Assignment #1 Predictive Analytics Basics by Joshua Troup

1. Give any one other term used for: (4 points)

(a) Input variable—**Alcohol (AKA-Independent variable or predictor variable)**

(b) Target variable—**Type (AKA-Outcome variable or output variable)**

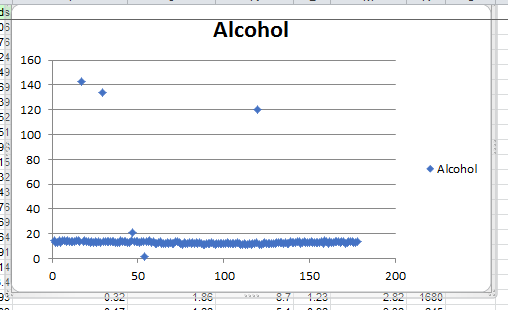
(c) Attribute—**Malic\_Acid (AKA-Cell contents or titles of columns)**

(d) Row.—**2.43 (Ash) (D2) (AKA-Row header horizontal)**

2. Find records that have missing values in the "wine" data sheet. Give the cell # where missing value exists. (2 points)

**I118**

3. Find possible "outliers" in the alcohol content variable. What are the possible causes of outliers (atleast 3 reasons). Write row numbers of outlier cases and mention against each row the possible cause by indicating the reasons 1,2,3 (4 points)



**Outliers include:**

**Point A-17,143 (B18)**

**Point B-29,133.87 (B30)**

**Point C-120,120 (B121)**

**Point D-47,21.38 (B48)**

**Point E-54,1.377 (B55)**

**The possible outcome of outliers include typing non numerical values, misplaced decimal point, or genuine case of outlier.**