**IDS462 HW6**

Download the SAS datasets attached for this assignment, and see the problem descriptions below. The are 5 problems some with multiple parts in this assignment.

Your SAS code/program should start with the following:

/\* IDS462 HW6\*/

/\* Name of student1 \*/

and if this is a team of more than 1 student contributed, then add the names of all students as

/\* Name of student2 \*/

….

Submit only one SAS code to solve all problems given, and make sure to put comments.

**Problem 1)**

Monthly order data for the first half of the year for Orion Star is stored in the **orders\_midyear** data set. This data set is updated monthly and can contain data for 1 to 12 months.

* 1. The company decided to create a Frequent Shopper Program (FSP) if the customer has order data for at least three months. To be eligible for the invitation to the program, the customer has to have placed order in 50% of the months to date and has to have spent at least $1,000 since the beginning of the year. Example: If data file has 6 months of orders from a customer, the customer must have spent over $1,000 in three or more months. If the data file has 3 months of orders, the customer has to have spent over $1,000 in two or more months to be eligible for the Frequent Shopper program.
* Open the input data set and verify that it contains at least three months of data.
* If there is less than three months, write a message to the SAS log and stop the DATA step immediately. Use the STOP statement, i.e. stop; if the condition holds, and the PUT statement to display the message (for ex: put ‘insufficient data for FSP’;)
* If there are at least three months of data in the data set, create an array, **Mon**, to access the **Month*n*** variables, regardless of how many **Month*n*** variables exist in the data set.   
  Hint: Consider using a SAS variable list to list the array elements.
* Use a DO loop to examine each customer’s data to determine whether the customer qualifies for the Frequent Shopper Program.
* Create a new data set, **fsp**, that contains an observation for each qualifying customer.
* The report should only display variables Customer\_ID Months\_Ordered Total\_Order\_Amount.

Hints:, the DIM function (i.e. dim(amt)<3 if the array is named as amt)

* 1. Print the resulting data set with **an appropriate title and formats**, and verify your results.

**Problem 2)**

The Public Safety Department at Orion Star wants all employees to be aware of the new polices and procedures regarding customer incidents in their retail stores.

* Each employee must participate in web-based training and then take a multiple-choice test that consists of 10 questions.
* Each question has five choices (A through E).
* The test results from each testing session are entered into the SAS dataset **testAnswers**.
* Each observation in **testAnswers** contains a single person’s answers.
  1. Use the answer key 'A','C','C','B','E','E','D','B','B','A' and read **testAnswers** and determine whether each person passed or failed the test.
* Compute a variable **Score** that contains the total correct answers for each person.

**🖉** Create a **temporary array** for the answer key.

* If an employee scores 7 or higher, write the observation to a data set named **passed**.
* If an employee scores less than 7, write the observation to a data set named **failed**.
  1. Print the **passed** data set to verify that it contains 12 observations.

**Problem 3)**

Two cholesterol-lowering medications (statin) and a placebo were given to teach of 10 volunteers with total cholesterol readings of 240 or higher. After 6 weeks, the following total cholesterol values were recorded:

Statin A: 220 190 180 185 210 170 178 200 177 189

Statin B: 160 168 178 200 172 155 159 167 185 199

Placebo: 240 220 246 244 198 238 277 255 190 188

Create a SAS dataset by reading these data. That is, the program create the treatment variable (call it TREAT) with values of A, B, and Placebo. Next run a one-way ANOVA followed by a Student-Newman-Keuls multiple-comparison test.

**Problem 4)**

Two group of patients, those with genetic deficiency and the other a “normal” group, are given either a drug to cure depression or a placebo. A standardized depression survey is given to all the subjects with the following results. Perform a two-way analysis of variance to test for drug & genetic differences.

|  |  |  |
| --- | --- | --- |
|  | Anti-depression drug | placebo |
| Genetic deficiency | 9 | 9 |
| 11 | 6 |
| 10 | 6 |
| 10 | 7 |
| normal | 5 | 12 |
| 4 | 11 |
| 7 | 10 |
| 7 | 11 |

**Problem 5)**

Each of 12 patients wet on the South Beach diet for 5 weeks. Their weights before & after the diet are listed below

300 290 350 331 190 200 400 395 244 240 321 300

330 332 250 242 190 185 160 158 260 256 240 220

Perform a paired t-test comparing before & after values. Analyzed these data as if the values weren’t paired. How the p-values compare?