SAS VIYA Exercise 10

Logistic Regression

(7/22/2020)

**Sources**

Steve Nolan, Ron Freeze, Elizabeth Keiffer, Michael Gibbs, Jorge Moreda

Enterprise Systems, Sam M. Walton College of Business, University of Arkansas, Fayetteville

SAS® Visual Analytics. Release 8.5

SAS® Viya® release V.03.05

Copyright © 2018 *For educational uses only - adapted from sources with permission. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, or otherwise, without the prior written permission from the author/presenter.*

# Use Case – Logistic Regression

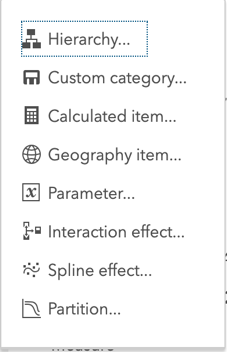
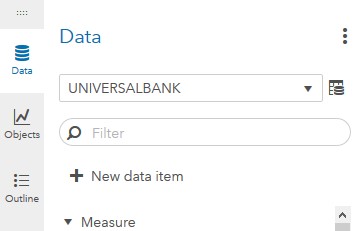
Universal Bank is a banking dataset with the goal of identifying if a customer accepted a personal loan from a marketing campaign.

## Step 0: Import Data

Following the guide of importing data and import your chosen dataset. See **Viya 02a – Importing data** for specific steps. In this tutorial the dataset used is UNIVERSALBANK.sashdat.

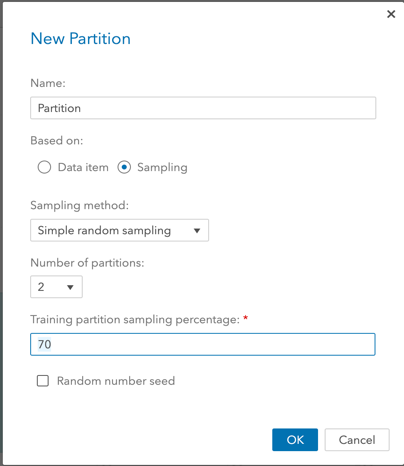
**Step 1: Create Partition**

Refer back to **Viya 07 – Create Partitions** or follow the steps below to create a partition.



1. Go the **Data** tab on the left-hand pane and
2. Click on **+ New data item**

At the bottom of the drop-down menu that will appear,

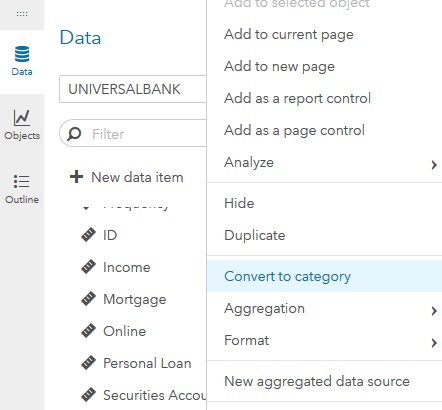


1. Select **Partition...**

A new window will open where you can give the partition a name and designate how many partitions you wish to have. In our use case, we will have 2, one for the training and one for validation. Also, we want to set our training data to account for 70% of the data:

1. Select 2 as **Number of partitions**
2. Write 70 under **Training partition sampling percentage**
3. Click **Ok**

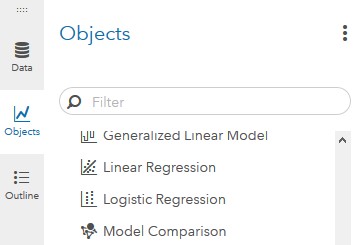
## Step 2: Category Conversion



1. Click on **Data**
2. Right-click on **Personal Loan**
3. Left-click on **Convert to category**

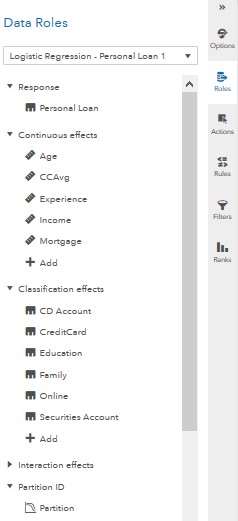
Perform the same right-click and left-click actions to similarly convert the following measures into categories: **CD Account**, **CreditCard**, **Education**, **Family**, **Online**, and **Securities Account**.

## Step 3: Logistic Regression Object



1. On the left-pane, click **Objects**
2. Scroll down to find **Logistic Regression**
3. Drag and drop it to the working space in the middle-pane

## Step 4: Variables



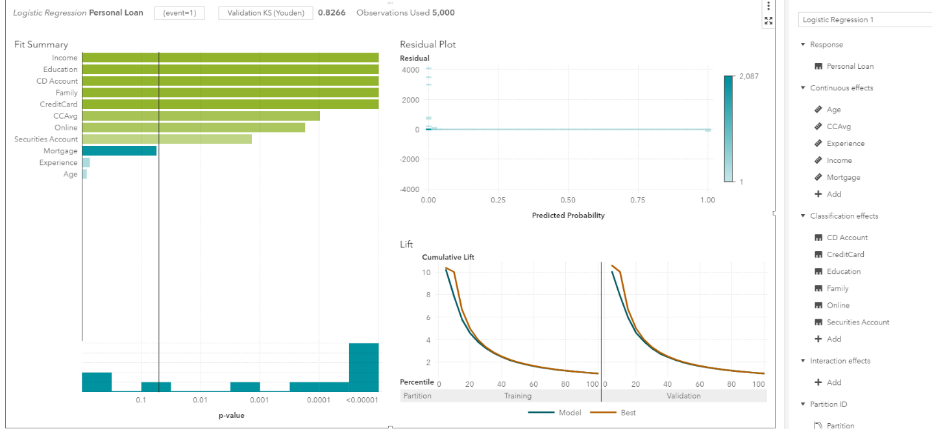
The next step is to select the variables we want to work with.

1. On the right-pane, click on **Roles** (make sure you have clicked on the working space before doing this)

A new window named **Data Roles** will appear for you to add variables.

1. Click on **+ Add** under **Response**
2. Add **Personal Loan** (dependent variable)
3. Click on **+ Add** under **Continuous effects**
4. Add **Age, CCAvg, Experience, Income, Mortgage** (continuous independent variables)
5. Click on **+ Add** under **Classification Effects**
6. Add **CD Account, CreditCard, Education, Family, Online, Securities Account** (classification independent variables)
7. Click on **+ Add** under **Partition ID**
8. Add the partition you created in **Step 1**

## Step 5: Regression Display

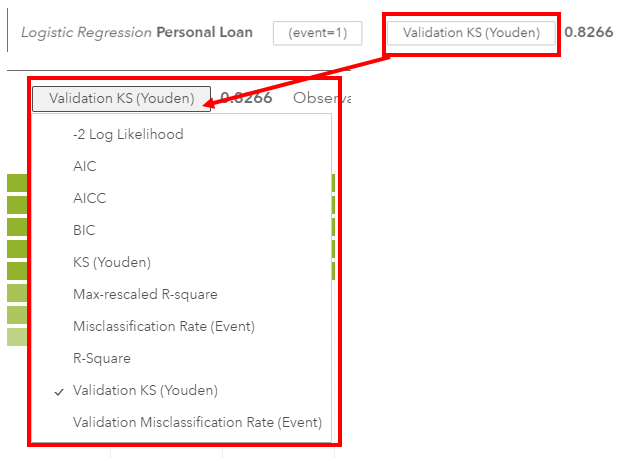
Once you have selected your data, the logistic regression will update on the screen and display several outputs:

## Step 6: Error Metrics

Here are some key things to note:



The event is currently set to 1.

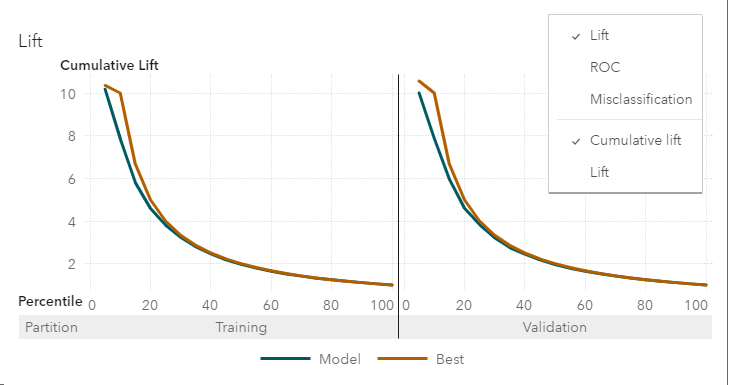
One of the key results is set to determine which error metric you would like to view.

The default value is “Validation KS(Youden)”.

This is a clickable button that you can change error metrics.

Our recommendation is to toggle between Misclassification Rate (Event) and Validation Misclassification Rate (Event) to observe both the values and their relationship to each other.

You can change the lift chart into the ROC curve by hovering over the graph and,



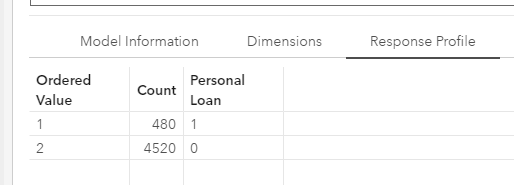
1. Selecting the three vertical dots in the upper right-hand corner
2. Click the **ROC** option

## Step 7: Statistics

1. Click on the maximize button in the upper-right portion of the visuals

This will bring up the evaluations table at the bottom of your screen.

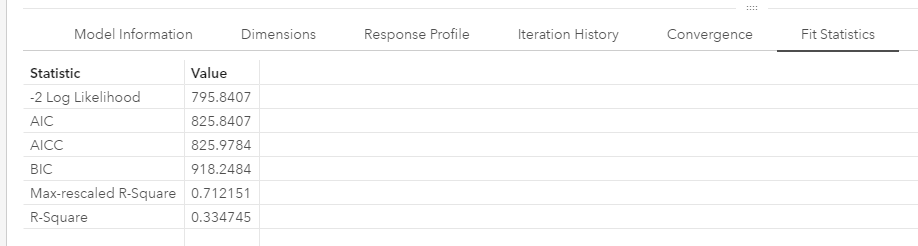
1. Click **Response Profile**



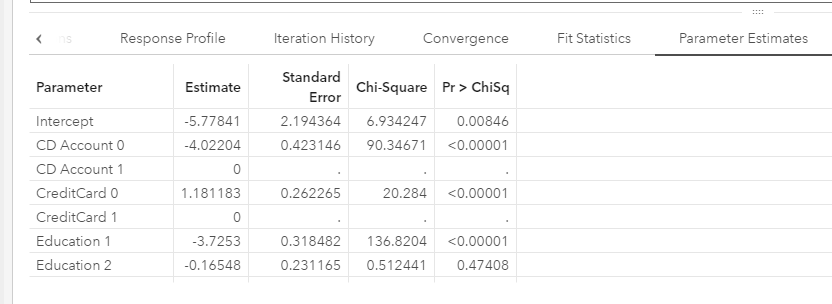
Here you can view the response profile of how individuals answered to the target variable.

1. Click on **Fit Statistics**

Here you can view the fit statistics.



1. Click on **Parameter Estimates**



Here you can evaluate the parameter estimates and develop the logistic regression formula for classification.

1. Click on the right arrow until **Misclassification** is visible



1. Click on **Misclassification**

Here you can also view the confusion matrix.



Congratulations, you have created a logistic regression in VIYA!