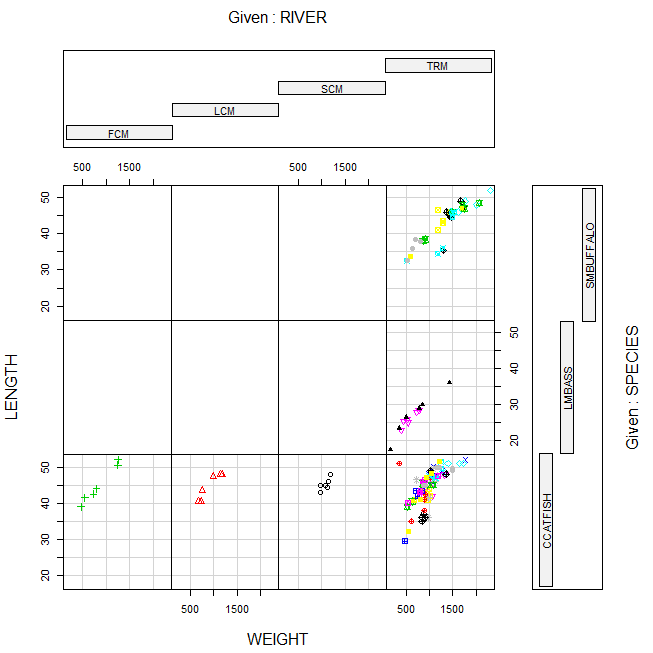
Clayton Glenn

Assignment #1

2/9/2018

**Answered 13/15 problems fully**

1. There are 4 assignments that equal 15% of my grade and all work must be shown to receive full credit. There are 16 Labs in the class that equal 10% of my grade. I cannot deposit my lab into the drop box after time is up. The 1 project I have is 10% of my total grade. The project is over Simple Linear Regression and needs to be submitted in the outlined format provided on canvas. Clickers are done in class and worth 10% of my total grade. Missing a couple class will most likely not change my grade. Chapter quizzes are online and equal 5% of my total grade. They are usually 10 questions and are graded automatically. I have 2 midterm exams worth a total of 20% of my grade. The midterms should not overlap in content, but the final exam will. My final in this class is worth 30% of my grade and is cumulative. The breakup of the final exam will be about 1/3 Exams 1 and 2, and 2/3 from chapters 8 and 10. The grading scale in this class is as follows: A(90-100), B(80-89), C(60-79), D(50-59), F(0-49) without the possibility of a curve of the total grade, so what you earn is what you get.



* 1. This coplot shows the broken-up data of each river and species. The lower far left plot shows the number and size of catfish in the FCM River. The Lower Left Mid plot shows the number and size of catfish in the LCM River, and the Lower Third from the Left plot shows the number and size of catfish that are in the SCM River.
  2. Line A shows the Each independent mile number in an array of numbers.
  3. Line B shows the Length of the array of independent mile numbers.
  4. The top six plots are empty due to the fact that FCM, LCM, and SCM Rivers do not contain SMBUFFALO or LMBASS, so no data is shown for the coplots.
  5. The mean value of DDT in CCATFISH Caught in the FCM river is 45.

1. 1. Length of Maximum Span(Feet) = Quantitative
   2. Number of Vehicle Lanes = Quantitative
   3. Toll Bridge = Qualitative
   4. Average Daily Traffic = Quantitative
   5. Condition of Deck(good, fair, or poor) = Qualitative
   6. Bypass or Detour Length(Miles) = Quantitative
   7. Route Type(interstate, U.S., state country, or city) = Qualitative
2. 1. Simple Random Sampling(Simple)
      1. Randomly chose units out of a population.
   2. Stratified Random Sampling(Complex)
      1. Sampling used when units can be separated into strata or groups by characteristics.
   3. Cluster Sampling(Complex)
      1. Sampling used to break down large samples into clusters and then compare them.
   4. Systematic Sampling(Complex)
      1. Sampling used by selecting every Kth element in a population for a random sample.

pH SpConduct DissOxy RoadsPct IndPct UrbanPct DevPct WellClass Aquifier Depth SafeYld Distance MTBE.Detect MTBE.Level HouseDen PopDen

2 8.63 225.9 0.84 0.72 0 10.43 10.43 Private Bedrock 36.5760 NA 3667.692 Below Limit 0.20 47.69 24.52

210 8.32 547.1 0.35 1.35 0 16.03 16.03 Public Bedrock 61.8744 NA 3909.5721 Below Limit 0.20 61.54 20.78

6 8.36 198.2 0.18 1.48 0 22.49 29.19 Private Bedrock 115.8240 NA 2396.746 Below Limit 0.20 206.73 59.47

95 8.24 694.3 3.27 3.50 0 49.35 49.35 Private Bedrock 91.4400 NA 1711.4695 Detect 0.39 174.77 85.67

101 7.61 608.1 0.61 2.05 0 38.58 38.58 Private Bedrock 121.9200 NA 2685.91101 Below Limit 0.20 62.87 29.94

* 1. The standard deviation of Depth of Bedrock Wells is 56.45357.



YEAR MONTH DAY HOUR MINUTE MAGNITUDE

2129 1994 1 28 14 0 2.2

2772 1994 2 4 18 33 1.9

2683 1994 2 3 20 52 2.0

297 1994 1 18 15 51 4.0

1036 1994 1 21 23 47 1.8

1234 1994 1 22 22 37 2.0

1629 1994 1 25 6 55 1.4

1085 1994 1 22 5 46 1.8

2442 1994 2 1 14 28 1.8

2487 1994 2 2 3 34 1.5

649 1994 1 20 8 37 1.8

1290 1994 1 23 6 14 1.6

112 1994 1 17 17 35 2.9

690 1994 1 20 12 18 2.2

1012 1994 1 21 21 15 2.0

2807 1994 2 5 1 54 1.8

1201 1994 1 22 19 45 2.3

1529 1994 1 24 14 48 2.2

300 1994 1 18 16 1 2.3

2224 1994 1 29 5 38 2.1

1757 1994 1 25 23 27 1.8

1032 1994 1 21 23 2 2.5

447 1994 1 19 12 38 3.5

1110 1994 1 22 9 0 1.8

2319 1994 1 30 13 22 2.6

1072 1994 1 22 3 55 2.7

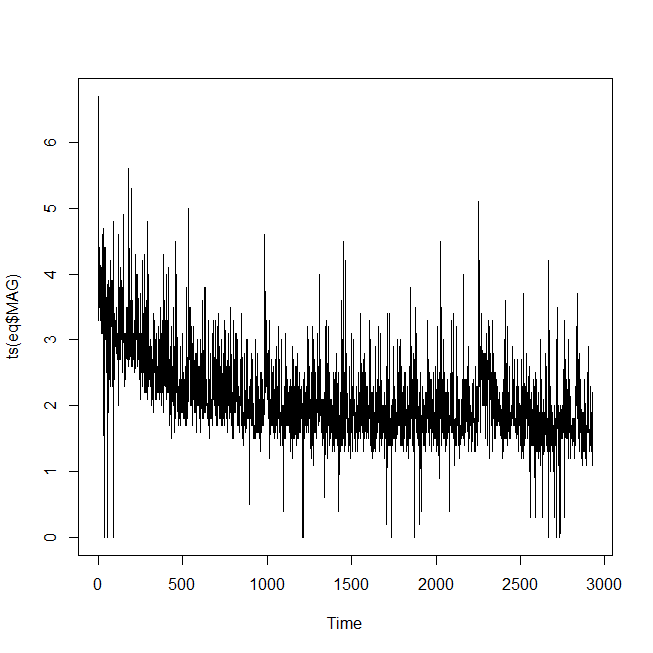
725 1994 1 20 16 13 2.5

1870 1994 1 26 16 9 1.3

2161 1994 1 28 19 50 2.1

899 1994 1 21 10 13 1.9

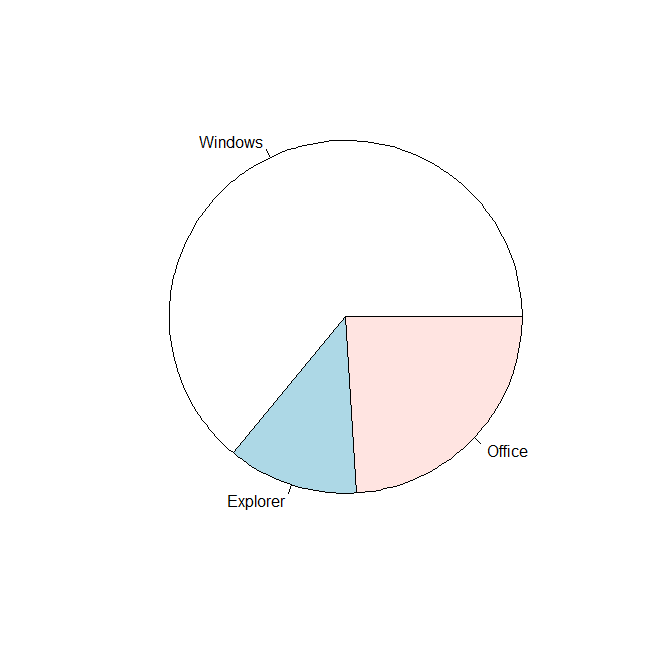




* 1. The median of the whole Earthquake data file based on magnitude is 2.
  2. The scientists used a stratified sample.
  3. The Population was all fish in the Tennessee River and its tributaries.
  4. The qualitative variables in DDT file is River and Species.

1. 1. A bar graph describes the data.
   2. Number of Robots
   3. According to the graph, the most used robot design is Legs Only.
   4. The Class Relative Frequencies are r/N
      1. None Frequency is 0.1415094.
      2. Both Frequency is 0.0754717.
      3. Legs ONLY Frequency is 0.5943396.
      4. Wheels ONLY Frequency is 0.1886792.
   5. Not Done

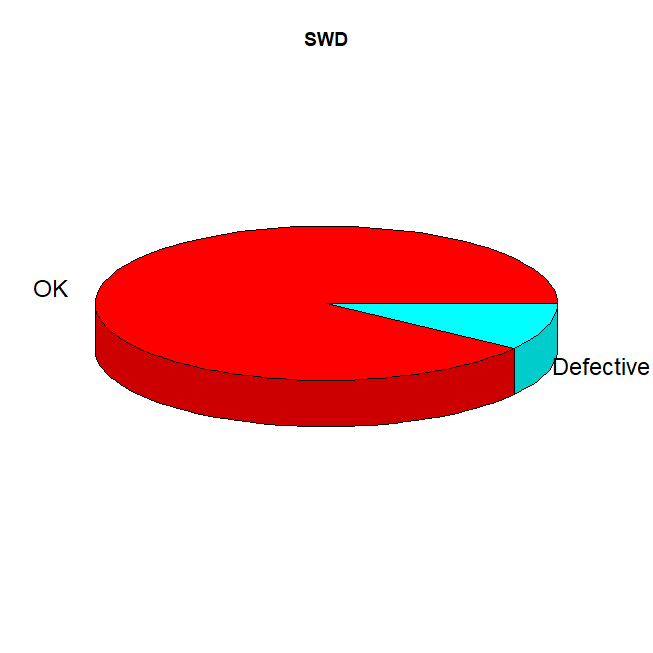




* 1. Based on the pie chart, Explorer has the lowest proportion of security issues.

**PARETO – Not Done**

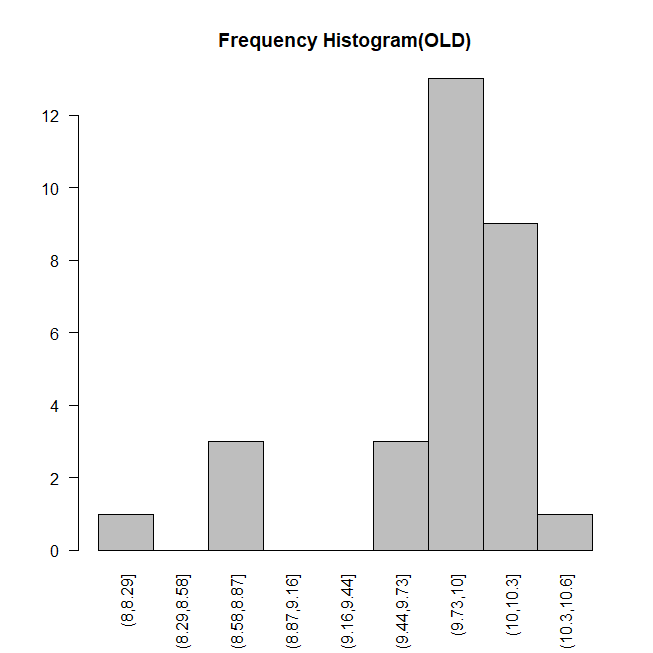
* 1. Based on the pareto graph, Windows should be the most focused on by Microsoft.



The likelihood of software code being defective is 10%. The probability of OK software is 10:1.



|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Class | Class Interval | Data Tabulation | Frequency | Relative Frequency |
| 1  2  3  4  5  6  7  8  9 | 8.0000-8.2889  8.2890-8.5779  8.5780-8.8669  8.8670-9.1559  9.1560-9.4449  9.4450-9.7339  9.7340-10.0229  10.0230-10.3119  10.3120-10.6000 | I  III  III  XIII  IX  I | 1  0  3  0  0  3  13  9  1 | 0.0333  0.0000  0.1000  0.0000  0.0000  0.1000  0.4333  0.3000  0.0333 |
| Total |  | XXX | 30 | 1 |



The Histogram more accurately shows the voltage reading because the classes are smaller in size and more precise.



The decimal point is at the |

8 | 1

8 | 778

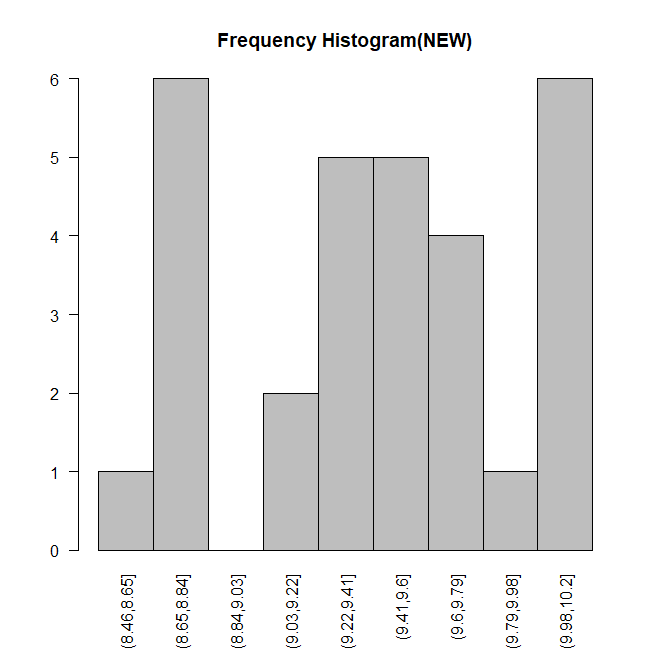
9 |

9 | 6778888999

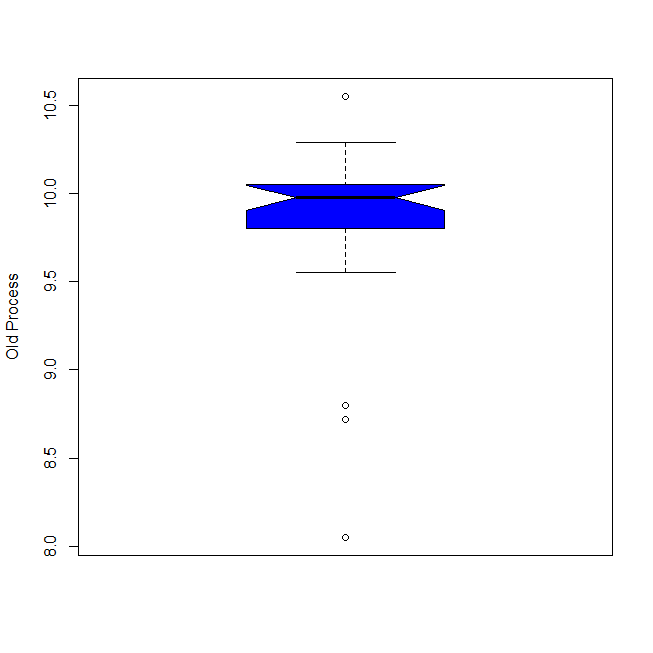
10 | 000000011122333

10 | 6





* 1. The Old Process is better than the new process due to less influx of voltage.
  2. We can use the median for the central tendency because the median is close enough to mean and is a better looking number.
     1. The mean, median, and mode of the old Process is 9.803667, 9.975, and 7 respectively
     2. The mean, median, and mode of the new Process is 9.422333, 9.455, and 8 respectively
  3. The Z Score for 10.5 at the old location is 1.287324.
  4. The Z score for 10.5 at the new location is 2.25041.
  5. Based on parts F and G, 10.5 Voltage will more likely occur at the old process. This is due to 10.5 being closer to mean+sd

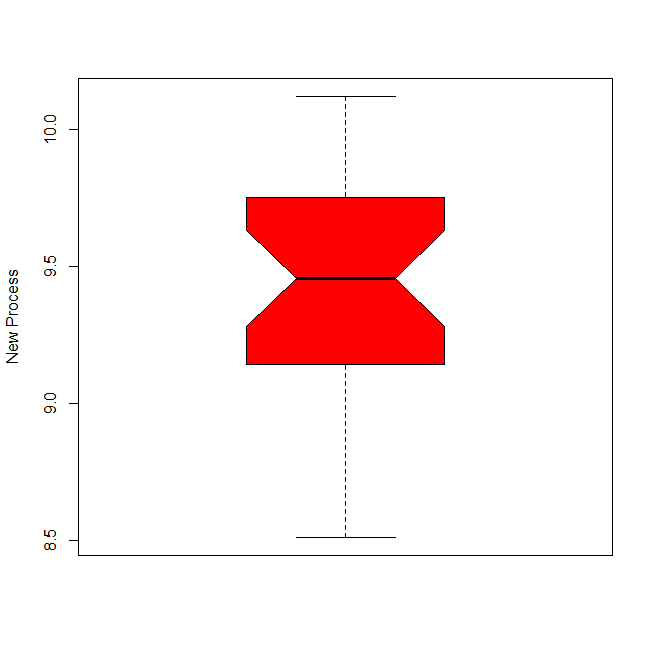


There are 4 outliers in the old process data.



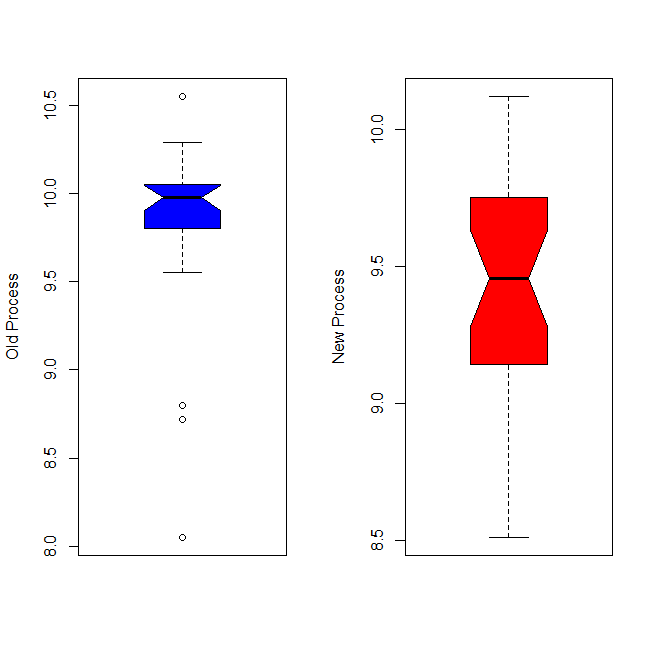
| **VOLTAGE**  <dbl> | **LOCATION**  <fctr> |  |  |
| --- | --- | --- | --- |
| 6 | 8.05 | OLD |  |  |
| 20 | 8.72 | OLD |  |  |
| 28 | 8.72 | OLD |  |  |





There are no outliers in the New Process.

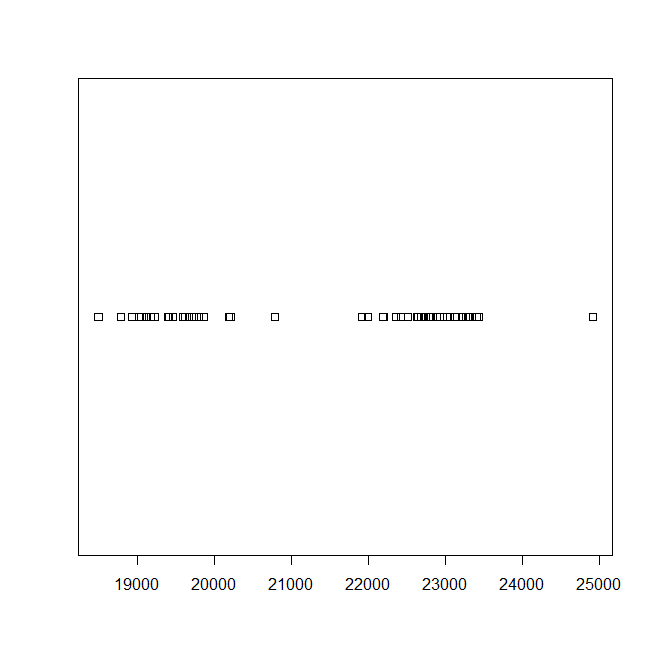
* 1. The Z-Score Method contains zero outliers.



1. The interval in which 95% of all data lies within can be 2 standard deviations. The lower bounds is 0.8331772, and the upper bounds is 2.928823.
2. 1. 1. Mean is 12.81818
         1. The Average Ant Species in all 11 sites.
      2. Median is 5
         1. The Amount of Ant Species at the Site that has the exact middle amount of Ant Species.
      3. Mode is 3
         1. The Most common number of Species at the sites.

* 1. The mean value best suits the data, due to the high volume of species at few sites.
  2. For Dry Steppe
     1. Mean is 40.4
     2. Median is 40
     3. Mode is 2
  3. For Gobi Desert
     1. Mean is 28
     2. Median is 26
     3. Mode is 4
  4. The ant species seems more abundant with less plant cover, so the Dry Steppe is more bountiful for Ants.





* 1. Yes, there are two clusters. One cluster is between 19000 and 20000, and the other cluster is between 22000 and 23000.
  2. 1. First Cluster
        1. Mean is 19462.24
        2. Standard Deviation is 532.2868
     2. Second Cluster
        1. Mean is 22838.47
        2. Standard Deviation is 560.9767
  3. The galaxy Velocity of 20000 would fit within A1775A because the velocity is much closer to A1775A's Mean + SD than A1775B's Mean - SD.

