| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 2.366517 | **Std Deviation** | 0.27434 |
| **Median** | 2.360000 | **Variance** | 0.07526 |
| **Mode** | 2.280000 | **Range** | 1.87000 |
|  |  | **Interquartile Range** | 0.35000 |

**1)**

**a)**

Descriptive Statistics for Alkalinity

|  |
| --- |
|  |

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 1.590899 | **Std Deviation** | 0.57236 |
| **Median** | 1.555000 | **Variance** | 0.32759 |
| **Mode** | 1.350000 | **Range** | 3.17000 |
|  |  | **Interquartile Range** | 0.70000 |

Descriptive Statistics for Color

Descriptive Statistics for Hue

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 5.058090 | **Std Deviation** | 2.31829 |
| **Median** | 4.690000 | **Variance** | 5.37445 |
| **Mode** | 2.600000 | **Range** | 11.72000 |
|  |  | **Interquartile Range** | 2.99000 |

Descriptive Statistics for Magnesium

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 19.49494 | **Std Deviation** | 3.33956 |
| **Median** | 19.50000 | **Variance** | 11.15269 |
| **Mode** | 20.00000 | **Range** | 19.40000 |
|  |  | **Interquartile Range** | 4.30000 |

Typical values for alkalinity are between 2 and 2.7 and have the smallest variance of the four variables. Values for color are usually between .8 and 2.2. Normal hue values range between 2 and 8. Typical values of magnesium are between 15 and 24 and have the highest variance of the four variables.

**b)**

Descriptive Statistics of Alkalinity for Alcohol 1

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 2.455593 | **Std Deviation** | 0.22717 |
| **Median** | 2.440000 | **Variance** | 0.05160 |
| **Mode** | 2.360000 | **Range** | 1.18000 |
|  |  | **Interquartile Range** | 0.33000 |

Descriptive Statistics of Alkalinity for Alcohol 2

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 2.244789 | **Std Deviation** | 0.31547 |
| **Median** | 2.240000 | **Variance** | 0.09952 |
| **Mode** | 1.920000 | **Range** | 1.87000 |
|  |  | **Interquartile Range** | 0.42000 |

Descriptive Statistics of Alkalinity for Alcohol 3

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 2.437083 | **Std Deviation** | 0.18469 |
| **Median** | 2.380000 | **Variance** | 0.03411 |
| **Mode** | 2.480000 | **Range** | 0.76000 |
|  |  | **Interquartile Range** | 0.30500 |

On average, the alkalinity values for alcohol 2 are the lowest and have the highest variance. Overall, alcohol 1 has the highest alkalinity values by a slight margin over alcohol 3.

Descriptive Statistics of Color for Alcohol 1

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 1.899322 | **Std Deviation** | 0.41211 |
| **Median** | 1.870000 | **Variance** | 0.16983 |
| **Mode** | 1.660000 | **Range** | 1.71000 |
|  |  | **Interquartile Range** | 0.48000 |

Descriptive Statistics of Color for Alcohol 2

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 1.630282 | **Std Deviation** | 0.60207 |
| **Median** | 1.610000 | **Variance** | 0.36249 |
| **Mode** | 1.350000 | **Range** | 3.17000 |
|  |  | **Interquartile Range** | 0.55000 |

Descriptive Statistics of Color for Alcohol 3

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 1.153542 | **Std Deviation** | 0.40884 |
| **Median** | 1.105000 | **Variance** | 0.16715 |
| **Mode** | 0.940000 | **Range** | 2.15000 |
|  |  | **Interquartile Range** | 0.50000 |

Alcohol 1 has the lowest values of color by a wide margin. Alcohol 2 has the second highest values of color and also has the largest variance. On average, alcohol 3 has the largest values of color

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 5.528305 | **Std Deviation** | 1.23857 |
| **Median** | 5.400000 | **Variance** | 1.53406 |
| **Mode** | 5.100000 | **Range** | 5.38000 |
|  |  | **Interquartile Range** | 1.75000 |

Descriptive Statistics of Hue for Alcohol 1

Descriptive Statistics of Hue for Alcohol 2

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 3.086620 | **Std Deviation** | 0.92493 |
| **Median** | 2.900000 | **Variance** | 0.85549 |
| **Mode** | 2.600000 | **Range** | 4.72000 |
|  |  | **Interquartile Range** | 0.90000 |

Descriptive Statistics of Hue for Alcohol 3

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 7.396250 | **Std Deviation** | 2.31094 |
| **Median** | 7.550000 | **Variance** | 5.34045 |
| **Mode** | 5.000000 | **Range** | 9.15000 |
|  |  | **Interquartile Range** | 3.82500 |

Alcohol 3 has the largest values of hue by a wide margin and has the highest variance. Alcohol 2 has the lowest values of hue by a wide margin.

Descriptive Statistics of Magnesium for Alcohol 1

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 17.03729 | **Std Deviation** | 2.54632 |
| **Median** | 16.80000 | **Variance** | 6.48376 |
| **Mode** | 16.00000 | **Range** | 13.80000 |
|  |  | **Interquartile Range** | 2.80000 |

Descriptive Statistics of Magnesium for Alcohol 2

| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 20.23803 | **Std Deviation** | 3.34977 |
| **Median** | 20.00000 | **Variance** | 11.22096 |
| **Mode** | 18.00000 | **Range** | 19.40000 |
|  |  | **Interquartile Range** | 4.00000 |

Descriptive Statistics of Magnesium for Alcohol 3

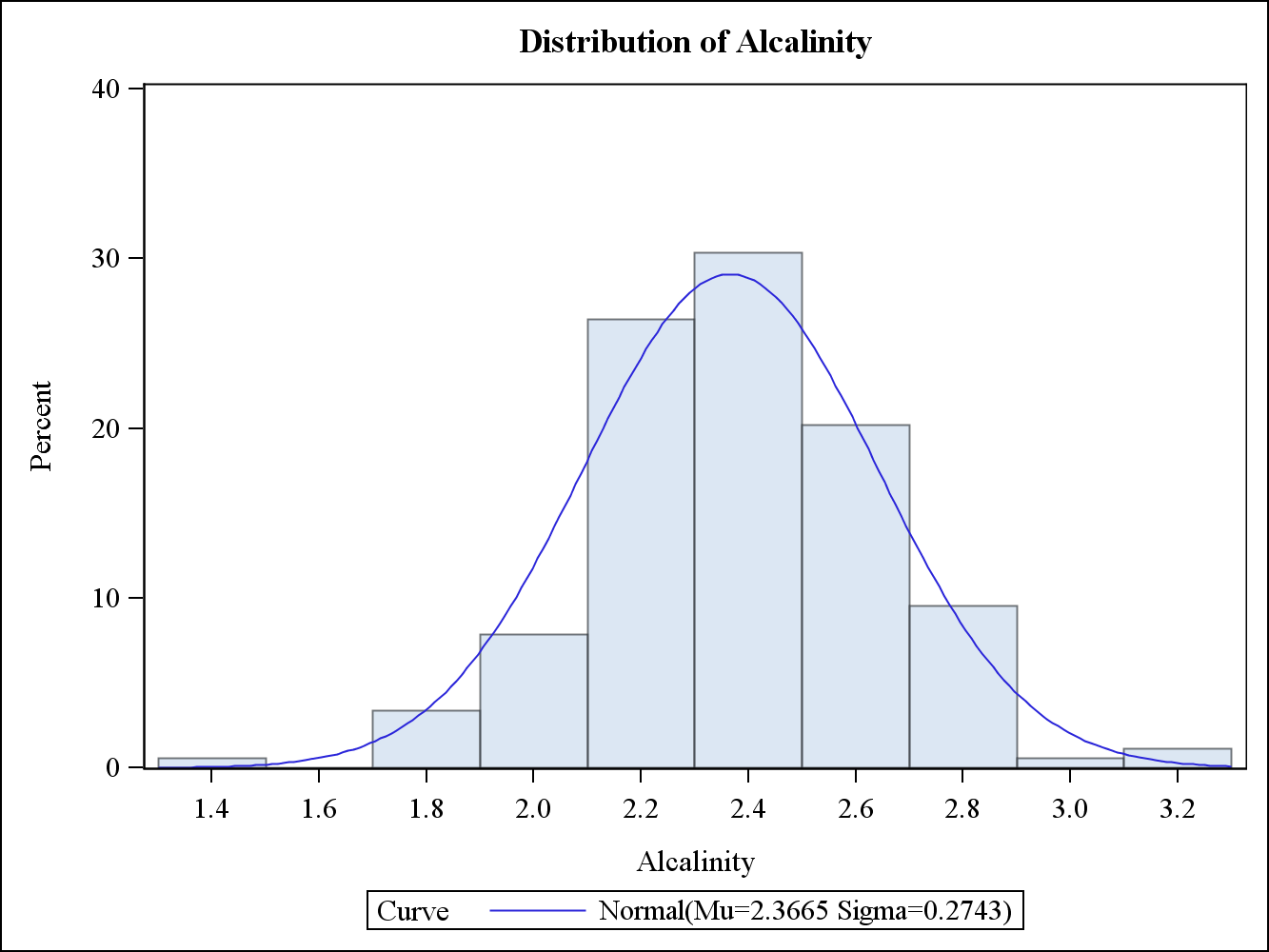
| **Basic Statistical Measures** | | | |
| --- | --- | --- | --- |
| **Location** | | **Variability** | |
| **Mean** | 21.41667 | **Std Deviation** | 2.25816 |
| **Median** | 21.00000 | **Variance** | 5.09929 |
| **Mode** | 20.00000 | **Range** | 9.50000 |
|  |  | **Interquartile Range** | 3.00000 |

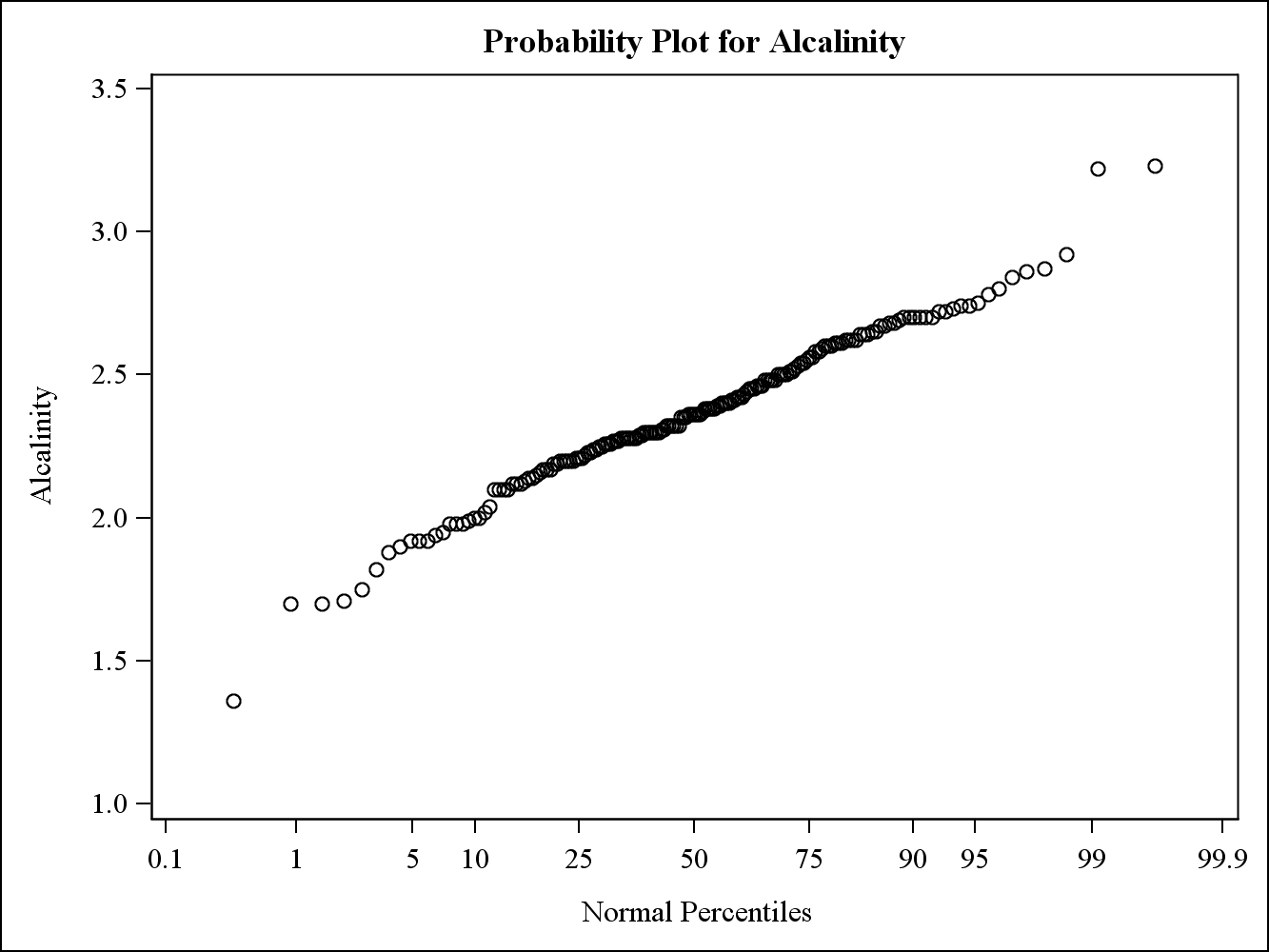
Alcohol 1 has the lowest values for magnesium, while alcohol 2 and alcohol 3 have very similar magnesium values. Although these magnesium values are similar, alcohol 2 has a much larger variance than alcohol 3.

**c)**

Normality Tests for Alkalinity

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.983947 | **Pr < W** | 0.0387 |
| **Kolmogorov-Smirnov** | **D** | 0.058454 | **Pr > D** | 0.1409 |
| **Cramer-von Mises** | **W-Sq** | 0.100029 | **Pr > W-Sq** | 0.1142 |
| **Anderson-Darling** | **A-Sq** | 0.678432 | **Pr > A-Sq** | 0.0792 |



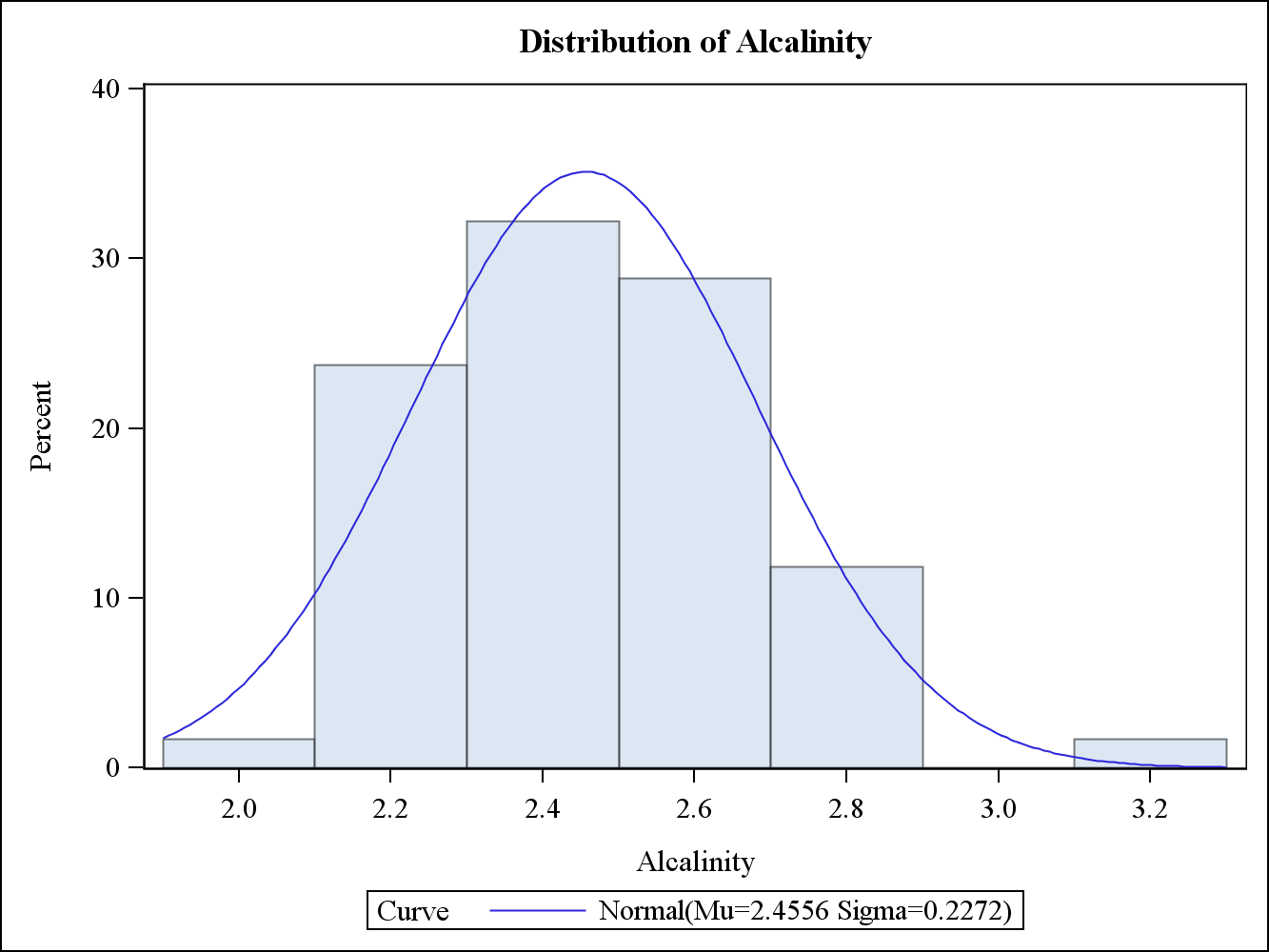


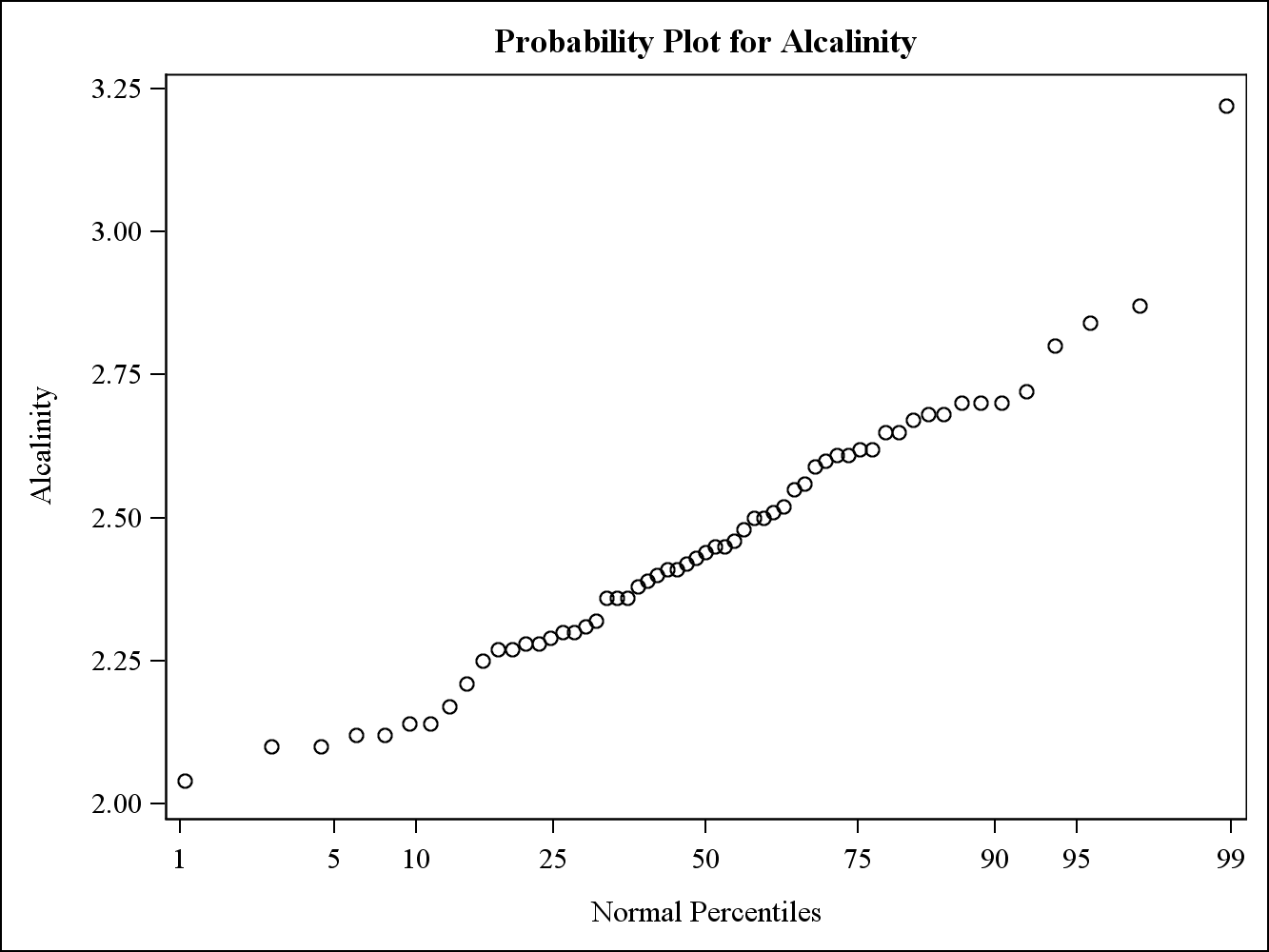
The normal histogram and probability plot both indicate the assumption of normality may be valid for alkalinity. The quantitative results testing for normality are mixed. The Shapiro-Wilk test indicates the assumption of normality is not valid at a 5% level of significance, while the other three normality tests do not reject the null hypothesis of normality. Overall, this indicates the assumption of normality for alkalinity values is reasonable, although we should proceed with caution when relying on this assumption to use statistical tests.

**d)**

Normality Tests for Alkalinity for Alcohol 1

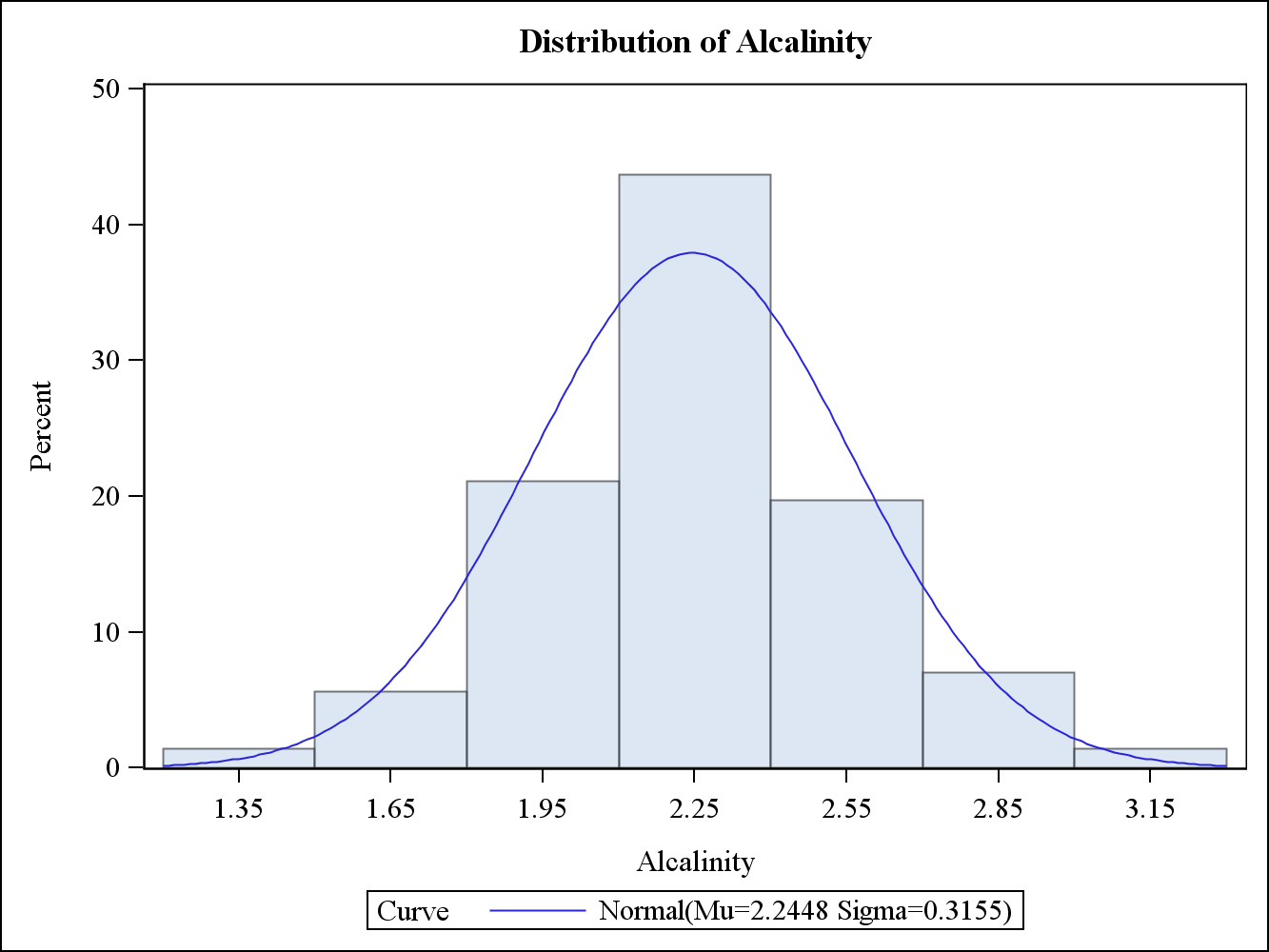
| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.970156 | **Pr < W** | 0.1556 |
| **Kolmogorov-Smirnov** | **D** | 0.056241 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.032657 | **Pr > W-Sq** | >0.2500 |
| **Anderson-Darling** | **A-Sq** | 0.290814 | **Pr > A-Sq** | >0.2500 |

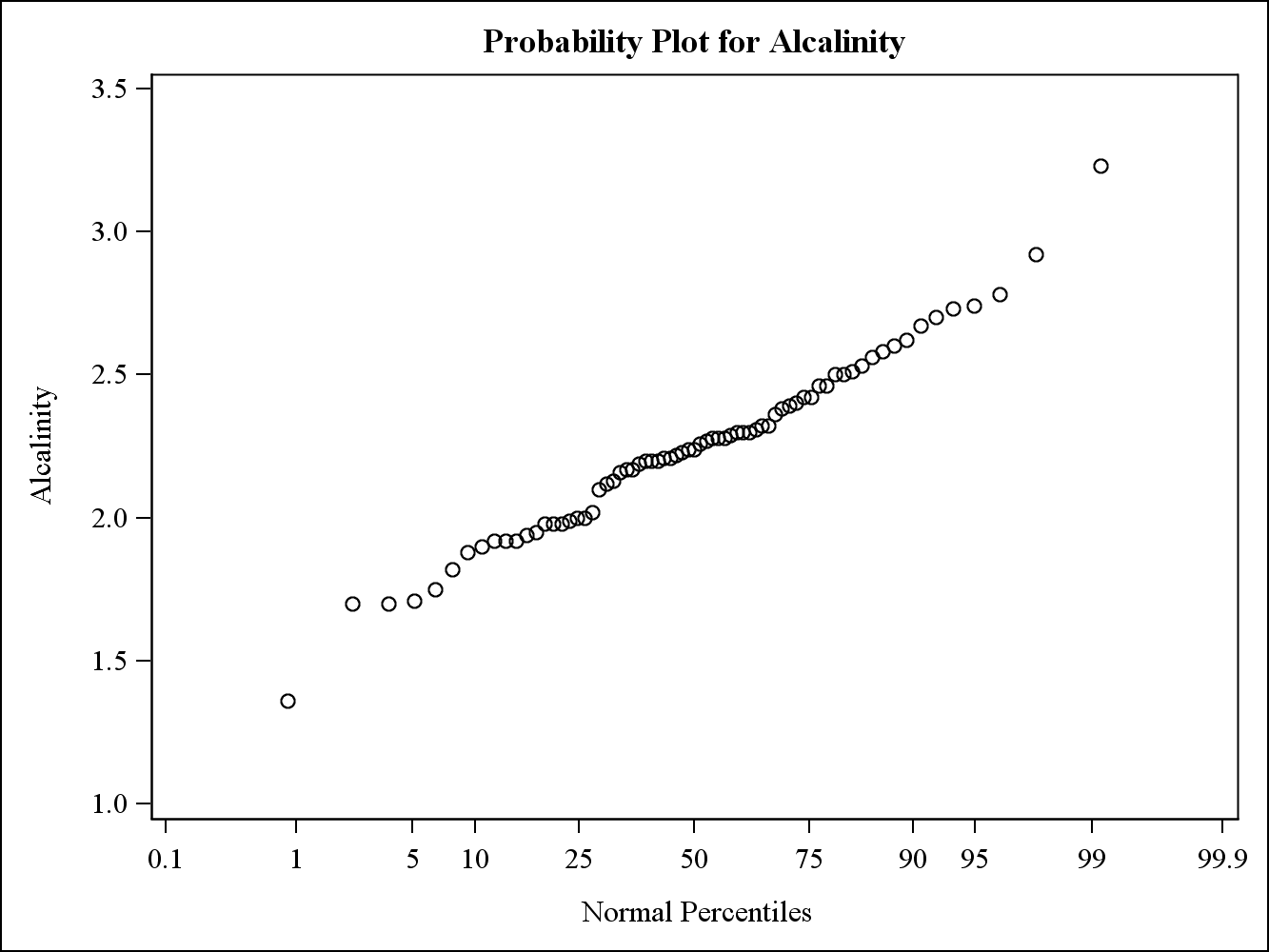




The histogram and probability plot both appear to be reasonably normal. Furthermore, the quantitative tests for normality do not reject the null hypothesis. Therefore, the assumption of normality for alkalinity values for alcohol 1 is reasonable.

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.986039 | **Pr < W** | 0.6198 |
| **Kolmogorov-Smirnov** | **D** | 0.081837 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.066376 | **Pr > W-Sq** | >0.2500 |
| **Anderson-Darling** | **A-Sq** | 0.364875 | **Pr > A-Sq** | >0.2500 |

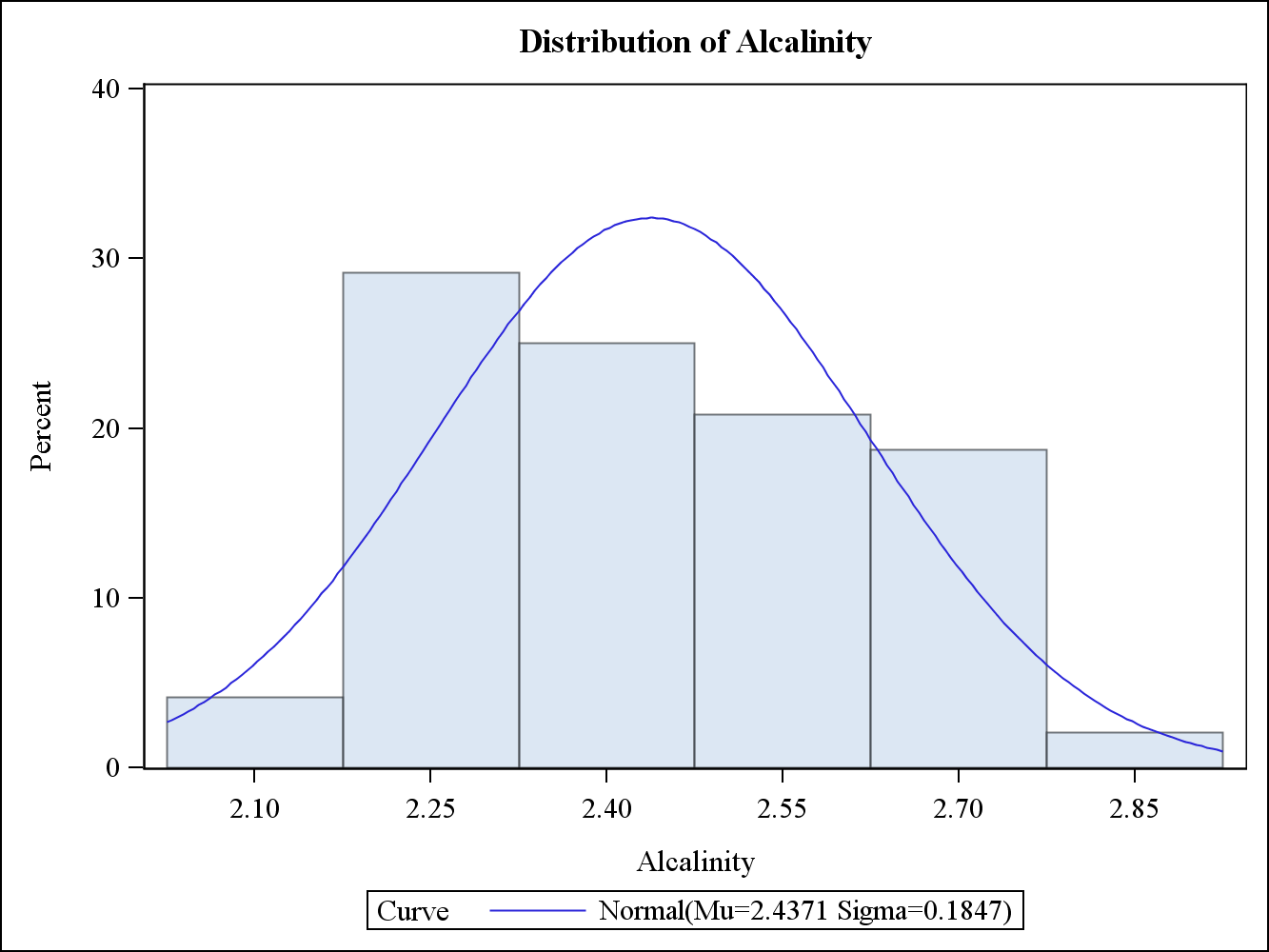
Normality Tests for Alkalinity for Alcohol 2

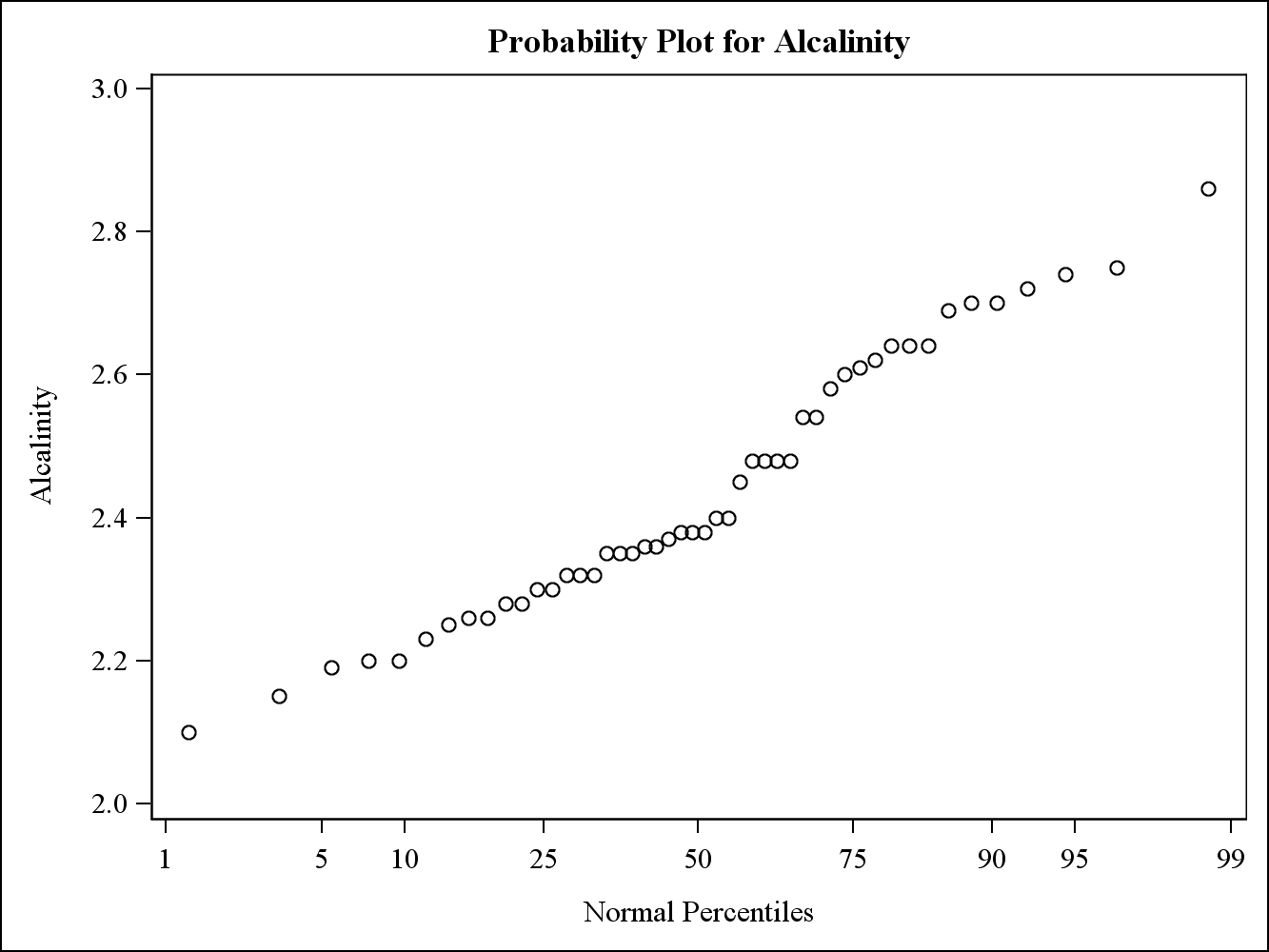


Based on the histogram and probability plot, the alkalinity values for alcohol 2 appear to be approximately normal. These visual tests are confirmed by the quantitative tests that indicate the assumption of normality is reasonable.

Normality Tests for Alkalinity for Alcohol 3

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.960846 | **Pr < W** | 0.1092 |
| **Kolmogorov-Smirnov** | **D** | 0.142202 | **Pr > D** | 0.0163 |
| **Cramer-von Mises** | **W-Sq** | 0.145257 | **Pr > W-Sq** | 0.0264 |
| **Anderson-Darling** | **A-Sq** | 0.775688 | **Pr > A-Sq** | 0.0424 |





Based on the histogram and probability plot, the alkalinity values for alcohol 3 do not appear to be normally distributed. Furthermore, based on the quantitative tests, three out of the four tests reject the null hypothesis of normality at a 5% level of significance. Therefore, the assumption of normality is not valid for alkalinity values for alcohol 3.

**2)**

**a)**

T-test for Alkalinity for Alcohol 2

| **DF** | **t Value** | **Pr > |t|** |
| --- | --- | --- |
| 70 | -3.08 | 0.0030 |

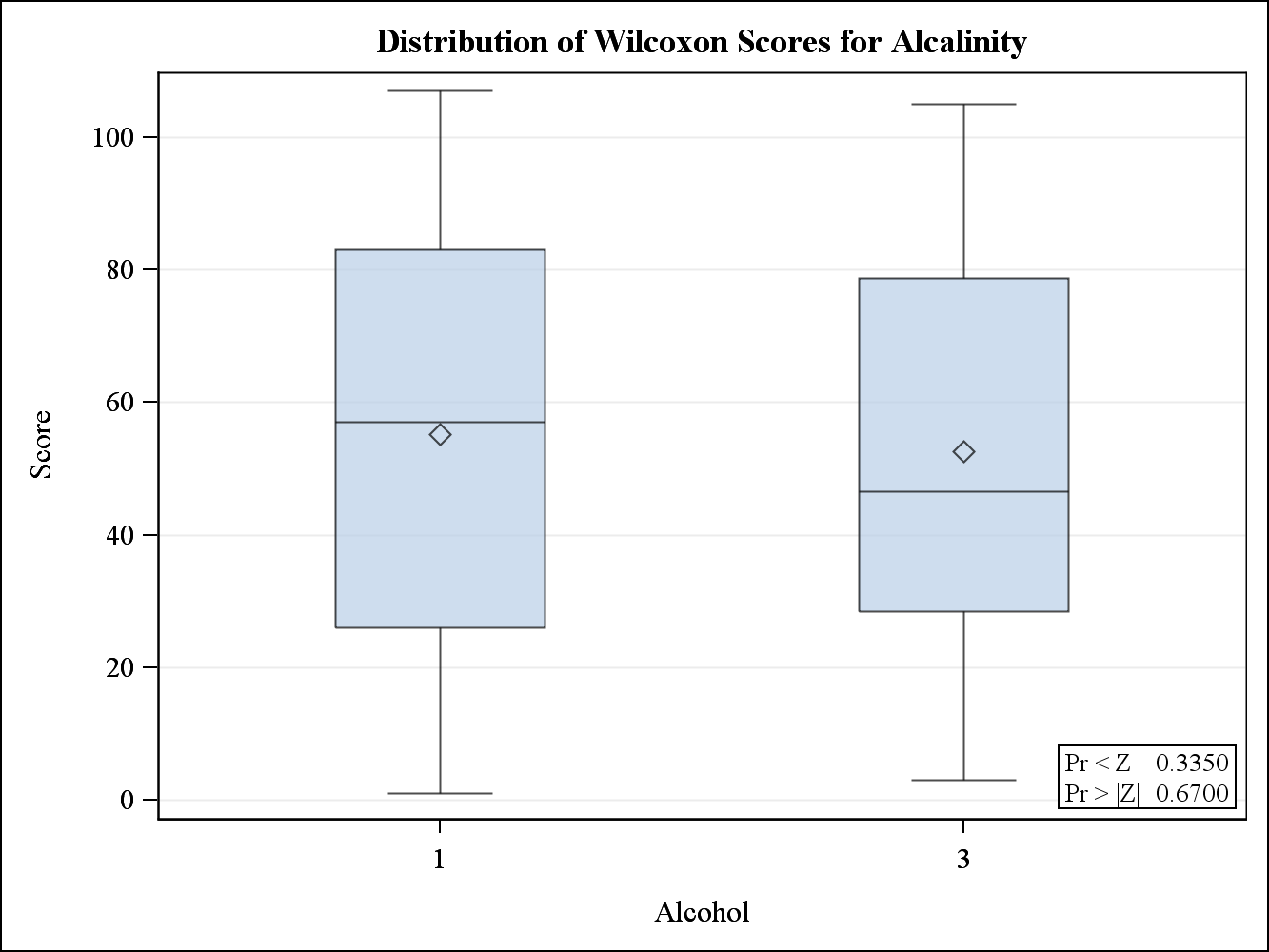
Because the assumption of normality was reasonable for alkalinity values for alcohol 2, the t-test can be used. The null hypothesis of this test is the mean alkalinity values for alcohol 2 are 2.36, the typical ash alkalinity value for wines in this region. The null hypothesis is strongly rejected at a 5% level of significance, indicating the mean alkalinity values for alcohol 2 are statistically significantly different from 2.36.

**b)**

Wilcoxon Ranked Sum Test for Alkalinity values of Alcohol 1 vs Alcohol 3

| **Wilcoxon Scores (Rank Sums) for Variable Alcalinity Classified by Variable Alcohol** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Alcohol** | **N** | **Sum of Scores** | **Expected Under H0** | **Std Dev Under H0** | **Mean Score** |
| **1** | 59 | 3254.50 | 3186.0 | 159.594476 | 55.161017 |
| **3** | 48 | 2523.50 | 2592.0 | 159.594476 | 52.572917 |
| **Average scores were used for ties.** | | | | | |

| **Wilcoxon Two-Sample Test** | |
| --- | --- |
| **Statistic** | 2523.5000 |
|  |  |
| **Normal Approximation** |  |
| **Z** | -0.4261 |
| **One-Sided Pr < Z** | 0.3350 |
| **Two-Sided Pr > |Z|** | 0.6700 |
|  |  |
| **t Approximation** |  |
| **One-Sided Pr < Z** | 0.3355 |
| **Two-Sided Pr > |Z|** | 0.6709 |
| **Z includes a continuity correction of 0.5.** | |



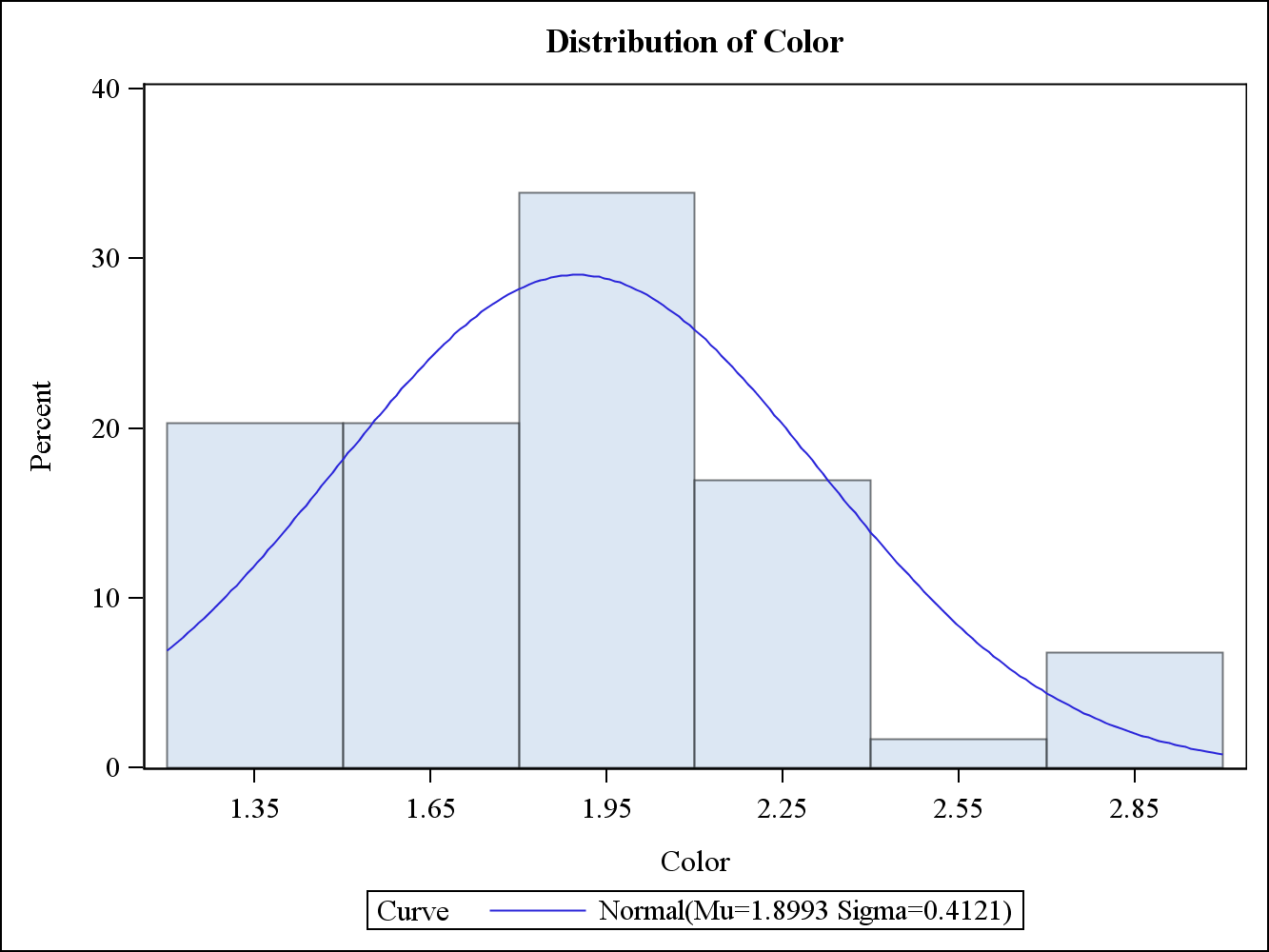
Since the assumption of normality was not reasonable for alkalinity values for alcohol 3, the t-test is not appropriate to use and the Wilcoxon ranked sum test is used instead. The null hypothesis of this test is the alkalinity values for alcohol 1 and alcohol 3 have stochastically equal distributions. Based on the results, the null hypothesis is not rejected at a 5% level, indicating the assumption alcohol 1 and alcohol 3 have similar alkalinity values is appropriate.

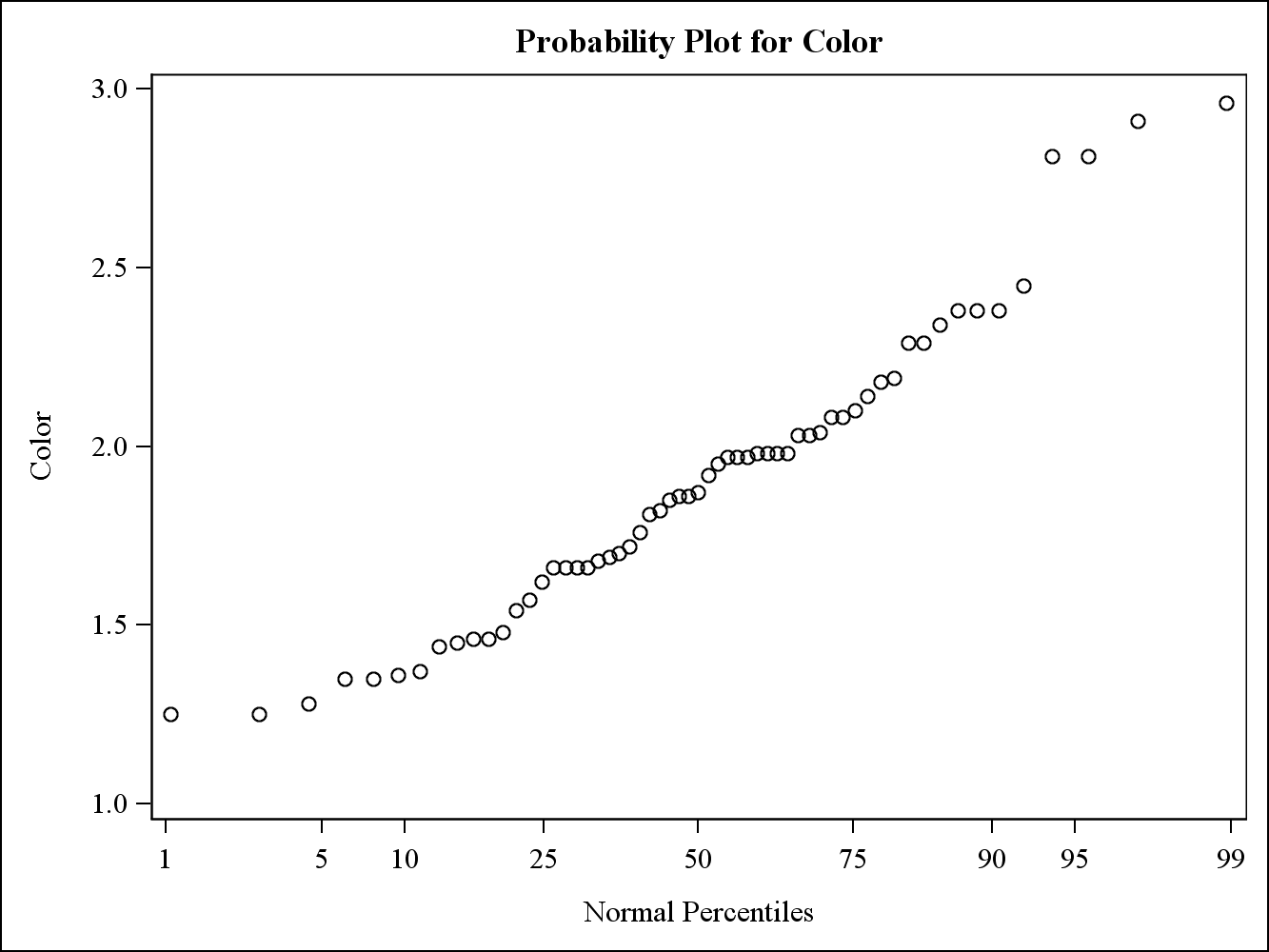
**3)**

**a)**

Normality Tests for Color for Alcohol 1

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.955786 | **Pr < W** | 0.0315 |
| **Kolmogorov-Smirnov** | **D** | 0.083413 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.067573 | **Pr > W-Sq** | >0.2500 |
| **Anderson-Darling** | **A-Sq** | 0.575064 | **Pr > A-Sq** | 0.1345 |

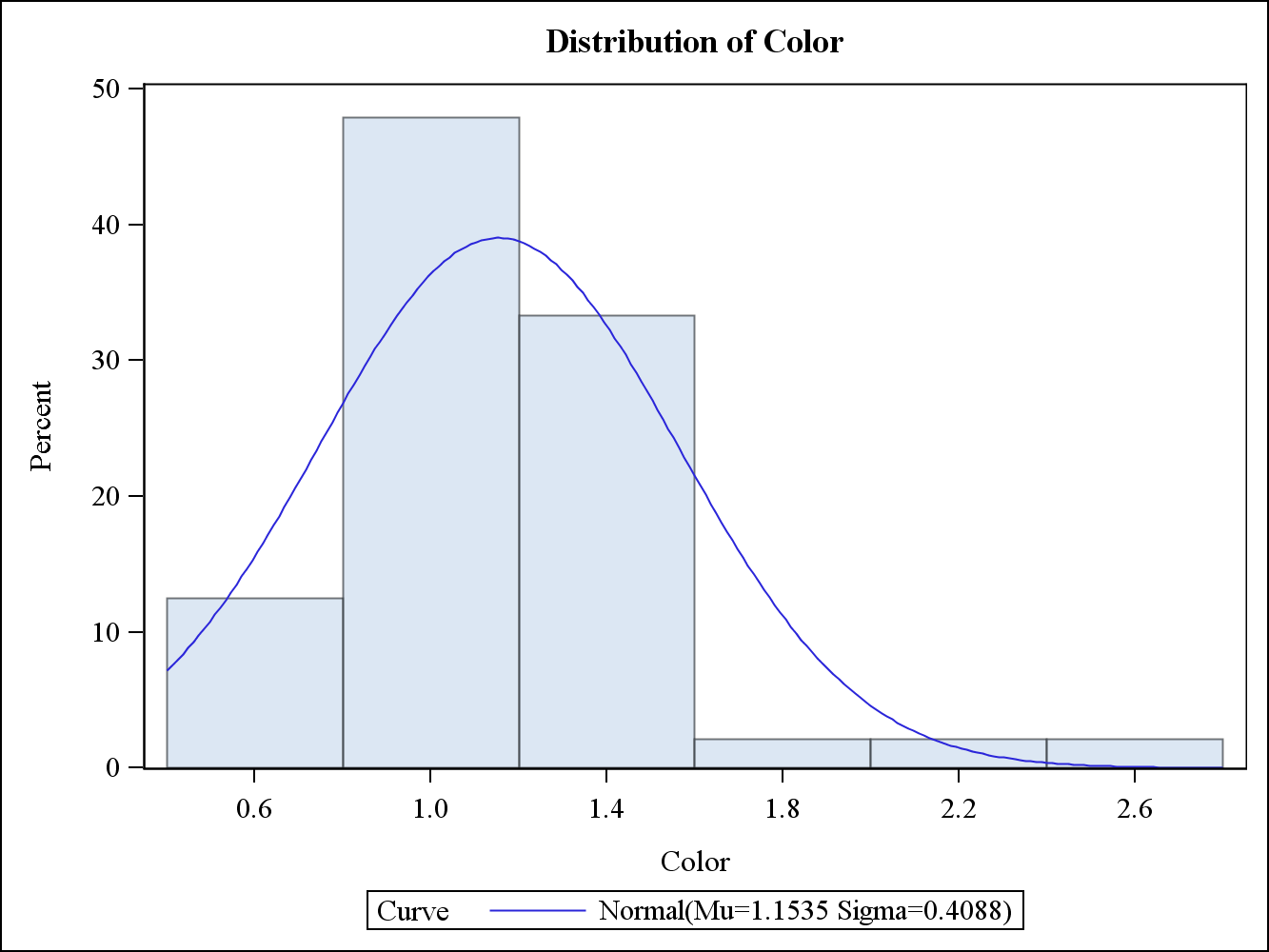


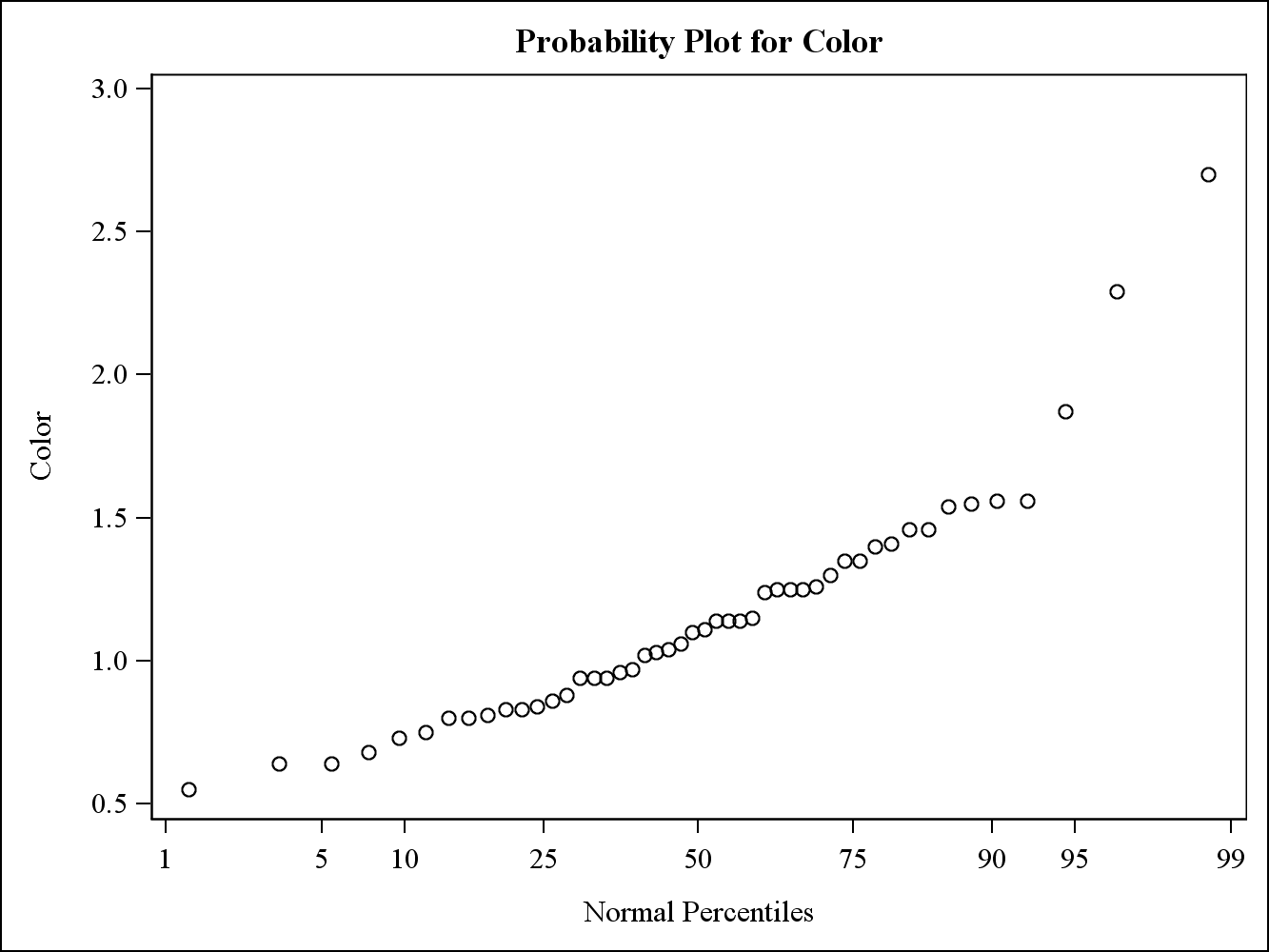


Based on the results from the histogram and probability plot, it is difficult to discern whether the assumption of normality is appropriate. The quantitative results all fail to reject the null hypothesis at a 5% level of significance except for the Shapiro-Wilk test. Overall, these results indicate the assumption of normality for color values for alcohol 1 is not too unreasonable.

Normality Tests for Color for Alcohol 3

| **Tests for Normality** | | | | |
| --- | --- | --- | --- | --- |
| **Test** | **Statistic** | | **p Value** | |
| **Shapiro-Wilk** | **W** | 0.887202 | **Pr < W** | 0.0002 |
| **Kolmogorov-Smirnov** | **D** | 0.107623 | **Pr > D** | >0.1500 |
| **Cramer-von Mises** | **W-Sq** | 0.144202 | **Pr > W-Sq** | 0.0276 |
| **Anderson-Darling** | **A-Sq** | 1.09874 | **Pr > A-Sq** | 0.0067 |





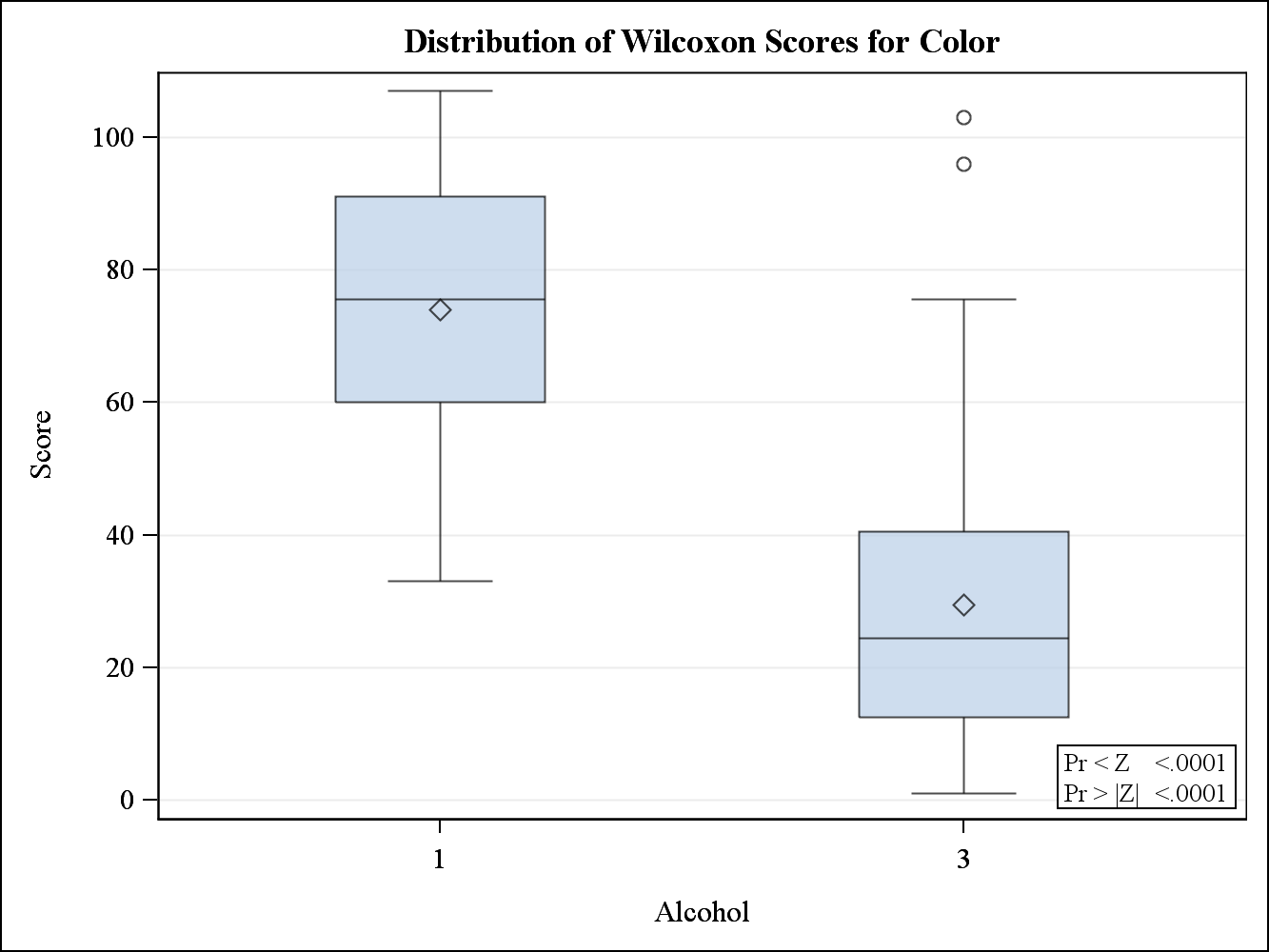
Based on the histogram and probability plot, it appears the alkalinity values for alcohol 3 are not normally distributed. Furthermore, every quantitative test except for the Kolmogorov-Smirnoff test rejects the null hypothesis of normality at a 5% level of significance. Therefore, the assumption of normality for alkalinity values for alcohol 3 is not reasonable.

**b)**

Wilcoxon Ranked Sum Test for Color values of Alcohol 1 vs Alcohol 3

| **Wilcoxon Scores (Rank Sums) for Variable Color Classified by Variable Alcohol** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Alcohol** | **N** | **Sum of Scores** | **Expected Under H0** | **Std Dev Under H0** | **Mean Score** |
| **1** | 59 | 4367.0 | 3186.0 | 159.614423 | 74.016949 |
| **3** | 48 | 1411.0 | 2592.0 | 159.614423 | 29.395833 |
| **Average scores were used for ties.** | | | | | |

| **Wilcoxon Two-Sample Test** | |
| --- | --- |
| **Statistic** | 1411.0000 |
|  |  |
| **Normal Approximation** |  |
| **Z** | -7.3959 |
| **One-Sided Pr < Z** | <.0001 |
| **Two-Sided Pr > |Z|** | <.0001 |
|  |  |
| **t Approximation** |  |
| **One-Sided Pr < Z** | <.0001 |
| **Two-Sided Pr > |Z|** | <.0001 |
| **Z includes a continuity correction of 0.5.** | |



Since the assumption of normality was not valid for color values for alcohol 3, the t-test is not appropriate to use and the Wilcoxon ranked sum test should be used instead. The null hypothesis of the test is the color values for alcohol 1 and 3 have stochastically equal distributions. This test strongly rejects the null hypothesis at a 5% level, indicating the assumption of stochastically equal distributions is not reasonable. Furthermore, it’s clear from the results that color values for alcohol 1 tend to be greater than color values for alcohol 3.

**4)**

**a)**

Pearson and Spearman correlation between Alkalinity, Color, Hue and Magnesium

| **Pearson Correlation Coefficients, N = 178 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | 0.00965 0.8983 | 0.25889 0.0005 | 0.44337 <.0001 |
| **Color** | 0.00965 0.8983 | 1.00000 | -0.02525 0.7380 | -0.19733 0.0083 |
| **Hue** | 0.25889 0.0005 | -0.02525 0.7380 | 1.00000 | 0.01873 0.8040 |
| **Magnesium** | 0.44337 <.0001 | -0.19733 0.0083 | 0.01873 0.8040 | 1.00000 |

| **Spearman Correlation Coefficients, N = 178 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | 0.02438 0.7466 | 0.28305 0.0001 | 0.36637 <.0001 |
| **Color** | 0.02438 0.7466 | 1.00000 | -0.03095 0.6818 | -0.25370 0.0006 |
| **Hue** | 0.28305 0.0001 | -0.03095 0.6818 | 1.00000 | -0.07378 0.3277 |
| **Magnesium** | 0.36637 <.0001 | -0.25370 0.0006 | -0.07378 0.3277 | 1.00000 |

Alkalinity and hue both have a statistically significant positive correlation at a 5% level, meaning increases in Hue correspond to an increase in Alkalinity. Alkalinity and magnesium also have a statistically significant positive correlation. Color and magnesium have a statistically significant negative correlation, meaning as values of color increase values of magnesium decrease and vice versa.

**b)**

Pearson and Spearman correlation between Alkalinity, Color, Hue and Magnesium for Alcohol 1

| **Pearson Correlation Coefficients, N = 59 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | -0.14547 0.2716 | -0.12422 0.3486 | 0.54933 <.0001 |
| **Color** | -0.14547 0.2716 | 1.00000 | 0.42470 0.0008 | -0.17363 0.1885 |
| **Hue** | -0.12422 0.3486 | 0.42470 0.0008 | 1.00000 | -0.21095 0.1088 |
| **Magnesium** | 0.54933 <.0001 | -0.17363 0.1885 | -0.21095 0.1088 | 1.00000 |

| **Spearman Correlation Coefficients, N = 59 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | -0.18306 0.1652 | -0.04686 0.7245 | 0.45607 0.0003 |
| **Color** | -0.18306 0.1652 | 1.00000 | 0.47933 0.0001 | -0.14022 0.2895 |
| **Hue** | -0.04686 0.7245 | 0.47933 0.0001 | 1.00000 | -0.17478 0.1855 |
| **Magnesium** | 0.45607 0.0003 | -0.14022 0.2895 | -0.17478 0.1855 | 1.00000 |

Similar to the results from part a, alkalinity and magnesium have a statistically significant positive correlation. However, there is no statistically significant correlation between alkalinity and hue and between color and magnesium as in part a. Additionally, there is a very strong positive correlation between color and hue for alcohol 1 that is not present in the overall data.

Pearson and Spearman correlation between Alkalinity, Color, Hue and Magnesium for Alcohol 2

| **Pearson Correlation Coefficients, N = 71 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | 0.04296 0.7221 | 0.06025 0.6177 | 0.69526 <.0001 |
| **Color** | 0.04296 0.7221 | 1.00000 | -0.07376 0.5410 | 0.10884 0.3663 |
| **Hue** | 0.06025 0.6177 | -0.07376 0.5410 | 1.00000 | -0.08586 0.4765 |
| **Magnesium** | 0.69526 <.0001 | 0.10884 0.3663 | -0.08586 0.4765 | 1.00000 |

| **Spearman Correlation Coefficients, N = 71 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | 0.04338 0.7194 | -0.01031 0.9320 | 0.61752 <.0001 |
| **Color** | 0.04338 0.7194 | 1.00000 | -0.04360 0.7181 | 0.07740 0.5211 |
| **Hue** | -0.01031 0.9320 | -0.04360 0.7181 | 1.00000 | -0.21612 0.0703 |
| **Magnesium** | 0.61752 <.0001 | 0.07740 0.5211 | -0.21612 0.0703 | 1.00000 |

Similar to the results from part a, alkalinity and magnesium have a statistically significant positive correlation. There are no other statistically significant correlations between the variables for alcohol 2.

Pearson and Spearman correlation between Alkalinity, Color, Hue and Magnesium for Alcohol 3

| **Pearson Correlation Coefficients, N = 48 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | 0.19383 0.1868 | 0.12515 0.3967 | 0.75852 <.0001 |
| **Color** | 0.19383 0.1868 | 1.00000 | 0.68491 <.0001 | 0.26340 0.0705 |
| **Hue** | 0.12515 0.3967 | 0.68491 <.0001 | 1.00000 | 0.16062 0.2755 |
| **Magnesium** | 0.75852 <.0001 | 0.26340 0.0705 | 0.16062 0.2755 | 1.00000 |

| **Spearman Correlation Coefficients, N = 48 Prob > |r| under H0: Rho=0** | | | | |
| --- | --- | --- | --- | --- |
|  | **Alcalinity** | **Color** | **Hue** | **Magnesium** |
| **Alcalinity** | 1.00000 | 0.10648 0.4713 | 0.08644 0.5591 | 0.74458 <.0001 |
| **Color** | 0.10648 0.4713 | 1.00000 | 0.61218 <.0001 | 0.25821 0.0764 |
| **Hue** | 0.08644 0.5591 | 0.61218 <.0001 | 1.00000 | 0.12577 0.3943 |
| **Magnesium** | 0.74458 <.0001 | 0.25821 0.0764 | 0.12577 0.3943 | 1.00000 |

Similar to results from part a and every other alcohol, there is a statistically significant positive correlation between alkalinity and magnesium. Additionally, there is a statistically significant positive correlation between color and hue that also occurred in alcohol 1, but not in the overall data.