

Stat 404 Lab 6

Renae L. Shrum

Friday 09 October 2015

SAS

1. Create a library called Aquarius and associate it with the following folder address: 'S:\Courses\stat-renaes\Stat404'.
2. Use PROC PRINT to visualize the data set called shoes_tracker. Look for missing values anywhere.
3. The variable Product Category must not be missing and Supplier Country must be either GB or US.
4. Add a WHERE statement to PROC PRINT to find observations that do not meet the requirements in #3.
5. Use PROC PRINT and use the VAR statement to view Product Category, Supplier Name, Supplier Country and Supplier ID.
How many observations have missing Product Category?
How many observations have invalid Supplier Country?
6. Add a PROC FREQ with a TABLES statement to create frequency tables for Supplier Name and Supplier ID. Include the NLEVELS option (and I would use the NOPRINT option also). The data should meet the following requirements:
 - a. Supplier_Name must be 3TOP Sports, or Greenline Sports Ltd.
 - b. Supplier_ID must be 2963 or 14682

What invalid data exist for for Supplier Name and Supplier ID?

7. I am not collecting this lab but for up to 5 points of extra credit, you will make the corrections to this dataset. Call the new dataset newshoes. Make the following corrections to this data. An example of your DATA step could look like the following:

```
DATA newshoes; SET aquarius.shoes_tracker; WHERE<insert SAS statements here>;  
RUN;
```

Corrections:

- a. Supplier_Country must be either GB or US.
- b. Product_Category should be Shoes
- c. Supplier_Name must be 3TOP Sports, or Greenline Sports Ltd.
- d. Supplier_ID must be 2963 (3TOP Sports) or 14682 (Greenline Sports Ltd.)

Make sure you use PROC PRINT to verify that your changes have been made. For submission, please copy and paste the log and your program code into the body of an email (not the output – only the Log and your program code). No attachments please.