

I have personally translated this task description from Persian. Please note that due to my NDA I will not be able to attach the dataset (which is about Divar Android app usages).

Explanation of available features:

- **adid** - unique id associated with each user
- **network_name** - Name of the network that each user was in when Divar advertisement was shown and resulted in an installation by that user. Values include:
 - Organic
 - Tapsell
 - Adword UAC Installs
 - Cafebazaar searchad
- **installed_at** - Date Time of installation
- **activity_kind** - Type of activity carried out by a user. Can be *event* or *install*.
- **event_name** - Name of the event. Can be *get_contact* or *new_post_submit*. *get_contact* means touching the button -contact method- for an ad. *new_post_submit* means successfully finishing the process of submitting a post (an ad)
- **created_at** - Timestamp of performed activity

First, perform necessary EDA to get an understanding of data and drop unnecessary/faulty values.

Question 1:

A few metrics are used in order to compare the quality of users coming from each network. Two of them are defined below, measure each one for this dataset:

- Percentage of users that have performed at least one important activity (*get_contact* or *new_post_submit*) in their first three days after installing the app.
- Daily retention rate of said users for the period of one month after the first important activity.

Question 2:

In order to detect fraud in the system, perform two kinds of analysis:

- Distribution of installation times within 24-hour period
- Period between consecutive events for each user

Question 3:

Assume the cost of acquisition is the same for all the networks. Also, we define the value of each *get_contact* to be \$10 and the value of each *new_post_submit* \$100. Compare the performance of each network in terms of their total value. Which network would you choose for future deals?

Question 4:

The context of ads on each network can be classified into two categories:

- Ads aimed at increasing sells, which leads to more *new_post_submit* activities.
- Ads aimed at increasing buys, which leads to more *get_contact* activities.

Assume that for the network called *Adwords*, we have chosen the first ad type and for the network called *Tapsell*, the second ad type was used. Check if these content types have a meaningful effect on how users will behave (sell and buy) or users will behave independently from ad type.