SASVIYA Exercise

Linear Regression

(4/2/2019)

**Sources**

Steve Nolan, Ron Freeze, Elizabeth Keiffer, Michael Gibbs

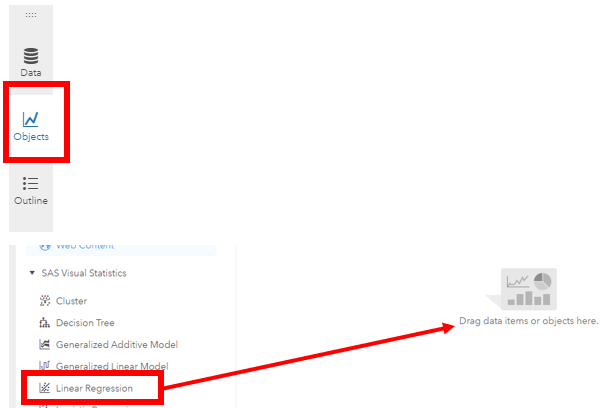
Enterprise Systems, Sam M. Walton College of Business, University of Arkansas, Fayetteville

SAS® VIYA 8.2 Release V03

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 – Linear Regression**

Razorback Stores is a local department store serving a metropolitan area. As a department store, they offer a wide variety of items and services and track sales through a point of sale system. Over the past several months, Razorback Stores performed a marketing campaign designed to promote and incentivize a loyalty program.

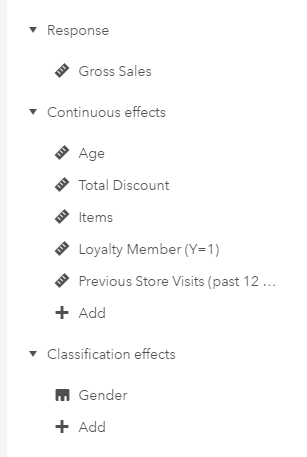
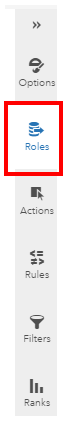


## Step 1: Select Data

Select the RAZORBACK STORES.SASHDAT dataset for use. There is an assumption that the Razorback Stores has already been loaded into memory. Where is the tutorial that did this?

## Step 2: Select Model

Select New Report and use the object panel to select Objects>SAS Visual Statistics>Linear Regression and drop it into the middle space



## Step 3: Select Variables

From here, you will have the data **Roles** on the right to work. Add a **Response** variable (dependent variable) and the **Effect** variables (independent variables). Note that your effect variables are separated by **Continuous** and **Classification** effects.

## Step 4: Review Results

Once you have selected your response variable and “effects”, the linear regression will update and display. Key things to keep in mind: There is a Training and Validation data set. Where were these selected?



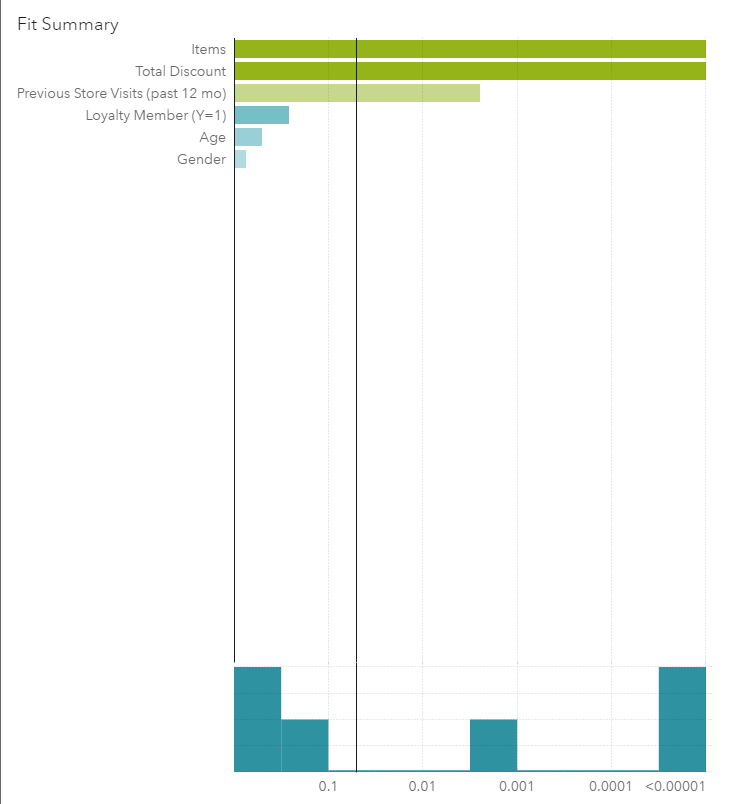
### Residual Plots

Residual plots are displayed in the upper-right hand part of the display. This is used to check to ensure residuals do not have a pattern.

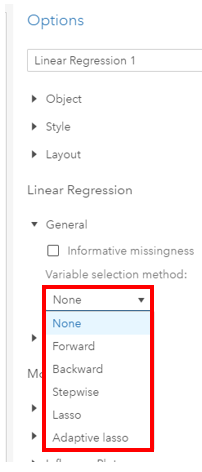
### Lift Chart

Lift chart is used to evaluate overfitting. The chart is in the bottom-right hand corner of your screen. Ensure that the validation data is not overfitting.

### Fit Summary



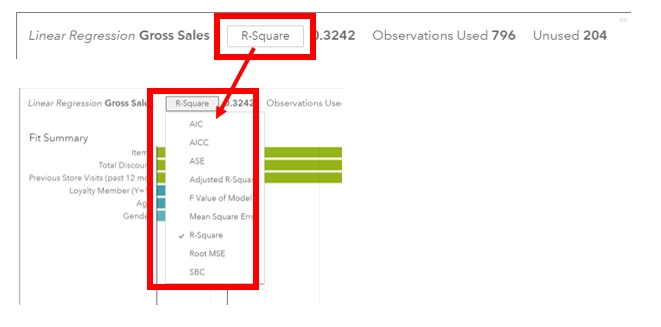
The fit summary shows you variable importance and the p-value threshold for statistical significance. You can adjust the p-value significance by using sliding the vertical bar.



### Options

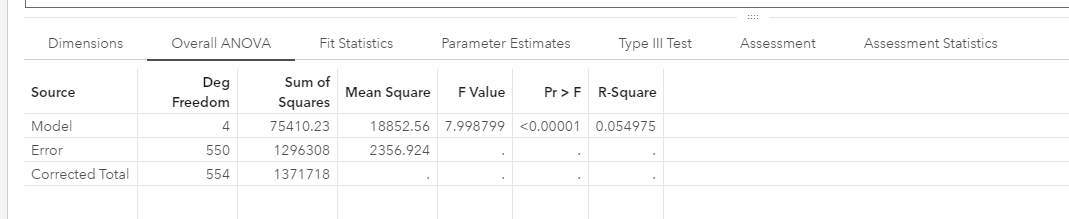
Once you have this, you can then select from a variety of **Options** in the **Options** panel on the right. A key option to note is under the **Linear Regression**>**General** option. You can select various regression methods (forward, backward, stepwise, etc.). (Note: Make sure you have selected the report in order for the Options to be available.)

### Parameters

As part of the review process you can click on the textbox at the top of the screen and change the shown parameter to several different options:

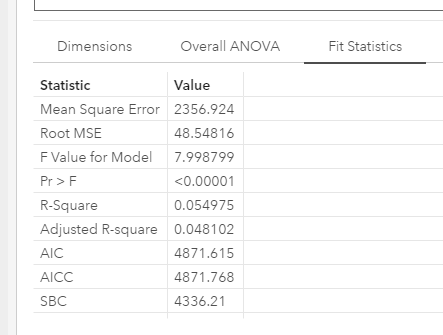
## Step 5: Evaluate Model

Now that the settings are in place, we can evaluate the predictive model. Click on the icon below that is located in the upper-right corner of the modeling screen: This will bring up the evaluations table at the bottom of your screen.



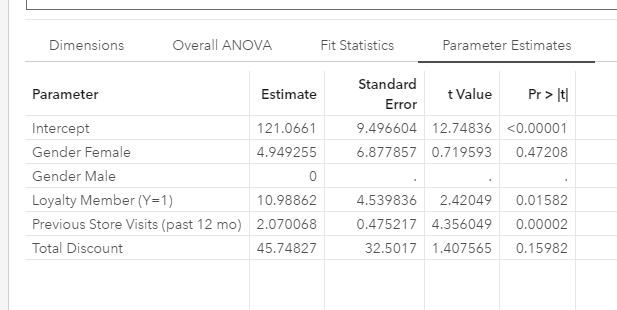
The overall ANOVA will display your **R-Square** value as well as the **Sum of Squares**, **p-Value**, and **Mean Square**.

## Step 6: Fit Statistics



You can also evaluate your model on the **Fit** **Statistics** tab. This has values such as your **root mean squared error** (**RMSE**) and **Adjusted R-square**.

### Parameter Estimates



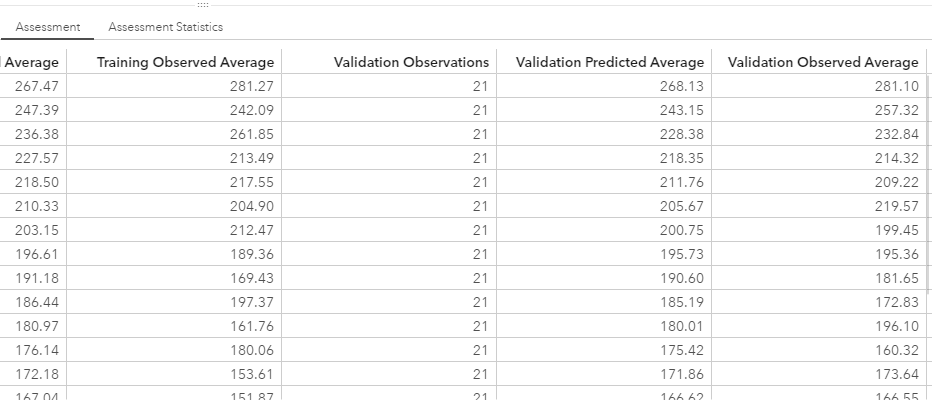
On the **Parameter Estimates** tab, you can review the linear regression formula. Be sure to check t-value significance of the variables.

### Regression Equation

The above example would produce this formula

Gross Sales = 121.0661 + 4.949255\*(Female Gender) + 0\*(Male Gender) + 10.98862\*(Loyalty Member) + 2.070068\*(Previous store visits) + 45.74827\*(Total Discount)

### Assessment



On the assessment tab, you can review the predicted average of an observation against the observed average. This shows a visual of your predictive model and how well it is performing.