##===================Domain Understanding

# SalesLead\Lead : a person\organization who may eventually become a client or purchanse the prodcut

#Several factors determine the quality of sales leads, such as whether the individual or business had the incentive to offer contact information, the accuracy of the data provided, and the validity of the sales lead.

Lead generation is the way that you attract and get people to give you their contact information to you.

COLUMNS :

Product\_ID : Unique identifier which the determines the product i.e product tracebility

Source : Source from where lead info is aquired i.e email,social media,website,blogs ..etc

Sales\_Agent : Person who invloved in getting the lead info

Location : Location to which lead beong to

Status : Status of the lead

Delivery\_Mode : already clean

* FicZon Inc is an IT solution provider with products ranging from on- premises products to SAS based solutions. FicZon major leads generation channel is digital and through their website.
* FicZon business is majorly dependent on the sales force effectiveness. As the market is maturing and more new competitors entering the market, FicZon is experiencing the dip in sales.
* Effective sales is dependent on lead quality and as of now, this is based on manual categorization and highly depended on sales staff. Though there is a quality process, which continuously updates the lead categorization, it’s value is in for post analysis, rather than conversation.
* FicZon wants to explore Machine Learning to pre-categorize the lead quality and as result, expecting significant increase in sales effectiveness.

##====================Tasks and assmptions

* # Sales dataset with shape 7422 entries , and 7 columns
* # Mobile and Email : columns doesnot add any significancec to dataset in predicting the lead potential we are dropping these columns
* #====Product\_ID : There are 58 rows whose Product\_ID='' ,replace it with Mode
* #====Status : Converting mutilclass values to binary i.e into high potential(1) n low potential leads(0)

li\_high =['Potential','In Progress Positive','CONVERTED','converted','Long Term']

li\_low=['Open','Not Responding' ,'Junk Lead','Just Enquiry','In Progress Negative','LOST']

* #====SOurce : forward filling the 7 empty string values
* #====Sales\_Agent : replace empty string with the mode value
* #====Location : replace 6 empty string with the mode value