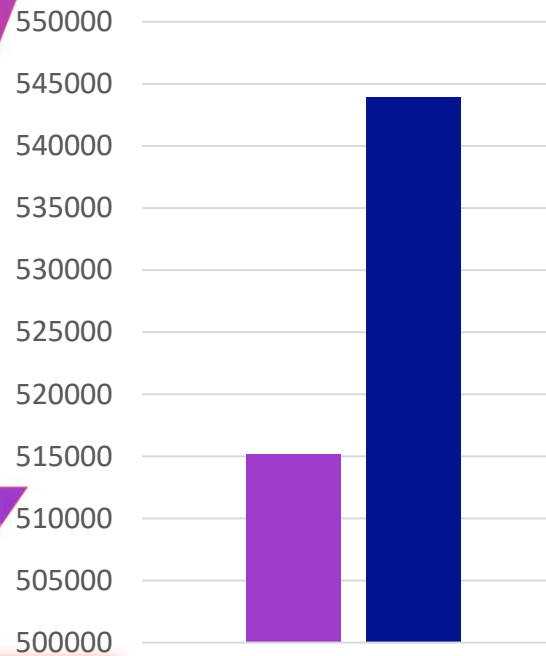


# Разведывательный анализ

- Датасет содержит данные страхования КАСКО у водителей разного пола, возраста, стажа и т.д.
- Изначально в датасете содержится 148476 строк.
- После первичной очистки данных осталась 25891 строка, в связи с совпадающими номером полиса, полом и возрастом.

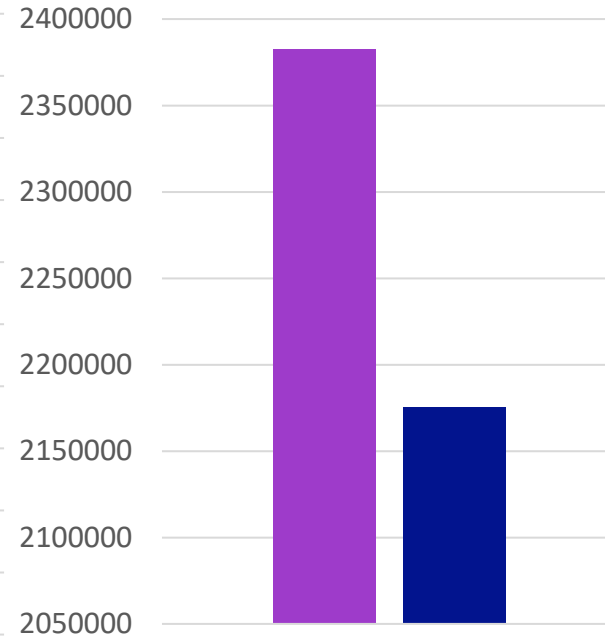


# Разведывательный анализ



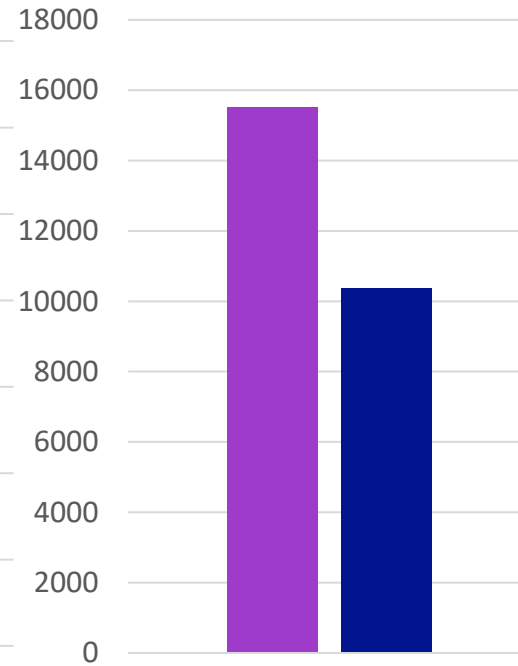
Средняя стоимость страхования КАСКО по полу

■ Мужчины ■ Женщины



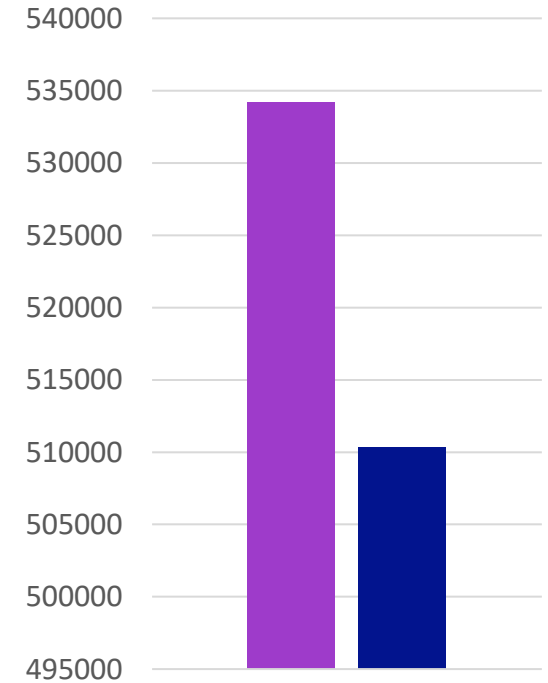
Средняя стоимость автомобиля по полу

■ Мужчины ■ Женщины



Количество страховок

■ Центр ■ Регионы

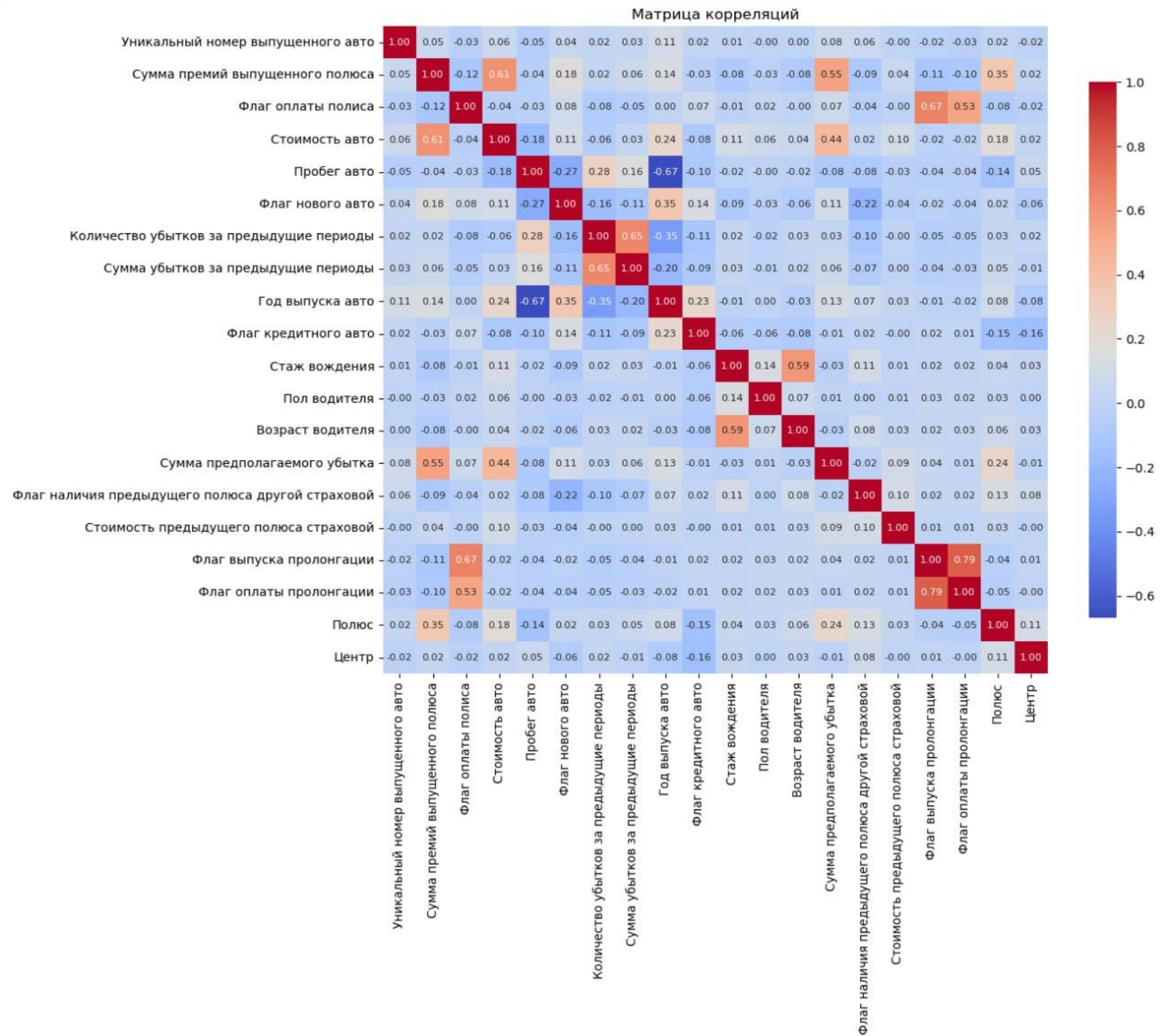


Средняя стоимость страховки

■ Центр ■ Регионы

# Методы

1. Использовали метод линейной регрессии и обучили ее с помощью метода наименьших квадратов.
2. Построили матрицу корреляций
3. С помощью матрицы корреляций и коэффициентов, полученных в линейной регрессии, очистили модель от незначимых переменных





<b>Название переменных</b>	<b>Результаты модели</b>
Сумма премий выпущенного полюса	3.518e-07
Флаг оплаты полиса	-0.0208
Стоимость авто	-3.725e-08
Пробег авто	-2.261e-06
Флаг нового авто	--0.0112
Количество убытков за предыдущие периоды	0.0179
Сумма убытков за предыдущие периоды	5.917e-09
Год выпуска авто	-0.0035
Флаг кредитного авто	-0.1210
Стаж вождения	0.0011
Пол водителя	0.0277
Возраст водителя	0.0018
Сумма предполагаемого убытка	1.04e-07
Флаг наличия предыдущего полюса другой страховой	0.1375
Стоимость предыдущего полюса страховой	-2.293e-10
Флаг выпуска пролонгации	0.0242
Флаг оплаты пролонгации	--0.307
Центр	0.0559

### Результаты модели

Чем больше стоимость КАСКО, тем больше большую сумму старховки выплачивает страховщик и несет большую ответственность за авто

Чем меньше стоимость авто, тем больше людей хотят его застраховать по более дорогому виду КАСКО

Чем больше пробег авто, тем меньше его хотят застраховать по более дорогому КАСКО

Чем больше убытков, тем люди больше берут более дорогой КАСКО

Чем больше сумма убытков, тем больше человек хочет купить более дорогое КАСКО

Чем новее авто, тем люди больше хотят более дорогое КАСКО

Если авто кредитное, то для него покупают менее дорогое КАСКО

Если за рулём опытный водитель , то он в большей вероятности купить более дорогое КАСКО

Если за рулём мужчина, то он с большей вероятностью купит более дорогое КАСКО

Если за рулём возрастной водитель , то он в большей вероятности купить более дорогое КАСКО

Чем больше предполагаемая сумма убытков по полису, то человек с большей вероятностью возьмет более дорогое КАСКО

Если человек уже брал КАСКО, то он с большей вероятностью возьмет более полный тариф

Если человеку уже продлевали КАСКО, то он с большей вероятностью продлит его еще раз

Если пролонгация не оплачена, то человек с большей вероятностью возьмет более дорогое КАСКО

Если в центре, то с большей вероятностью человек возьмет более полное покрытие

# Обсуждение

Основные выводы, применение результатов, ограничения:

