**Problem Statement**

Assuming you’re the data scientist working in the office of alumni of a university. Every year, the university receives donations from a set of people, who might be staff, alumni, external organization, someone not directly related, … To help with the fund-raising, the university decides to reach out to individuals by sending greetings privately and other actions.

The following hypothetical dataset simulates the demographics information from two population, folks who had donated money to the university in the past and folks who never.

The problem statement is as follows – **among those who do not have donation histories, who may be the potential donors and be converted to become the next donors once got in touch?**

Instead of prescribing a strict approach we would leave it up to you to decide what kind of machine learning approach you would want to take.

At the core of it, this is a predictive modelling challenge. Although, you may decide if you want to treat it as a classification/regression problem.

**Data**

Per mentioned above, there are two population provided, which can be distinguished by the column **donor\_or\_not**. Please make necessary assumptions and call them out in your code/submission.

**NOTE**: The data is masked to avoid any potential confusion with actual transactions. As such you will not be able to join any other relevant information you might find in the open-source domain. We request you to confine your analysis to the data provided.

**Data dictionary**

|  |  |
| --- | --- |
| Columns | **Description** |
| name |  |
| location |  |
| latest\_position | latest position in latest work |
| current\_company |  |
| latest\_edu\_place | latest education affiliation |
| latest\_edu\_qual | latest education qualification |
| yrs\_exp | year of working experience |
| volunteer\_flag | is there volunteer experience |
| children\_flag | whether has children |
| Age |  |
| Has 1 Kid |  |
| Has 2-3 Kids |  |
| Has More Than 3 Kids |  |
| Alum Family | whether belongs to alumni family |
| Stay with Parents |  |
| Is Alumnus |  |
| Is Staff |  |
| Is Student |  |
| Is Parent |  |
| IsAddressAvailable |  |
| IsEmailAvailable |  |
| IsPhoneAvailable |  |
| District | the district of address |
| Marital status\_label\_encoded |  |
| Gender\_label\_encoded |  |
| Nationality\_label\_encoded |  |
| donor\_or\_not | whether donated in the past |

**Submission**

Please submit a compressed zip file containing all the artifacts from your analysis/modelling. There should be only **ONE** main Jupyter notebook/R markdown. You may have other additional scripts supporting the analysis.