Grouped Data Analysis of Rockwell Type A Hardness Values

Introduction:
A raw data set was provided from the manufacturing floor. A statistical analysis was performed in order for quality control to know if the manufacturing process is capable.

Measure of Central Tendency:
The mean, median, and mode are considered measures of central tendency. The mean was used in this analysis to find the value that the measurements were grouped around. This gives a starting point for a normal distribution.

Measures of Dispersion:
Since variation and quality go hand in hand the data needed to be analyzed for its variance from the mean. Measures of dispersion include range and sample standard deviation. What these values are measuring is the width of the normal curve. For example the larger the standard deviation the more dispersion there is in the data making the process harder to control. The smaller the standard deviation the more consistent and predictable the process is. The standard deviation for this sample was 11.1 which is a high variance from the mean.

Kurtosis & Skewness:
The kurtosis and skew are calculated to characterize variability and location of the data set. Kurtosis is a measure of whether the data is peaked (Leptokurtic) or flat (Platykurtic) relative to the normal distribution curve. A normal distribution curve has a kurtosis of zero. A positive kurtosis indicates a peaked distribution and a negative kurtosis indicates a flat distribution. The kurtosis for this data set was -0.71 indicating a platykurtic distribution. The skewness of the data shows how symmetrical the normal distribution curve is. A perfectly symmetrical normal distribution has a skewness of zero. The value of -0.08607 signals the normal distribution curve for this data set has a long tail to the left indicating lower values.

Tolerance Limits:
Limits created for a process usually by the customer in order to fulfill the design needs. These tolerance limits create a boundary for the values. If the value is above the Upper Limit or below the Lower Limit then the part can be determined as out – of – control. Thus given a data set, a normal distribution, and confidence interval it can be determined the percent of the measurement that is in tolerance.

Process Capability:
$C_p$ and $C_{pk}$ are both measurements of determining whether or not a process is capable of meeting specifications. $C_p$ is an indicator of the processes capability to be consistent with minimum variation. The $C_{pk}$ is a measurement of how close a process is running to its specification limits relative to the variability of the process. It is a value that determines how capable your process is. The larger the $C_{pk}$ the better your process is. The $C_{pk}$ for this data set is far too low at 0.419. It does not meet six sigma and will not be a reliable capable process.
Equations:

Mean: $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$

Sample Standard Deviation: $S = \sqrt{\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{n-1}}$

Range: $R = x_{max} - x_{min}$

<table>
<thead>
<tr>
<th>Sample Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>59</td>
</tr>
<tr>
<td>70</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>46</td>
</tr>
<tr>
<td>68</td>
</tr>
<tr>
<td>61</td>
</tr>
<tr>
<td>51</td>
</tr>
<tr>
<td>56</td>
</tr>
<tr>
<td>59</td>
</tr>
<tr>
<td>50</td>
</tr>
<tr>
<td>34</td>
</tr>
<tr>
<td>44</td>
</tr>
<tr>
<td>64</td>
</tr>
<tr>
<td>52</td>
</tr>
<tr>
<td>39</td>
</tr>
<tr>
<td>32</td>
</tr>
<tr>
<td>47</td>
</tr>
<tr>
<td>71</td>
</tr>
<tr>
<td>51</td>
</tr>
<tr>
<td>67</td>
</tr>
<tr>
<td>70</td>
</tr>
<tr>
<td>46</td>
</tr>
</tbody>
</table>

Quantitative Data Analysis

| Sample Size (n) | 110 |
| Sample Average ($x_{\text{bar}}$) | 54 |
| Standard deviation of Sample (s) | 11.1 |
| Kurtosis | -0.71 |
| Skewness | -0.08 |
| Maximum Value | 75 |
| Minimum Value | 31 |
| Specification | 55+/-15 |
| Upper Tolerance Limit (UTL) | 0.924 |
| Lower Tolerance Limit (LTL) | 0.104 |
| % in Tolerance | 0.820 |
| Cp | 0.449 |
| Cpk | 1.589 |
Normal Distribution of Data