**Practical No. 13**

**Aim**: To summarize data by creating an accumulating total variable for a group of data and read raw data files with formatted Input.

**Prerequisite:**

1. Understanding of fundamental programming constructs of Base SAS

**Outcome:** After successful completion of this experiment students will be able to

1. Use RETAIN statement to retain the value of a variable in the PDV across iterations of the DATA step.
2. Use First and Last variables for each variable listed in the BY statement.
3. Use multiple INPUT statements to read a raw data file with multiple records per observation.

(TO BE COMPLETED BY STUDENTS)

|  |  |
| --- | --- |
| Roll No. A020 | Name: Nicole Michelle DSouza |
| Class: | Batch: |
| Date of Practical: | Date of Submission: |
| Grade: |  |

**Assignment 1:**

For the following program.

**data** work.mid\_q4;

set orion.order\_fact;

where **'01nov2008'd** <= Order\_Date <= **'14dec2008'd**;

**run**;

* Modify the program to create an accumulating total Sales2Dte, to display the sales-to-date total (add Sales2Dte and Total\_Retail\_Price)
* Also create an accumulating variable, Num\_Orders, indicating how many orders to date that total represents. Each observation counts as one order.
* Display only the columns Order\_ID, Order\_Date ,Total\_Retail\_Price , num\_orders and Sales2Dte with a DOLLAR10.2 format.
* Add a title 'Orders from 01Nov2008 through 14Dec2008'

**Code of the program:**

LIBNAME orion "D://PA\_2021\_22/Data\_sets";

**data** work.mid\_q4;

set orion.order\_fact;

retain Sales2Dte **0**;

Sales2Dte =sum(Sales2Dte,Total\_Retail\_Price);

retain num\_orders **0**;

num\_orders = sum(num\_orders,**1**);

where **'01nov2008'd** <= Order\_Date <= **'14dec2008'd**;

format Sales2Dte DOLLAR10.2;

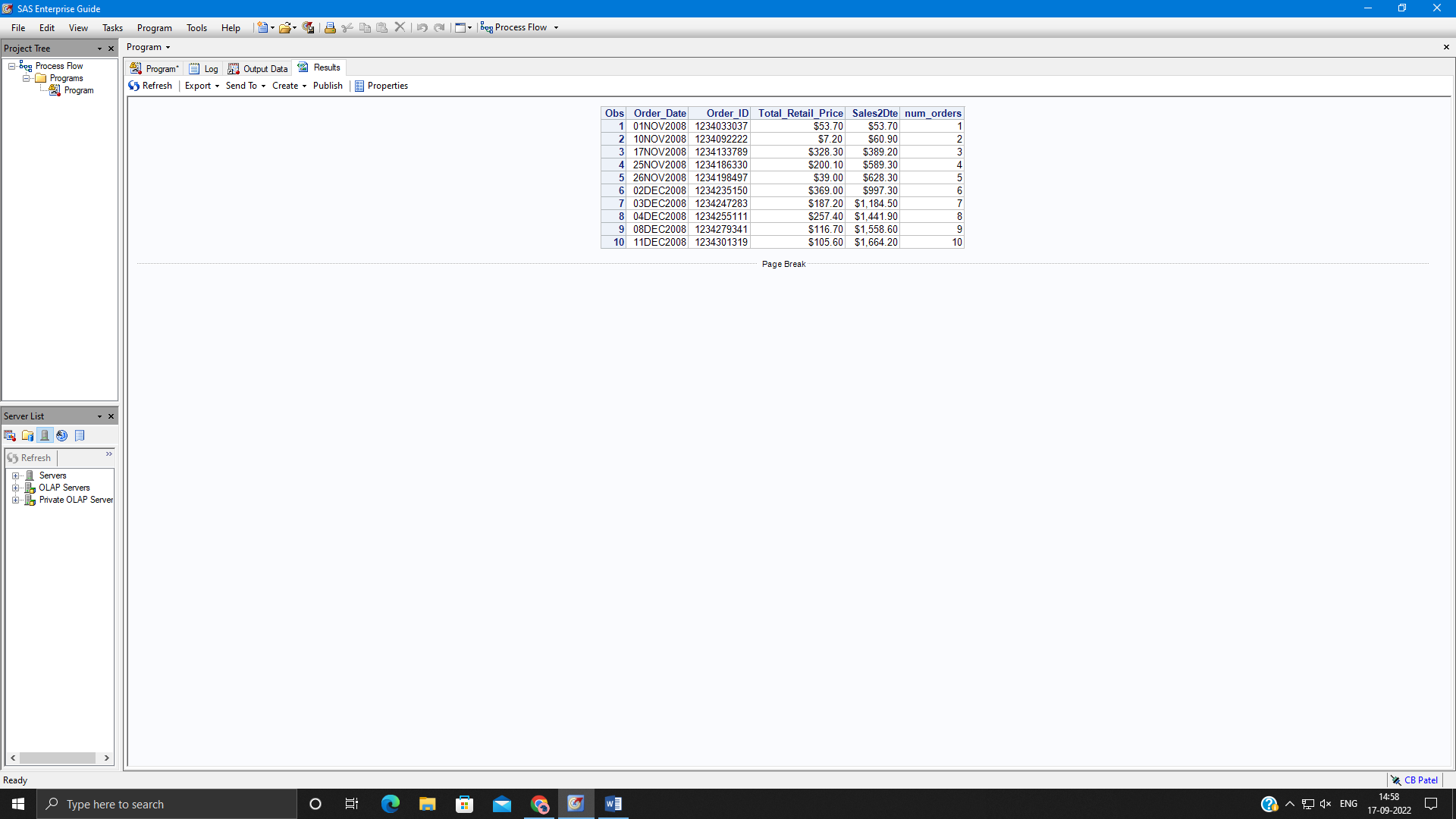
keep Order\_ID Order\_Date Total\_Retail\_Price num\_orders Sales2Dte;

**run**;

**proc** **print** data= work.mid\_q4;

**run**;

**Output of the Program:**



**Assignment 2:**

* Sort the input data set orion.order\_summary by customer\_id. and store the sorted data into work.sumsort.
* Create a new data set work.customers which will have a new variable total\_sales. This variable contains the total of sales across all months for each customer (add total\_sales and Sale\_Amt). The new data set should have only the columns customer\_id, Sale\_amt and total\_sales.
* Print the work.customers data set using DOLLAR11.2 format for the variable total\_sales.

**Code of the program:**

**proc sort data= orion.order\_summary out =work.sumsort;**

**by customer\_id;**

**run;**

**data work.customers;**

**set work.sumsort;**

**by customer\_id;**

**if first.customer\_id then total\_sales =0;**

**total\_Sales+sale\_amt;**

**if last.customer\_id;**

**keep customer\_id Sale\_amt total\_Sales;**

**format total\_sales dollar12.2;**

**run;**

**proc print data=work.customers;**

**run;**

**Output of the Program:**

****

**\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\***