

Do you remember “cust.dat” and “purchase.dat” from Spark? I have converted them to permanent SAS datasets (“cust.sas7bdat” and “purchase.sas7bdat”) but with an additional “birthdate” column in “cust”.

Your task is to print out all of the purchase data for the oldest customer (without using PROC SQL). Once you have the id of that person (assume it is a unique person) in a macro variable, you just need to use a PROC PRINT with a WHERE clause accomplish the task.

Turn in “QuickSASHW.sas”.

There are two basic approaches one might use to get the needed id:

- 1) Use, e.g., PROC MEANS to find the smallest birthdate, extract that value into a macro variable, then use a DATA step to extract the id of the person with that birthdate into a second macro variable.
- 2) Use a single data step to find the smallest birthdate and extract the id.

Use option number 2 for this assignment.

Remember to include a TITLE statement.

If you get stuck, ask Howard and/or see the hints on the back of this page.

Here is the results of PROC CONTENTS:

CUST:

#	Variable	Type	Len	Format
1	id	Char	9	
2	female	Num	8	
3	origin	Num	8	
4	birthdate	Num	8	MMDDYY10.

PURCHASE:

#	Variable	Type	Len	Format
1	id	Char	9	
2	store	Num	8	
3	purchase	Num	8	6.2
4	debit	Num	8	

General hints:

- 1) To have high efficiency, your goal should be to use `CALL SYMPUTX()` one time rather than multiple times.
- 2) Remember that we use `CALL SYMPUTX('myMacroVar', myDataSetColumnName);` to store a value from the current row into a macro variable.
- 3) Remember that to access a binary data file, you need to make a `LIBNAME` pointing to the directory that holds the files(s), and then use the `myLibname.myDataSetName` two-level syntax, where “myDataSetName” is the file name without the “.sas7bdat” extension.
- 4) Remember that you can use the form `"01Apr2019"D` as a date constant.
- 5) Remember that you can use `DATA _NULL_;` to run a data step without changing or re-writing any data.
- 6) Remember that you must use `RETAIN` to carry information across observations (rows of data). Also, you can use the form `RETAIN myVar myInitialValue;` to set an initial value. To retain multiple variables with different initial values, you need to use multiple `RETAIN` statements.
- 7) Remember that the `IF` statement has two forms:

```
a. IF myCondition THEN myAction;

b. IF myCondition THEN DO;
    myAction1;
    ...
    myActionN;
END;
```

If you need more help, read these specific hints:

- 1) If “manual code” would include quotes around the information in the macro variable, then your code that uses the macro variable will also need quotes.
- 2) You can make a retained variable to hold “the smallest birthdate seen so far” and one to hold the id of the person with the smallest birthdate seen so far. As your `DATA` step processes each line of the dataset, update the retained variables if you see a smaller birthdate. It is easiest to initialize the retained variable to a date far into the future.
- 3) You can use the `END=myVar` clause in a `SET` statement just like for an `INFILE` statement. Then an `IF myVar` (or equivalently `IF myVar=1`) clause triggers only after the entire data is processed.