مشتریان وام بانکی

This problem is about a bank whose management wants to explore ways of converting its liability customers to personal loan customers. A campaign that the bank ran last year for liability customers showed a healthy conversion rate of over 9% success. This has encouraged the retail marketing department to devise campaigns with better target marketing to increase the success ratio with a minimal budget.

The file train_data.csv contains data on 4500 customers. You can download the file from here: https://www.dropbox.com/s/pt6bleedtpo3h5c/train_data.csv?dl=0 The data include customer demographic information (age, income, etc.), the customer's relationship with the bank (mortgage, securities account, etc.), and the customer response to the last personal loan campaign (yes/no). Among these customers, only about 9% accepted the personal loan that was offered to them in the earlier campaign.

Columns:

- ID: Customer ID
- Age: Customer's age in completed years
- Experience: Number of years of professional experience
- Income: Annual income of the customer ($000)
- ZIP Code: Home Address ZIP code.
- Family: Family size of the customer
- CCAvg: Avg. spending on credit cards per month ($000)
- Education: Education Level. 1: Undergrad; 2: Graduate; 3: Advanced/Professional
- Mortgage: Value of house mortgage if any. ($000)
- Personal Loan: Did this customer accept the personal loan offered in the last campaign?
- Securities Account: Does the customer have a securities account with the bank?
- CD Account: Does the customer have a certificate of deposit (CD) account with the bank?

- Online: Does the customer use internet banking facilities?
- CreditCard: Does the customer use a credit card issued by this Bank?

Your job is to do EDA and build a classification model to predict whether a customer accepts the Personal Loan or not? In your jupyter notebook, per cell, you need to explain (in natural language) why you are doing that part. Also, you need to explain what you have gained/understood from that part. If you only provide code, you will not get any mark (I need reasons). I've shuffled and separated a portion of the data for the private assessment of your task.

Submit your jupyter notebook as a zip file. Please only use the libraries that we have discussed in the class. Please Please Please, DO NOT copy code from anywhere on the internet. It should be your work and I will check similarity of your work with the existing codes out there.