

Первое задание

Необходимо проанализировать использование фильтров пользователями на страницах городов. вопросы, на которые стоит ответить:

1. Сколько людей пользуются фильтрами?

Как отследить людей, которые пользуются фильтрами?

Попытаемся определить, что такое "пользоваться фильтром".

Событие "открытие кнопки фильтры/сортировки/категории" еще не значит, что фильтром в итоге воспользовались, так же как и прожатие чекбоксов или другой процесс внутри фильтра. Если процесс не завершился применением фильтров, считаем, что фильтры не были использованы. Так как у нас нет данных по каждому отдельному пользователю, мы можем только предполагать количество пользователей, которые пользуются фильтрами.

Будем считать, что в рамках одного события `unique_events` - число уникальных пользователей, которые совершили действие (допускаем, что 1 сессия = 1 пользователь).

Соответственно, количество кликов по "применить фильтры" = число людей, использующих фильтры.

62760

2. В каких городах фильтрами пользуются больше? в каких меньше?

Ответ на этот вопрос в абсолютных цифрах будет неинформативным, т.к. разница в количестве посещений страниц разных городов неизбежно влечет разницу в количестве использований фильтра.

Ответ можно привести в долях (применение фильтра/посещение страницы города), перед этим оставив в выборке города, для которых количество применений фильтров больше 5 для того, чтобы доли получились более правдоподобные.

Получены следующие результаты

Ларнака (Кипр), Любляна (Словения), Белград (Сербия), Пафос (Кипр), Айя-Напа (Кипр), Протарас (Кипр), Будва (Черногория), Верона (Италия), Ереван (Армения), Кипр

city	visit	open	submit	close	fil_usage
Larnaca	917	209	179	13	19.52
Ljubljana	124	29	23	3	18.55
Belgrade	1066	230	185	18	17.35
Paphos	694	131	112	6	16.14
Ayia Napa	1881	386	293	43	15.58
Protaras	663	102	91	5	13.73
Budva	549	76	69	6	12.57
Verona	65	10	8	1	12.31
Yerevan	6305	974	767	137	12.16
Cyprus	4197	649	494	77	11.77

где visit - количество посещений страницы города, open - количество открытий фильтров, submit - количество применений фильтра, close - количество закрытий окошка с фильтрами, fil_usage - процент людей, использовавших фильтр по отношению к общему числу посетителей страницы города.

Интересный факт - 5 из 10 объектов с наибольшим процентным использованием фильтров это Кипр 😊

Меньше всего в процентном соотношении используют фильтры для следующих городов:

Хабаровск (Россия), Симеиз (Россия), Абрау-Дюрсо (Россия), Новый Афон (Абхазия), Гоа (Индия), Избербаш (Россия), Хаммаммет (Тунис), Астана (Казахстан), Варадеро (Куба), Чернобыль (Украина)

city	visit	open	submit	close	fil_usage
Khabarovsk	708	15	8	1	1.13
Simeiz	700	13	7	2	1.00
Abrau-Durso	4148	76	40	13	0.96
New Athos	2155	26	20	2	0.93
Goa	920	11	8	1	0.87
Izberbash	697	14	6	2	0.86
Hammamet	702	10	6	1	0.85
Astana	1648	33	13	4	0.79
Varadero	994	23	7	8	0.70
Chernobyl	3334	44	21	6	0.63

4/10 объектов с наименьшим процентным использованием фильтров это Россия

3. Какие разделы фильтров наиболее востребованы? “фильтры”, “сортировка”, “категории”?

Расчеты приведены в приложенном ноутбуке. Результат следующий:

Раздел фильтров	Количество кликов
Фильтры	19246
Категории	45157
Сортировка	17382

Получается, большей популярностью пользуются категории.

Также мы можем, например, посмотреть, какой тип сортировки пользуется большей популярностью

Тип сортировки	Количество кликов
По рейтингу	1778
Сначала популярные	6159
Сначала дешевые	9445

Можно увидеть, что чаще всего выбирают сортировку по дешевизне

4. Как часто люди пользуются выбором цены?

Посчитаем отношение количества кликов по цене (price_first, price_second, price_third) к общему количеству кликов по применению фильтров (price_submit).

Получилось, 14.8% из кликов по применению фильтров приходится на выбор ценового диапазона.

Второе задание

Предположим, что Спутник хочет добавить на сайт новый фильтр по времени начала экскурсий. Мы хотим провести А/Б-тест, чтобы узнать, увеличит ли этот новый фильтр количество бронирований.

Вопросы по тесту:

1. Опиши, какой результат мы хотим достичь с помощью нового фильтра, на какие метрики стоит смотреть.

Пока на сайте не реализован функционал выбора начала экскурсии, пользователь, которому это важно, отобрав по всем остальным необходимым фильтрам какое-то количество экскурсий, скроллит сайт в поисках экскурсии на подходящее время (плашки на нижнем краю карточки с экскурсией).

ИНДИВИДУАЛЬНАЯ 12 3 ч.

Знакомство с Петропавловском-Камчатским

Познакомиться с героической историей
восточного форпоста нашей Родины.

завтра в 06:00

завтра в 07:00



Предположу, что чем больше времени пользователь тратит на поиск подходящего времени, тем больше он устает и тем больше снижается мотивация для бронирования. Таким образом, стоит посмотреть на метрику *время нахождения на странице поиска*. Чем меньше время, тем быстрее пользователь находит подходящую экскурсию - повышаем счастье пользователя и лояльность к нашему сервису. Также, чем быстрее он находит подходящий вариант - тем вероятнее он забронирует экскурсию. То есть стоит также посмотреть на *конверсию в бронирование*.

Возможно, стоит еще последить за *показателем отказов*. Есть вероятность, что добавление новых фильтров - соответственно необходимость делать новые выборы - может оттолкнуть пользователей. Время нахождения на странице может уменьшиться за счет того, что пользователь будет быстрее закрывать страницу. В данном случае показатель отказов стоит считать, если пользователь посетил страницу сайта и не перешел дальше.

Допустим, после проведения теста на протяжении недели, мы получили следующие результаты:

- В группе А (без нового фильтра) из 5000 пользователей бронирование совершили 300 человек.
- В группе Б (с новым фильтром) из 5000 пользователей бронирование совершили 450 человек.

2. Подсчитай конверсию в бронирование для каждой группы. Конверсия в бронирование (доля людей, совершивших бронирование)

- Группа А: 0.06
- Группа Б: 0.09

3. Оцени, насколько новый фильтр повлиял на увеличение бронирований.

В абсолютных цифрах 450 человек значительно больше, чем 350 человек. В долях разница как будто не такая большая: 0.06 против 0.09.

Для точного ответа на вопрос, повлиял ли новый фильтр на число бронирований, нужно провести статистический тест на равенство долей. Подробные расчеты приведены в прикрепленном скрипте. Зададим заранее максимальную вероятность отвергнуть основную гипотезу при ее верности равную 5%.

По результатам проведенного теста достигнутый уровень значимости равен 0, что меньше 0.05 => у нас есть основания отвергнуть нулевую гипотезу о равенстве долей => доли в выборках не равны.

Соответственно, внесенные изменения значимо повлияли на конверсию в бронирование

4. Стоит ли внедрять данный фильтр на основании полученных результатов?

На мой взгляд, внедрять эффект, опираясь только на одну метрику, не стоит. Как минимум, нужно промониторить метрики, указанные выше - например, показатель отказов, ведь если он статистически значимо увеличивается, то потребуются дополнительные тесты и спешить с внедрением не стоит.

Третье задание

Перед тобой лежат 4 карточки:



Определи, какие карточки обязательно необходимо перевернуть, чтобы проверить, выполняется ли для этих четырех карточек правило:

- **“Если на одной стороне карточки – гласная буква, то на другой стороне этой карточки должно быть четное число”**

Максимально подробно обоснуй свой ответ.

Решение:

В задаче не сказано ничего про то, как вообще устроены карточки - нет уточнения, что нарисовано на каждой стороне карточки. В связи с этим в решении я хочу рассмотреть два случая:

- Обозначим допущение, которое заключается в том, что на каждой карточке с одной стороны буква, с другой число. Тогда начнём перебирать карточки по порядку:
 - Карточка А. Согласно допущению, на оборотной стороне написано число. Если оно чётное, тогда правило, которое мы проверяем, выполняется, так как А - гласная буква. Если нечётное - то правило не выполняется. Следовательно, чтобы проверить выполнение правила, её **необходимо перевернуть**.
 - Карточка F. Какое бы число не было на оборотной стороне карточки, для неё выполняется проверяемое правило, так как F - согласная, а значит на неё нет ограничений. Её переворачивать **нет смысла**.
Здесь я хочу сделать замечание: правило, которое мы проверяем, работает только в одну сторону. То есть, если буква гласная, то на оборотной стороне обязательно должно быть чётное число. Но вот если на одной стороне чётное число, то для соответствия правилу совсем необязательно, чтобы на оборотной стороне была гласная буква. Связь односторонняя так как нет связки “тогда и только тогда”.
 - Карточка 2. Из замечания, обозначенного выше, можно сделать вывод, что для этой карточки правило будет выполняться в любом случае, а значит переворачивать её **не имеет смысла**.

- Карточка 7. Если на обороте согласная буква - правило выполняется. Если гласная - то не выполняется. Значит, **её нужно перевернуть.**

Таким образом, нужно перевернуть карточки А, 7.

- В случае, когда этого допущения нет, и на карточках может быть что угодно (буквы на обеих сторонах карточки или числа или даже картинки), карточки А и 7 всё ещё нужно перевернуть по обозначенным выше причинам. Рассмотрим карточки F и 2.
 - Карточка F. Правило может не выполниться, если на оборотной стороне написана гласная буква, в любом другом случае правило выполняется. Значит, эту карточку **нужно перевернуть.**
 - Карточка 2. Что бы ни было на другой стороне - любая буква или любое число, правило будет выполняться, поэтому переворачивать **её не будем.**

В таком случае переворачиваем карточки А, 7, F.

Доказательство также можно записать в терминах матлогики, запишем для варианта с допущением:

Утверждение “**Если на одной стороне карточки – гласная буква, то на другой стороне этой карточки должно быть четное число**” можно записать через импликацию:

$a \rightarrow b$, где a означает гласную букву, b обозначает четное число

Таблица истинности для импликации имеет следующий вид:

a	b	$a \rightarrow b$
0	0	1
0	1	1
1	0	0
1	1	1

Правило не выполняется, если результат импликации равен 0. Как можно видеть, это происходит только в ситуации, когда из истины следует ложь. То есть, в терминах задачи, когда на обороте гласной буквы написано нечетное число - тогда правило нарушается.

Таким образом, нужно проверить карточки, на которых написана гласная буква и те, на которых стоит нечетное число, т.е. карточки А и 7. Пришли к такому же выводу, как выше.

Можно подойти с другой стороны и раскрыть формулу импликации

$$a \rightarrow b = \bar{a} \vee b$$

Получается, что правило точно выполняется, если на карточке согласная буква либо четное число. Таким образом, получается, что нам не нужно переворачивать карточки F и 2.

Ссылка на collab

<https://colab.research.google.com/drive/1lg03myQnCObjFrOoGNArcqK8pn8aYtwL#scrollTo=OL6vUcTLmZYh>