**Use Univariate statistics on at least the first 4 attributes to describe the data. Discuss the results obtained, highlighting any result which you consider particularly useful. Use visualisations if needed.**

|  |  |
| --- | --- |
| **Some Attributes** | **Description** |
| n | Number of samples |
| sd | Standard Deviation |
| trimmed | Trimmed mean |
| mad | Mean Absolute Deviation |

**Score**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
| 292 | 6.39 | 2.39 | 6.5 | 6.48 | 2.67 | 0 | 15 | 15 | -0.19 | -0.24 | 0.14 |

**Score2**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
| 292 | 6.41 | 2.4 | 6.5 | 6.49 | 2.74 | 0 | 14 | 14 | -0.25 | -0.34 | 0.14 |

**Age**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
| 292 | 4.2 | 1.95 | 4 | 4.08 | 1.48 | 1 | 9 | 8 | 0.54 | -0.48 | 0.11 |

**Cost**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| n | mean | sd | median | trimmed | mad | min | max | range | skew | kurtosis | se |
| 292 | 1951.24 | 778.2 | 1920 | 1958.19 | 896.97 | -30 | 3840 | 3870 | -0.03 | -0.81 | 45.54 |

The standard deviation shows that there is little spread amongst the values for the data samples for score, score2 and age, cost has a large distribution between the values. The min and max show the range of values for the data column. The median and the mean show a much normal distribution in the data, with a normal bell curve. There is a little skew towards the right side of the data, but the difference between the mean and the median is pretty small.

Below shows a normal distribution plot of score, it shows a bell curve for the data distribution, with a little skew at the tail. Score2, age and cost follow similar patterns in their data distribution.

