User Referral Program

Goal

Almost all sites have a user referral program: you can invite new users to try a given product. Typically, after the new user completes a transaction, you get rewarded with a certain amount of money or credit to be used on the site.

The goal of this challenge is to analyze the data from a referral program and draw conclusions about its effectiveness.

Challenge Description

Company XYZ has started a new referral program on Oct, 31. Each user who refers a new user will get 10$ in credit when the new user buys something.

The program has been running for almost a month and the Growth Product Manager wants to know if it's been successful. She is very excited cause, since the referral program started, the company saw a spike in number of users and wants you to be able to give her some data she can show to her boss.

- Can you estimate the impact the program had on the site?
- Based on the data, what would you suggest to do as a next step?
- The referral program wasn't really tested in a rigorous way. It simply started on a given day for all users and you are drawing conclusions by looking at the data before and after the test started. What kinds of risks this approach presents? Can you think of a better way to test the referral program and measure its impact?

Data

We have just 1 table downloadable by clicking **here**.

The table is:

```
    "referral" - provides information about each transaction that happens on
the site and whether the user came from the referral program or not.
```

Columns:

- **user_id** : the id of the user
- **date** : date of the purchase
- **country** : user country based on ip address
- **money_spent** : how much the item bought cost (USD)
- **is_referral** : whether the user came from the referral program(1) or not (0)
- **device_id** : It is an identifier of the device used to make the purchase. You can assume here that for a given physical device, its id never changes

Example

```
    Let's check one purchase event
```

head (referral, 1)

Column Name	Value	Description
user_id	2	this is the user id
date	2015-10-03	she bought something on Oct, 3
country	FR	she is in France
money_spent	65	her purchase cost 65$
is_referral	0	she didn't come via the referral program
device_id	EVDCJTZMVMJDG	to make the purchase, she used a device identified by this id

---

```
    Let's now also check one purchase event from a user who came via the
referral program
```

head(subset(referral,is_referral==1), 1)

Column Name	Value	Description
user_id	5016	this is the user
date	2015-10-31	she bought on Oct, 31
country	DE	she is in Germany
money_spent	45	her purchase cost 45$
is_referral	1	she joined the site via another user referral
device_id	OGAZARJAGCPUQ	this is the device she used to complete the purchase