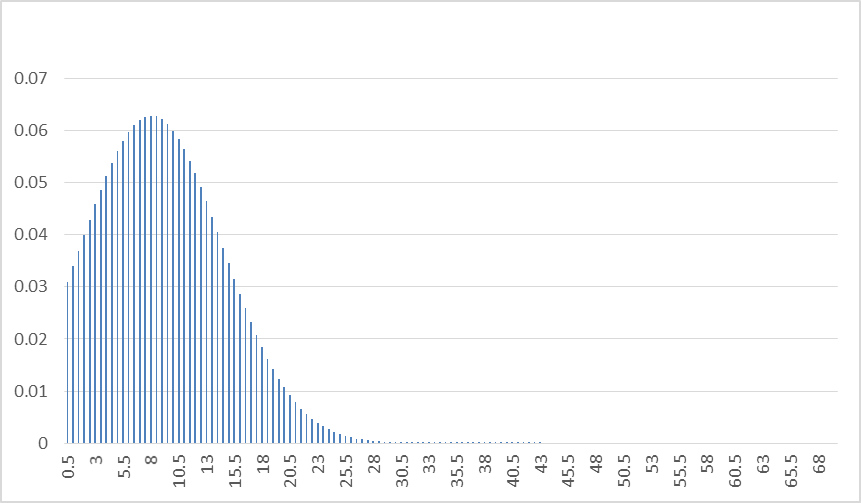
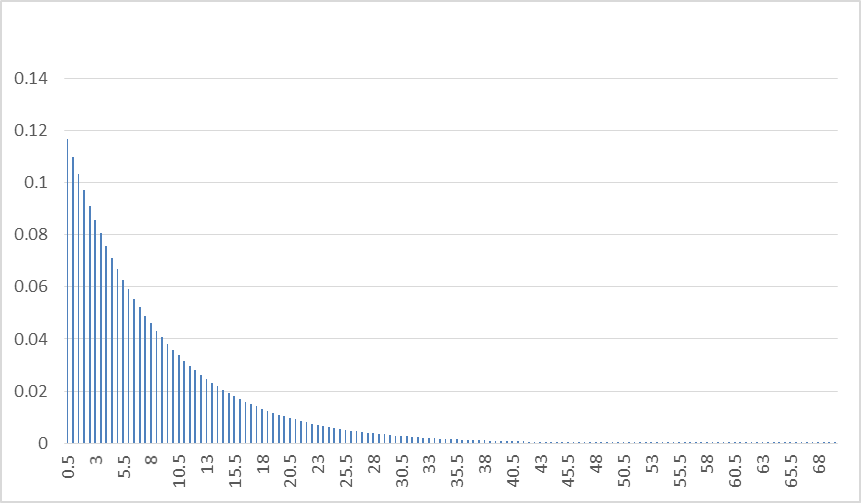
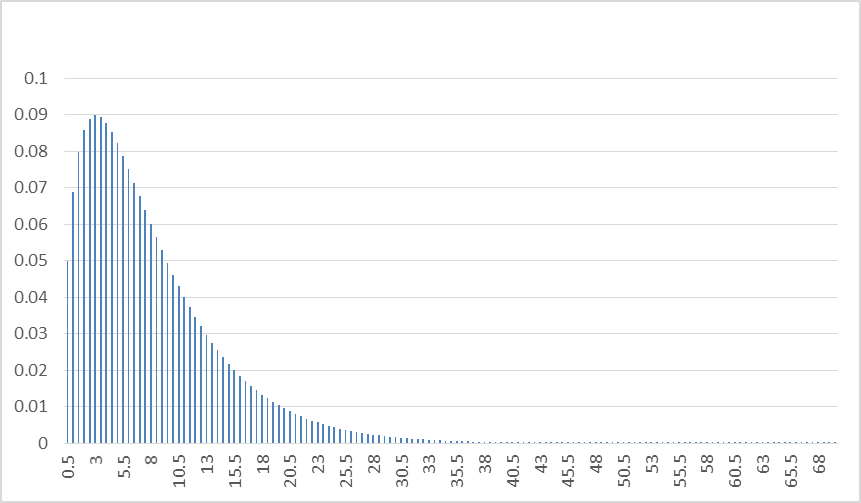
2. According to the excel function, AVERAGE = 8.048, and VAR. S = 40.356.
3. Normal distribution based on the mean and sample variance from (a) will show the distribution as below：



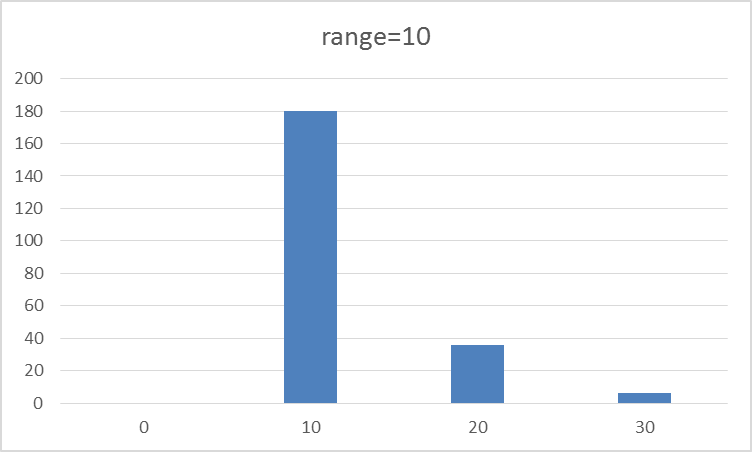
1. For mean = , , but for variance = , which means that , doesn’t match. Take , the distribution will be show as below:

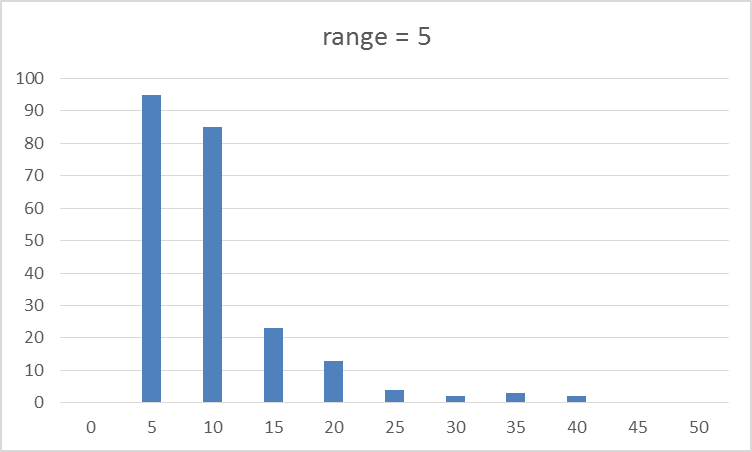


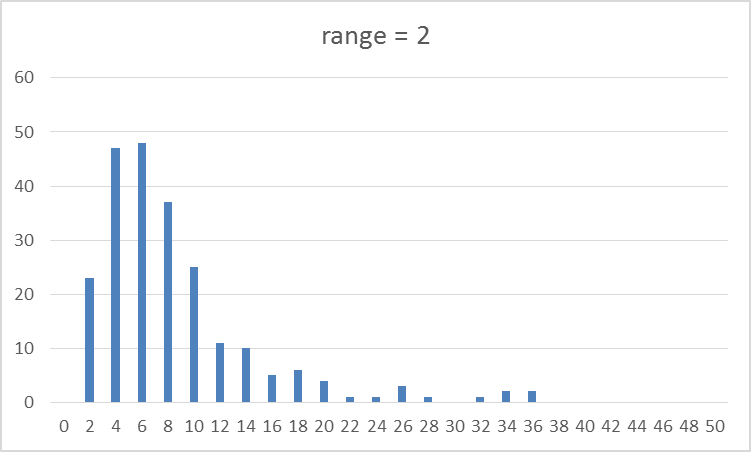
1. For Gamma distribution, since and , we can have and . the distribution according to the sample mean and variance will be like:

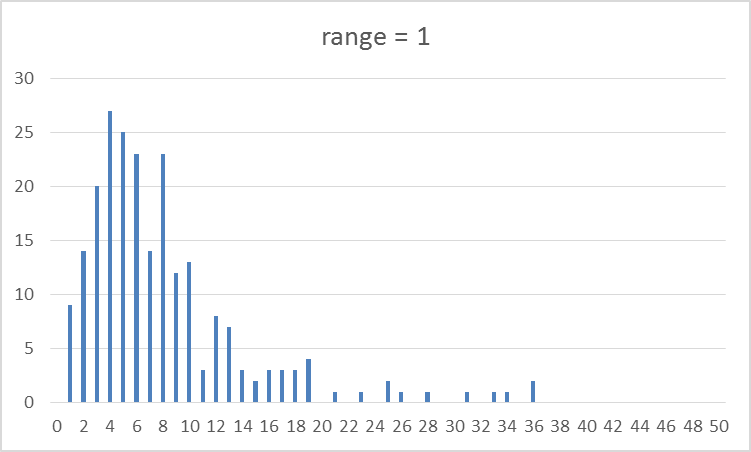


1. The histogram of the waiting time with range = 10, 5, 2, 1 is shown in below:

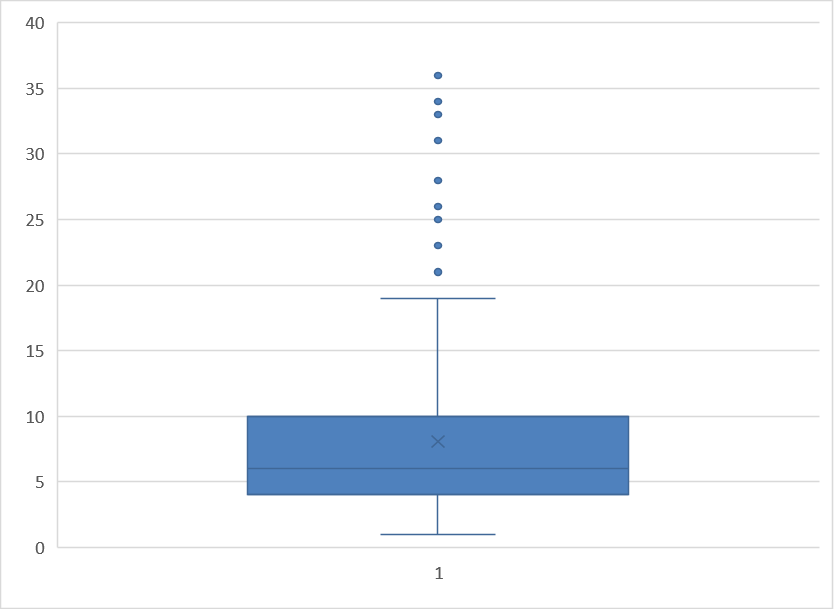








1. The box plot of the samples are shown as below. Note that the sample minimum = 1, 1st quartile = 4, median = 6, 3rd quartile = 10, and maximum = 36.



1. Compare box plot with histogram:



Compare box plot with different distribution:



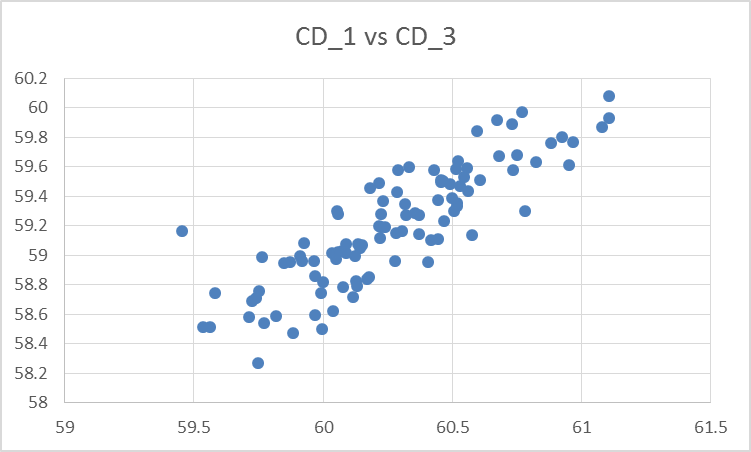
Observe from the comparison, we can tell that normal distribution can describe the randomness of the waiting time.

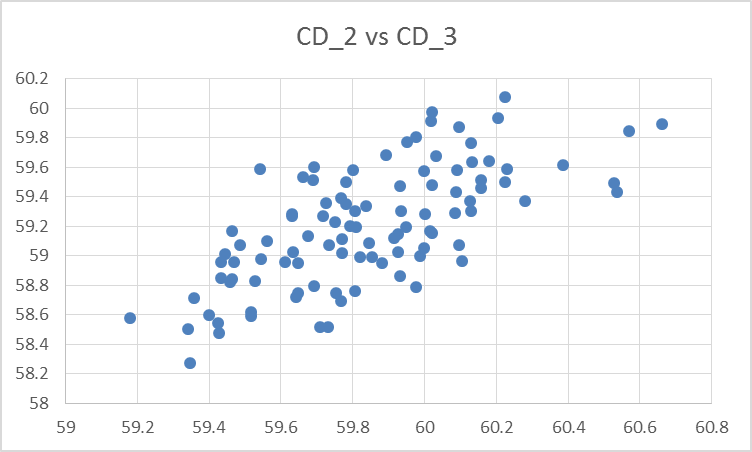
2. The average and mean of each CD are:

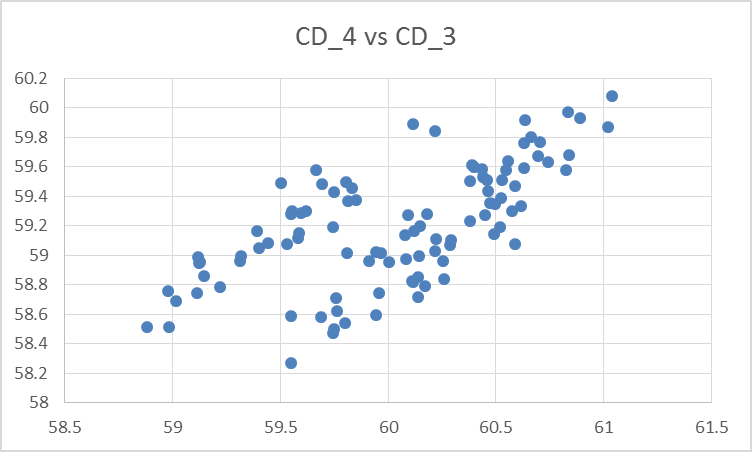
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| average | 60.270114 | 59.84056 | 59.19759 | 60.03255 | 59.5364 |
| stdev | 0.365786218 | 0.295817 | 0.396881 | 0.526366 | 0.297992 |

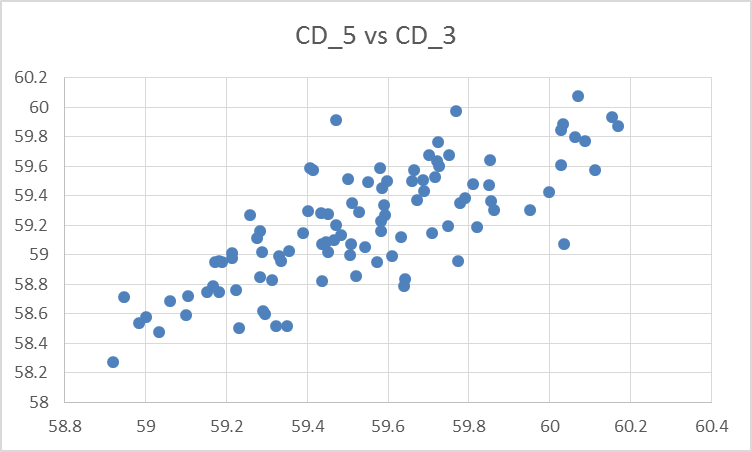
The covariance and correlation coefficient of 4 pair are:

|  |  |  |
| --- | --- | --- |
|  | covariance | correlation |
| 1 vs 3 | 0.124034253 | 0.854386 |
| 2 vs 3 | 0.08006673 | 0.681976 |
| 4 vs 3 | 0.136675279 | 0.654247 |
| 5 vs 3 | 0.091652257 | 0.774957 |

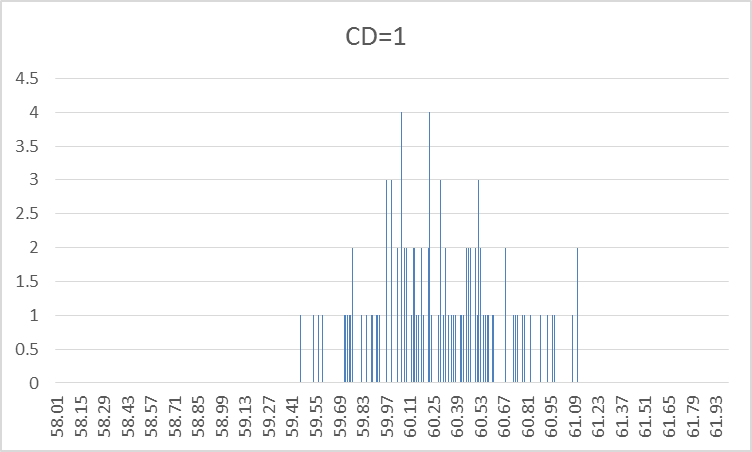


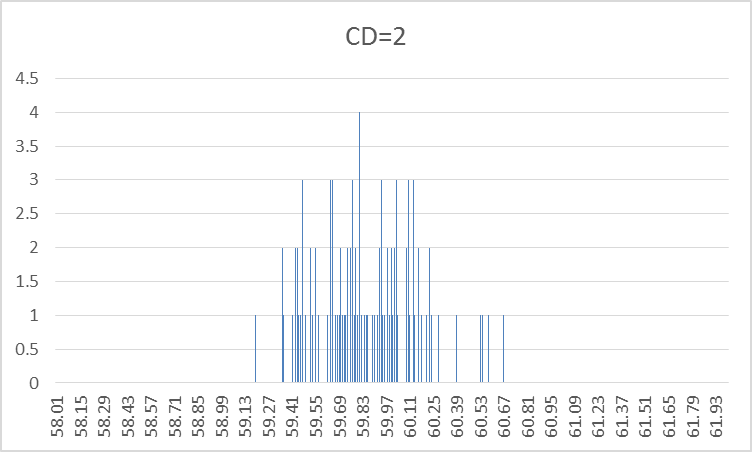


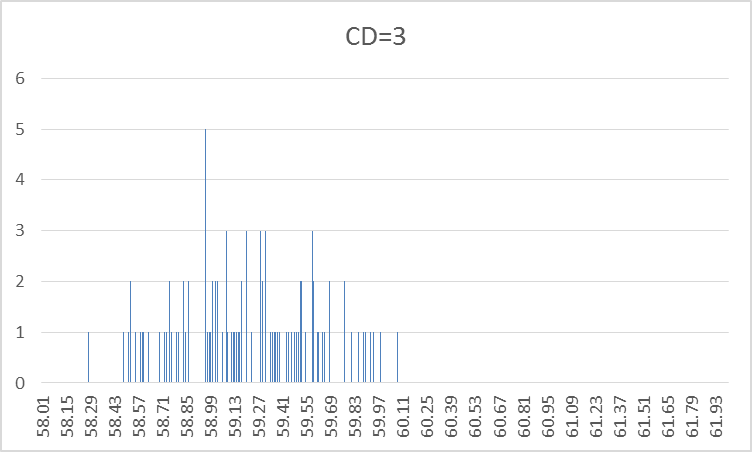


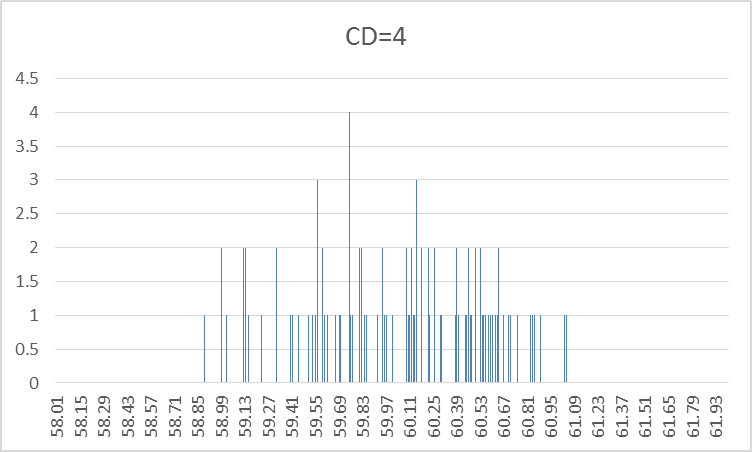


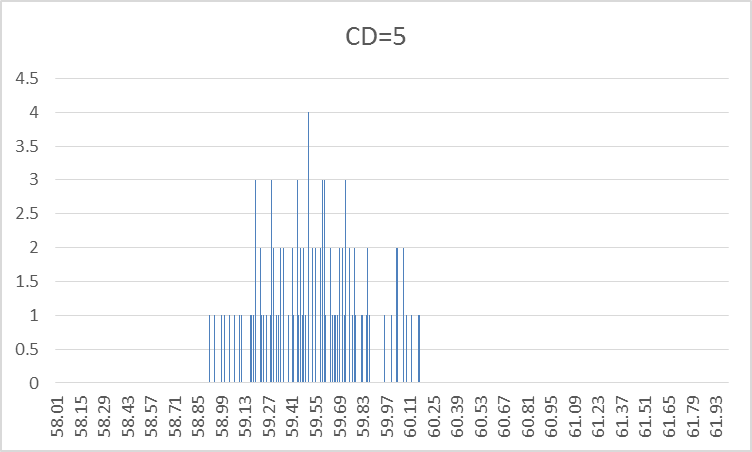
1. Histograms for CDs:



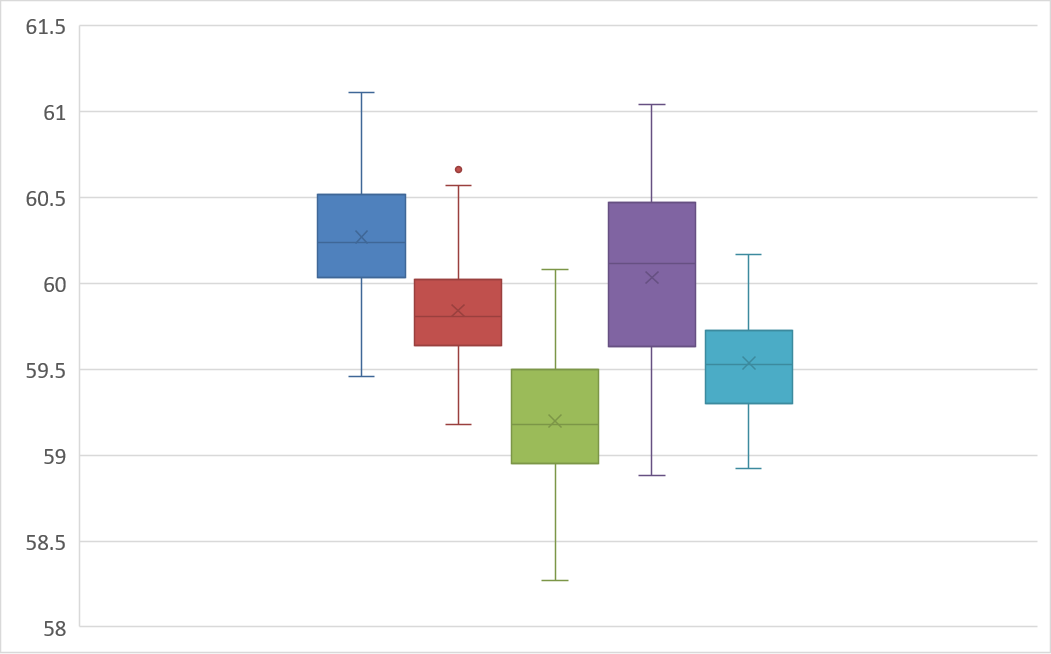








1. The box plot for CD\_1-CD\_5 are is shown in below. From the box plot, we can see that five distribution are following the same trend, which looks as similar as a normal distribution model.



1. The sample means are calculated and showed in the table below:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | CD\_1 | CD\_2 | CD\_3 | CD\_4 | CD\_5 |
| average | 60.2701 | 59.8406 | 59.1976 | 60.0325 | 59.5364 |
| median | 60.2343 | 59.8079 | 59.1776 | 60.1124 | 59.5249 |
| extreme average | 60.2807 | 59.9213 | 59.1741 | 59.9609 | 59.5447 |
| trim average(10%) | 60.2662 | 59.8305 | 59.1966 | 60.0406 | 59.5352 |
| stdev | 0.3658 | 0.2958 | 0.3969 | 0.5264 | 0.2980 |
| sample range | 1.6525 | 1.4816 | 1.8072 | 2.1551 | 1.2502 |
| IQR(Q3-Q1) | 0.4825 | 0.3865 | 0.5479 | 0.8399 | 0.4261 |

Difference among the 4 measures aren’t so obvious. We can say that because of the distribution of the sample are nearly symmetric, and either of the measure of location are able to find the center of sample, so the difference are small.

For difference among the 3 measures of dispersion, it describe how discrete the samples are distributed. The larger the measure of dispersions are, the larger the samples range are. We can say that any two of the measurement are positive related. It is totally different in three measurement, but they all tells the same message about the distribution.