|  |  |  |
| --- | --- | --- |
| **Predictor** | **Expected Sign of effect** | **Rationale** |
| Income | + | As income increases, we expect the credit rating to increase and vice-versa due to the increased capability of a person to pay off his debts on time. |
| Limit | + | As limit increases, we expect the credit ratings to decrease and vice-versa because limit of a person is increased only when he has been a good timely paying customer for the bank. |
| Balance | - | When balance is high, the credit rating decreases and vice-versa. Whenever balance is maintained less than 30% of utilization ratio, the credit score increases. |
| Cards | -/+ | Depending on how they use their cards, the amount of credit cards a person possesses will impact their credit score. Having multiple credit cards can be advantageous if used correctly, but having too many cards and a lot of debt can be detrimental. |
| Age | None | Age does not impact a direct impact credit score of a person. |
| Education | None | Education does not have any direct impact on the credit score. |
| Gender | None | Gender does not have any impact on the credit score. |
| Student | None | Being a student does not impact the credit score. |
| Married | None | Marital status does not impact the credit score. |
| Ethnicity | None | Ethnicity does not impact the credit score. |
| Credit Utilization (Calculated) | - | When credit utilization is high, credit score goes down and visa-versa. 30% is the cut-off limit and anything beyond that can depreciate the credit-score. |

**Model 1-**

This can serve as a starting point. Based on our justification, the predictors selected for this model suggest that Rating is affected in some way. Limit and Balance were left out because utilization already captures their ratio. In this model, the predictor variables are Income and Utilization (Balance / Limit \* 100).

Graphical user interface, text

Description automatically generated

* Income: When there is an increase in the Income of a person by 1000 dollars, we expect the credit rating to increase by approximately 3.18 points when the utilization is kept constant.
* Utilization: When there is an increase in the credit utilization ratio of a person by 1 percent, we expect the credit rating to increase by approximately 12 points when the Income is kept constant.
* The coefficient for the level "Cards2" is -9.69, which means that if a person changes from having 1 credit card (the reference level) to 2 credit cards, the prediction for their "Rating" is expected to decrease by 9.69 units.
* "Cards2": -9.69, which means that if a person changes from having 1 credit card (the reference level) to 2 credit cards, the prediction for their "Rating" is expected to decrease by 9.69 units.
* "Cards3": -2.76, which means that if a person changes from having 2 credit cards to 3 credit cards, the prediction for their "Rating" is expected to decrease by 2.76 units.
* "Cards4": 0.44, which means that if a person changes from having 3 credit cards to 4 credit cards, the prediction for their "Rating" is expected to increase by 0.44 units.
* "Cards5": 4.18, which means that if a person changes from having 4 credit cards to 5 credit cards, the prediction for their "Rating" is expected to increase by 4.18 units.
* "Cards6": -25.66, which means that if a person changes from having 5 credit cards to 6 credit cards, the prediction for their "Rating" is expected to decrease by 25.66 units.
* "Cards7": 25.04, which means that if a person changes from having 6 credit cards to 7 credit cards, the prediction for their "Rating" is expected to increase by 25.04 units.
* "Cards8": -18.79, which means that if a person changes from having 7 credit cards to 8 credit cards, the prediction for their "Rating" is expected to decrease by 18.79 units.
* "Cards9": -11.03, which means that if a person changes from having 8 credit cards to 9 credit cards, the prediction for their "Rating" is expected to decrease by 11.03 units.

It's important to note that these interpretations assume that all other variables in the model are held constant.

**Model 2**- Let us add Gender and Ethnicity to see whether there is a bias in these predictors.

Text

Description automatically generated

In this model, we included how Gender and Ethnicity would affect the credit rating and we observe the following-

* Income: 1000 dollars increase in Income is associated with an increase in the "Rating" by 3.18579 units, holding all other independent variables constant.
* Utilization: for a unit increase in the utilization rate, the expected credit rating score increases by 12.06495 points, holding all other predictor variables constant.
* EthnicityCaucasian: If an individual is Caucasian, the expected credit rating score is 4.63 points lower than an African American, holding all other predictors constant.
* EthnicityHispanic: If an individual is Hispanic, the expected credit rating score is 9.43 points lower than an African American, holding all other predictors constant.
* GenderMale: If an individual is Male, the expected credit rating score is 0.73 points higher than a female, holding all other predictors constant.
* "Cards2" : The estimate is -10.36, which means if an individual has 2 credit cards, the expected credit rating score is 10.36 points lower than an individual with only one credit card, holding all other predictors constant.

Holding all other variables constant, being male is associated with a credit card rating that is 0.72594 units higher on average than being female. However, the t-value of 0.147 indicates that this estimate is not statistically significant at the conventional significance levels (p > 0.05), and the high standard error also suggests that the estimate is not very precise. Therefore, based on the output of this regression model, we cannot conclude that there is a gender bias in the credit card rating. We would have said there is a bias favoring men, if the p value was statistically significant. Similarly, for the Ethnicity, if the p-values were significant, we can state there was a bias favoring African Americans. However, it's important to note that these coefficients don't necessarily indicate a bias. Other factors could be affecting the credit ratings and it's difficult to determine causality based solely on the coefficients. Further analysis and investigation would be necessary to determine if there is a bias.

**Factors that affect the credit rating are-**

**Payment history (35%)**- Late payments can decrease the credit rating. Once your credit cards are paid off, your credit score may rise by 10 to 50 points. The amount by which you paid off the bills and how you manage other credit accounts are among the variables that will determine how much your score will rise.

**Credit Utilization ratio (30%)** – (Balance / Credit Limit) \* 100. Experts advise use no more than 30% of your credit line. The highest scorers typically use a lot less than that. You can attempt methods like setting balance alerts or making extra payments throughout the month to keep your credit utilization low and hence improving the credit score.

**Length of credit history (15%)** - The number of years that you've had credit accounts open in your name determines your credit history. Your credit score may increase with a longer credit history. It makes sense to keep using a credit card properly if you've had it open for a while to have a good score.

**Types of credits you use (10%)**- If you appropriately employ various credit kinds, including installment and revolving loans, your score may rise. However, to have a decent score, you don't necessarily need to have a wide variety of credit.

**New credit you apply for (10 %)** - Your rating drops each time you request for credit. There is one exception: If you compare rates from several lenders within a 14- to 45-day window when looking for a mortgage, student loan, or vehicle loan, credit scoring algorithms will only consider one query.

Stargazer output- Uploaded separately.