

1. Есть ряд кампаний с определенными настройками (таргет, частота показа креатива). Рекламодатель создает кампанию с целью получить конверсию, но оплачивает клики. Существует гипотеза, что CTR по кампании снижается пропорционально возрасту кампании. Каким образом и какими методами вы бы подтверждали/опровергали данную гипотезу? На какие нюансы обратили бы внимание?
2. Есть 4 формата трафика (a, b, c, d). Существуют определенные индикаторы X и Y, которые описывают источник (source) трафика. Также, каждому формату соответствует определенный threshold, при пересечении которого источник, с которого пришел трафик, считается фродовым, то есть (в общем виде)

$threshold_X = list(a = tx_a, b = tx_b, etc)$
 $threshold_Y = list(a = ty_a, b = ty_b, etc)$

`source_a = if(x > tx_a & y > ty_a){`
 `“fraud”`
`} else if (x > tx_a | y > ty_a) {`
 `“suspicious”`
 `}else{`
 `“not_fraud”`
 `}`

Также существует data.frame() вида

source	X	Y	format
foo	x_1	y_1	a
bar	x_2	y_2	b
baz	x_3	y_3	b
qux	x_4	y_4	c
quux	x_5	y_5	a
quuz	x_6	y_6	d
corge	x_7	y_7	d
grault	x_8	y_8	c

Необходимо написать код, который будет присваивать значение фрода источнику в зависимости от его формата и индикаторов и сохранять в отдельный столбец датафрейма

3. Предположим есть база данных со следующими полями date, publisher (char), advertiser (char), ip (char), event (enum)
Event - это событие, которое происходит. Например,
0 - ошибка при открытии рекламы,

- 1 - показ
- 2 - клик
- 3 - конверсия

Необходимо составить псевдо SQL-запросы следующего типа -

А. Вывести список из 10 паблишеров, у которых наилучшее соотношение конверсия / клик за последнюю неделю

Б. Вывести уникальные сочетания паблишеров и адвертайзеров, у которых хотя бы единожды за день соотношение ошибок и показов было больше, чем среднее значение соотношения ошибок и показов за все время