**MATH 3210**

**Assignment #2 - Sommelier Project**

**Scenario:** The dataset Wine.xls contains the values of thirteen attributes for three classes of wines. This data is the result of a chemical analysis of wines grown in the same region in Italy but derived from three different cultivars. The three classes of wine do not occur in equal portions and result from a combination of the grape cultivar, the growing conditions and the seasonal weather.

Human wine experts (Sommelier) are not always consistent in their classification and are subject to environmental issues such as illness and allergies. These inconsistent classification results contribute classification noise to the historical dataset.

Our customer wants to use an automated chemical analysis device to classify the wine by type. This process is intended to make the classification process more stable. The Sommeliers challenge this goal. They contend that the classification noise is, in fact, due not to incorrect classification but subtle differences in the wine, such as the bouquet, the body, or the legs, that are not subject to chemical analysis. These experts assert that the combination of attributes is assessed more finely by the senses than by any chemical analysis device.

Analyze the Wine.xls dataset using descriptive statistics and recommend a rule set that can be used by an automated chemical analysis device to classify the wine type.

Your report must address:

* A clear description of the problem
* A reliable model to chemically classify the wine
* Justification for your choice of the model that you used
* Evidence that your proposed model is accurate and reliable
* How your model will be implemented in a wine chemical analysis device
* Reconciliation of your model’s results with the Sommeliers’ concerns

**Write a COMPLETE Analysis Report using the course format describing your analysis and recommendations.**