

# Анализ базы данных водителей

“я реально офигел”



**Команда 24**  
**(г. Нижний Новгород**  
**МАОУ Лицей №36)**

# Исследовательский вопрос

“круто, да это ж круто”



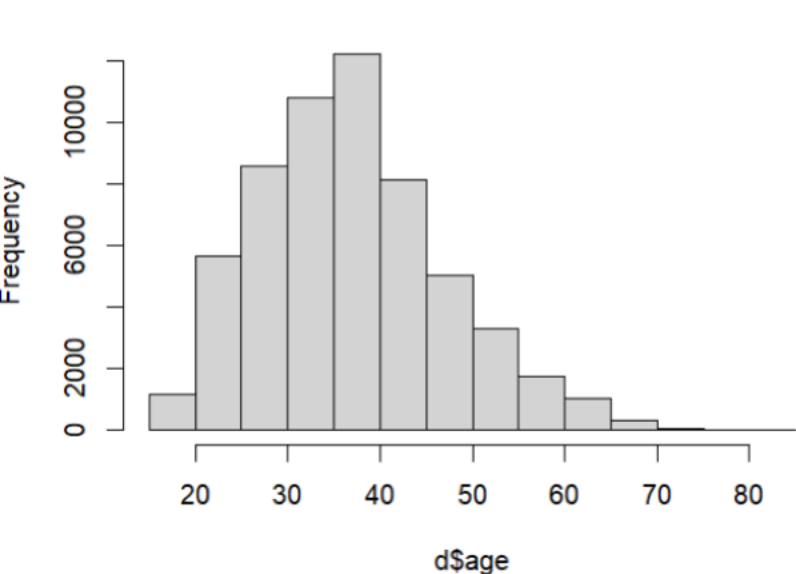
**Взаимосвязаны ли характеристики клиентов и их автомобилей с характеристиками правонарушений?**



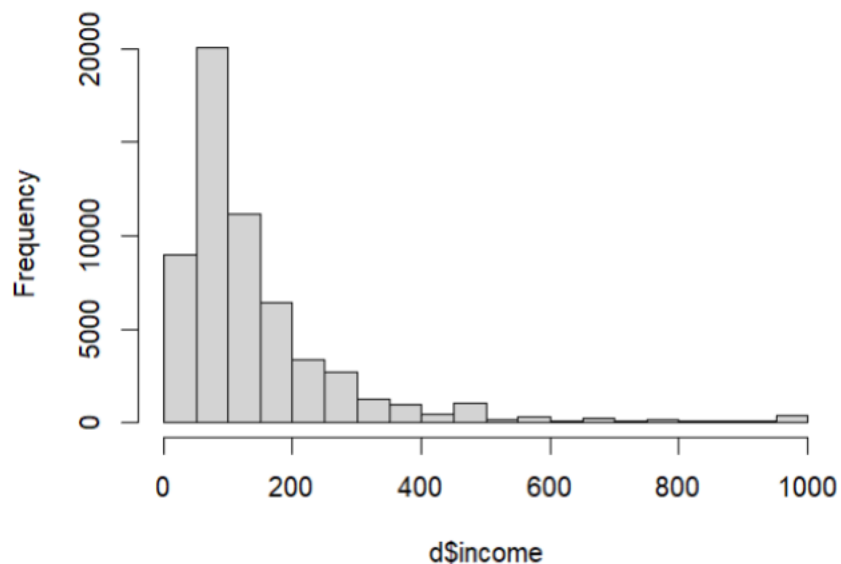
3

# Предварительный анализ

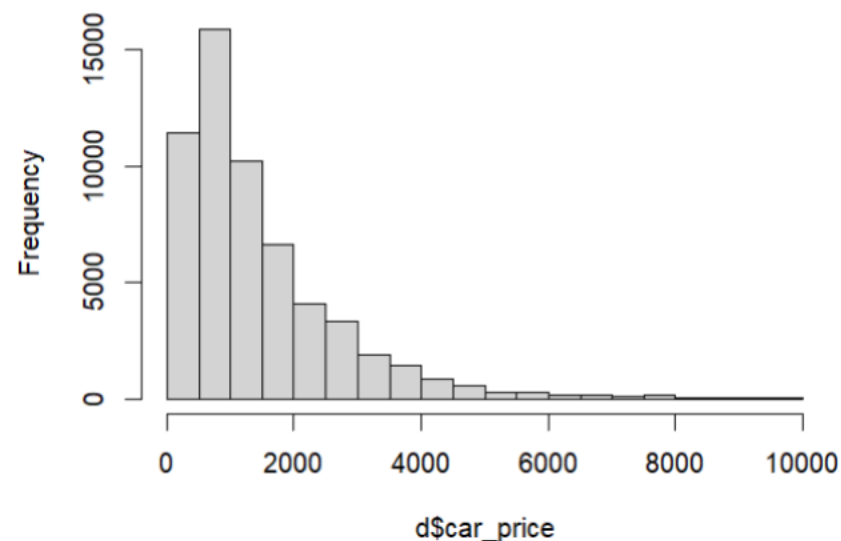
“ну это капец какой-то просто! ну, сколько можно”



**Гистограмма  
возраста**



**Гистограмма  
дохода**



**Гистограмма  
цены машины**

# Чистка данных

“круто, да это ж круто”

- **убрали ошибочные данные**
  - **разные значения стоимости автомобиля у одного водителя**
- **оценка выбросов**

# Гипотеза

“сомнительно, но окЭй”

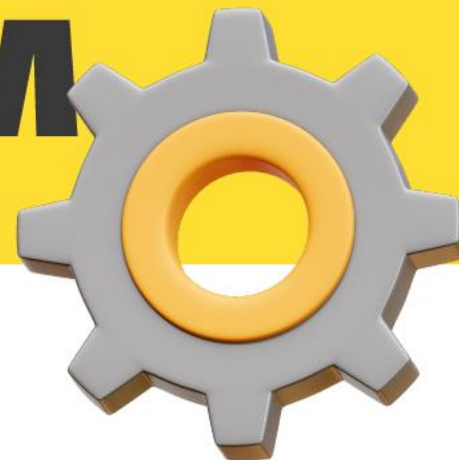
**Сумма штрафов  
зависит от  
финансового  
положения**





# Механизм

“блин, я заплакал”



Человек имеет большой  
ежемесячный доход

Из-за этого человек  
ведет себя менее  
ответственно за рулем

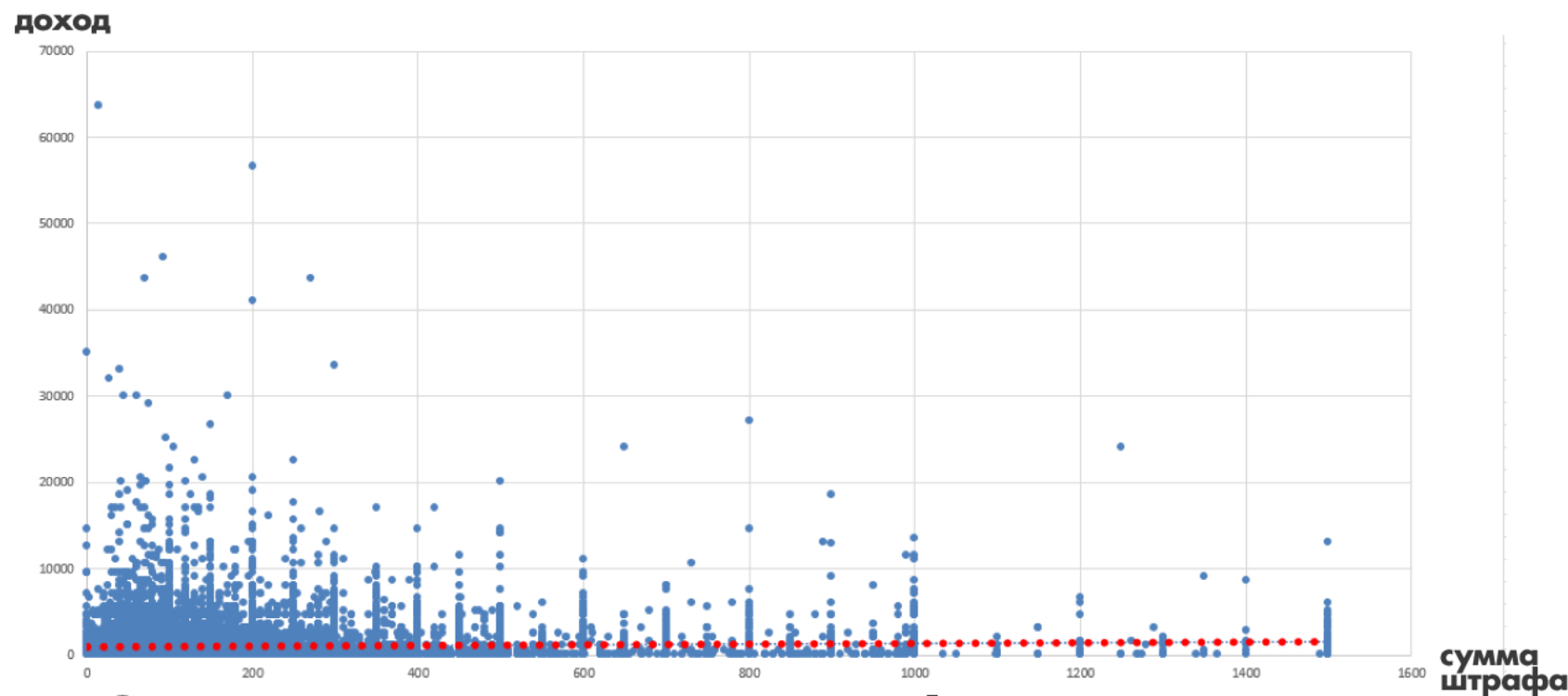
Он меньше беспокоится о  
влиянии штрафов на его  
бюджет

Получает штрафы на  
большую сумму

7

# Разведывательный анализ

“Миш, мне пофиг, я так чувствую”



$R^2=0,0027$

# Найденная взаимосвязь

“Миш, мне пофиг, я так чувствую”

сумма  
штрафа



средний  
взвешенный = 119.9  
ДОХОД

■ выше среднего  
■ ниже среднего



# Группировка по сумме штрафа

**0**

**Сумма штрафов = 500**

**1**

**Сумма штрафов > 500**

**Описание  
переменных  
ных**

**ВОЗРАСТ**

---

**СТОИМОСТЬ  
МАШИНЫ**

---

**ДОХОД**

---

**ПОЛ**

---

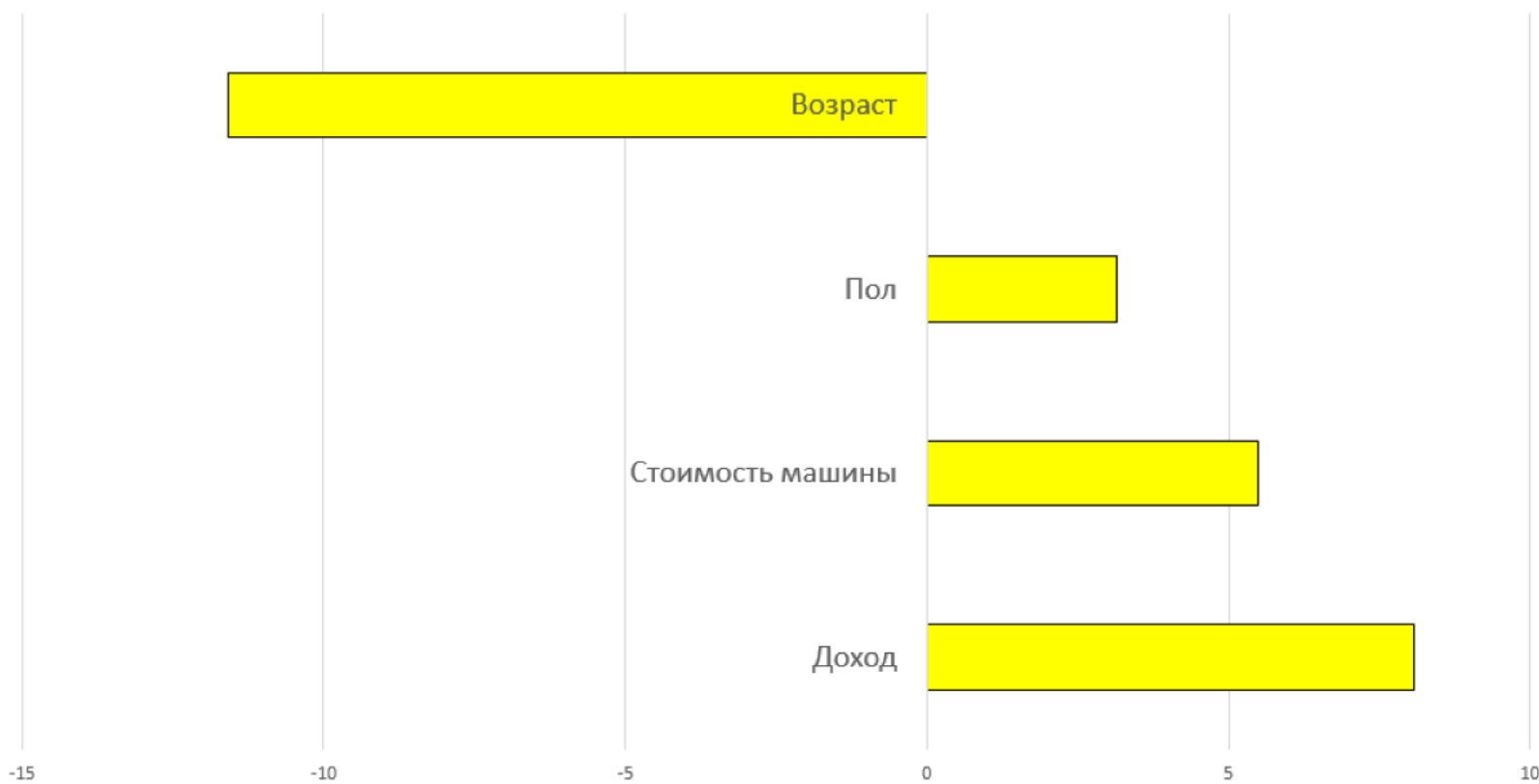
# Методы анализа

“круто, да это ж круто”

- **Логистическая регрессия  
(logit-model)**
  - **Метод наименьших  
квадратов**

# Математическая модель

“Что такое отпуск? У меня отпуска как такового нет”

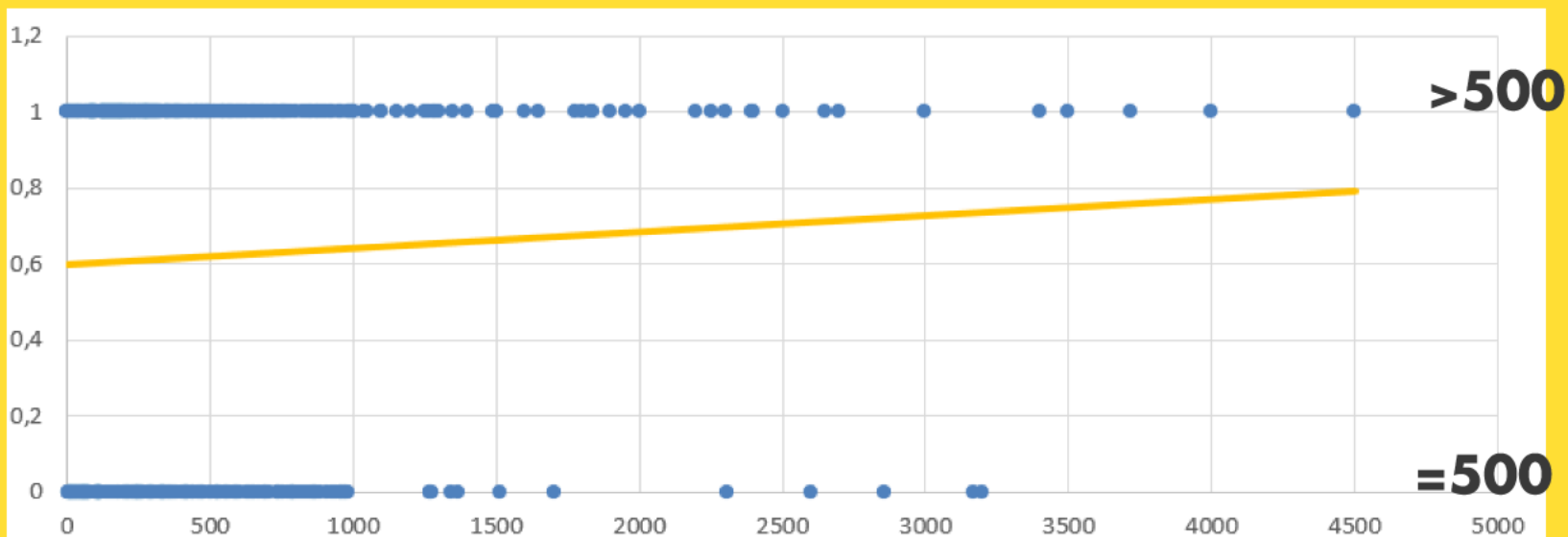


Влияние на вероятность попасть в категорию “>500”



# Математическая модель

“Что такое отпуск? У меня отпуска как такового нет”

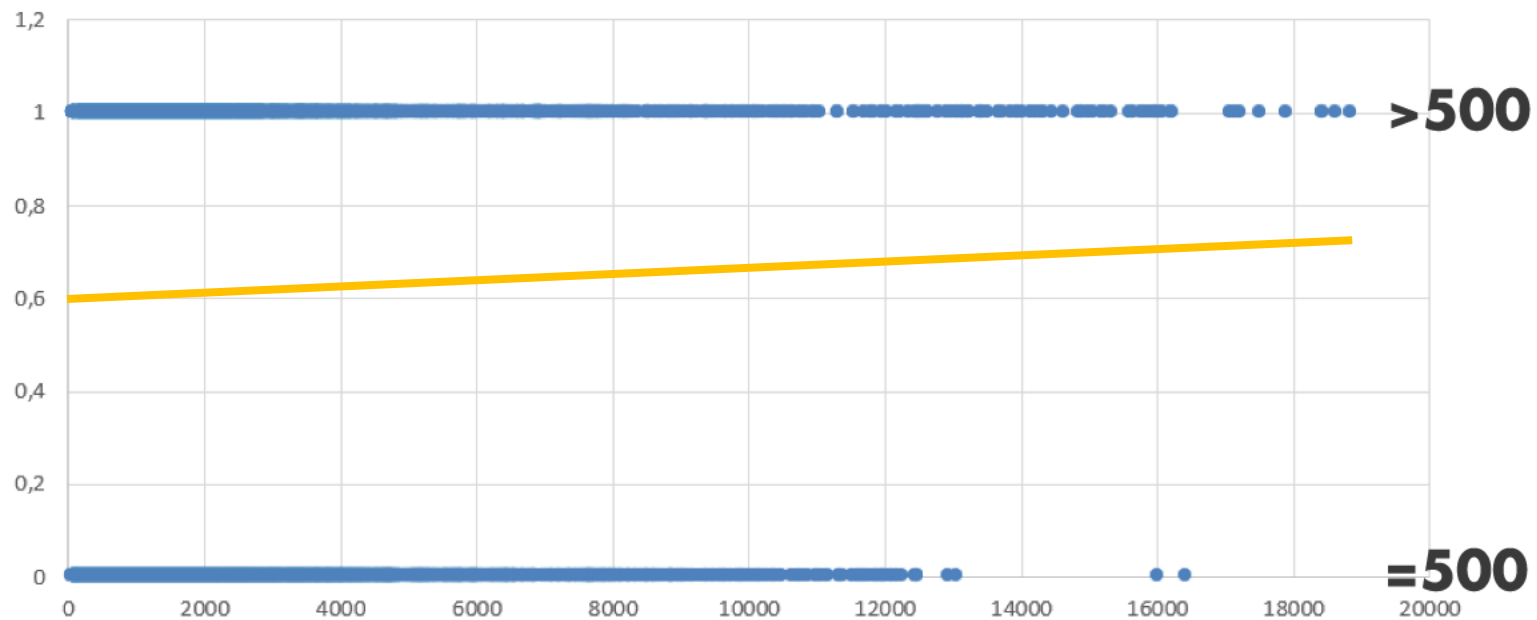


Соотношение доходов и уровня нарушений



# Математическая модель

“Что такое отпуск? У меня отпуска как такового нет”

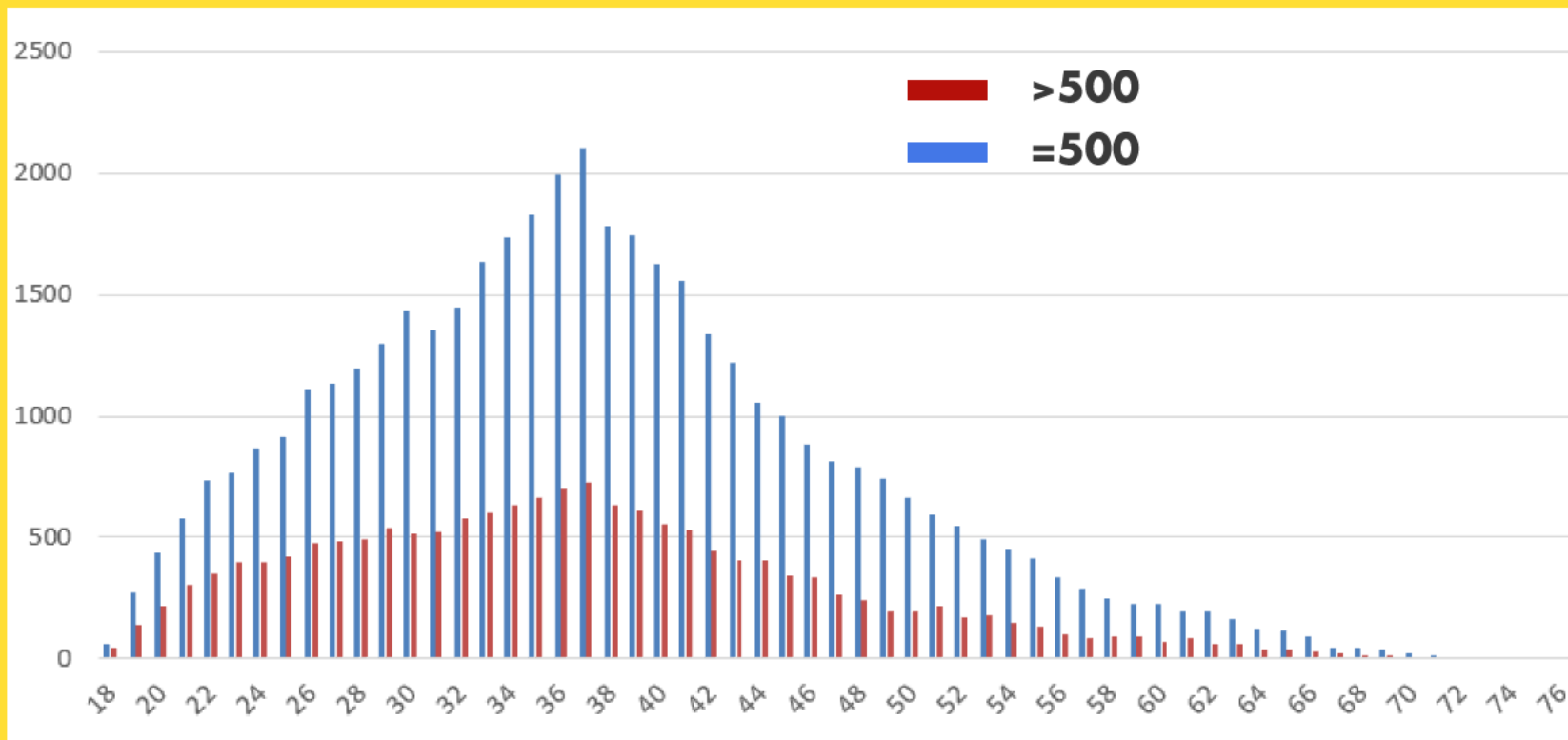


Соотношение стоимости  
машины и уровня нарушений

# 3

# Математическая модель

“Что такое отпуск? У меня отпуска как такового нет”



Влияние возраста на уровень нарушений



# Policy improvement



**“В моей фразе есть много логики, если про неё хорошо подумать”**

**Ввести прогрессивную шкалу штрафов относительно критериев связанных с доходом, такие как стоимость машины и заработок**



# Вывод

“Да, мне было страшно, но я это сделал”

**Гипотеза  
подтверждена.  
Действительно,  
финансовое  
положение  
влияет на сумму  
штрафов**

