Week3_Part1

September 13, 2019

1 GWU STAT 4197/6197

1.1 Week 3 SAS Code Examples (Part 1): Working with Formats and Informats

- User-Defined Formats
  - Creating, Storing, Accessing, and Maintaining Formats
  - Grouping Data Values Using Formats
  - Removing Formats, and Labels from SAS Data Sets
- Picture Formats and User-Defined Informats

1.1.1 Defining character formats for discreate character values

The FORMAT procedure enables you to define your own formats for variable values. Formats determine how variable values are printed in the PROC FREQ output below. Note the following:

- Character format name (user-defined)
- Character data values
- labels
- Format applied to the PROC FREQ step

```
In [1]: * *Ex1_Numeric_Character_Formats.sas (Part 1);
        Title 'Format for character values';
        options nocenter nodate nosource;
        proc format;
        value $regionfmt
            'AFR' = 'Africa'
            'AMR' = 'Americas'
            'EUR' = 'Europe'
            'EMR'  ='Eastern Mediterranean'
            'SEAR' = 'South-East Asia'
            'WPR' = 'Western Pacific';
              run;
        proc freq data=sashelp.demographics;
          tables region;
          format region $regionfmt.;
        run;
```

```
SAS Connection established. Subprocess id is 5572
```

```
Out[1]: <IPython.core.display.HTML object>
```

1.1.2 Defining numeric formats for ranges of numeric data values

The FORMAT procedure enables you to define your own formats for ranges of numeric data values. Formats determine how variable values are printed in PROC FREQ output below. Note the following:

- Numeric format name (user-defined)
- Ranges of numeric data values with keywords LOW and OTHER
- labels
- Format applied to the PROC FREQ step

The special keyword LOW is used to define the lowest data value. Because it is a numeric format, LOW does not format missing values. In contrast, for character formats, LOW includes missing or blank values.

```
In [6]: *Ex1_Numeric_Character_Formats.sas (Part 2);
        options nocenter nodate nosource;
        proc format;
          value numfmt
                    Low - <0  = "Nonresponse"
                    0="Never"
                    1-5 = "Within past 5 years"
                    6-High = "More than 5 years ago"
                    . ="Missing" ;
           value $charfmt
                    Low-<'0'  = "Nonresponse"
                   '0' = "Never"
                    '1'-'5' = "Within past 5 years"
                    '6'-High = "More than 5 years ago" ;
         run;
        data work.have;
        input id $ 1 Colonoscopy 3-4 c_Colonoscopy $6-7;
        datalines;
        A -1 -1
        B  .
        C 3   3
        D -9 -9
        F 3  3
        G 5  5
        H 6  6
        I .
        J 7  7
        ;
```