

Описание проекта

Описание данных

Данные для первой части

Файл `/datasets/hypothesis.csv`. [Скачать датасет](#)

- `Hypothesis` — краткое описание гипотезы;
- `Reach` — охват пользователей по 10-балльной шкале;
- `Impact` — влияние на пользователей по 10-балльной шкале;
- `Confidence` — уверенность в гипотезе по 10-балльной шкале;
- `Efforts` — затраты ресурсов на проверку гипотезы по 10-балльной шкале. Чем больше значение `Efforts`, тем дороже проверка гипотезы.

Данные для второй части

Файл `/datasets/orders.csv`. [Скачать датасет](#)

- `transactionId` — идентификатор заказа;
- `visitorId` — идентификатор пользователя, совершившего заказ;
- `date` — дата, когда был совершён заказ;
- `revenue` — выручка заказа;
- `group` — группа A/B-теста, в которую попал заказ.

Файл `/datasets/visitors.csv`. [Скачать датасет](#)

- `date` — дата;
- `group` — группа A/B-теста;
- `visitors` — количество пользователей в указанную дату в указанной группе A/B-теста

Описание проекта

Контекст

Вы — аналитик крупного интернет-магазина. Вместе с отделом маркетинга вы подготовили список гипотез для увеличения выручки.

Приоритизируйте гипотезы, запустите A/B-тест и проанализируйте результаты.

Часть 1. Приоритизация гипотез.

В файле `/datasets/hypothesis.csv` 9 гипотез по увеличению выручки интернет-магазина с указанными параметрами Reach, Impact, Confidence, Effort.

Задача

- . Примените фреймворк ICE для приоритизации гипотез. Отсортируйте их по убыванию приоритета.
- . Примените фреймворк RICE для приоритизации гипотез. Отсортируйте их по убыванию приоритета.
- . Укажите, как изменилась приоритизация гипотез при применении RICE вместо ICE. Объясните, почему так произошло.

Часть 2. Анализ A/B-теста

Вы провели A/B-тест и получили результаты, которые описаны в файлах `/datasets/orders.csv` и `/datasets/visitors.csv`.

Задача

Проанализируйте A/B-тест:

- . Постройте график кумулятивной выручки по группам. Сделайте выводы и предположения.
- . Постройте график кумулятивного среднего чека по группам. Сделайте выводы и предположения.
- . Постройте график относительного изменения кумулятивного среднего чека группы В к группе А. Сделайте выводы и предположения.
- . Постройте график кумулятивного среднего количества заказов на посетителя по группам. Сделайте выводы и предположения.
- . Постройте график относительного изменения кумулятивного среднего количества заказов на посетителя группы В к группе А. Сделайте выводы и предположения.
- . Постройте точечный график количества заказов по пользователям. Сделайте выводы и предположения.
- . Посчитайте 95-й и 99-й перцентили количества заказов на пользователя. Выберите границу для определения аномальных пользователей.

- . Постройте точечный график стоимостей заказов. Сделайте выводы и предположения.
- . Посчитайте 95-й и 99-й перцентили стоимости заказов. Выберите границу для определения аномальных заказов.
- . Посчитайте статистическую значимость различий в среднем количестве заказов на посетителя между группами по «сырым» данным. Сделайте выводы и предположения.
- . Посчитайте статистическую значимость различий в среднем чеке заказа между группами по «сырым» данным. Сделайте выводы и предположения.
- . Посчитайте статистическую значимость различий в среднем количестве заказов на посетителя между группами по «очищенным» данным. Сделайте выводы и предположения.
- . Посчитайте статистическую значимость различий в среднем чеке заказа между группами по «очищенным» данным. Сделайте выводы и предположения.
- . Примите решение по результатам теста и объясните его. Варианты решений: 1. Остановить тест, зафиксировать победу одной из групп. 2. Остановить тест, зафиксировать отсутствие различий между группами. 3. Продолжить тест.