Unit 1-5 Exercises

1. Subsetting Observations and Variables Using the WHERE and KEEP Statements

a. Retrieve and submit the starter program **p105e01**.

What is the variable name that contains gender values? _____

What are the two possible gender values? _____

b. Add a DATA step before the PROC PRINT step to read the data set **orion.customer_dim** to create a new data set called **Work.youngadult**.

c. Modify the PROC PRINT step to refer to the new data set.

d. Submit the program and confirm that **Work.youngadult** was created with 77 observations and 11 variables.

e. Add a WHERE statement to the DATA step to subset for female customers.

f. Submit the program and confirm that **Work.youngadult** was created with 30 observations and 11 variables.

g. Modify the WHERE statement to subset for female customers whose **Customer_Age** is between 18 and 36.

h. Submit the program and confirm that **Work.youngadult** was created with 15 observations and 11 variables.

i. Modify the WHERE statement to subset for female 18- to 36-year-old customers who have the word `Gold` in their **Customer_Group**.

j. Submit the program and confirm that **Work.youngadult** was created with 5 observations and 11 variables.

k. Modify the DATA step so that **Work.youngadult** contains only **Customer_Name**, **Customer_Age**, **Customer_BirthDate**, **Customer_Gender**, and **Customer_Group**.

l. Submit the program and confirm that **Work.youngadult** was created with 5 observations and 5 variables.

2. Subsetting Observations and Variables Using the WHERE and DROP Statements

a. Write a DATA step to read the data set **orion.product_dim** to create a new data set called **Work.sports**.

Work.sports should include only those observations with **Supplier_Country** from Great Britain (`GB`), Spain (`ES`), or Netherlands (`NL`) and **Product_Category** values that end in the word `Sports`.

Work.sports should not include the following variables: **Product_ID**, **Product_Line**, **Product_Group**, **Supplier_Name**, and **Supplier_ID**.

b. Write a PROC PRINT step to create the following report:

Partial PROC PRINT Output (First 10 of 30 Observations)

Obs	Product_ Category	Product_Name	Supplier_ Country
1	Children Sports	Butch T-Shirt with V-Neck	ES
2	Children Sports	Children's Knit Sweater	ES
3	Children Sports	Gordon Children's Tracking Pants	ES
4	Children Sports	O'my Children's T-Shirt with Logo	ES
5	Children Sports	Strap Pants BBO	ES
6	Indoor Sports	Abdomen Shaper	NL
7	Indoor Sports	Fitness Dumbbell Foam 0.90	NL
8	Indoor Sports	Letour Heart Bike	NL
9	Indoor Sports	Letour Trimag Bike	NL
10	Indoor Sports	Weight 5.0 Kg	NL

3. Using the SOUNDS-LIKE Operator and the KEEP= Option

a. Write a DATA step to read the data set **`orion.customer_dim`** to create a new data set called **`Work.tony`**.

b. Add a WHERE statement to the DATA step to subset the observations with the **`Customer_FirstName`** value that sounds like `Tony`.

Documentation on the SOUNDS-LIKE operator can be found in the SAS Help and Documentation from the Index tab by typing **`sounds-like operator`**.

c. Add a KEEP= data set option in the SET statement to read only the **`Customer_FirstName`** and **`Customer_LastName`** variables.

Documentation on the KEEP= data set option can be found in the SAS Help and Documentation from the Contents tab. (Select **SAS Products** ⇨ **Base SAS** ⇨ **SAS 9.2 Language Reference: Dictionary** ⇨ **Dictionary of Language Elements** ⇨ **SAS Data Set Options** ⇨ **KEEP= Data Set Option**.)

d. Write a PROC PRINT step to create the following report:

Obs	Customer_ FirstName	Customer_ LastName
1	Tonie	Asmussen
2	Tommy	Mcdonald

4. Adding Permanent Attributes to `Work.youngadult`

a. Retrieve and submit the starter program **p105e04**.

Notice the format and labels stored in the descriptor portion of **`Work.youngadult`**.

b. Add a LABEL statement and a FORMAT statement to the DATA step to create the following PROC PRINT report:

Obs	Gender	Customer Name	Date of Birth	Member Level	Customer Age
1	F	Sandrina Stephano	July 9, 1979	Orion Club Gold members	28
2	F	Cornelia Krahl	February 27, 1974	Orion Club Gold members	33
3	F	Dianne Patchin	May 6, 1979	Orion Club Gold members	28
4	F	Annmarie Leveille	July 16, 1984	Orion Club Gold members	23
5	F	Sanelisiwe Collier	July 7, 1988	Orion Club Gold members	19

The labels need to be changed for **`Customer_Gender`**, **`Customer_BirthDate`**, and **`Customer_Group`**.

The format needs to be changed for **`Customer_BirthDate`**.

Hint: Do not forget the option in the PROC PRINT step that enables the labels to appear.

Why do **`Customer_Name`** and **`Customer_Age`** appear with a space in the column header but do not need labels? _____

5. Adding Permanent Attributes to `Work.sports`

a. Retrieve the starter program **p105e05**.

b. Add a LABEL statement to the DATA step and a LABEL option to the PROC PRINT step to add the following labels:

Variable	Label
`Product_Category`	Sports Category
`Product_Name`	Product Name (Abbrev)
`Supplier_Name`	Supplier Name (Abbrev)

c. Add a FORMAT statement to the DATA step to display only the first 15 letters of **`Product_Name`** and **`Supplier_Name`**.

d. Submit the program to create the following PROC PRINT report:

Partial PROC PRINT Output (First 10 of 30 Observations)

Obs	Sports Category	Product Name (Abbrev)	Supplier Country	Supplier Name (Abbrev)
1	Children Sports	Butch T-Shirt w	ES	Luna sastreria
2	Children Sports	Children's Knit	ES	Luna sastreria
3	Children Sports	Gordon Children	ES	Luna sastreria
4	Children Sports	O'my Children's	ES	Luna sastreria
5	Children Sports	Strap Pants BBO	ES	Sportico
6	Indoor Sports	Abdomen Shaper	NL	TrimSport B.V.
7	Indoor Sports	Fitness Dumbbel	NL	TrimSport B.V.
8	Indoor Sports	Letour Heart Bi	NL	TrimSport B.V.
9	Indoor Sports	Letour Trimag B	NL	TrimSport B.V.
10	Indoor Sports	Weight 5.0 Kg	NL	TrimSport B.V.

Add a PROC CONTENTS step to the end of the program to verify that the labels and formats are stored in the descriptor portion.

6. Using the $UPCASE*w*. Format and the SPLIT= Option

a. Retrieve the starter program **p105e06**.

b. Add a FORMAT statement to display **`Customer_FirstName`** and **`Customer_LastName`** in uppercase values.

Documentation on the $UPCASE*w*. format can be found in the SAS Help and Documentation from the Contents tab (**SAS Products** ⇨ **Base SAS** ⇨ **SAS 9.2 Language Reference: Dictionary** ⇨ **Dictionary of Language Elements** ⇨ **Formats** ⇨ **$UPCASE*w*. Format**).

c. Add a LABEL statement to add the following labels:

Variable	Label
`Customer_FirstName`	CUSTOMER*FIRST NAME
`Customer_LastName`	CUSTOMER*LAST NAME

d. In the PROC PRINT statement, replace the LABEL option with the SPLIT= option and reference the asterisk as the split character.

Documentation on the SPLIT= option can be found in the SAS Help and Documentation from the Contents tab (**SAS Products** ⇨ **Base SAS** ⇨ **Base SAS 9.2 Procedures Guide** ⇨ **Procedures** ⇨ **The PRINT Procedure**).

e. Submit the program to create the following PROC PRINT report:

Obs	CUSTOMER FIRST NAME	CUSTOMER LAST NAME
1	TONIE	ASMUSSEN
2	TOMMY	MCDONALD