# **BAN210- Workshop 2**

| Total Mark: | 10 marks |
| --- | --- |
| Submission file(s): | * WS02.docx * *WS2-pg5-76.lst* * WS2-pg5-38.lst |

Please work in **groups** to complete this lab. This workshop is worth 5% of the total course grade. Please submit the submission file(s) through Blackboard. Only one person must submit for the group and only the last submission will be marked.

## **Part I: Video Tutorial**

1. Watch the second video in Getting Started with SAS Enterprise Miner playlist:

* Video (Getting Started with SAS Enterprise Miner: Exploring Input Data and Replacing Missing Values): <https://youtu.be/EBFtrPKQ0zg>

## **Part II: Association Rules**

1. Find the workshop instructions on pages 5-74 to 5-79 of the SAS Advanced Business Analytics course notes (available under BB> Course Information> Resources). As you go through the steps, answer the following questions.

(On page 5-74, before getting to step 3, first create the ABA data library, as as in the previous workshop.)

**Answer the following questions:**

* At step 17 on page 5-76,
  + Save the output window as *WS2-pg5-76.lst* and submit with this workshop.
  + Based on the Statistics Plot,
  + which two rules have the highest support percentage?

CKING == > SVG (Support 54.17)

SVG == > CKING (Support 54.17)

CKING == > ATM (Support 36.19)

* + What are the confidence percentages for these rules?
  + Although the rules are based on the same itemsets, why are the confidence values different?

Chart, scatter chart

Description automatically generated

* + What is the maximum lift? Which rule has this lift value? What does this rule say (use page 5-73)?
* At step 19 of page 5-78, what does Rule 10 say?

Map Rule

RULE10 HMEQLC & CKING ==> CCRD

* At step 20 of page 5-79, by looking at the rule matrix,
  + What is the most common right hand of rules?
  + What is the left hand of rules shown on the second row from the top?

A screenshot of a computer

Description automatically generated with medium confidence

* List five possible actionable decisions from the analysis (Page 5-79, idea exchange).

## **Part III: Clustering**

1. Find the workshop instructions on pages 5-26 to 5-45 of the SAS Advanced Business Analytics course notes (available under BB> Course Information> Resources). As you go through the steps, answer the following questions. (On page 5-26, step 3, look for ABA library)

**Answer the following questions:**

* At step 7 of page 5-28, replace the following image with a screenshot of the Data Source Wizard, showing the summary information:

Text

Description automatically generated

* At step 4 of page 5-29, paste an image of your histogram here.

Chart, histogram

Description automatically generated

* At step 6 of page 5-31, paste an image of your histogram here.

Chart, histogram

Description automatically generated

* Paste an image supporting each of the observations made on page 5-32:
  + The zero average household size seems to be evenly distributed across the longitude and latitude variables.
  + It seems concentrated on low incomes and populations,
  + And also makes up the majority of the missing observations in the distribution of region density.

Graphical user interface, application, Word

Description automatically generated

* At step 10 of page 5-38, save the output window as *WS2-pg5-38.lst* and submit with this workshop.
* What standardization method is used (page 5-40 to 5-41)? What does it do and why?

A picture containing text

Description automatically generated

Chart

Description automatically generated

* At step 1 of page 5-42, paste the segment size diagram here.

Chart, pie chart

Description automatically generated

* At step 2 of page 5-45, which segment has the highest median household income? And what is the average household size in this segment?

Table

Description automatically generated

* Paste the Tree on page 5-46 here.

Diagram

Description automatically generated

## **Part IV: Group Work**

1. Add this declaration to your file:

We, ------------ (mention your names), declare that the attached assignment is our own work in accordance with the Seneca Academic Policy. We have not copied any part of this assignment, manually or electronically, from any other source including web sites, unless specified as references. We have not distributed our work to other students.

1. Specify what each member has done towards the completion of this work:

|  | Name | Task(s) |
| --- | --- | --- |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |