

1 A/B–тестирование

1.1 Условие

Одной из основных задач аналитика в нашей команде является корректное проведение экспериментов. Для этого мы применяем метод A/B–тестирования. В ходе тестирования одной гипотезы целевой группе была предложена новая механика оплаты услуг на сайте, у контрольной группы оставалась базовая механика. В качестве задания Вам необходимо проанализировать итоги эксперимента и сделать вывод, стоит ли запускать новую механику оплаты на всех пользователей.

1.2 Входные данные

В качестве входных данных Вы имеете 3 csv-файла:

- **groups.csv** : файл с информацией о принадлежности пользователя к контрольной или экспериментальной группе (А – контроль, В – целевая группа)
- **active_studs.csv** : файл с информацией о пользователях, которые зашли на платформу в дни проведения эксперимента.
- **checks.csv** : файл с информацией об оплатах пользователей в дни проведения эксперимента.
- Скачать файлы можно [тут](#).

1.3 Вопросы

Предлагаем Вам ответить на следующие вопросы:

- На какие метрики Вы смотрите в ходе анализа и почему?
- Имеются ли различия в показателях и с чем они могут быть связаны?
- Являются ли эти различия статистически значимыми?
- Стоит ли запускать новую механику на всех пользователей?

Данный список вопросов не является обязательным, и Вы можете при своём ответе опираться на собственный план.

1.4 Требования к ответу

- При выполнении тестового задания необходимо использовать язык программирования Python.
- В файле обязательно должны быть комментарии к тем действиям, которые Вы выполняете с данными.
- Файл должен представлять собой законченный отчёт с выводами, сделанными в ходе исследования.

2 SQL

2.1 Очень усердные ученики.

2.1.1 Условие

Образовательные курсы на uchi.ru состоят из различных уроков, каждый из которых состоит из нескольких маленьких заданий. Каждое такое маленькое задание называется "горошиной". Назовём очень усердным учеником того пользователя, который хотя бы раз за текущий месяц правильно решил 20 горошин за час.

2.1.2 Задача

Дана таблица **peas**:

Название атрибута	Тип атрибута	Смысловое значение
st_id	int	ID ученика
timest	timestamp	Время решения карточки
correct	bool	Правильно ли решена горошина?
subject	text	Дисциплина, в которой находится горошина.

Необходимо написать **оптимальный запрос**, который даст информацию о количестве очень усердных студентов за март 2020 года.

2.2 Оптимизация воронки

2.2.1 Условие

Uchi.ru предлагает пройти студентам курсы по модели trial: студент может решить бесплатно лишь 30 горошин в день. Для неограниченного количества заданий в определённой дисциплине студенту необходимо приобрести полный доступ. Команда uchi.ru провела эксперимент, где был протестирован новый экран оплаты.

2.2.2 Задача

Дана таблицы: **peas** (см. выше), **studs**:

Название атрибута	Тип атрибута	Смысловое значение
st_id	int	ID ученика
test_grp	text	Метка ученика в данном эксперименте

и **checks**:

Название атрибута	Тип атрибута	Смысловое значение
st_id	int	ID ученика
sale_time	timestamp	Время покупки
subject	text	Дисциплина, на которую приобрели полный доступ
money	int	Цена, по которой приобрели данный курс

Необходимо в **одном запросе** выгрузить следующую информацию о группах пользователей:

- ARPU
- ARPAU
- CR в покупку
- CR активного пользователя в покупку
- CR пользователя из активности по математике (`subject = 'math'`) в покупку курса по математике

3 Python

3.1 Задача

1. Создайте класс `BaseGraph` со следующими атрибутами:
 - Источник данных (`pandas.DataFrame` или `csv`)
 - Название графика
 - Подписи осей
 - Другие параметры, которые, возможно, захочет увидеть пользователь
2. Создайте метод `plot`, который будет строить `line graph`.
3. Создайте класс `BarGraph`, который будет наследником класса `BaseGraph`. Переопределите метод `plot`, который будет адаптирован для построения `bar chart`.
4. Проиллюстрируйте работу написанных классов на собственных примерах, проверив, что Ваш класс корректно обрабатывает все краевые случаи.