

Введение в статистику

Проект 6

Бои без правил

Кейс, проблема, идея



Кейс

- Платформа показывает пользователям интерактивный видеоконтент
- Платформа планирует повысить вовлечённость пользователей
(сколько шт. видео пользователь смотрит в месяц)

Отдел А разработал продукт А — специальное предложение, дающее доступ к определённому контенту при определённом поведении пользователя

Отдел В разработал продукт В аналогичный продукту А

- Директор решил запустить только один продукт
- Руководитель каждого из отделов считает, что нужно запускать именно его продукт
- Директор поставил задачу провести тестовый запуск: месяц использования, 200 пользователей на каждый продукт
- Для теста взяли с запасом по 500 пользователей на продукт на случай, если нужно будет очищать данные или часть клиентов откажется в процессе
- Аналитик выбирает по 200 пользователей и готовит презентацию для директора с результатами тестового запуска



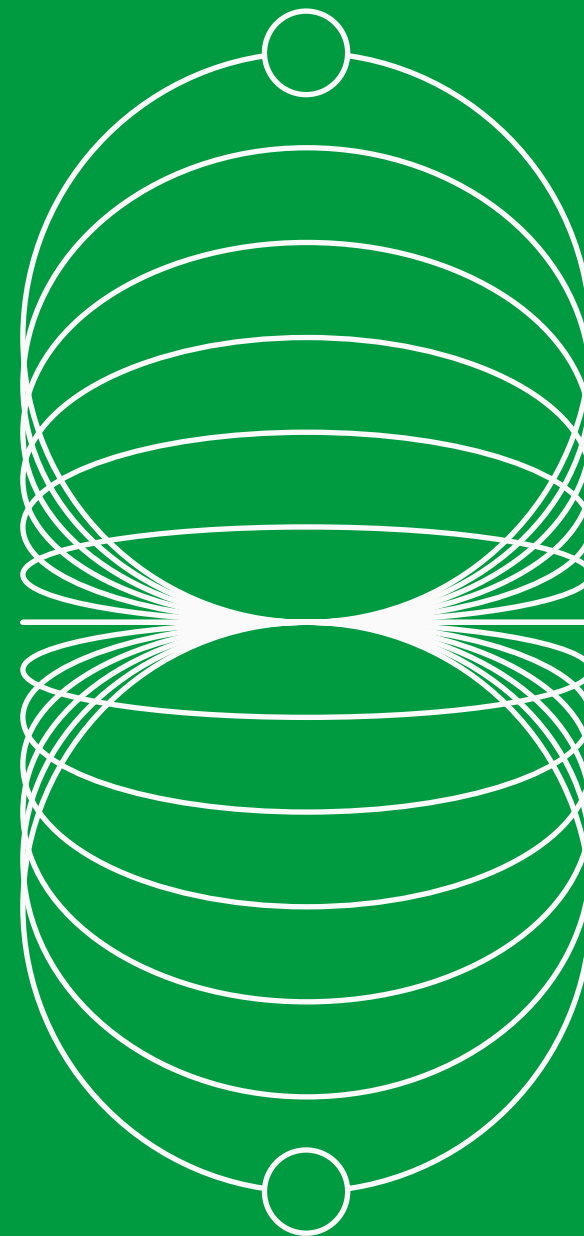
Проблема

Аналитик дружит с руководителем одного отдела и враждует с руководителем другого



Идея решения

Отобрать 200 пользователей так, чтобы вывод в слайдах был в пользу нужного руководителя



Роль, задача, результат



Роль

Аналитик отдела качества продукта



Задача

- Отобрать выборку
- Подготовить слайды, в которых «доказано с помощью данных»: вовлечённость выше у пользователей продукта, предложенного другом, а не врагом



Доступные данные

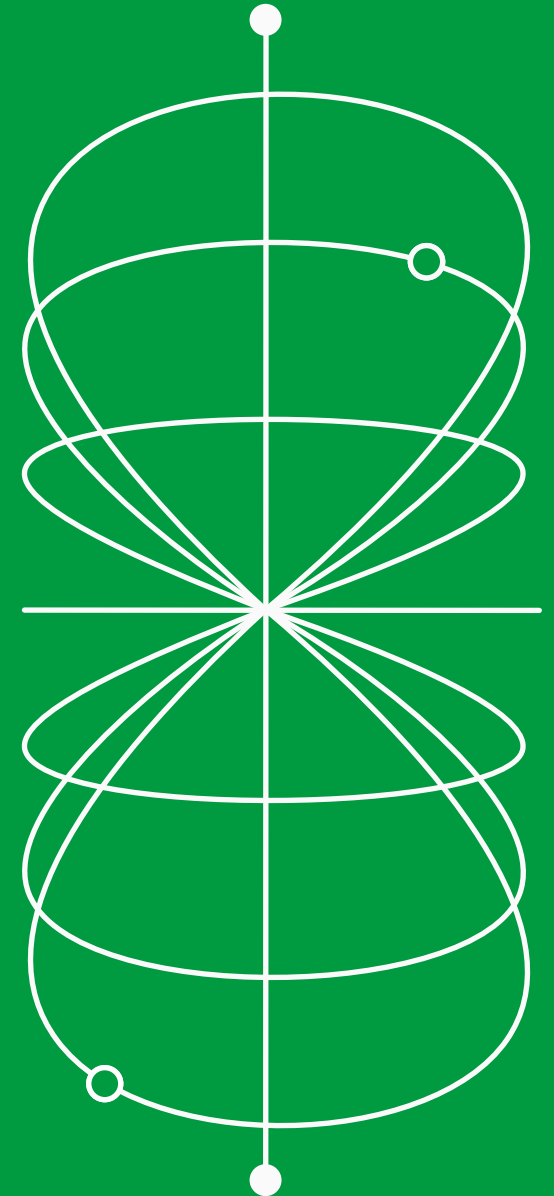
Датасет [without_rules.csv](#), который содержит данные о клиентах:

- возраст
- пол
- получили ли они специальное предложение А или В
- количество просмотренного контента



Ожидаемый результат

- Презентация на слайдах в формате pdf
- Jupyter Notebook в Google Colab с расчётами



Требования к слайдам

Если слайды или Jupyter Notebook не приложен, решение кейса **оценивается в 0 баллов**

Понятность/внешний вид

- Внешний вид презентации
- не мешает воспринимать информацию
- Понятно на какие вопросы отвечает каждый слайд
- Содержимое таблиц, графиков понятно из слайда без необходимости открывать исходный датасет
- Выводы явно сформулированы

МАКСИМУМ 4 БАЛЛА

Обоснованность

- Выводы основаны на таблицах, графиках, показателях, полученных из данных
- Таблицы и графики получены скриншотом или картинкой из Jupyter Notebook, поэтому их можно перепроверить
- Выводы явно сформулированы

МАКСИМУМ 4 БАЛЛА

Реакция заказчика

- **0 баллов**
Не принимает, ищет другого исполнителя
- **1 балл**
Частично принимает, считает необходимым отдать на доработку текущему исполнителю
- **2 балла**
Принимает, готов пересылать слайды от своего имени, под свою ответственность

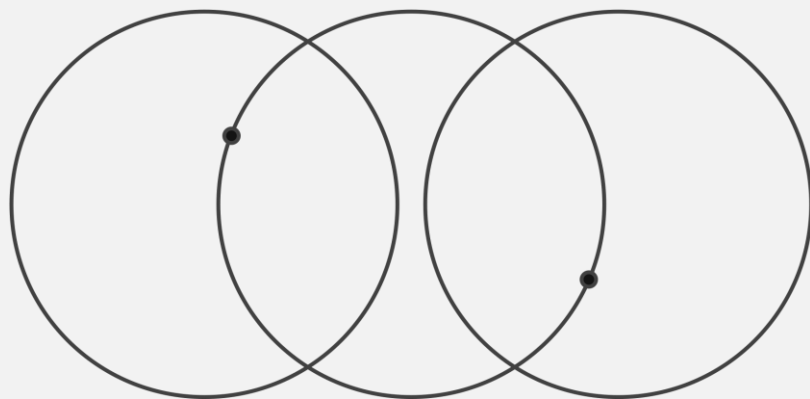
МАКСИМУМ 2 БАЛЛА

Максимальная оценка — **10 баллов**

Бонусные баллы могут поднять основную оценку, но не выше **10 баллов**, даже если в сумме получится 11 или 12 баллов

Синий уровень

Подход к решению



01

Подготовить 2 слайда:

- Руководитель А — друг, руководитель В — враг
- Руководитель В — друг, руководитель А — враг

На каждом из слайдов:

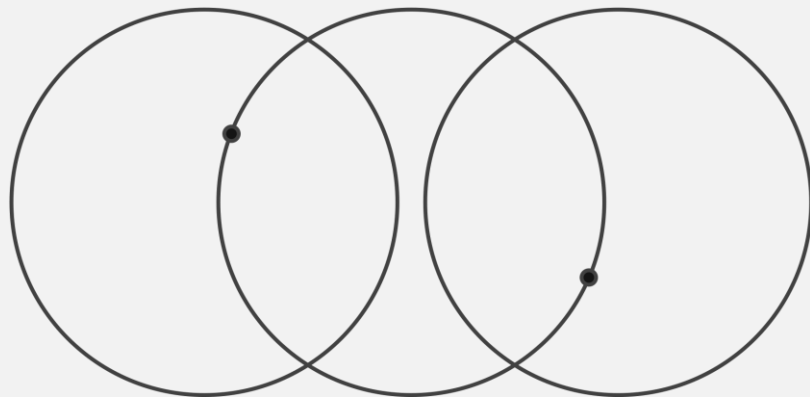
- Показать график или таблицу с вовлечённостью 200 пользователей А и 200 пользователей В
- Сделать вывод на основе графика или таблицы в пользу друга

В Jupyter Notebook с расчётами:

- к исходному датафрейму должны быть добавлены столбцы "Selection Winner A", "Selection Winner B",
- в каждом из них ровно по 400 значений «1»,
- значение «1» в "Selection Winner A" описывают выборку 200+200 пользователей для победы продукта А,
- аналогично в "Selection Winner B",
- нет ограничений на то, какие 400 человек выбирать.

Красный уровень

Подход к решению



01

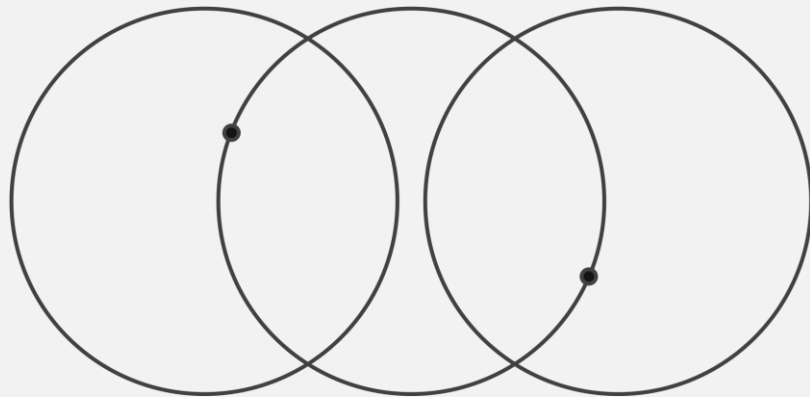
Сделать слайды синего уровня

02

Подготовить ещё 2 слайда: для победы А и В соответственно
Теперь нужно «доказать на основе данных» победу дружественного продукта отдельно для каждой категории пола и возраста 2

Чёрный уровень

Подход к решению



01

Сделать слайды синего и красного уровня

02

Добавить к предыдущим слайдам или показать на отдельных слайдах распределение пола и возраста в каждой из четырех выборок 200 участников

Требуется, чтобы оно было таким же, как и среди всех 500 участников тестового запуска каждого продукта

Что и когда нужно сдать



Что сдавать?

- Презентация в слайдах в формате pdf
- Jupyter Notebook в Google Colab с расчётами



Когда сдавать?

Сроки сдачи указаны в информационной системе

