# DSO 510 Assignment 3

Siqin Yang

**Problem Statement**

The goal is to help Costco solicit more members like their highest value members. To do so, I would first build a response model to predict prospects' response to the solicitation and then use a customer value management (CVM) model to further predict high value members among all responsers.

**Model Design**

Response Model

- ✧ Model: As I have access to data of past solicitations, I would consider the response model as a supervised model, specifically, a binary classification model, to predict if someone will respond to solicitations.
- ✧ Input: Since all data Costco has about past solicitations is PII and response, I'd like to buy some external data like demographic and/or psychographic data from a data broker as inputs of my model.
- ✧ Output: The output of the model would be the probability of response.

CVM Model

- ✧ Model: As I have access to data of current members, I would consider the CVM model as a supervised model, specifically, a regression model, to predict someone's customer value.
- ✧ Input: Since all data Costco has about current members is PII and their transaction details, I'd like to buy some external data like demographic and/or psychographic data from a data broker as inputs of my model.
- ✧ Output: The output of the model would be customer value of a person, which is built on transaction data.

**Proposed Framework**

- ➢ I would do the following steps to build the **response model**.
- ● First I would randomly select several thousands of people solicited by Costco in past 3 years, 20% of whom responded to the solicitation and 80% didn't.
- ● By sending their PII to the data broker, I could buy external data about them to help build the model.
- ● After building the model, I would send the model to the data broker to let them run and return probabilities of response of prospect universe.
- ● Then I would ask the data broker to save a ranked list *S* of those with a probability higher than criterion *A*.
  - ■ About criterion *A*, I would discuss it with management of Costco. Since my final goal is to solicit high value members, I wouldn't set *A* too high, in case of ruling out too many prospects. So probably I would use 0.7 as

the criterion.

- About list *S*, I would further send a list *s* to the data broker to let them remove those on list *s* and save the sub-list *S'* instead.
  - Costco's current members are on list *s*, because I don't want to waste money to solicit them.
  - People did not respond to past solicitations, especially those did not respond to solicitations launched in past 3 years, are on List *s*, too. Because they are not likely to respond this time either.

➢ Then, I would do the following steps to build the **CVM model**.

- First, I would talk with the management of Costco and build a CVM.
  - The likely formulation would be: CVM = a core value * some expert factors. By talking to a few senior stakeholders, I would get their inputs in my CVM design.
    - These inputs should all be accessible within transaction data.
    - These inputs should use recent year's transaction data.
  - The core value might be total revenue a member contributes, while the expert factors could be duration of membership, frequency of shopping, diversity of spend….
- With concrete formulation of CVM, I could calculate CVMs for each member and used them as *y*s of my model.
- Then I would randomly select a list *L* consisted of several thousands Costco members. By sending PII of those on *L* to the data broker, I could buy their external data.
  - CVMs of those on *L* should be normally distributed.
- With external data as my inputs, CVM as my output, I could build the regression model.

➢ After building the response model and the CVM model, I would combine these two to solicit high value members.

- I would send my CVM model to the data broker and let them run on the list *S'*, instead of the prospect universe. And I would ask for a list *H* ranked by predicted CVM.
  - Specifically, I would set a criterion *B* for CVM and let the data broker get me people whose predicted CVMs are higher than *B*.
    - Criterion *B* is decided by management of Costco, as their definition for "high" value members.
    - If the number of people who meet with *B* is not enough, I would ask for a certain number *N* of people instead. Specifically, the first one to the *N*-th one, since the list is ranked by predicted CVMs.
  - The list *H* should has address as a filed in it, so that Costco could mail them.