1. The data set customer\_dim is in the subdirectory STA5066 on the SAS cloud.
2. Create a library called “customer” that points to the directory containing the data set customer\_dim.
3. Using the data set customers.customer\_dim, create a temporary data set youth that contains only observations meeting all of the following conditions:
4. Female customers
5. Customer\_Age is between 18 and 36
6. Have the word Gold in their Customer\_Group
7. The data set youth should contain only that variables Customer\_Name, Customer\_Age, Customer\_BirthDate, Customer\_Gender, and Customer\_Group.
8. The data set product\_dim is in the STA5066 directory on the SAS cloud.
9. Create a library called prg1 that points to the directory containing the data set product\_dim.
10. Use the data step to create a temporary data set sports that includes only observations with Supplier\_Country from Great Britain (GB), Spain (ES), or Netherlands (NL) and Product\_Category values that end in the word Sports.
11. The data set work.sports should not include the variables Product\_ID, Product\_Line, Product\_Group, and Supplier\_ID
12. The data set work.sports should contain the following labels:

|  |  |
| --- | --- |
| Variable | Label |
| Product\_Category | Sports Category |
| Product\_Name | Product Name (Abbrev) |
| Supplier\_Name | Supplier Name (Abbrev) |

1. The data set work.sports should include formats to assure that only the first 15 letters of Product\_Name and Supplier\_Name are displayed.
2. Include a PROC CONTENTS step and verify that the labels and format specifications are included in the descriptor portion.
3. Include a PROC PRINT step to display 14 observation from the data set work.sports

1. The data set adult is in the STA5066 subdirectory on the SAS cloud.
2. Create a library nh3 that points to the subdirectory containing the data set adult.
3. Create a temporary SAS data set, Demographics that contains the following variables (For categorical variables, the table provides the meaning of the numeric codes):

|  |  |
| --- | --- |
| Variable | Description |
| seqn | Identification Number |
| dmarethn | Race-Ethnicity  1=Non-Hispanic White  2=Non-Hispanic Black  3=Mexican American  4=Other |
| dmaracer | Race  1=White  2=Black  3=Other  8=Mexican-American of unknown race |
| dmaethnr | Ethnicity  1=Mexican American  2=Other Hispanic  3=Not Hispanic |
| hssex | Gender  1=Male  2=Female |
| hsageir | Age in years at Interview |

Include labels in the data set that provide the coding of the categorical variables: dmarethn, dmaracer, dmaethnr, and hssex as part of their definition.

1. Use a proc print step to display the first 14 observations of the data set demographics.
2. The sas data set exam in in the 5066 subdirectory on the SAS cloud.

Define a library, Nhanes3, that points to the subdirectory containing the exam data set.

1. Use a data step to create a temporary data set called, examsub1, that contains the following subset of variables from the exam data set. The names that these variables should have in the data set examsub1 is provided in the last column.

|  |  |  |
| --- | --- | --- |
| Variable | Description | Name on examsub1 |
| hsageir | Age at Interview | age |
| hssex | Gender | gender |
| dmaracer | Race | race |
| bmpwt | Weight in kg | wt\_kg |
| bmpht | Height in cm | ht\_cm |
| pep6g1 | Systolic Blood Pressure, 1st reading | sbp1 |
| pep6h1 | Systolic Blood Pressure, 2nd reading | sbp2 |
| pep6i1 | Systolic Blood Pressure, 3rd reading | sbp3 |
| pep6g3 | Diastolic Blood Pressure, 1st reading | dbp1 |
| pep6h3 | Diastolic Blood Pressure, 2nd reading | dbp2 |
| pep6i3 | Diastolic Blood Pressure, 3rd reading | dbp3 |
| sppfvc | Forced Vital Capacity (ml) | fvc |
| sppfev1 | Forced Vital Capacity, 1st second (ml) | fvc1 |

1. For the variables pep6g1, peph1, pep6i1, pep6g3, pep6h3, and pep6i3 include formats to assure that, when printed, these variables will be displayed as integers (no decimal).
2. Add a step to print the first 7 observations of the data set examsub1 to make sure the printing is done correctly.
3. Add a step to display the entire descriptor portion of the data set examsub1.

1. The SAS data set lab is in the STA5066 subdirectory on the SAS cloud.
2. Create a library, NH, that points to the location of the data set lab.
3. Create a temporary SAS data set, labsub1, a subset of the NH.lab data set.
4. work.labsub1 should contain only the following variables.

|  |  |
| --- | --- |
| Variable | Description |
| seqn | sequence number |
| hgp | hemoglobin (g/dl) |
| htp | hematocrit (%) |
| tcp | cholesterol (mg/dl) |
| tgp | triglycerides (mg/dl) |
| lcp | low density lipoprotein (mg/dl) |
| hdp | high density lipoprotein (mg/dl) |
| fbpsi | fibrinogen (mg/dl) |
| crp | C-reactive protein (mg/dl) |
| sgp | plasma glucose (mg/dl) |
| urp | urinary creatinine (mg/dl) |

1. Use a proc contents step to verify that the data set labsub1 has the correct variables.
2. Use a proc print step to print the first 5 observations on the data set labsub1.

1. The data set mortality is in the 5066 subdirectory on the SAS cloud.
2. Create a library, nh3, that points to the location of the mortality data set.
3. Use a data step to create a temporary SAS data set, mortsub1 that contains only those observations on the nh3.mortality data set for which the variable eligstat is equal to 1.
4. The data set work.mortsub1 should contain only the variables

SEQN and MORTSTAT. The variable mortstat should have a label that includes the meaning of the categorical numerical codes: 0=alive at end of follow-up 1=died during the follow-up period.

1. Use a proc contents step to examine the descriptor portion of work.mortsub1.
2. Use a proc print step to display the first 100 observations on the data set mortsub1.