**Name:**

**Assignment 6 - SAS #1 – Introduction to Working with SAS**

**Guidelines**

* Follow the steps below, using the in-class exercise as a guide.
* You must submit your answers electronically in a single Word document. Fill in the answer sheet at the end of this document.
* You must include your name at the top of the document.
* Your answers should be emailed, as an attachment, to your instructor with the subject:
**MIS2502: Intro to SAS**
* The email must be sent by the start of class the day the assignment is due.

***If you do not follow these instructions, your assignment will be counted late.***

**MAKE SURE YOU HAVE COMPLETED THE IN-CLASS EXERCISE FIRST BEFORE YOU DO THIS ASSIGNMENT. AND MAKE SURE YOU’VE CHANGED THE SAMPLING DEFAULTS AS SPECIFIED ON PAGE 30 OF THE IN-CLASS EXERCISE! OTHERWISE, YOU WILL GET THE WRONG ANSWERS!**

Initial Data Exploration

A supermarket is offering a new line of organic products. The supermarket's management wants
to determine which customers are likely to purchase these products.

The supermarket has a customer loyalty program. As an initial buyer incentive plan, the supermarket provided coupons for the organic products to all of the loyalty program participants and collected data that includes whether these customers purchased any of the organic products.

The **ORGANICS** data set contains 13 variables and over 22,000 observations. The variables in the data set are shown below with the appropriate roles and levels:

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **ModelRole** | **Measure-mentLevel** | **Description** |
| **ID** | ID | Nominal | Customer loyalty identification number |
| **DemAffl** | Input | Interval | Affluence grade on a scale from 1 to 30 |
| **DemAge** | Input | Interval | Age (years) |
| **DemCluster** | Rejected | Nominal | Type of residential neighborhood |
| **DemClusterGroup** | Input | Nominal | Neighborhood group |
| **DemGender** | Input | Nominal | M = male, F = female, U = unknown |
| **DemRegion** | Input | Nominal | Geographic region |
| **DemTVReg** | Input | Nominal | Television region |
| **PromClass** | Input | Nominal | Loyalty status: tin, silver, gold, or platinum |
| **PromSpend** | Input | Interval | Total amount spent (dollars) |
| **PromTime** | Input | Interval | Time as loyalty card member (months) |
| **TargetBuy** | Target | Binary | Organics purchased? 1 = Yes, 0 = No |
| **TargetAmt** | Rejected | Interval | Number of organic products purchased |

**🖉** Although two potential target variables are listed, this assignment concentrates on the binary variable **TargetBuy**.

1. Create a new diagram named **Organics**.
2. Define the data set **AAEM.ORGANICS** as a data source for the project.
3. Use the Data Source Wizard defaults, but set the roles for the analysis variables as shown in the table on the first page.

(You can go back and modify variable roles even after you complete the wizard by right-clicking on the **Organics** data source and selecting **Edit Variables…**)

The variable **DemClusterGroup** contains collapsed levels of the variable **DemCluster**. Presume that, based on previous experience you believe that **DemClusterGroup** is sufficient for this type of modeling effort. Set the model role for **DemCluster,** along with **TargetAmt,** to Rejected.

**Include a screen shot showing DemCluster and TargetAmt is rejected (as on page 11 or 13 of the in-class exercise). (place on the answer sheet)**

1. Finish the Organics data source definition.
2. Examine the distribution of the target variable **TargetBuy**. You can do this by right-clicking on the data source and selecting **Edit Variables…** Then right-click on **TargetBuy** and then click the **Explore** button.

What is the proportion of individuals who purchased organic products (hint: take a look at the Means column in the “Sample Statistics” window)?

**ANSWER:** \_\_\_\_\_\_\_\_ (use the answer sheet)

1. Create a pie chart of those who purchased organics versus those who did not.

**Include a screen shot displaying the pie chart (place on the answer sheet)**
2. Add the **AAEM.ORGANICS** data source to the **Organics** diagram workspace.

Explore the **Organics** data source and answer the following questions:

1. Are there more males or more females in the sample?
[Hint: Create a bar chart.]

**ANSWER:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (use the answer sheet)

**Include a screen shot displaying the bar chart (place on the answer sheet).**
2. Plot **DemAge** using a histogram with 8 “X” bins.
3. Based on this histogram, the largest of those eight groups consists of people between…

**ANSWER:** \_\_\_\_\_\_\_ and \_\_\_\_\_\_\_\_ years old (use the answer sheet)

**Include a screen shot displaying the histogram (place on the answer sheet).**
4. How many people in the sample have no data for **DemAge**?
[Hint: You may have to edit the graph properties to show the “missing bin”.]

**ANSWER**: \_\_\_\_\_\_\_\_\_\_ (use the answer sheet)

1. What is the maximum, minimum, and median values for DemAffl?

**ANSWER: Max \_\_\_\_\_\_\_ Min \_\_\_\_\_\_\_\_\_ Mean \_\_\_\_\_\_\_\_\_\_\_\_** (use the answer sheet)
2. Plot **DemAffl** using a histogram with 30 “X” bins. Based on that, is it more common to have a high income or a low income (note: wealthier households have greater affluence scores)?

**ANSWER:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (use the answer sheet)

**Include a screen shot displaying the histogram (place on the answer sheet).**
3. Are older people, on average, wealthier than younger people?
[HINT: Create a bar chart of **DemAge** (category) versus **mean** **DemAffl** (response)]

**ANSWER:** \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ (use the answer sheet)

**Include a screen shot displaying the bar chart (place on the answer sheet).**

Answer Sheet for Assignment 6

Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

 *Note that the item numbers correspond to the assignment instructions.*

|  |  |
| --- | --- |
| Item Number | Answer |
| 3 (screen shot) |  |
| 5 (answer) |  |
| 6 (answer and screen shot) |  |
| 8 (answer and screen shot) |  |
| 10 (answer and screen shot) |  |
| 11 (answer) |  |
| 12 (answer) |  |
| 13 (answer and screen shot) |  |
| 14 (answer and screen shot) |  |