SASVIYA Exercise 07

Setting a Data Partition

(4/13/2020)

**Sources**

Steve Nolan, Ron Freeze, Elizabeth Keiffer, Michael Gibbs, Matias Delay

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

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.

Before we attempt to answer any business questions, we need to create a **data partition** in preparation for data mining. The partition will randomly separate our data into groups to train our supervised models and check for accuracy.

## Step 0: Getting Started

Following the guide of importing data and import your chosen dataset. See **Viya 02 – Importing data** for specific steps. In this tutorial we will be using Razorbacks Stores dataset.

## Step 1: Navigation

Once your data is imported, you will be presented with the working report screen. The screen is broken out into three key areas:

* Left pane (Red) – This is where you can manipulate your data elements, choose objects to work with, and see an outline of the work you are creating.
* Middle pane (Blue) – This is the workspace. This is where you drag objects to and begin to build your visualizations and models
* Right pane (Green) – this is where you set various roles (for data mining) and manipulate and enhance chart features (options). You can also filter pieces of data as well.



## Step 2: Create Partition

Go the **Data** tab on the left-hand pane and click on **+ New Data Item**. At the bottom of the drop-down menu that will appear, select **Partition...**.


## Step 3: Modify Partition Settings

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 3: Apply Partition

When you create a partition, a new categorical variable is created. This will be used in modeling when it asks for a partition variable once you begin your modeling. For instance, if we were working with decision trees, under **Roles** in the right pane, we would have the option to add a partition under **Partition ID**.

1. Click on **+ Add** under Partition ID.
2. Select the partition you created in **step 2** and **step 3**.

