SASVIYA Exercise 02b

Importing Local File Through

“Manage Data”

(2/5/2020)

**Sources**

Steve Nolan, Ron Freeze, Elizabeth Keiffer, Michael Gibbs, Evelyn Day

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 – Importing Data

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.

It is time to import the Razorback Stores data set (or a dataset of your choosing). This guide will help us import the point of sale data provided by Razorback Stores.

## Step 1: Import Local File

1. Click on the three stacked horizontal lines in the upper left corner of the screen.

1. Choose “Manage Data”.

The next window you will see shows three different options: Available, Data Sources, and Import.

1. Select “Import”.

After clicking on “Import”, select “Local File” and import the file from your user drive. For this tutorial, we will use Sales.csv.

After importing the local file, define a Target Destination.

1. Click the folder icon next to the target destination field.

1. Click on the arrow next to “cas-shared-default”.

1. Click on the arrow next to “CASUSER(username)”

Here, you can see that the username of this individual is marked out. In place of this will be YOUR own username. For example, if your username is “abc123”, this screen would show “CASUSER(abc123)”.

1. Click the “Select” button on the bottom when you are inside the username CASUSER folder.

After defining the target destination, make sure to select the appropriate file delimiter depending on the type of .csv file you import.

1. Select “Import Item” at the top of the screen in the righthand corner.

\*Note: Importing the file may take up to 15-30 minutes depending on the size of the file. SAS Viya is not a particularly slow software, but with large files, importing items will not be instantaneous.It is **important to** **be patient**, and do not exit the software during importing.

## Step 2: Exploring the Data

1. Click on “Data Sources” and find your newly imported file through the steps described above.

In the pathway that ends with CASUSER(username), you can find your newly imported file.

In this example, the imported file is SALES with the lightning bolt and SALES.sashdat with the snowflake. The snowflake signifies that the file is on disc. The lightning bolt signifies that the file has been lifted into memory. This means that any analytics that are run are finished faster than if the file were just on disc.

If the file has not yet been lifted into memory:

1. Click on the file with the snowflake, and in the top right-hand corner you will see the lighting bolt symbol. Click this symbol. This process may take a while, so it is important to be patient.


## Step 2a: The “Details” Tab

1. Click on the “Details” tab.

The “Details” tab shows an overview of each data column. The column number, name, label, type of data, raw length of the data, formatted length of the data, format, and tags are all included in the Details tab.

In this example, there are eight columns, seven of which are double data type and the other a variable character data type. The name of each column is included in the screenshot below.

## Step 2b: The “Sample Data” Tab

1. Click on the “Sample Data” tab.

The “Sample Data” tab selects a simple random sample from the dataset. In this case, the sample is of 100 rows.

The number of rows included in the sample can be changed in the top right-hand corner of the screen after selecting the “Sample Data” tab.

Columns in this tab include Item ID, Location ID, Date ID, etc. This will change depending on your dataset. These columns are columns directly copied from the .csv file.

## Step 2c: The “Profile” Tab

1. Click on the “Profile” tab. You will see a message that says profiles are available.

1. Click “Run Profile and Save”.

You will see this pop-up. Name the table and select a location. The location must be the same as the location of the imported file (in this case, CASUSER(username)).

1. Click “Run”.

Running the profile should not take longer than a few minutes.

The “Profile” tab now has descriptive statistics from the dataset. The columns include column name, unique null, blank, pattern count, mean, median, mode, standard deviation, standard error, minimum, maximum, data type, actual type, data length, etc.

The “unique” column shows how unique that particular variable is. Results that show 100% unique are called **primary keys** and are usually ID numbers. When a dataset includes ID number, most often the ID number is completely unique for each row of data, rendering ID numbers great primary keys for a dataset.

In this particular dataset, “Location ID” is 43.76% unique, making the Location ID variable the most unique in this dataset. However, there is no variable that is 100% unique, meaning there are no primary keys in this dataset.

\*Note: If you click on one of the column names such as “Actual Type”, you can sort the given variables by that parameter. If sorted by “Actual Type”, the variables would be reorganized to put all similar data types together.

Congratulations! You have successfully imported data into SAS Viya via the “Manage Data” button and looked at descriptive statistics!