Math 475 Homework 2

Due Wednesday 7th September at the beginning of lecture

The first part of this homework is from Chapter 2 of Gentle's book, and the remainder is related to Chapters 1 and 2 of the Little SAS Book.

Exercises from Gentle's *Computational Statistics*

1. Exercise 2.1

2. Exercise 2.3

3. Exercise 2.9

4. Exercise 2.12 (`R` exercise). Print and submit the code and plots.

SAS Programming

5. The body mass index (BMI) is a measure used as an indicator of an individual's body fat. The following program computes BMI using weight (in pounds) and height (in inches).

   ```
   ** Compute BMI using pounds and inches ;
   DATA bodymass;
           Weight = 150;
           Height = 68;
           BMI = (Weight / Height ** 2) * 703;
   RUN;
   ```

   This program creates a SAS data set called BODYMASS in the WORK library.

   (a) Find the BMI variable value.
   (b) Find the types and lengths for the variables (look at the properties of the data set).
   (c) Choose another pair of values for Weight and Height in the program. Add a `PROC PRINT` to list the data in the BODYMASS data set and write down the values.

6. The following `DATA` step attempts to create a SAS data set containing information about a city.

   ```
   OPTIONS NONUMBR;
   DATA info;
           City = 'Sao Paulo';
   ```

```
                  Country = 'Brazil'
                  CountryCode = 55;
                  CityCode = 11;
          RUN;
```

(a) Type this program into the editor and submit it. Review the information in the SAS log. Write down the number of notes, warnings, and errors produced by this code (not including any start-up messages).

(b) How should this code be fixed?

(c) Fix the code and submit the program again. Write down how many observations and variables were created in this data set.

SAS Multiple Choice

7. What type of data files are not considered raw data?

   (a) ASCII files

   (b) CSV files

   (c) Text files

   (d) SQL tables

8. The data in the following program are an example of what type of data?

```
          DATA readme;
                  INPUT Place $ Code $;
                  DATALINES;
          AG 5678
          SLO 1234
          PB 3456
                  ;
          RUN;
```

   (a) Character

   (b) Instream

   (c) Internal raw data

   (d) All of the above

9. Which SAS statement allows you to refer to an external raw data file?

   (a) DATALINES

   (b) DATA

   (c) INFILE

(d) `INPUT`

10. Which of the following data types cannot be read with list input?

    (a) Missing data indicated by a period

    (b) Character data with embedded blanks

    (c) Standard numeric data

    (d) All of the above

11. Given this note in the SAS log, what could you add to fix the INPUT statement so that the ID variable would be read correctly including all digits and hyphens?

    ```
    INPUT ID GPA Age;
    NOTE: Invalid data for ID in line 1 1-9.
    RULE          - - - - + - - - - 1 - - - - + - - - - 2
    1                  5437          - 2212    3.84    21
    ID=. GPA=3.84 AGE=21 _ERROR_=1 _N_=1
    ```

    (a) A dollar sign

    (b) A column range

    (c) An informat

    (d) None of the above

12. Which of the following data values would not require an informat?

    (a) 44.5E2

    (b) $1,689

    (c) 08/18/1920

    (d) 4,928

13. Which informat would be appropriate to read the value 07/04/1776.

    (a) MMDDYY8

    (b) MMDDYY10

    (c) DATE8

    (d) DATE10

14. Select the INPUT statement that would be appropriate for reading data values for the variables Name, Salary, and Age in the following raw data.

    ```
    - - - - + - - - - 1- - - -+ - - - - 2 - - - - + - - - - 3
    Sally            $64, 350    41
    Marian           $55, 500    38
    Oprah            $75, 000, 000    59
    ```

(a) `INPUT Name $ Salary DOLLAR11.   Age;`

(b) `INPUT Name $ Salary :DOLLAR11.   Age;`

(c) `INPUT Name $ 10 Salary DOLLAR11.   Age;`

(d) `INPUT Name $ '$' Salary Age;`

15. Which of the following tells SAS to go to the next line when reading in data?

   (a) @

   (b) @@

   (c) /

   (d) +n

16. If your raw data file contains only data and no variable names, which `PROC IMPORT` option should you use?

   (a) `DELIMITER=`

   (b) `GUESSINGROWS=`

   (c) `GETNAMES=`

   (d) `OUT=`

You should also look at the website for the Little SAS Book (you probably need to register for a SAS account). Try reading in some data sets with variables of different types to get some practice with reading data into SAS. This is not to be submitted as part of this homework, but a familiarity with reading data into SAS will be needed for future homework assignments.