

Seneca

Academic Year	2022 – 2023		
Semester	<input type="checkbox"/> Fall	<input checked="" type="checkbox"/> Winter	<input type="checkbox"/> Summer
Course Code - Name	BAN110		
Instructor	Dr. Razi Iqbal		
Assessment	Assignment 4		

Student ID **176938215**

Student Name **Rongzhao Yi**

Assignment 4

The main purpose of this lab is to get students familiarize with PROC SQL and Joins.

Instructions:

- You are required to submit your answers in this document by pasting your SAS code under the solutions heading below.
- Please do not submit .sas files. Submit only this word document with your code inside it.
- Total Marks for this assignment are 5 marks.
- Students having exactly similar code will get a straight 0.
- You are required to complete these exercises using SAS.
- The deadline for submission of this assignment is April 4, end of the day.

Question

You are provided with the dataset 'Superstore.xlsx' file. It is an excel file with three sheets, 'Orders', 'Returns' and 'People'. You are required to perform the following tasks using this dataset in SAS:

- Import the data from the excel sheet in SAS.
- Make sure you create at least 2 datasets in SAS namely Orders which would get all the data from Orders sheet and Returns which would get all the data from the Returns sheet in the excel.
- Use Proc SQL to perform the following tasks:
 - Get only the records for the orders that have been returned by joining Orders and Returns datasets.
 - Print only the returned orders from **California** under **Technology** category with **Quantity** greater than 5.
 - Make sure your results are sorted by Order_Date in ascending order (you might have to do some research on how to do this in PROC SQL).
- Do not use any other procedure except PROC IMPORT and PROC SQL. You cannot use a DATA module for this assignment.
- Below is the expected output of this program:

Order Date	Customer Name	City	Postal Code	Product Name	Sales	Quantity
10/19/2014	Eugene Hildebrand	Lakewood	90712	RCA Visys Integrated PBX 8-Line Router	321.552	6
12/29/2014	Anthony Johnson	San Francisco	94110	DYMO CardScan Personal V9 Business Card Scanner	767.952	6
11/20/2015	Lena Creighton	Oakland	94601	Gear Head AU3700S Headset	72.744	7
12/30/2015	Clay Ludtke	Los Angeles	90049	Zebra GX420t Direct Thermal/Thermal Transfer Printer	2548.56	6
05/28/2016	Clay Ludtke	Los Angeles	90004	Imation Clip USB flash drive - 8 GB	131.6	7
02/26/2017	Bradley Talbott	Los Angeles	90036	Toshiba IPT2010-SD IP Telephone	889.536	8
03/03/2017	Ed Jacobs	Los Angeles	90008	Imation Bio 2GB USB Flash Drive Imation Corp	1049.44	8
04/17/2017	Pamela Stobb	Los Angeles	90008	Logitech Wireless Touch Keyboard K400	274.89	11
09/11/2017	Ken Heidel	Santa Ana	92704	OtterBox Defender Series Case - Samsung Galaxy S4	143.952	6
10/02/2017	Chad McGuire	San Francisco	94109	Ricoh - Ink Collector Unit for GX3000 Series Printers	369.16	11
11/07/2017	Deborah Brumfield	Brentwood	94513	Plantronics CS 50-USB - headset - Convertible, Monaural	761.544	7

Make sure to include only the columns shown in the screenshot above.

Solution

```
options validvarname=v7;
PROC IMPORT OUT= WORK.Orders DATAFILE= "/home/u63055836/BAN110ZAA/Superstore.xlsx"
           DBMS=xlsx REPLACE;
           SHEET="Orders";
           GETNAMES=yes;
RUN;

PROC IMPORT OUT= WORK>Returns DATAFILE= "/home/u63055836/BAN110ZAA/Superstore.xlsx"
           DBMS=xlsx REPLACE;
           SHEET="Returns";
           GETNAMES=yes;
RUN;

proc sql;
    select Order_Date, Customer_Name, City, Postal_Code, Product_Name, Sales, Quantity
    from Work.Orders as O INNER JOIN Work>Returns as R on O.Order_ID = R.Order_ID
    where State = 'California' and Category = 'Technology' and Quantity > 5
    order by Order_Date ASC;
quit;
```