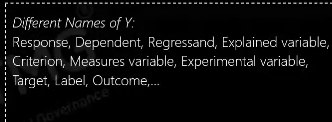
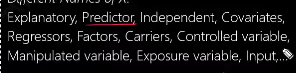
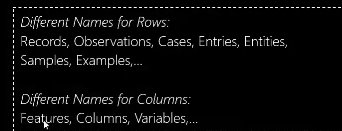
**VARIOUS NAMES OF OUTPUT VARIABLE NAMES IN INTERVIEW**

****

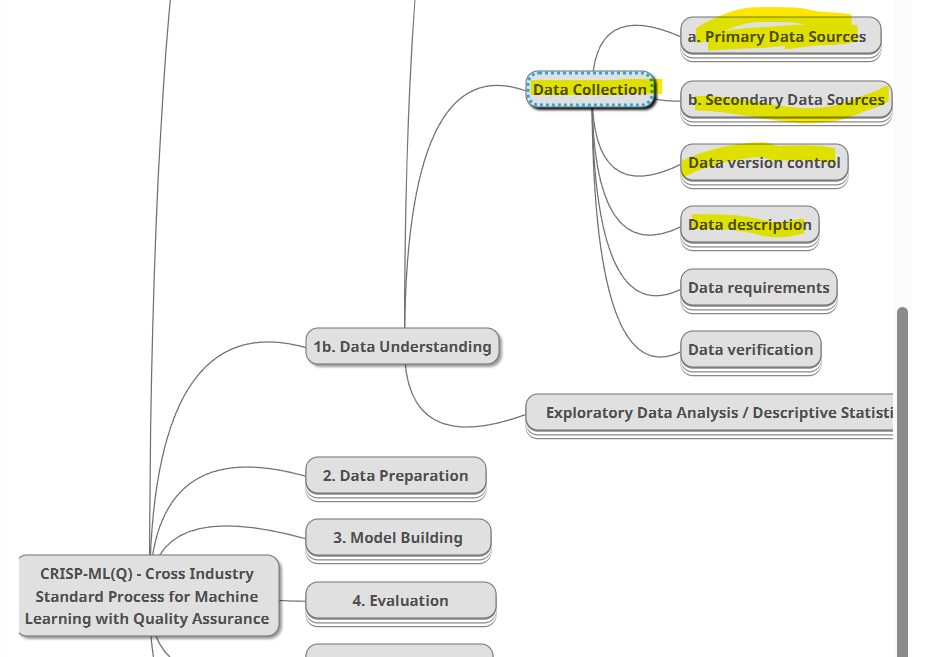
**VAROUS NAMES OF INPUT VARIABLES**

****

**VARIOUS NAMES OF ROWS**

****

**DATA COLLECTION**

****

**TYPES OF DATA SOURCE**

****

**Example for secondary data:**

To implement 5G network on a city we have to analysis the places of the city to put the tower, and that data can be analyzed by the google maps.

**Example for primary data:**

To build a model to predict if a person will pay the education loan or not, let’s assume we build model with already existing that are loan application, by analysis the application of the loaner we will came up with one prediction rate, now we want to improve our model’s accuracy, for that we planned to scrape the user’s google account or facebook account details, by analysis the facebook account of user we can now if the person will pay our loan or not.

example if a person is a tech youtuber guy who constantly buys tech gadgets and post it on facebook page, by analyze this person’s facebook data we we can that the person will not consider to pay the loan in first priority because he constantly buys gadgets to make youtube review, this is **called primary data**

Because the user’s facebook data is not already ready to make analysis related loan, some people will not have any single post. So we made analysis to grab the related information to predict if the person will pay loan or not.

This is called primary data.

**Difference between primary and secondary**

1. Primary – we will take the data we needed, we can get exact data to train]
2. Secondary – There will be irrelevant data exist we can’t get exact data to train

