**Appendix A**

**How to Use Teradata Warehouse Miner at the University of Arkansas**

Teradata University Network (TUN) members – faculty and students—can take advantage of **Teradata Warehouse Miner** (**TWM**) housed at the University of Arkansas (UA). This document illustrated using TWM by stepping through an association analysis example. The purpose to focus on how to do a association analysis data mining tasks using TWM.

You will access the TWM via a remote desktop platform. You can download instructions for this at

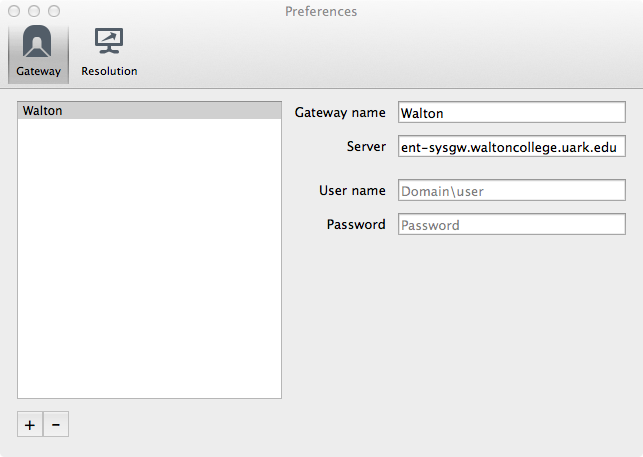
<http://enterprise.waltoncollege.uark.edu/Remote_Desktop_TUN_GW.pdf>

If you are using a **Mac**, the instructions are basically the same, except:

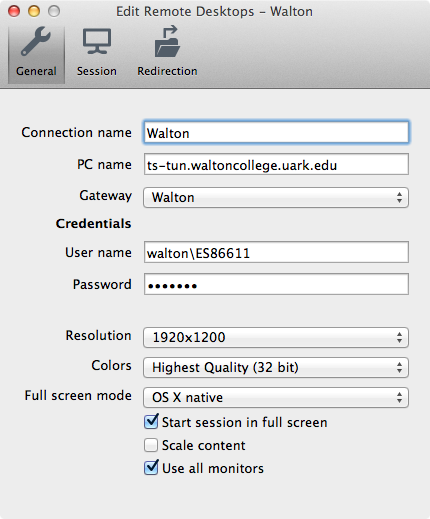
* Students should obtain the MS Remote Desktop Connection from AppStore
* The Gateway is set in "Preferences"
* Some screens are slightly different (as shown here)

Please contact Zero ( [itszero@gmail.com](file:///C:\Users\Otis\Desktop\itszero@gmail.com) ) if you have problems setting up the remote desktop for the Mac.

For Mac Desktop Setting

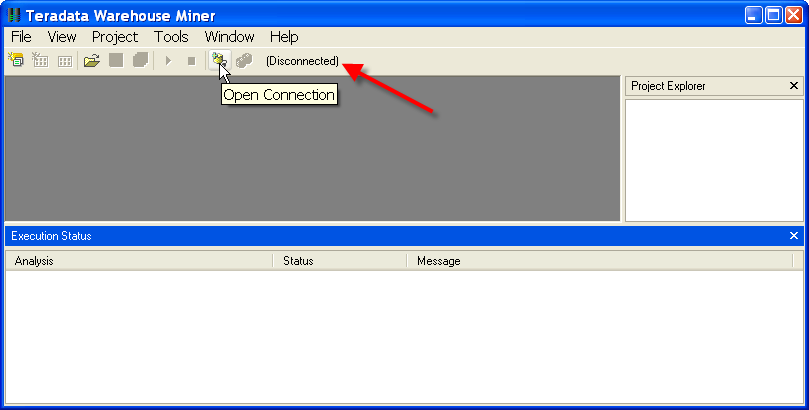


For Mac Desktop Setting

****

**Teradata Warehouse Miner**

Use the **icon on the Desktop** to open **Teradata Warehouse Miner 5.0.** Note that TWM is Disconnected when initially opened. Click the Open Connection icon as shown

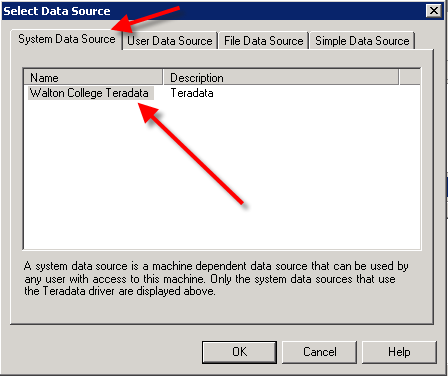


Select Data Source dialog.

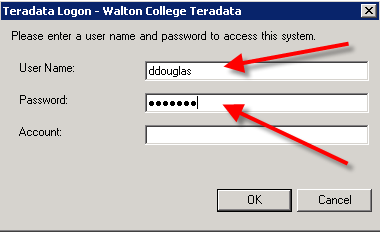
The System Data Source tab should be the default tab; if not, click it.

Then locate and select WaltonCollegeTeradata from the list of names—only name In this example.

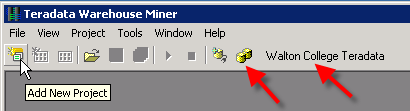
Click the OK button to open the logon dialog.



Enter your User Name and Password. Click the OK button.

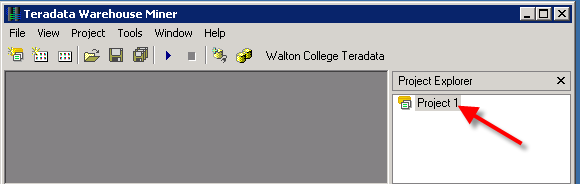


Note that the grayed out icon to the right of the Open Connection icon is now yellow and the word Disconnected has been replaced with the name of the data source to which the system is now connected. Also displayed is the tooltip for the Add New Project icon on the toolbar. The File menu can also be used to create a new project.

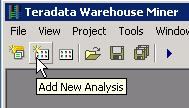
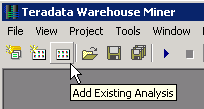


Also, note that there is a Project Explorer window to the right and an Execution Status window on the bottom.

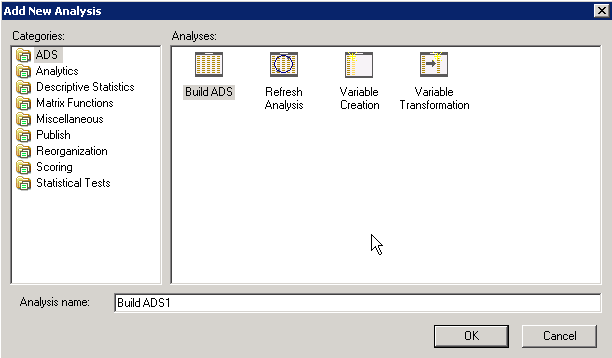
Open a new project by clicking the Add New Project icon—the default name of the new project is Project 1 and is in the Project Explorer window—you can change the project name by clicking it and then changing to a desired name.



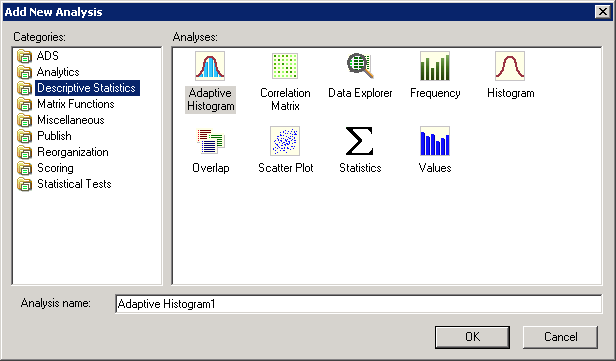
Having a project selected in the Project Explorer window activates additional icons on the toolbar—the standard windows icons for a folder, saving a file and save all. The other two icons are the Add New Analysis and Add Existing Analysis icons which are shown below.

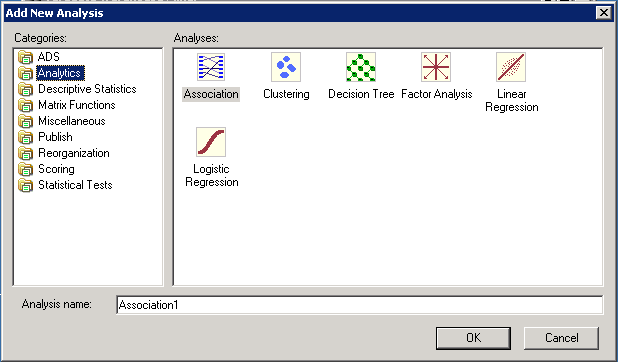
Click the Add New Analysis icon to open its dialog.



The types of analysis are listed in the Categories pane of the Add New Analysis dialog. Note the default **Analysis name** in the textbox near the bottom of the dialog. Because the type of analysis selected is ADS, the default name is BuildADS1—you can build an Analysis Data Set (ADS) if needed. Click the Descriptive Statistics category to view the types of analysis available in this category.



DM is the goal of this illustration so click Analytics in the **Categories** pane.



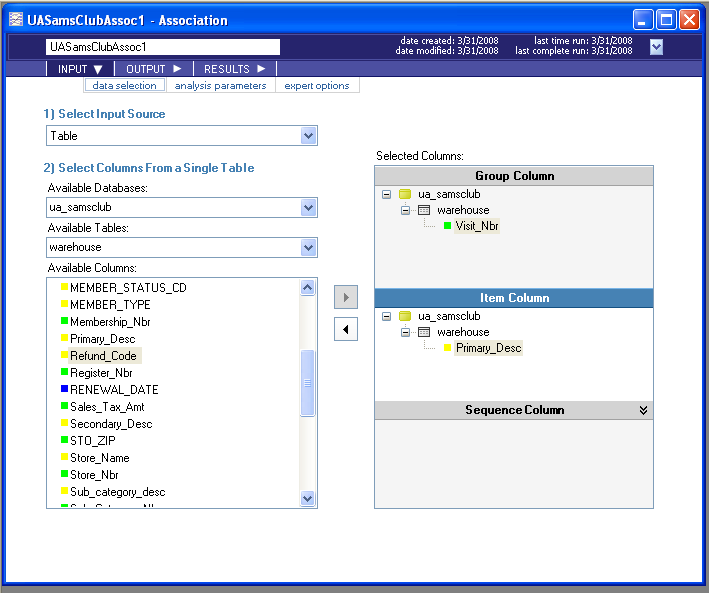
Six data mining techniques are available. Decision Trees, Linear Regression and Logistic Regression are directed or supervised data mining techniques. This means that they have a target variable (dependent variable) and thus are sometimes called predictive models. Association analysis and clustering are undirected or unsupervised data mining techniques and do not have a target or dependent variable.

The data for this example will be from a table (WAREHOUSE) in the ua\_samsclub database and consists of more than 4 million rows of data. This table was created by joining desired columns from a number of tables in the ua\_samsclub data.

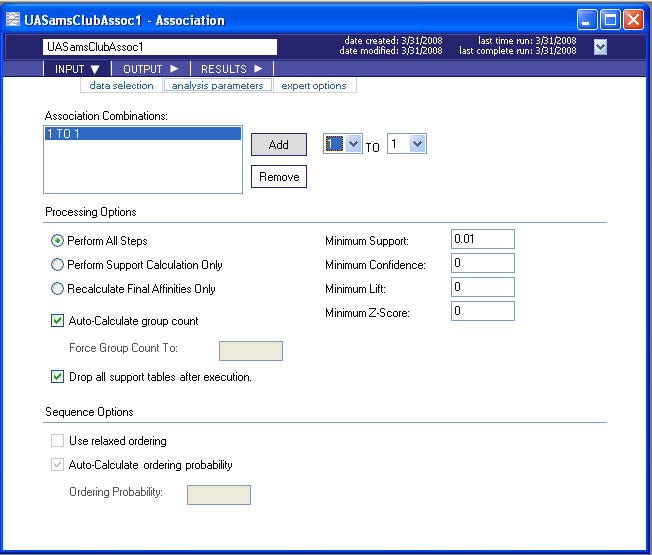
**Association Analysis**

The following example is for association analysis that uses the **warehouse** table of ua\_samsclub. The data format for **Association Analysis** is a transaction format—consisting of two columns; a customer and a product. Thus, if a customer purchases 4 items, then there would be 4 rows with the customer repeated.

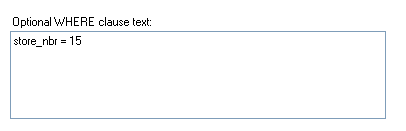
The equivalent of this for the **warehouse** table is visit\_nbr and Primary\_desc. See the setup below.



Next, click the **analysis parameters** tab toward the upper left. Note that the default setting is for one antecedent and one consequent.

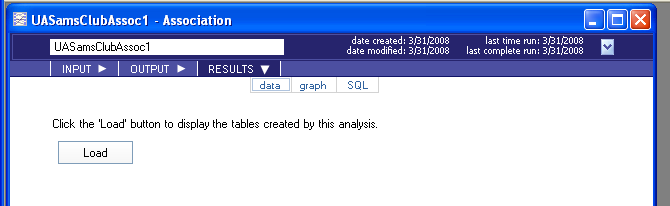


Click the **expert options** tab. Enter an SQL statement – in this case, only store number 15.

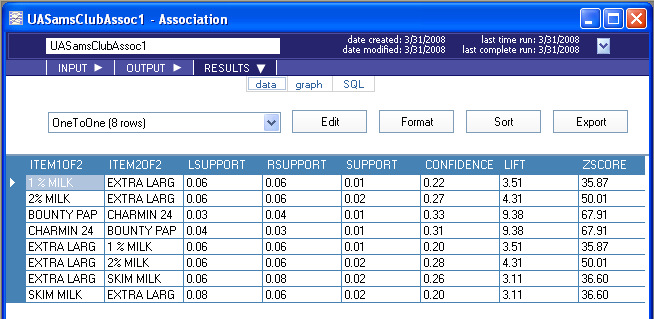


Click the run button.

When the run is completed, click results—a load table button should appear.



Click the Load button and review the table.



Click graph.

