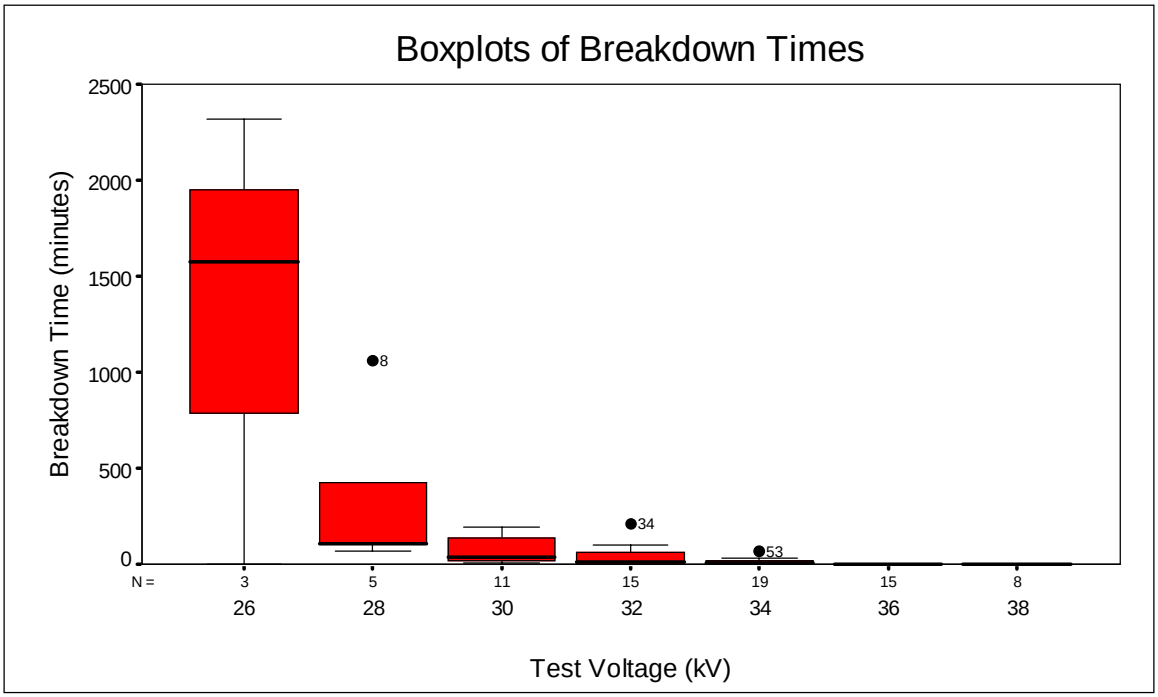
4. Displaying the Breakdown Times

Graphical displays of the data can be very helpful for understanding the information contained in the data. We will obtain the side-by-side boxplots of breakdown times for the seven voltage levels and a scatterplot of breakdown times against voltage levels.

4.1 Side-by-Side Boxplots of Breakdown Times on the Original Scale
4.2 Side-by-Side Boxplots of Breakdown Times on a Logarithmic Scale
4.3 Scatterplot of Breakdown Times versus Voltage Level
4.4 Scatterplot of Breakdown Times (Natural Logarithm Scale) versus Voltage Level

4.1 SPSS produces the following side-by-side boxplots of breakdown times for the seven experimental groups on the original scale of measurement:



The positions of medians indicate that the median breakdown time decreases exponentially as the voltage level increases. The observations obtained by subjecting the fluid to the voltages of 36 and 38 kV exhibit very small spread making their boxplots flat.

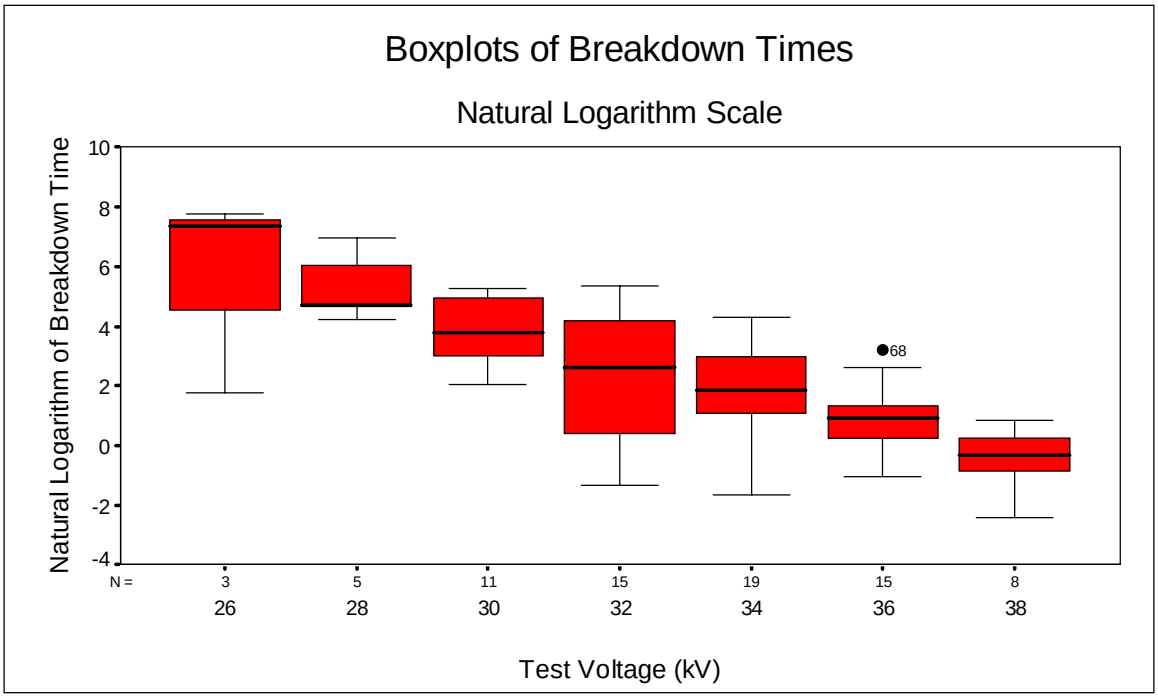
It is not difficult to notice that the variability within each group changes as the voltage level changes. More precisely, the variability decreases as the voltage level increases. The variability is very large for 26 kV, but it is very small for 36 and 38 kV.

The variability between groups decreases as voltage level increases. Indeed, there are very small overlaps among the boxplots for the groups obtained subjected to the low voltages and large overlaps for the high voltages.

As you can see there is some skewness in all the seven groups. The positions of medians and whiskers in the boxplots indicate that the distributions are skewed to the right.

Nevertheless, the skewness is exhibited in the groups consisting of a relatively small number of observations, it is difficult to detect nonnormality in such cases.

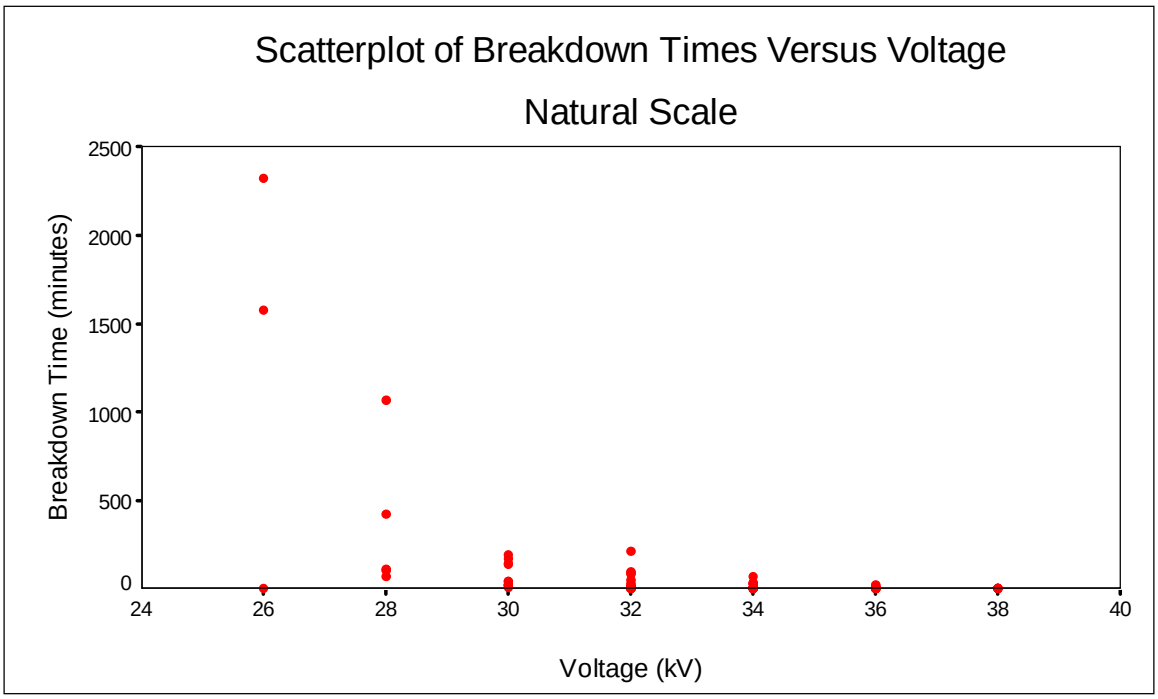
4.2 We will try to apply the natural logarithm transformation hoping to remove the skewness in the data. The following display shows a side-by-side boxplots of the breakdown times on a logarithmic scale:



From just a quick glance, we see that the logarithm transformation has helped compress the numbers. In general, the logarithm transformation tends to pull in the long tail of the distribution on the right, but stretch it out on the left. In other words, small numbers spread out more, while large numbers are squeezed more closely together.

The positions of the quartiles and whiskers indicate that on the logarithmic scale, the distributions for 30-38 kV are approximately symmetric, and have approximately the same spread. Thus, the logarithm transformation was very effective to remove the skewness in our data. Notice that the data for the groups 26 and 28 kV are still highly skewed. However, the groups consist of just 3 and 5 observations.

4.3 SPSS can also be used to obtain a scatterplot of breakdown times versus voltage levels:



The scatterplot of breakdown time versus voltage shows a non-linear relationship between breakdown time and voltage. More precisely, breakdown time decreases exponentially as the voltage level increases. This pattern suggests the need for transformation.

4.4 We will try to apply the natural logarithm transformation hoping to obtain a linear relationship between the transformed breakdown times and voltage. We believe that the transformation will be successful to linearize our data because the scatterplot of untransformed observations follows an exponential pattern.