IDS462 HW3

Download all the SAS datasets attached for this assignment, and see the problem descriptions below. The are 3 problems with multiple parts in this assignment.

Your SAS code/program should start with the following:

/\* IDS462 HW3 \*/

/\* 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/program to solve all problems given, and make sure to put comments within in the beginning of each problem/part to show what you’re solving, for ex **\*Prob1-Part1;**

**Problem 1:**

The data set **employee\_organization** contains information about employee job titles, departments, and managers.

**Part 1: Creating Multiple SAS Data Sets**

Use SELECT with conditional logic and explicit OUTPUT statements to write to these data sets depending on whether the value of **Department** is*Administration*, *Stock & Shipping*, or *Purchasing*, respectively. Ignore all other **Department** values. Name the data sets **work.admin**, **work.stock**, and **work.purchasing**.

**Part 2: Specifying Variables and Observations**

Create two data sets: one for the Sales Department and another for the Executive Department.

* Name the data sets **work.sales** and **work.exec**.
* Output to these data sets depending on whether the value of **Department** is *Sales* or *Executives*, respectively. Ignore all other **Department** values.
* The **work.sales** data set should contain three variables (**Employee\_ID,** **Job\_Title**, and **Manager\_ID)**
* The **work.exec** data set should contain two variables (**Employee\_ID** and **Job\_Title)**
* Print only the first six observations from **work.sales**. Use a data set option to display only the first six observations. Add an appropriate title.
* Print selected observations from **work.exec**. Use data set options to process only observations 2 and 3. Add an appropriate title.

**Problem 2)**

The data set **order\_fact** contains a group of orders across several years, sorted by **Order\_Date**.

**Part 1) Creating Accumulating Totals with Conditional Logic**

Orion Star would like to analyze 2009 data by creating accumulating totals for the number   
of items sold from retail, catalog, and Internet channels.

* The value of **Order\_Type** indicates whether the sale was retail (=*1*), catalog (=*2*),   
  or Internet (=*3*).
* Create a data set named **work.typetotals** with accumulating totals for **TotalRetail**, **TotalCatalog**, and **TotalInternet**, as described above.

**🖉** The variable **Quantity** contains the number of items sold for each order.

* For testing your program in this step, read only the first 10 observations that satisfy the WHERE statement.

**🖉** Remember to process only those rows where **Order\_Date** occurs in 2009.

* 1. Continue testing your program by printing the results from part **a**. Print all the variables and verify that the program is correctly calculating values for the accumulating totals.

**Part 2) Creating Accumulating Totals by Month**

Orion Star would like to generate the following report showing all orders in 2011 along with an accumulating total:

* The accumulating total should reset to zero at the start of each new month.
* Remember to process only those rows where **Order\_Date** occurs in 2011.

Partial PROC PRINT Output (148 Total Observations)

Accumulating Totals by Month in 2011

Total\_

Order\_ Retail\_

Obs Date Order\_ID Price MonthSales

1 02JAN2011 1241054779 $195.60 $195.60

2 03JAN2011 1241063739 $160.80 $356.40

3 04JAN2011 1241066216 $306.20 $662.60

4 06JAN2011 1241086052 $37.80 $700.40

5 13JAN2011 1241147641 $362.60 $1,063.00

**Use year(Order\_Date) & month(Order\_Date) to find year & month of the Order\_date.**

**Problem 3)**

1. Run the following program to create the data set PRINCIPAL:

DO SUBJ = **1** TO **200**;

X1 = ROUND(RANNOR(**123**)\***50** + **500**);

X2 = ROUND(RANNOR(**123**)\***50** + **100** + **.8**\*X1);

X3 = ROUND(RANNOR(**123**)\***50** + **100** + X1 - **.5**\*X2);

X4 = ROUND(RANNOR(**123**)\***50** + **.3**\*X1 + **.3**\*X2 + **.3**\*X3);

OUTPUT;

END;

1. Write statements necessary to run a factor analyses and produce a scree plot.
2. Looking at this plot, decide how many factors you want to maintain.
3. Add an option to the procedure to extract this many factors and to perform a VARIMAX rotation.
4. List the cluster of variables going into each factor.