

## How to Make the Best Credit Score Even Better: FairCreditScore GenIQ-enhanced

Bruce Ratner, Ph.D.

The [GenIQ Model](#)© is a machine learning alternative model to the statistical ordinary least squares and logistic regression models. GenIQ lets the data define the model – automatically data mines for new variables, performs variable selection, and then specifies the model equation – so as to "optimize the decile table," to fill the upper deciles with as much profit/many responses as possible. *As an extra-GenIQ application*, GenIQ can be used on a final regression model to let GenIQs data mining prowess enhance the results of the final model. The GenIQ-enhanced Regression Model: Enhance a final regression model by running GenIQ with only one predictor: the final regression equation score.

The purpose of this article is to apply the GenIQ-enhanced Regression Model to show “How to Make the Best Credit Score Even Better: FairCreditScore GenIQ-enhanced.” To make the best credit score (FairCreditScore) even better the GenIQ Model only uses the FairCreditScore variable. FairCreditScore (FCScore) is a number, based on statistical regression analysis of persons' credit reports, used to represent the creditworthiness of a person, i.e., the likelihood of a person will pay his/her debts. Specifically, FCScore is the primary score used throughout the money lending industry, chiefly based on credit report information from the three major credit bureaus. First, I provide a detailed *Illustration of the GenIQ-enhanced Regression Model* using a simple two-variable logistic regression model. If the GenIQ illustration piques your curiosity, I would be glad to [email](#) you, just for the asking: “How to Make the Best Credit Score Even Better.”

### Illustration of the GenIQ-enhanced Regression Model

**Objective #1:** Build a binary response model with data in Table 1. I build a logistic regression model (LRM) for predicting Response using predictors XX1 and XX2. The LRM equation is:

$$\text{Logit of Response (=Yes)} = 0.1978 - 0.0328 * \text{XX1} + 0.0308 * \text{XX2}$$

**Table 1.**

<b>ID</b>	<b>XX1</b>	<b>XX2</b>	<b>Response</b>
1	31	38	Yes
2	12	30	No
3	35	21	Yes
4	23	30	No
5	45	37	No
6	16	13	No
7	45	5	Yes
8	30	30	Yes
9	6	10	Yes
10	30	10	No

**LRM Results:** The Response ranking is not perfect. The ranking of Response based on the LRM probability score **Prob\_of\_Response**, which is obtained directly from the LRM logit score, is in Table 2, below. Clearly, the Yes-No ordering is not perfect. Notwithstanding the correct top position of ID #7, and the bottom two positions of ID #4 and #2, the ranking seems random!

**Table 2.**

<b>ID</b>	<b>XX1</b>	<b>XX2</b>	<b>Response</b>	<b>Prob_of_Response</b>
7	45	5	Yes	0.75472
10	30	10	No	0.61728
3	35	21	Yes	0.57522
5	45	37	No	0.53452
6	16	13	No	0.48164
8	30	30	Yes	0.46556
9	6	10	Yes	0.42336
1	31	38	Yes	0.41299
4	23	30	No	0.40913
2	12	30	No	0.32557

### **GenIQ Model: Its Output**

The GenIQ Model output consists of two parts: A visual display of the model, and the model “equation.”

1) Unsuspected Equation – Visual Display and Computer Code.

The visual display, called a parse tree, depicts the GenIQ Model. The GenIQ Model equation is actually a computer program/code. The regression modeler, anticipating an equation of the form –  $Y = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n$  – is stupefied when s/he unsuspectedly sees the GenIQ computer code.

2) Ungainly Interpretation – "Bit much to grasp."

The GenIQ parse tree and computer code can be a "bit much to grasp." The visual display provides the modeler with an ocular sense, albeit mote, of comfort and confidence for understanding and using GenIQ. The GenIQ tree, although not a “black box” like most other machine learning methods, gives the modeler a graphic, albeit Picasso-like, to make sense of. GenIQ for the everyday regression model produces a GenIQ tree defined by "branches" formed by yellow-boxed predictor variables attached to various numbers of turquoise-boxed functions, as displayed in Figure 1, below.

I run GenIQ with the final LRM score: The GenIQ-enhanced Regression Model tree is in Figure 1, below. Under the tree is the GenIQ-enhanced Regression Model computer code, blocked in light blue. The “enhanced” model score (aka Y-hat) is **GenIQvar**.

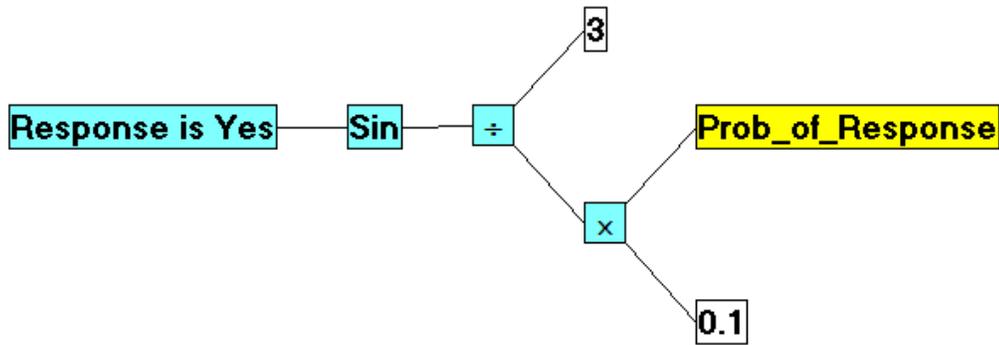


Figure 1. GenIQ-enhanced Regression Model

```

x1 = .1;
  x2 = Prob_of_Response;
  x1 = x1 * x2;
  x2 = 3;
  If x1 NE 0 Then x1 = x2 / x1; Else x1 = 1;
  x1 = Sin(x1);
  GenIQvar = x1;
  
```

**The GenIQ-enhanced LRM Results:** The Response ranking is perfect! The ranking of Response – based on the GenIQ Model with only the LRM score **Prob\_of\_Response** – produces a GenIQ score **GenIQvar**, in Table 3, yielding a perfect Yes-No ordering.

Table 3.

ID	XX1	XX2	Prob_of_Response	Response	GenIQvar
8	30	30	0.46556	Yes	0.99934
9	6	10	0.42336	Yes	0.98444
3	35	21	0.57522	Yes	0.95002
7	45	5	0.75472	Yes	0.88701
1	31	38	0.41299	Yes	-0.37436
5	45	37	0.53452	No	-0.41121
6	16	13	0.48164	No	-0.51791
2	12	30	0.32557	No	-0.86183
4	23	30	0.40913	No	-0.87665
10	30	10	0.61728	No	-0.99553