Assignment
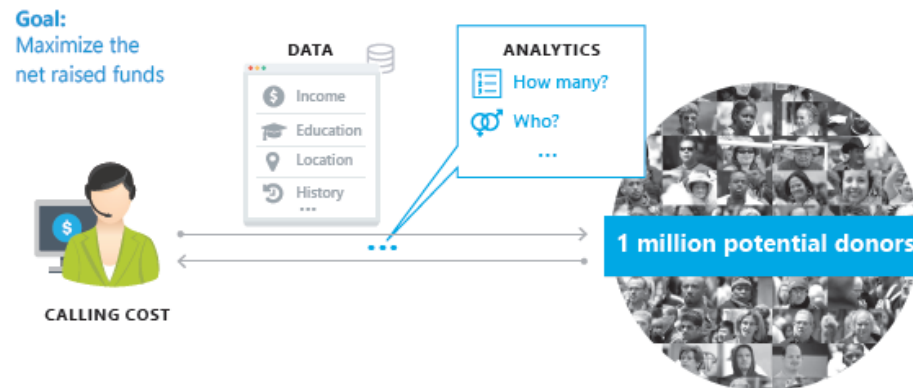
Cortex - A Simulation Game to Learn Predictive Analytics

Game Scenario

You will be working on a fundraising campaign for a not-for-profit organization with a million members. To aid their direct contact campaign, you will be using predictive modeling to **predict how many and which individuals to target**.

Objective: maximize operating surplus

Round 1 Recap

Your mission was to predict the amount that each member would give next year to the foundation. You then used this prediction to contact the most valuable members.

Round 2

Your mission is to increase the net amount of the donations by applying a more effective two-stage modeling approach.

§sas

Two-stage modeling

One way to improve your predictions is to adopt a two-stage modeling approach:

1. For all members, fit a model to predict the probability P that an individual will give

2. Keeping only data from members who gave, fit a model to predict the amount given M

3. Compute P*M to determine the expected donation for each member

MODEL

Probability
of giving

P

MODEL

Conditional amount
if they give

M

Expected Donation = P * M

§sas
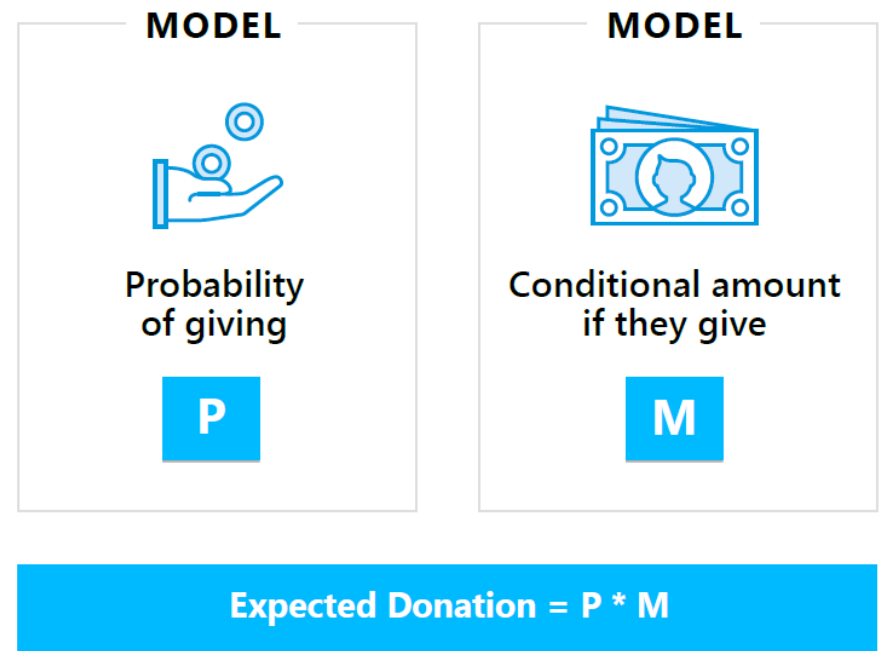
Uplift modeling

For uplift modeling, the following steps are required:

1. Predict the value if a member is contacted PV1

2. Predict the value if a member is not contacted PV2

3. Compute the difference to obtain the uplift generated by contacting a member

24
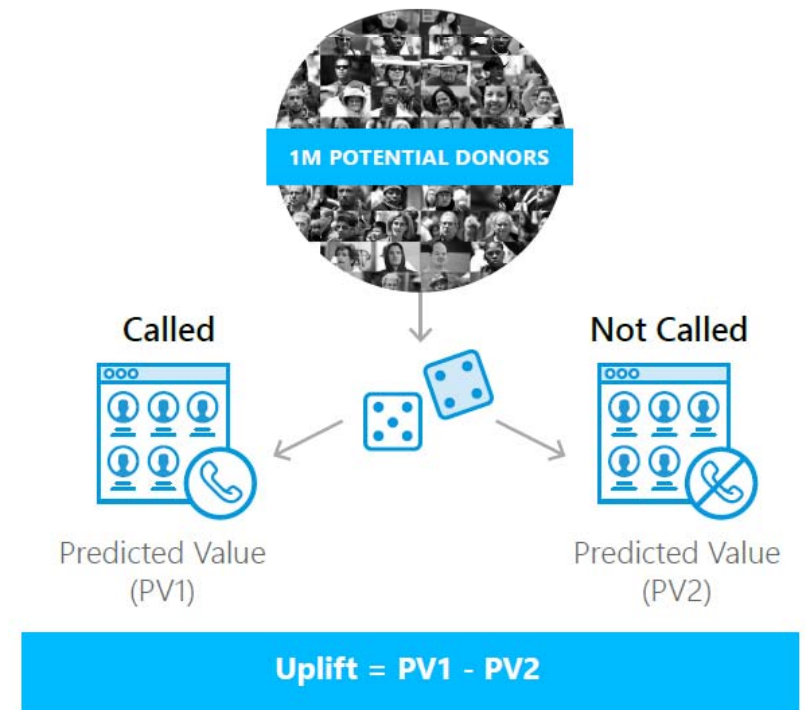
Uplift modeling: Who to Contact

For any targeted action (like a phone call inviting members to donate) there are different reactions:

1. **Persuadable:** are more likely to donate if contacted

2. **Sure Things:** are likely to donate whether contacted or not
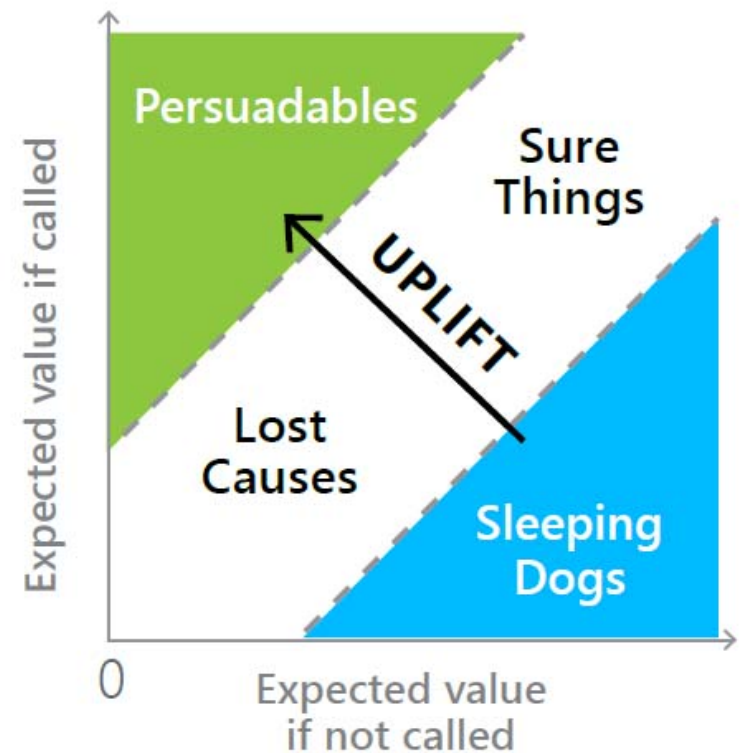
3. **Lost Causes:** are unlikely to donate whether contacted or not

4. **Sleeping Dogs:** are discouraged to donate when contacted

§sas.

Uplift modeling: Who to Contact

The challenge:

- contact the Persuadables

- avoid the **Sure Things** and **Lost Causes**

- avoid the Sleeping Dogs at all costs

§sas.

Stage 1

Goal: predict the **probability** that a member will donate next year → if contacted

→ if not contacted

Some people are more influenced than others. The member 2000010 could be categorized among the persuadables since the probability of giving increases

ID	Gave	AmtGave	PGivingContact	PGivingNoContact
2000001	0	0	0.308045343	0.252860157
2000002	0	0	0.203419312	0.203419312
2000003	0	0	0.203419312	0.203419312
2000004	0	0	0.203419312	0.203419312
2000005	0	0	0.203419312	0.203419312
2000006	0	0	0.203419312	0.203419312
2000007	0	0	0.203419312	0.203419312
2000008	0	0	0.203419312	0.203419312
2000009	0	0	0.203419312	0.203419312
2000010	0	0	0.427631752	0.187452942

§sas

Stage 2

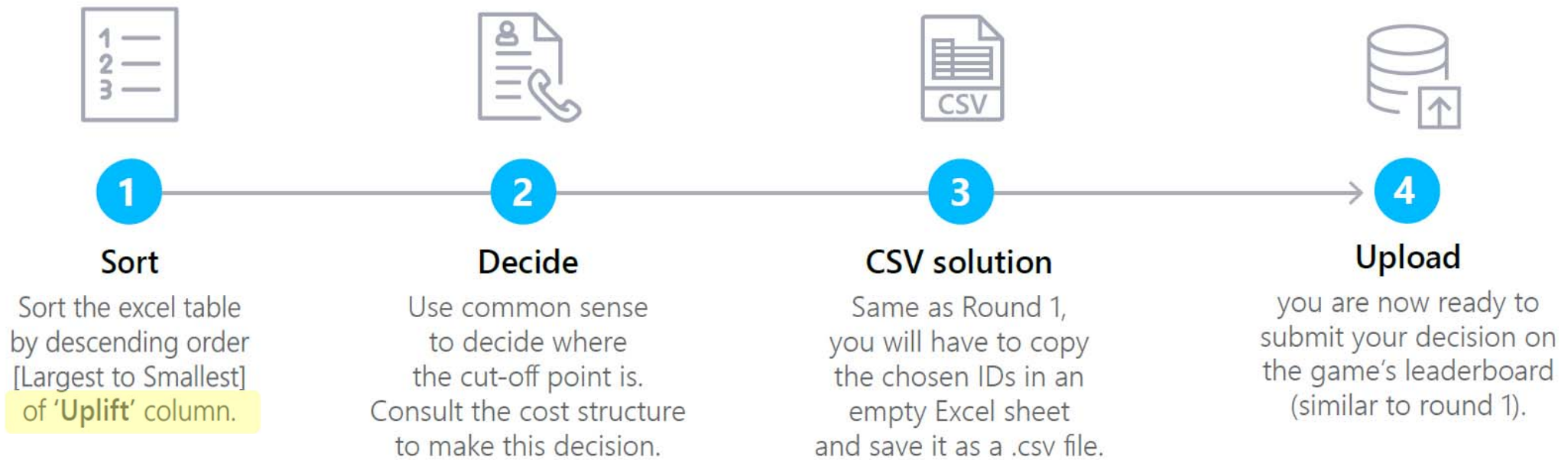**Goal:** predict the **amount** that a member will donate next year for the members who gave

if contacted

if not contacted

§sas.

Calculate Uplift

1. EC: Expected donation if contacted = PGivingContact * PredContact

2. ENC: Expected donation if not contacted = PGivingNoContact * PredNoContact

3. Uplift: Value created by contacting = EC - ENC

	A	B	C	D	E	F	G	H	I	J
1	ID	Gave	AmtGave	PredContact	PredNoContact	PGivingContact	PGivingNoContact	EC	ENC	Uplift
2	2422073	0	0	368,9250731	368,9250731	0,734872067	0,233744654	271,1127	86,23426	184,8785
3	2394415	1	40	267,5369606	267,5369606	0,851819193	0,227434215	227,8931	60,84706	167,0461
4	2201020	1	40	263,0737339	263,0737339	0,854196836	0,232647544	224,7168	61,20346	163,5133
5	2940637	1	20	283,3737778	283,3737778	0,735778966	0,162365502	208,5005	46,01013	162,4903
6	2561134	1	10	365,1596021	365,1596021	0,825811442	0,383343509	301,553	139,9816	161,5714
7	2387866	1	20	274,6629921	274,6629921	0,713165121	0,130874294	195,8801	35,94633	159,9337

§sas.

Submissions

1

Sort

Sort the excel table
by descending order
[Largest to Smallest]
of 'Uplift' column.

2

Decide

Use common sense
to decide where
the cut-off point is.
Consult the cost structure
to make this decision.

3

CSV solution

Same as Round 1,
you will have to copy
the chosen IDs in an
empty Excel sheet
and save it as a .csv file.

4

Upload

you are now ready to
submit your decision on
the game's leaderboard
(similar to round 1).

§sas

Leaderboard

To access the leaderboard:

- Navigate to: https://edge.hec.ca/play/

- Each group will create an account on the Leaderboard.

- Enter your SAS activation code to validate your account.

- Once validated, click "Join Game" and enter game code JBXOU.

§sas

Leaderboard

Leaderboard

1
Uploads

- The leaderboard will be populated as you start uploading decisions.
- You are allowed multiple uploads.
- After every upload you can see how well you performed in the leaderboard.

2
Rankings

- The ranking is based on 'Operating surplus' column.
- Your goal is to obtain a better result than the baseline. In this game, the baseline is a situation where no one is contacted.

3
Final Results

- It is important to choose your 'final decision' before the end of the round.
  Your final choice will be used for the 'Real-Life' leaderboard.
- At the end of the round, uploads are closed, and the "Real-Life" leaderboard is shown (a validation dataset is used for the final performance assessment).

§sas