Lab 8

Stat 426

Spring 2021

Instructions

Complete all questions. To prepare for the randomly collected lab, follow the instructions on the class website to prepare the work for submission. These submission rules will apply to all labs throughout the semester.

Data cleaning

(1) Problems 2-7 were originally a part of Lab 7; new problems start at number 8 but I thought I would leave the first parts here for reference so you do not have to switch back and forth between documents.

(2) Create a library called Aquarius and associate it with the `sasdatafiles` folder that is on the class website. It can be downloaded from here: sasdatafiles folder. Then just unzip and make sure to remember where you saved it (half the battle).

(3) Use PROC PRINT to visualize the data set called `shoes_tracker`. Look for missing values anywhere.

(4) The variable `Product_Category` must not be missing and `Supplier_Country` must be either GB or US.

(5) Add a WHERE statement to PROC PRINT to find observations that do not meet the requirements in #3.

(6) Add a VAR statement to create a PROC PRINT report that looks similar to Figure 1 below. How many observations have missing `Product_Category`? How many observations have invalid `Supplier_Country`?

Obs	Product_Category	Supplier_Name	Supplier_Country	Supplier_ID
1	Shoes	3Top Sports	us	.
2		3Top Sports	US	2963
5	Shoes	3Top Sports	UT	2963
10	Shoes	Greenline Sports Ltd	gB	14682

Figure 1: Al Bundy, LTD.

(7) Use PROC FREQ with a TABLES statement to create frequency tables for `Supplier_Name` and `Supplier_ID`. Include the NLEVELS option. What invalid data exist for `Supplier_Name and Supplier_ID`? The data should meet the following requirements:
  - `Supplier_Name` must be **3TOP Sports**, or **Greenline Sports Ltd.**
  - `Supplier_ID` must be **2963 14682**

(8) Make the corrections as stated in numbers 4 through 7 to a new dataset called newshoes. An example of your `DATA` step could look like the following:

```
DATA newshoes;
SET aquarius.shoes_tracker;
<insert IF-THEN, IF-THEN/ELSE, etc. statements here>
RUN;
```

Make sure you use PROC PRINT to verify that your changes have been made.

(9) Make the same changes using `Viewtable` in the `Explorer Window` in SAS.