**Assignment 1**

**Assignment questions:**

* Ratio of Good to Bad = 700/300=2.33
* Most of the missing values are observed in following attributes

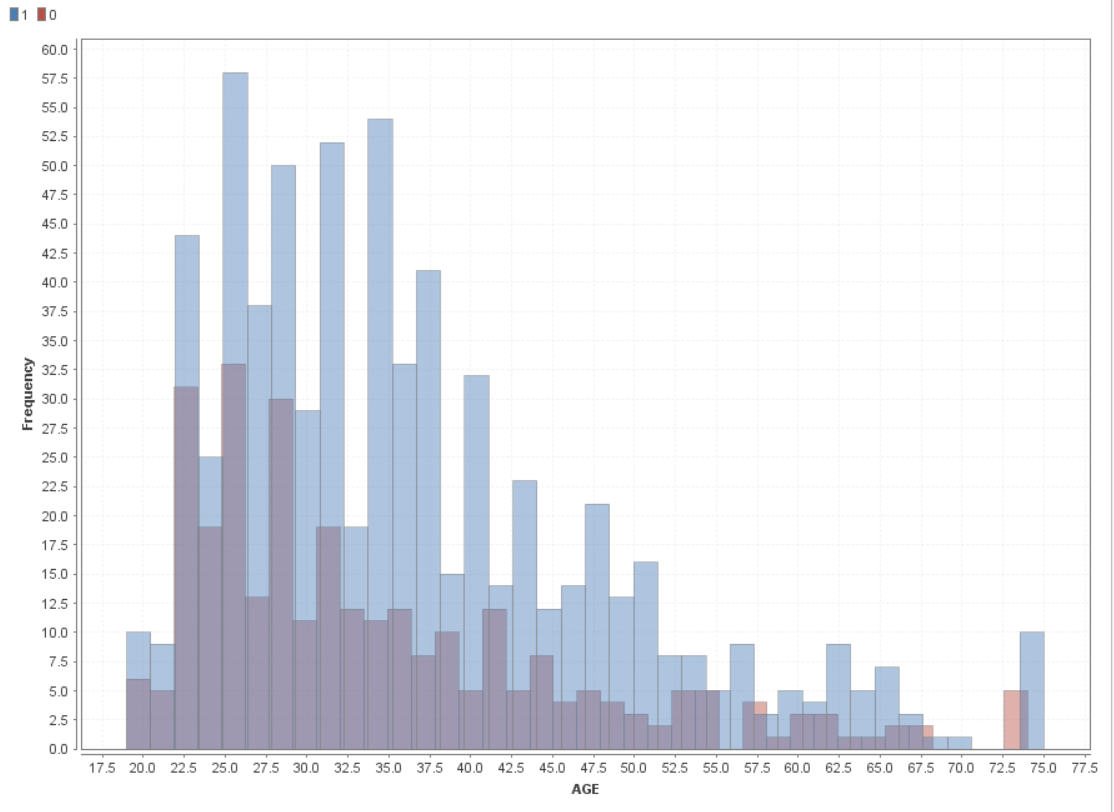
1. NEW\_CAR: Binominal variable, where Yes=1, No= Missing data, will be replace by 0
2. USED\_CAR**:** Binominal variable, where Yes=1, No= Missing data, will be replace by 0
3. FURNITURE**:** Binominal variable, where Yes=1, No= Missing data, will be replace by 0
4. RADIO\TV**:** Binominal variable, where Yes=1, No= Missing data, will be replace by 0
5. EDUCATION**:** Binominal variable, where Yes=1, No= Missing data, will be replace by 0
6. RETRAINING: Binominal variable, where Yes=1, No= Missing data, will be replace by 0

* Independent variables are:

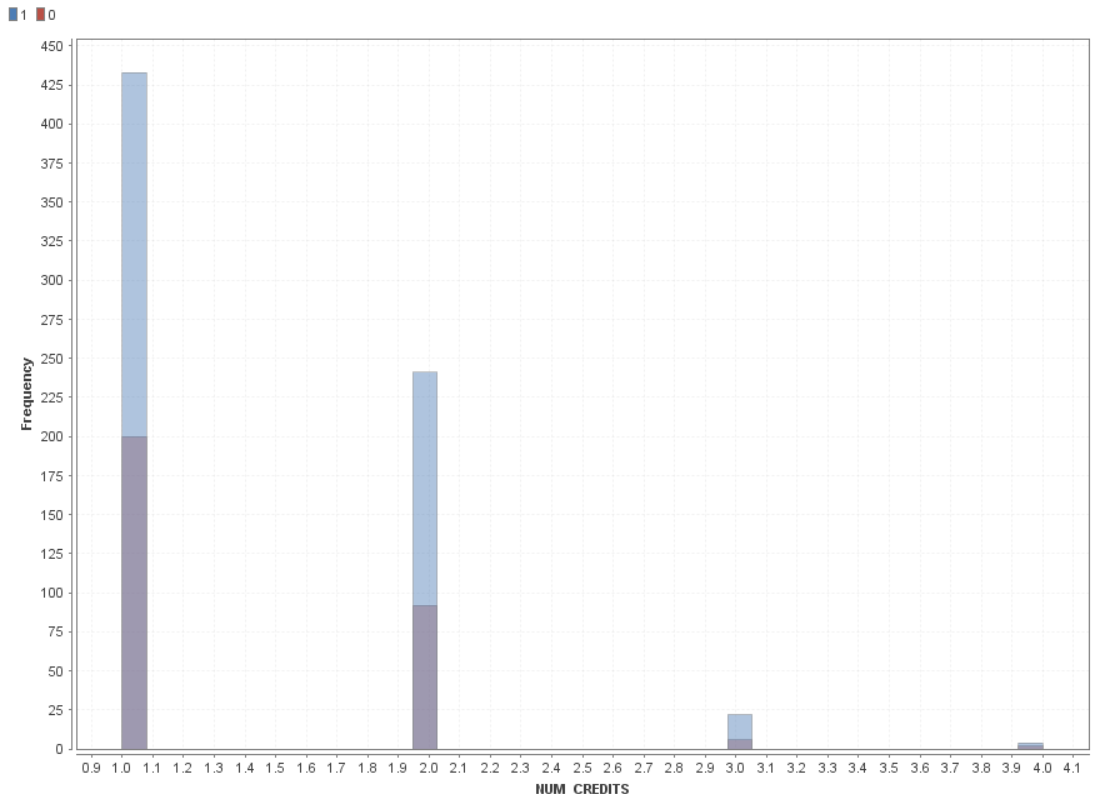
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable Name** | **Type** | **Mean** | **Standard Deviation** | **Freq** |
| CHK\_ACCT | Categorical | -- | -- | 0->274  1-> 269  2->63  3->394 |
| DURATION | Integer | 20.903 | 12.05 | --- |
| HISTORY | Categorical | -- | -- | 0->40  1->49  2->530  3->88  4->293 |
| AMOUNT | Integer | 32751.156 | 2822.625 | -- |
| SAV\_ACCOUNT | Categorical | -- | -- | 0->603  1->103  2->63  3->48, 4->183 |
| INSTALL\_RATE | integer | 2.973 | 1.119 | -- |
| EMPLOYMENT | Categorical | -- | -- | 0->62  1->172  2->339  3->174, 4->253 |
| PRESENT\_RESIDENT | Categorical | -- | -- |  |
| REAL\_ESTATE | Categorical | -- | -- | 0->718  1->282 |
| AGE | integer | 35.483 | 11.371 | -- |
| OTHER\_INSTALL | Categorical | -- | -- | 0->814  1->186 |
| OWN\_RES | Categorical |  |  | 1->713  0->287 |
| NUM\_CREDITS | integer | 1.407 | 0.578 | -- |
| JOB | Categorical | -- | -- | 0->22  1->200  2->630  3->148 |
| NUM\_DEPENDENTS | integer | 1.555 | 0.362 | -- |
| FOREIGN | Categorical | -- | -- | 1->963  0->37 |

1. Customer of Age range from 20 till 30 are more likely to turn to bad creditor

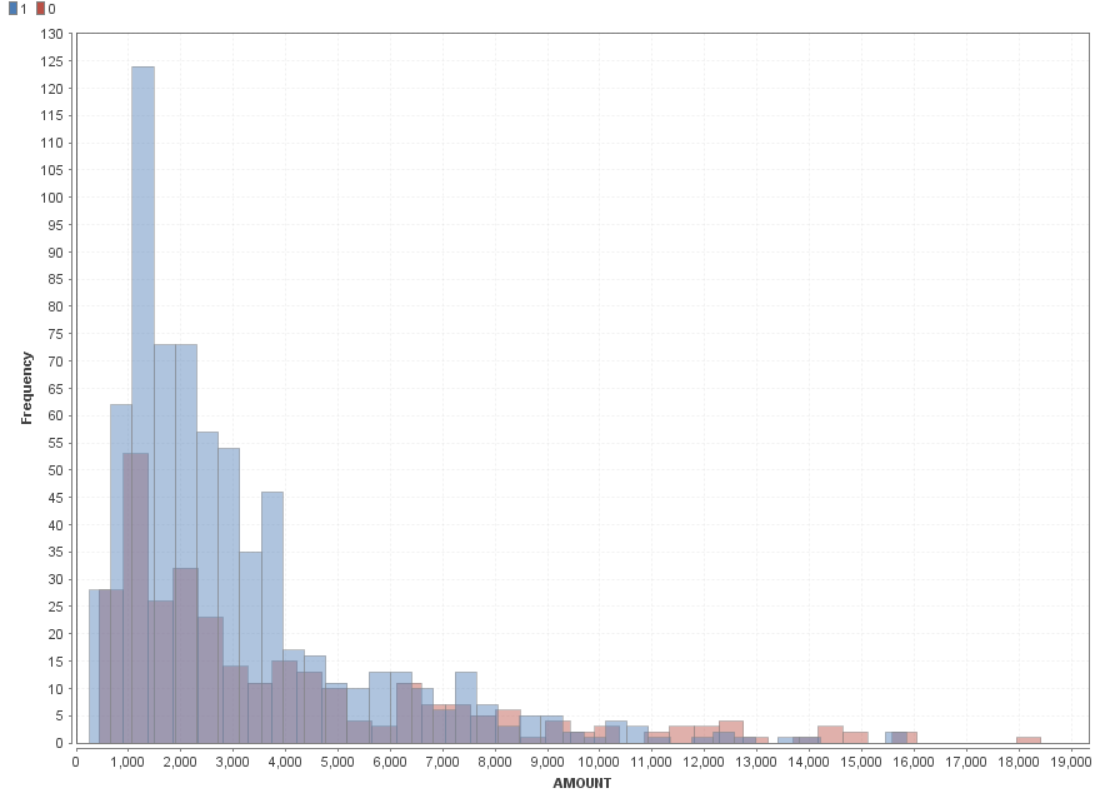
This is little surprising.



1. Low Credit Score more is bad creditor as expected .

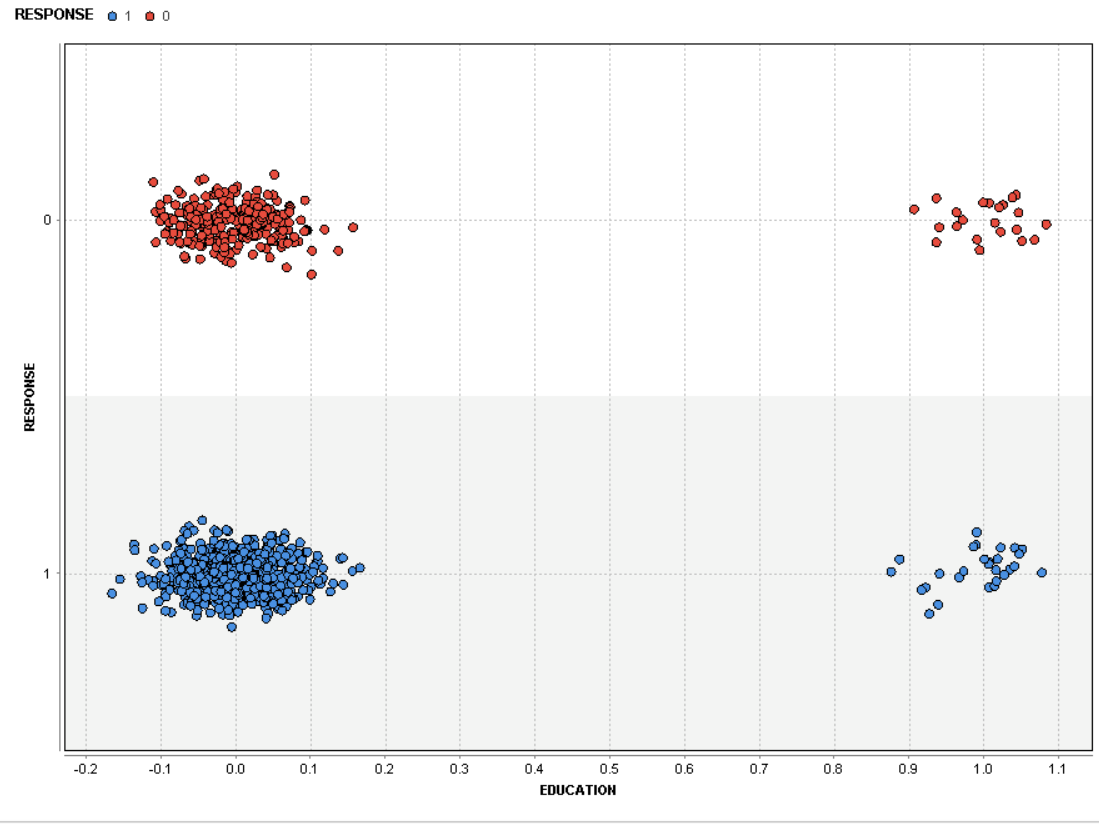


* Intresting Relation:
* Amount & Response : As amount increase ratio of BAD creditor increases



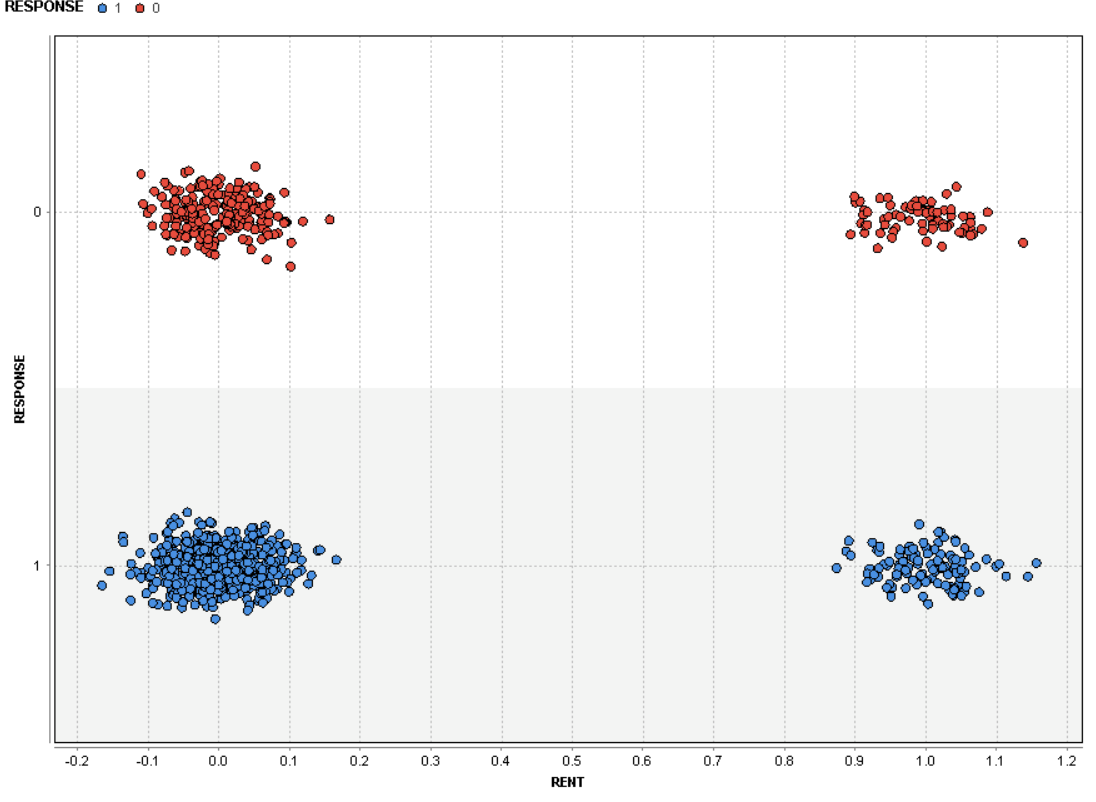
b)

* Education & Response; Non-Educated customers can be good creditor. it should be other way round



* RENT & RESPONSE: Customers that RENT are bad creditor

s



* Variables that are most relevant for the Outcome:

|  |  |
| --- | --- |
| **Variable Name** | **Why ?** |
| CHK\_ACCT | Discuss with TA ? |
| DURATION | Longer the duration, less like of BAD credit |
| HISTORY | Good credit history more possibility of GOOD Credit |
| AMOUNT | Less amount, less likelihood of BAD credit |
| SAV\_ACCOUNT | More balance in the saving account, more likelihood of GOOD credit |
| INSTALL\_RATE | More installment rate will lead to more BAD credits |
| EMPLOYMENT | Less duration of Employment, more chances of BAD credit |
| PRESENT\_RESIDENT | More duration at resident, less likelihood of credit turning BAD |
| REAL\_ESTATE | Applicant owns real state, less possibility of BAD credit |
| AGE | Older the age, less possibility of BAD credit |
| OTHER\_INSTALL | Applicant has other credit plan, more possibility of BAD credit |
| OWN\_RES | Applicant owns residence, less possibility of BAD credit |
| NUM\_CREDITS | GOOD credit, less possibility of BAD credit |
| JOB | More skilled the Job, less possibility of BAD credit |
| NUM\_DEPENDENTS | More number of dependent, more possibility of BAD loan |
| FOREIGN | Foreigner are less likely to default on the loan |
|  |  |

(2)

a) Best Model we get with