**Capstone #1:**

Refer:

<https://github.com/Vishnuvardhan777/datascience_playg/blob/master/bank_call_center/bank_mktg_capstone_one.ipynb>

*Data Wrangling:*

1. *Understand if there are any missing values .*

bnk\_mrktg\_raw.isnull().any()

Below is result:

age False

job False

marital False

education False

default False

balance False

housing False

loan False

contact False

day False

month False

duration False

campaign False

pdays False

previous False

poutcome False

y False

2. Understand if there are any categorical values.

# Identify all non integer parameters.

bnk\_mrktg\_raw.dtypes[bnk\_mrktg\_raw.dtypes == 'object']

Below is result:

job object

marital object

education object

default object

housing object

loan object

contact object

month object

poutcome object

y object

dtype: object

3. Replace y from no or yes to 0 and 1

bnk\_mrktg\_raw.y.replace({'no': 0, 'yes':1}, inplace=True)

4. Convert some of the categorical to set of booleans

# create hot encoding for below attributes.

bnk\_mrktg\_raw\_encoded = pd.get\_dummies(bnk\_mrktg\_raw,columns=['education','job','housing','day','marital', 'poutcome','default','loan','contact','month'])