

Введение в статистику

Проект #6

Бой без правил

Кейс

- "Платформа показывает пользователям интерактивный видео-контент
- Платформа планирует повысить вовлеченность пользователей
- Вовлеченность = сколько видео пользователь смотрит в месяц (штук)
- Отдел А разработал продукт А: специальное предложение, дающее доступ к определенному контенту при определенном поведении пользователя
- Отдел В разработал аналогичный продукт В.
- Директор решил запустить только один продукт.
- Руководитель каждого из отделов считает, что именно его продукт должен быть выбран.
- Директор поставил задачу провести тестовый запуск: месяц использования, 200 пользователей на каждый продукт
- Для теста взяли с запасом на продукт по 500 пользователей на случай:
 - - необходимости очищать данные
 - - часть клиентов откажется в процессе
- Аналитик выбирает по 200 пользователей и готовит презентацию для директора с результатами тестового запуска"

Проблема

Идея решения

Ваша роль

Ваша задача

- Аналитик дружит с одним руководителем отдела, и враждует с другим
-

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

- Аналитик отдела качества продукта
-

- "Отобрать выборку"
- Подготовить слайды, в которых ""доказать с помощью данных"", что вовлеченность выше у пользователей продукта, который предложил ваш друг, а не ваш недруг"

Доступные данные

Ожидаемый результат

- ??

https://raw.githubusercontent.com/zalig/cu-datasets/main/without_rules.csv

- Слайды в PowerPoint с презентацией результатов
- Jupyter Notebook с расчетами для слайдов

Требования к слайдам

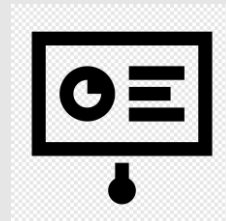
Если слайды или Jupyter Notebook не приложен, решение кейса оценивается в 0 баллов



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
- Расчеты корректны

Максимум: 4 балла



Реакция заказчика (руководителя)

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

Синий уровень: подход к решению

"Подготовьте два слайда:

1. Вы дружите с руководителем А, и враждуете с В
2. Вы дружите с руководителем В, и враждуете с А

На каждом из слайдов:

- - Покажите график или таблицу с вовлеченностью 200 пользователей А и 200 пользователей В
- - Сделайте вывод на основе графика или таблицы в пользу продукта вашего друга

В Jupyter Notebook с расчетами

- - 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

"Подготовьте два слайда:

1. Вы дружите с руководителем А, и враждуете с В
2. Вы дружите с руководителем В, и враждуете с А

На каждом из слайдов:

- - Покажите график или таблицу с вовлеченностью 200 пользователей А и 200 пользователей В
- - Сделайте вывод на основе графика или таблицы в пользу продукта вашего друга

В Jupyter Notebook с расчетами

- - 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»

Подготовьте еще два слайда: для победы А и В соответственно.

- - Теперь нужно "доказать на основе данных" победу дружественного продукта отдельно для каждой категории пола и возраста

Черный уровень: подход к решению

"Подготовьте два слайда:

1. Вы дружите с руководителем А, и враждуете с В
2. Вы дружите с руководителем В, и враждуете с А

На каждом из слайдов:

- - Покажите график или таблицу с вовлеченностью 200 пользователей А и 200 пользователей В
- - Сделайте вывод на основе графика или таблицы в пользу продукта вашего друга

В Jupyter Notebook с расчетами

- - 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»

Подготовьте еще два слайда: для победы А и В соответственно.

- - Теперь нужно "доказать на основе данных" победу дружественного продукта отдельно для каждой категории пола и возраста

"Добавьте к предыдущим слайдам или покажите на отдельных слайдах:

- - Распределение пола и возраста в каждой из четырех выборок двухсот участников такое же, как и среди всех пятисот участников тестового запуска каждого продукта"

Что и когда нужно сдать?

Что сдавать?

- Jupyter Notebook в Google Colab с расчетами
- PowerPoint презентация со слайдами

Когда сдавать?

- Сроки сдачи указаны в информационной системе

! Сдача кейса – необходимое условие прохождения курса