



Описание данных

В предложенном задании используются данные выпущенных полисов Каско Т-Страхования за несколько месяцев.

Что такое Каско?

Каско — это добровольное страхование, которое защищает интересы владельца транспортного средства. Даже если он сам станет виновником ДТП и его застрахованный автомобиль пострадает в аварии, выплаты по каско возместят этот урон. Владелец авто получит деньги, или машину отремонтируют за счет страховой компании.

В Каско могут быть включены следующие покрытия:

Ущерб включает в себя большинство событий, которые могут повредить автомобиль. (ДТП, столкновение с препятствиями, наезд, опрокидывание, самопроизвольное движение автомобиля, повреждение другим автомобилем на стоянке, падение в воду, пожар, взрыв, удар молнии, природные явления: град, землетрясение и другие, случайное попадание на автомобиль инородных предметов)

Хищение — это утрата автомобиля в результате угона, кражи, грабежа или разбоя.

Тоталь — повреждения автомобиля, при которых стоимость ремонта равна или больше 65% страховой суммы.

Клиент может выбрать один из следующих вариантов покрытий:

- 1 Полное Каско — хищение+ущерб+тоталь
- 2 Тоталь+угон
- 3 Только угон
- 4 Только ущерб

Ремонт, который выбран в полисе, также влияет на стоимость полиса:

1. Ремонт на СТОА официального дилера — самый дорогой ремонт
2. Ремонт на СТОА по направлению страховщика — среднее увеличение стоимости
3. Эконом ремонт на СТОА по направлению Страховщика — низкое увеличение стоимости
4. Отсутствует — для покрытия "Только Угон"

В датасете представлена информация о авто, убытках, водителе и о пролонгации(продлении) полиса на следующий год.

Название переменной	Описание	Единицы измерения
policy_rk	Уникальный номер выпущенного полиса	
total_premium_amt	Сумма премии выпущенного полиса	у.е.
coverage	Покрытие полиса	
repair_option	Выбранный ремонт по полису	
paid_flg	Флаг оплаты полиса	
	1 — оплачен	
	0 — не оплачен	
region	Регион использования авто	
auto_make	Марка авто	
auto_model	Модель авто	
auto_cost	Цена авто	у.е.
vehicle_mileage_km_amt	Пробег авто	
vehicle_new_flg	Флаг нового авто	
	1 — новое авто	
	0 — Б\у	
audatex_claim_cnt	Количество убытков за предыдущие периоды	
audatex_total_cost_amt	Сумма убытков за предыдущие периоды	у.е.
vehicle_issue_year	Год выпуска авто	
vehicle_credit_flg	Флаг кредитного авто	
	1 — кредитное авто	
	0 — не кредитное	
min_driver_experience_years_cnt	Стаж вождения	
insurant_gender_cd	Пол водителя	
	F — женский	
	M — мужской	
age	Возраст водителя	
issue_flg	Флаг выпуска полиса	
	1 — выпущен	
	0 — не выпущен	
expected_loss_amt	Сумма предполагаемого убытка по полису	у.е.

<u>Название переменной</u>	<u>Описание</u>	<u>Единицы измерения</u>
transition_flg	Флаг наличия предыдущего полиса другой страховой 1 — есть полис 0 — нет полиса	
premium_transition_amt	Стоимость предыдущего полиса страховой	у.е.
renewed_issue_flg	Флаг выпуска пролонгации 1 — пролонгация выпущена 0 — пролонгация не выпущена	
renewed_paid_flg	Флаг оплаты пролонгации 1 — оплачено 0 — не оплачено	

Задание:

Что может влиять на желание клиента застраховать авто?

Комментарии к заданию:

У каждого вида покрытия есть свои плюсы и минусы, которые можно выделить. Не каждому клиенту может подойти Полное Каско. Подумайте, какую проблему пытаются решить клиенты с помощью покупки страхового полиса. Какой вид покрытия подойдет под каждую из выделенных проблем? Каковы плюсы и минусы каждого вида покрытия?

Какие доступные в базе данных характеристики водителя и авто могут быть связаны с выбором страхового полиса? На что следует обращать внимание при выборе покрытия страхового полиса? Какие зависимости между переменными вы ожидаете?

Обратите внимание на процесс пролонгации(продления) полисов клиентами. Какие взаимосвязи могут быть между наличием пролонгации и другими переменными?

Учитывая все вышеперечисленное, проведите анализ данных и выявите ключевые факторы, влияющие на выбор вида покрытия страхового полиса.

Подведите итоги проведенного исследования. Что именно вы хотели выяснить? Какие результаты были получены? Как эти результаты могут быть использованы на практике? Какие ограничения в рамках своего исследования вы видите? Если бы у вас было больше времени и не было ограничений, связанных с доступностью данных, как вы могли бы дополнить и улучшить свое исследование?