Willis Allstead

STAT 352 – HW 1

9/8/17

1. Compute descriptive statistics for the dormant periods (all data)

Descriptive Statistics: Dormancy

Variable Eruption N N\* Mean SE Mean StDev Minimum Q1 Median

Dormancy Long 43 0 78.19 1.16 7.64 57.00 74.00 79.00

Short 17 0 53.47 1.53 6.32 42.00 50.50 53.00

Variable Eruption Q3 Maximum

Dormancy Long 84.00 93.00

Short 56.00 67.00

2. Find median, mean, and both quartiles for the lengths of the dormant periods.

Variable Mean Q1 Median Q3

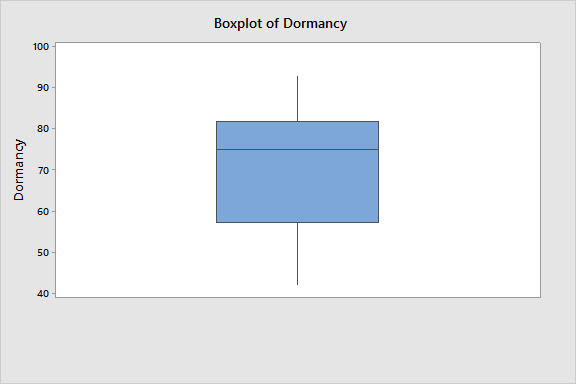
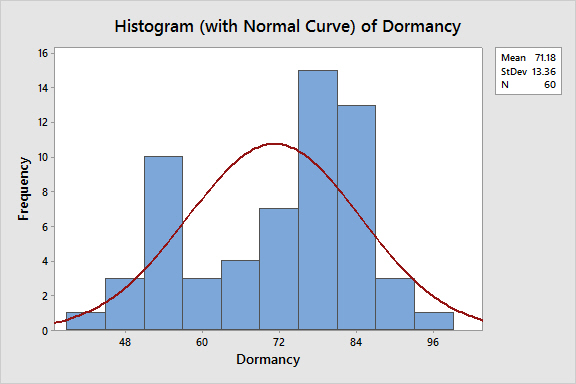
Dormancy 71.18 57.25 75.00 82.00 (in minutes)

3. What is the variance and standard deviation of the lengths of the dormant periods?

Variable StDev Variance

Dormancy 13.36 178.46 (in minutes)

4. Construct histogram and boxplot of the lengths of the dormant periods, and discuss the symmetry and possible outliers in this data.



The histogram is bimodal and not very symmetric. Neither the histogram or boxplot display any significant outliers.

5. Compute mean, median, both quartiles, variance and standard deviations of the lengths of the dormant periods separately for the long and short eruption. How do these statistics compare?

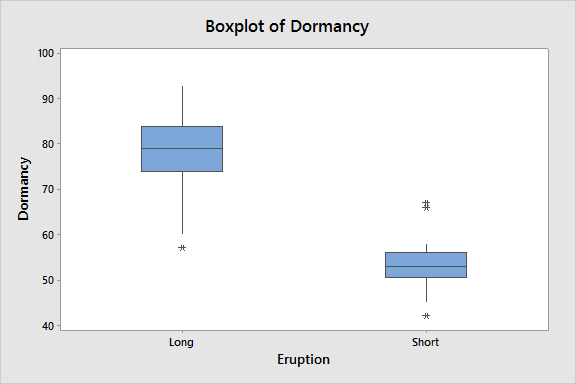
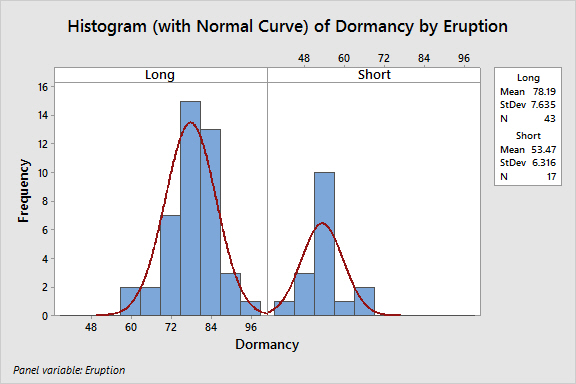
Variable Eruption Mean StDev Variance Q1 Median Q3

Dormancy Long 78.19 7.64 58.30 74.00 79.00 84.00

Short 53.47 6.32 39.89 50.50 53.00 56.00

The long eruptions display higher values across the board compared to the shorter ones. The standard deviation for the longer eruptions is larger by about 1.3 minutes. The mean Is higher by about 25 minutes for the longer eruptions. All quartiles are higher by about 24-26 minutes for the longer eruptions compared to the shorter eruptions. The variance for the longer eruptions was about 20 minutes higher than that of the shorter eruptions.

6. Construct histograms and side by side boxplots of the lengths of the dormant periods for long and short eruptions. Do you see any differences in the distribution of these data?



Both histograms are fairly symmetric, with the histogram of shorter eruptions having a slightly smaller spread than that of the larger eruptions histogram. There are outliers displayed in both boxplots. There are low outliers for both long and short eruptions but also high outliers for the short eruptions.