# **Structural Pyramid Analysis Plan (SPAP)**

1. S.M.A.R.T Goal:

We want to determine factors affecting clients’ decisions of buying term deposits from a bank.

2. Measures of the dependent variable:

1. The binary variable that shows whether a client subscribes to a term deposit. (Data.world database, binary values summed by dates) Based on it, we calculate the ratio of 1 (yes to the term deposit).
2. The binary variable that shows whether a client subscribe a term deposit. (Data.world database, binary values summed by names of clients) Based on it, we calculate the ratio of 1 (yes to the term deposit).
3. The binary variable that shows whether a client subscribe a term deposit. (Data.world database, binary values summed by names of deposit products) Based on which, we calculate the ratio of 1 (yes to the term deposit).

3. Propose potential aspects for independent variables:

* 1. Do clients’ characteristics specified above influence their decisions of buying term deposits?
  2. Do clients’ previous contacts information specified above influence their decisions of buying term deposits?
  3. Does the performance of market specified above influence their decisions of buying term deposits?

4. Determine independent variables:

1. From hypothesis 1 and the first question above, we can tell that independent variables may contain age, type of job, marital status, default record, ongoing housing loan, ongoing personal loan, and education level.
2. From hypothesis 2 and the second question above, we can tell that independent variables may contain number of days that passed by after the client was last contacted from a previous campaign, contact communication type, last contact month of year, last contact duration, number of contacts both before and during the campaign, and outcome of previous campaign on clients.
3. From hypothesis 3 and the third question above, we can tell that independent variables may contain employment variation rate, consumer price index, consumer confidence index, EUIRBOR 3-month rate, and number of employees.

5. Evaluate the accessibility of the data:

1. We know the location: consumer price index, consumer confidence index, EUIRBOR 3-month rate, type of job, age, marital status, default record, ongoing housing loan, ongoing personal loan.
2. We only know the existence: number of employees, contact communication type, outcome of previous campaign on clients, number of days that passed by after the client was last contacted from the last campaign, number of contacts both before and during the campaign, last contact month of year,.
3. The existence is unclear: employment variation rate, last contact duration, education level.

6. For independent variables that the locations are known, here are the table of primitive analysis:

Since the y axis for every independent variable is the ratio of yes in purchasing the term deposits, we omit the column of y axis.

|  |  |  |
| --- | --- | --- |
| independent variable | Chart type | x-axis |
| consumer price index | line | consumer price index |
| consumer confidence index | line | consumer confidence index |
| EUIRBOR 3-month rate | line | EUIRBOR 3-month rate |
| last contact month of year | bar | last contact month of year |
| Number of days that passed by after the client was last contacted from a previous campaign | line | Number of days that passed by after the client was last contacted from a previous campaign |
| Number of contacts both before and during the campaign | bar | Number of contacts both before and during the campaign |
| age | bar | age |
| type of job | bar | type of job |
| marital status | bar | marital status |
| default record | bar | number of defaults |
| ongoing housing loan | bar | number of housing loans |
| ongoing personal loan | bar | number of personal loans |