

ХАКАТОНОВЫЕ МОНСТРЫ - 21 команда

В главных ролях:
Игнатенко Никита,
Удалов Макар,
Керский Лев,
Абрамов Артем,
Яковлев Тимофей

Структура Данных и подготовка данных

Информация

Типы нарушений

97307 нарушений

85 регионов

Май 2024 года

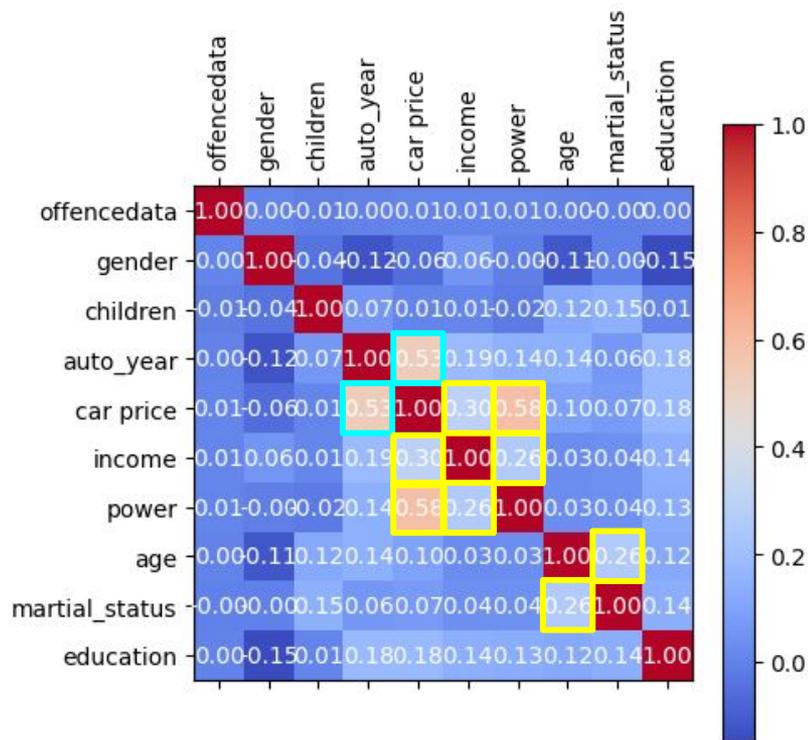
Подготовка данных

Очистка от выбросов

Заполнение пустых ячеек

Перевод текстовых значений к эквивалентным числовым

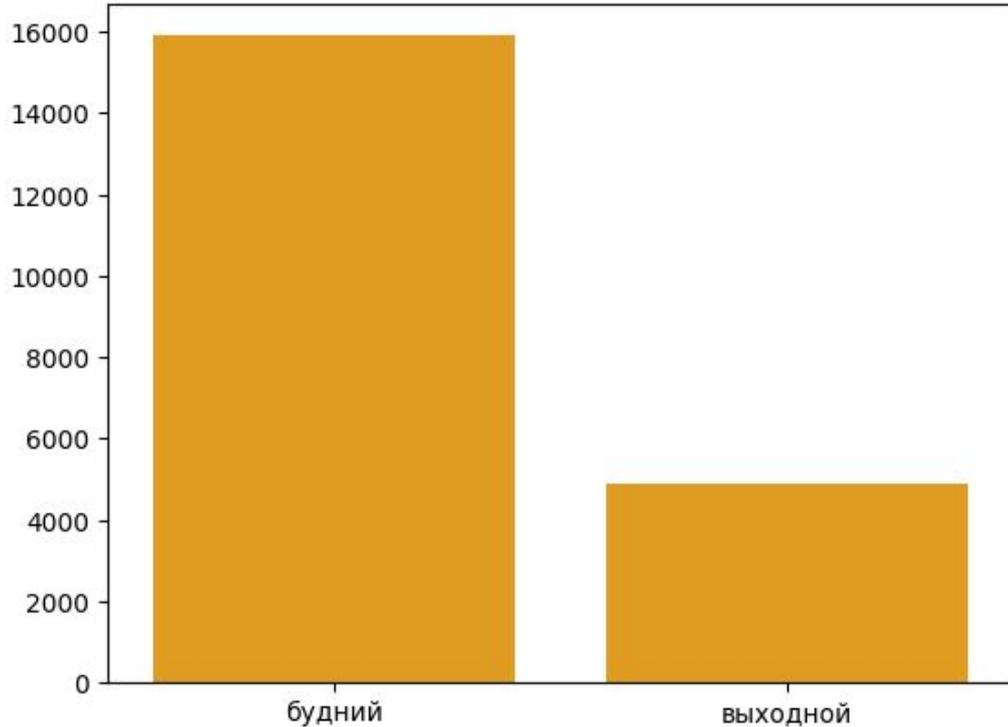
Первичный анализ



замечены корреляции между столбцами:

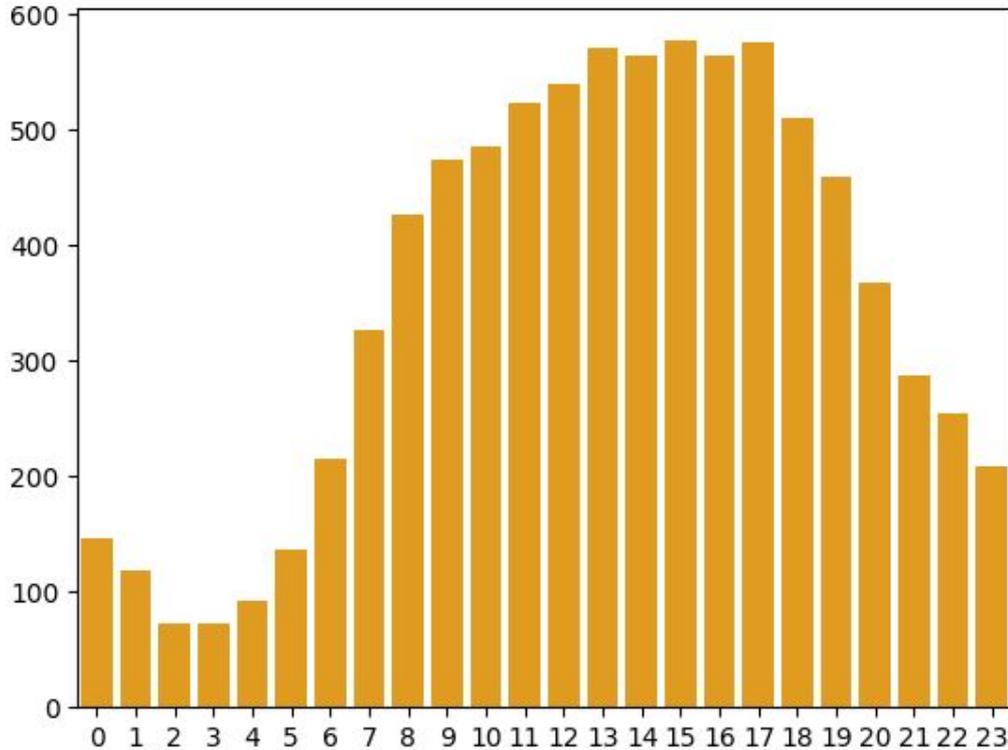
1. ценой машины и мощностью двигателя (0,58)
2. ценой машины и годом выпуска(0,53)
3. ценой машины и зарплатка(0,3)
4. заработок и мощность двигателя(0,26)
5. Семейный статус водителя и возраст(0,26)

Первичный анализ



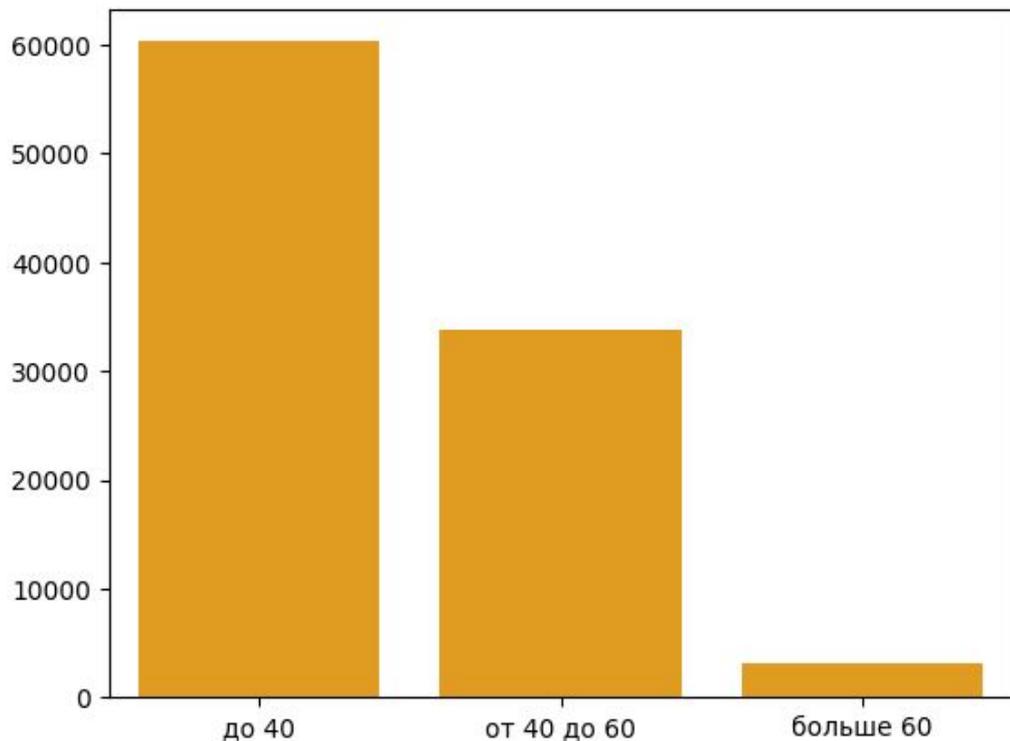
В какие дни больше всего было зафиксировано штрафов?

Первичный анализ



**В какие часы
больше всего было
зафиксировано
штрафов?**

Первичный анализ



Как возраст влияет на кол-во нарушений?

Исследовательский вопрос

Взаимосвязаны ли характеристики клиентов и их автомобилей с характеристиками правонарушений (время/дата совершения правонарушения, статья правонарушения)

Гипотеза:

Правонарушения
чаще всего
происходят в рабочее
время у людей с
низкой социальной
ответственностью.

*Социальная ответственность:

Должность

Наличие детей

Семейный статус

Механизм гипотезы



Факторы

Время

`offencetime(час)`

День недели (Выходной ли?)

`day_of_week`

Должность (Зарплата)

`person_monthly_income_amt`

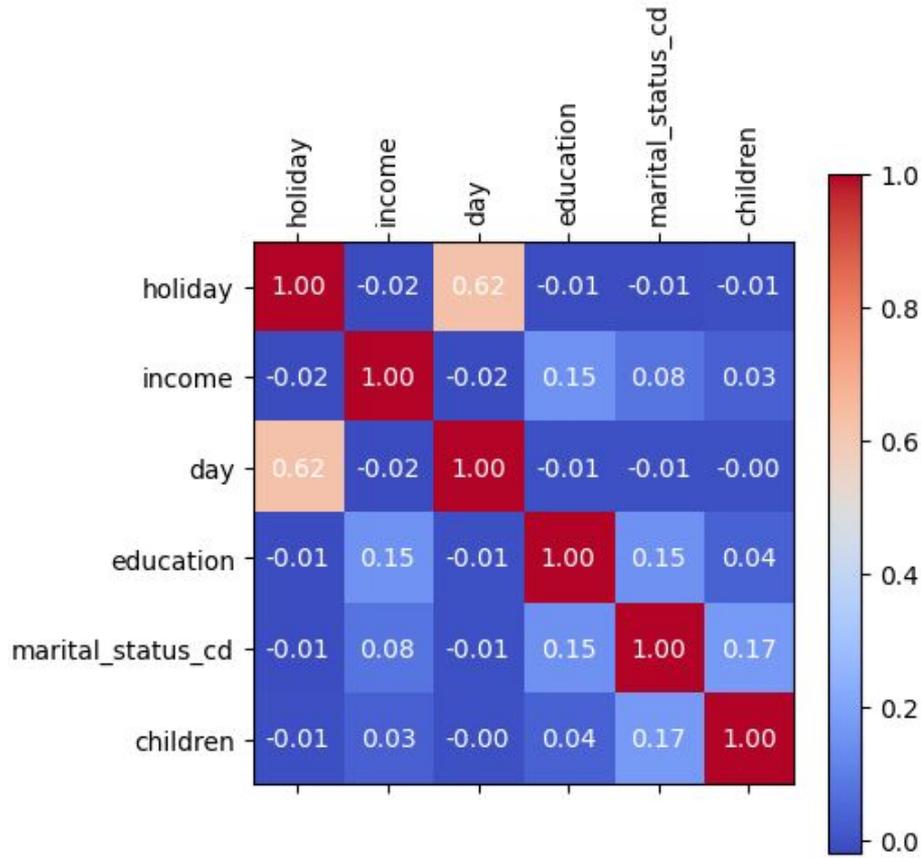
Родительство (Количество детей)

`children_cnt`

Семейное положение

`marital_status_cd`

Множественная регрессия/ Мультиколлинеарность



Построим матрицу корреляций
с параметрами

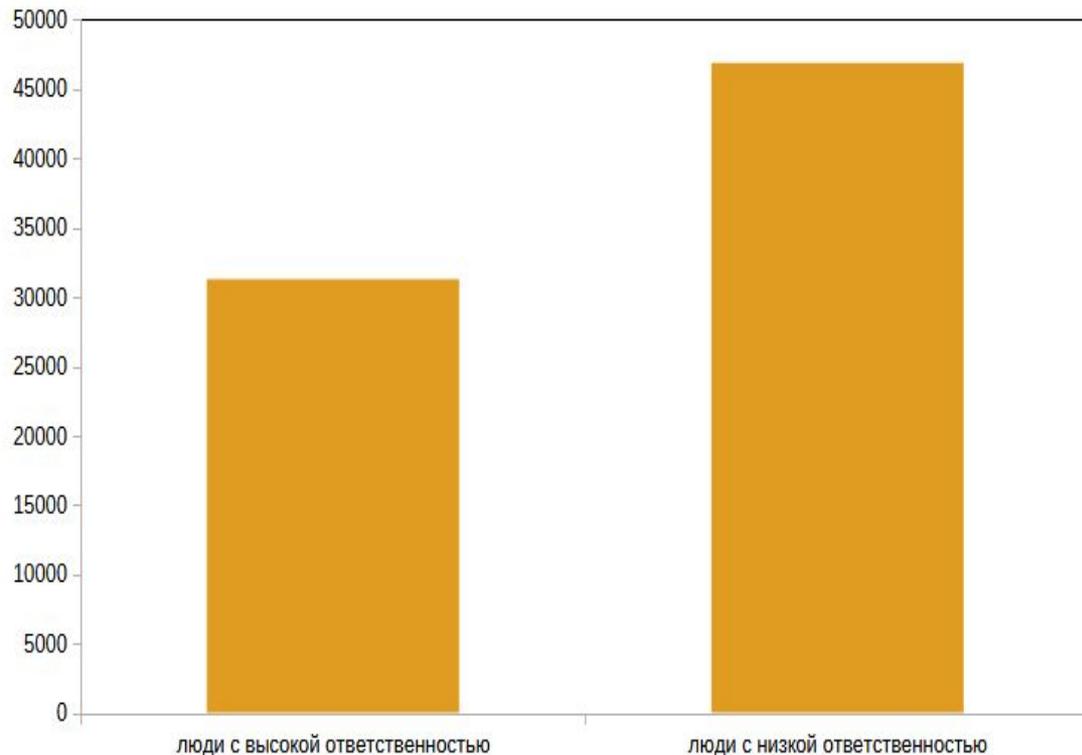
∨

Мультиколлинеарности не
наблюдается, так что можно
сделать регрессию

Множественная линейная регрессия/ Результат

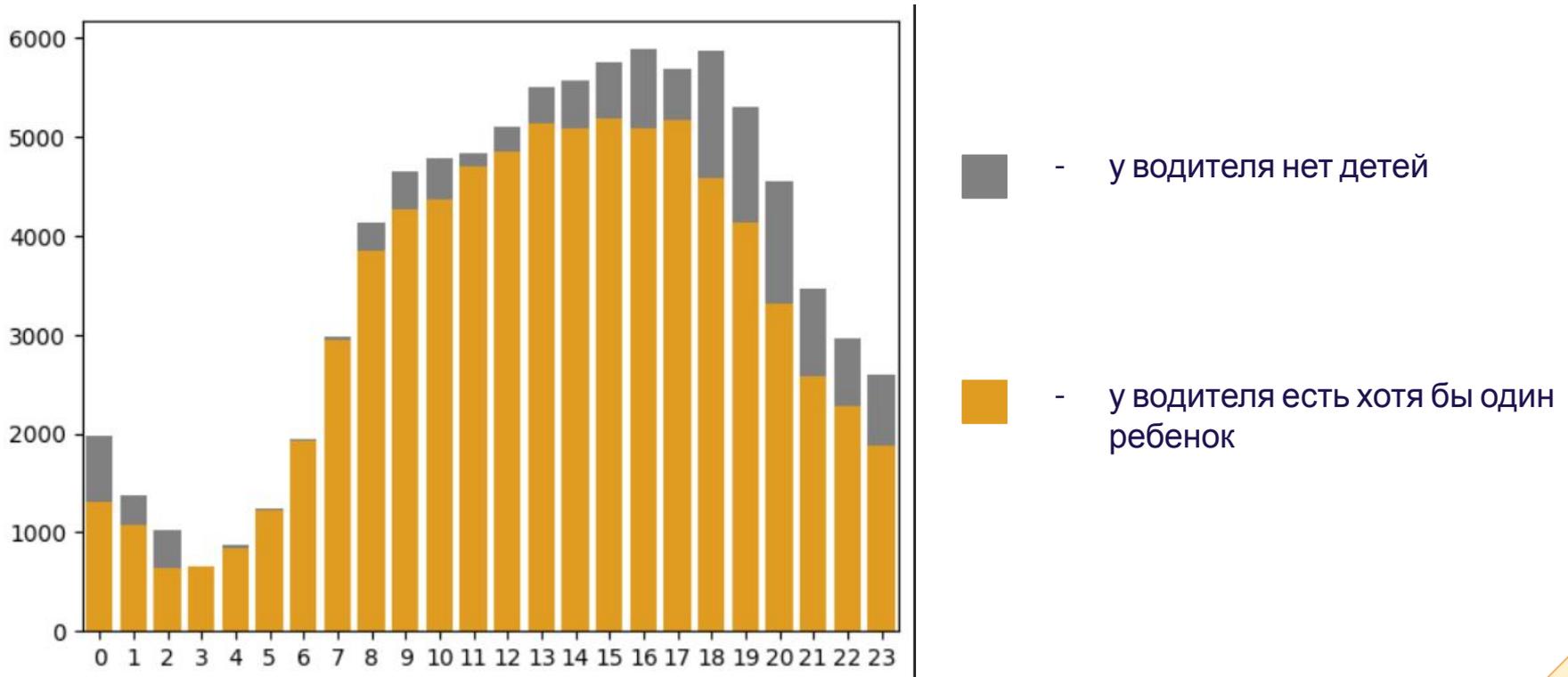
	Переменная	Коэфф.	p-value
✓	Время наруш.	0,3	<0,01
✓	Дата(выходной ли)	-0,406	<0,01
✗	Должность(зарплата)	0,15	0,371
✓	Родительство	1,02	<0,01
✗	Семейное положение	-0.003	0,8

Подтверждение регрессии



**Влияние
социальной
ответственности
на кол-во
штрафов**

Как наличие детей влияет на штрафы?



Ограничения и перспективы

Ограничения

- Рассмотрены данные только за месяц
- Мало количественных переменных
- Предоставлено мало данных о водителе

Чего не хватило

- Использовать данные, полученные за несколько лет
- Получить данные о мере наказания за нарушение

Вывод

Люди с низкой социальной ответственностью чаще нарушают правила, в то время как люди с высокой социальной ответственностью склонны к соблюдению этих норм и, как следствие, реже совершают нарушения.

Рекомендации

Рекомендация №1

Мотивировать компании переходить на свободный график

Рекомендация №2

Мотивировать компании вводить дистанционный формат работы для сотрудников

Рекомендация №3

Ввести программы для новичков-автомобилистов и людей с низкой социальной ответственностью

Время вопросов

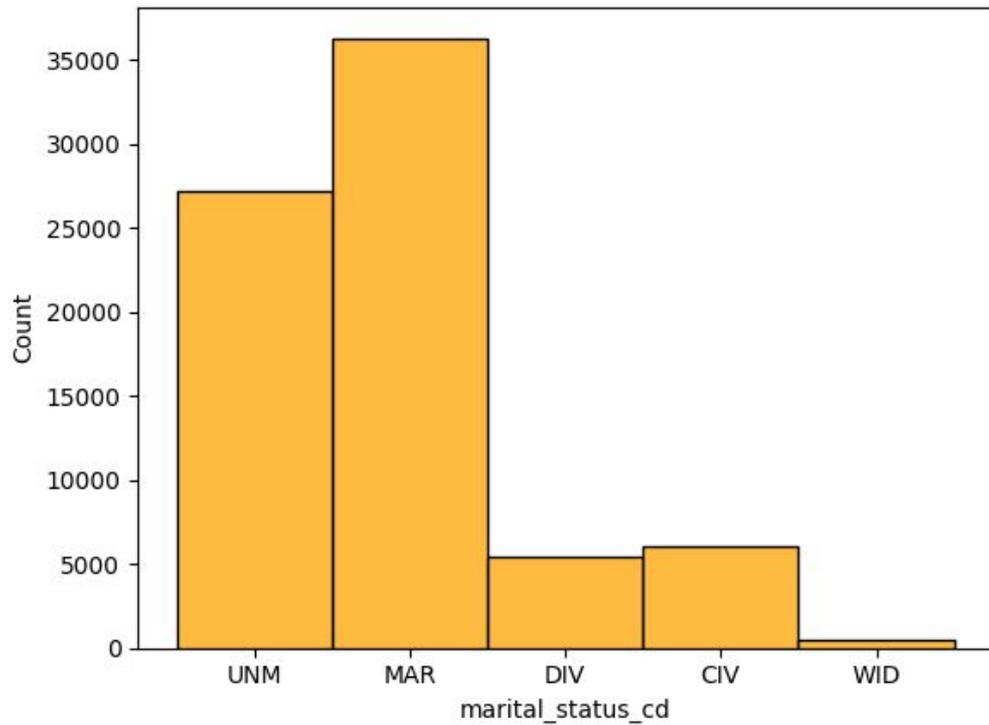


GitHub нашего проекта

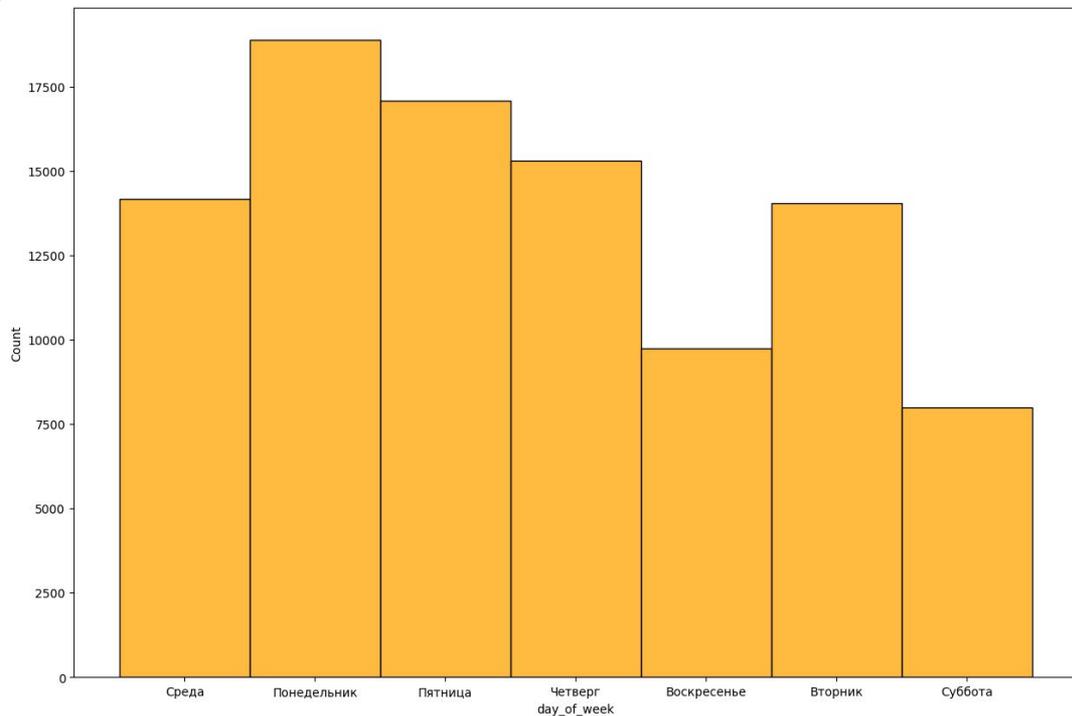


По всем вопросам:





**Как семейное
положение влияет на
кол-во нарушений?**



**Как день
недели
влияет на
количество
штрафов.**