

|  |  |  |  |
| --- | --- | --- | --- |
| Academic Year | 2022 – 2023 | | |
| Semester | Fall | Winter | Summer |
| Course Code - Name | BAN110 | | |
| Instructor | Dr. Razi Iqbal | | |
| Assessment | Assignment 3 |  | |

**Student ID 176938215**

**Student Name Rongzhao Yi**

**Assignment 3**

The main purpose of this lab is to get students familiarize with dealing with missing values.

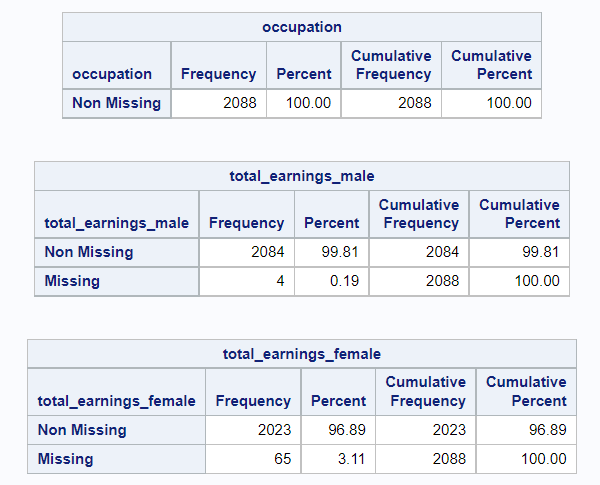
**Instructions:**

* You are required to submit your answers in this document by pasting your SAS code under the solutions heading below.
* Please do not submit .sas files. Submit only this word document with your code inside it.
* Total Marks for this assignment are 5 marks.
* Students having exactly similar code will get a straight 0.
* You are required to complete these exercises using SAS.
* The deadline for submission of this assignment is March 21, 2023 end of the day.

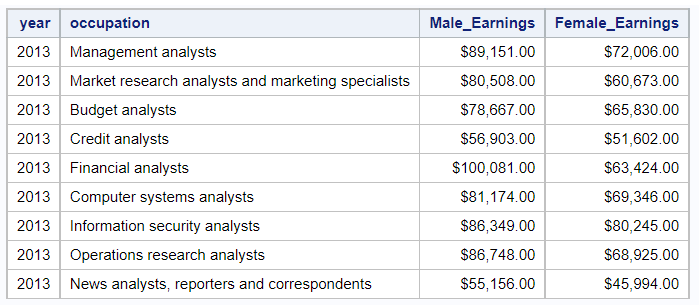
**Question**

You are provided with the dataset ‘JobsGender.xlsx’ file. You are required to perform the following tasks using this dataset in SAS:

* Import the data from the excel sheet in SAS
* Create a dataset in SAS namely JobsGender which would get all the data from Excel sheet.
* Once the data is imported you need to create a new dataset called ‘Demography’ which would bring only the following columns from the original Orders dataset: Year Occupation Total\_Earnings\_Male Total\_Earnings\_Female.
* There are some values in Total\_Earnings\_Male Total\_Earnings\_Female columns that are ‘NA’. SAS does not recognize these values as missing values. Write a proc format so that all ‘NA’ values from these two columns are recognized as missing or non-missing as shown in the screenshot below: (You can show occupation in the output as well)

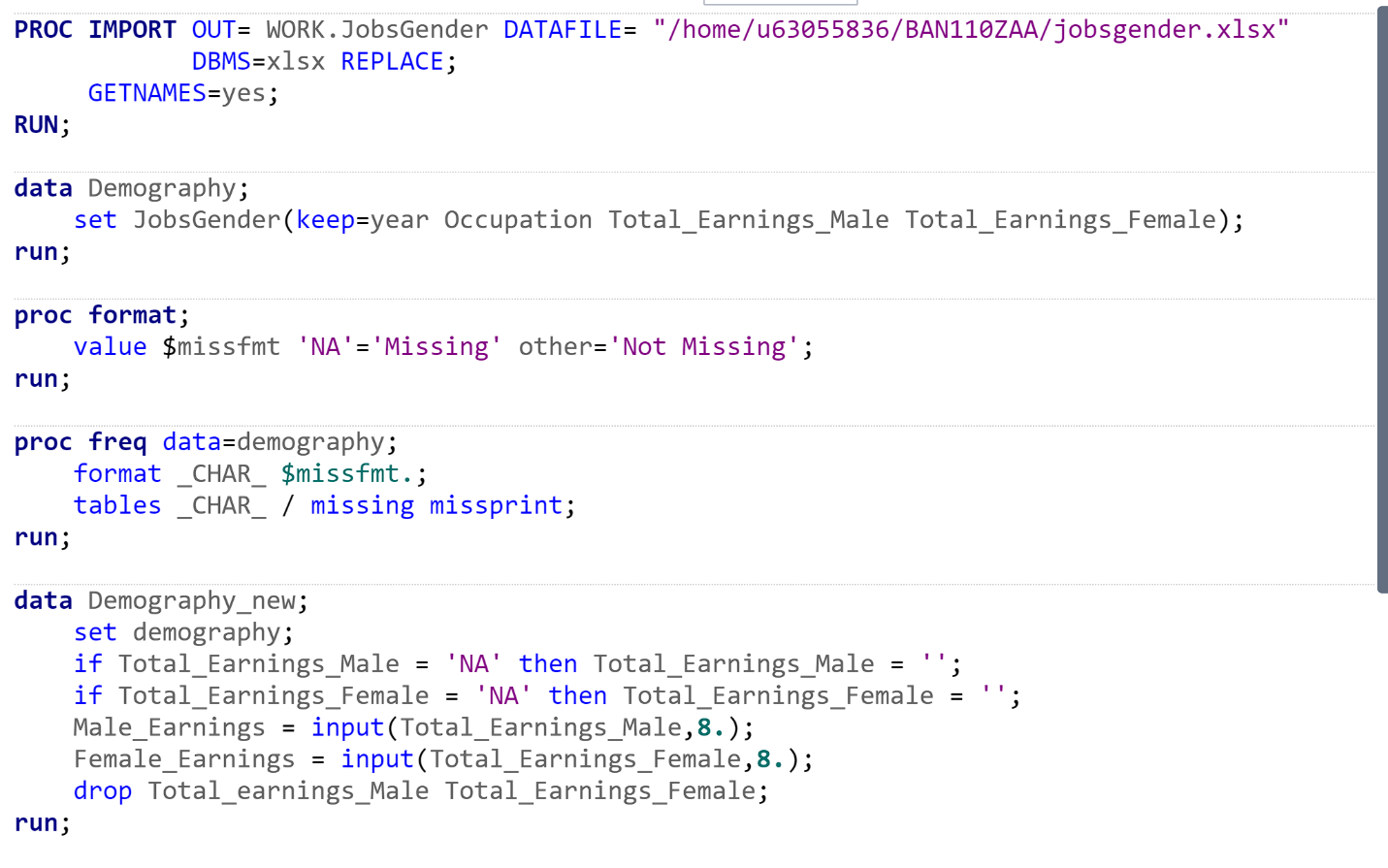


* Once you know how many missing and non-missing values you have, go ahead and replace those ‘NA’ values by SAS understandable missing values.
* After SAS could recognize those missing values, use Proc Standard to replace all the missing values in those 2 columns by the mean of those columns. (Note that those two columns are character columns and you might want to convert them to numeric columns first.)
* Finally show a list of all the analysts in Year 2013 as shown below. Make sure to format earnings:



**Make sure to include only the columns shown in the screenshot above.**

**Solution**

****