**Scenario: This is the same as week 1.**

“Healthcare for All”, is a not-for-profit organization that provides financial help to people with who are not able to afford healthcare. “Healthcare for All” raises funds through donations from all across the country. They have an in-house database of over 13 million donors. “Healthcare for All” has been consistently raising money through various campaigns to request

people to contribute to their cause. They reach out to the donors through various channels including personal mails, emails, fund raising events, reaching out to other businesses and corporations, and other philanthropists. They have sent their donors with some gift amounts with the promotions mails.

You are working for “Healthcare for All” as an analyst. You will be analyzing the results of one of their recent direct mail fund raising appeals.

**Objective:**

With this case study the senior management wants to analyze their gifting patterns. As mentioned before, the company has been sending gift amounts to their donors along with the promotions mails. They want to augment/optimize their operations by using some AI/machine learning in the process of deciding the right amount of gifting amount to the prospective customers based on some data that they have collected over the years.

**Task:**

Your task as a data analyst is to help improving the process by developing machine learning models to predict the gift amount based on X features. Follow the data analysis workflow to build the model.

**Data:**

The data is available as a csv document in the folder here for 4.01

Below you will find the description of the features used in the data.

'STATE'. State abbreviation (a nominal/symbolic field)

PVASTATE EPVA State or PVA State

Indicates whether the donor lives in a state

served by the organization's EPVA chapter

P = PVA State

E = EPVA State (Northeastern US)

DOB Date of birth (YYMM, Year/Month format.)

RECP3 P3 File Flag

\_ = Not a P3 Record

X = Donor has given to PVA's P3 program

MDMAUD The Major Donor Matrix code

The codes describe frequency and amount of

giving for donors who have given a $100+

gift at any time in their giving history.

An RFA (recency/frequency/monetary) field.

The (current) concatenated version is a nominal

or symbolic field. The individual bytes could separately be

used as fields and refer to the following:

First byte: Recency of Giving

C=Current Donor

L=Lapsed Donor

I=Inactive Donor

D=Dormant Donor

2nd byte: Frequency of Giving

1=One gift in the period of recency

2=Two-Four gifts in the period of recency

5=Five+ gifts in the period of recency

3rd byte: Amount of Giving

L=Less than $100(Low Dollar)

C=$100-499(Core)

M=$500-999(Major)

T=$1,000+(Top)

4th byte: Blank/meaningless/filler

'X' indicates that the donor is not a major donor.

'GENDER', Gender of the donor

DOMAIN DOMAIN/Cluster code. A nominal or symbolic field.

could be broken down by bytes as explained below.

1st byte = Urbanicity level of the donor's neighborhood

U=Urban

C=City

S=Suburban

T=Town

R=Rural

2nd byte = Socio-Economic status of the neighborhood

1 = Highest SES

2 = Average SES

3 = Lowest SES (except for Urban communities, where

1 = Highest SES, 2 = Above average SES,

3 = Below average SES, 4 = Lowest SES.)

HOMEOWNR Home Owner Flag

H = Home owner

U = Unknown

'INCOME' HOUSEHOLD INCOME

HV1 Median Home Value in hundreds

HV2 Average Home Value in hundreds

HV3 Median Contract Rent in hundreds

HV4 Average Contract Rent in hundreds 'IC1'

IC1 Median Household Income in hundreds

IC2 Median Family Income in hundreds

IC3 Average Household Income in hundreds

IC4 Average Family Income in hundreds

IC5 Per Capita Income

'VETERANS', VETERANS (Y/N)

'RFA\_2', Donor's RFA status as of 97NK promotion date

CARDPROM Lifetime number of card promotions received to

date. Card promotions are promotion type FS, GK,

TK, SK, NK, XK, UF, UU.

MAXADATE Date of the most recent promotion received (in

YYMM, Year/Month format)

NUMPROM Lifetime number of promotions received to date

CARDPM12 Number of card promotions received in the last

12 months (in terms of calendar months translates

into 9603-9702)

NUMPRM12 Number of promotions received in the last 12

months (in terms of calendar months translates

into 9603-9702)

'NGIFTALL', Number of lifetime gifts to date

'TIMELAG']. Number of months between first and second gift

neighborhood demograohics

AGE901 Median Age of Population

AGE902 Median Age of Adults 18 or Older

AGE903 Median Age of Adults 25 or Older

'AVGGIFT'