Part 4.3: Unusual Points

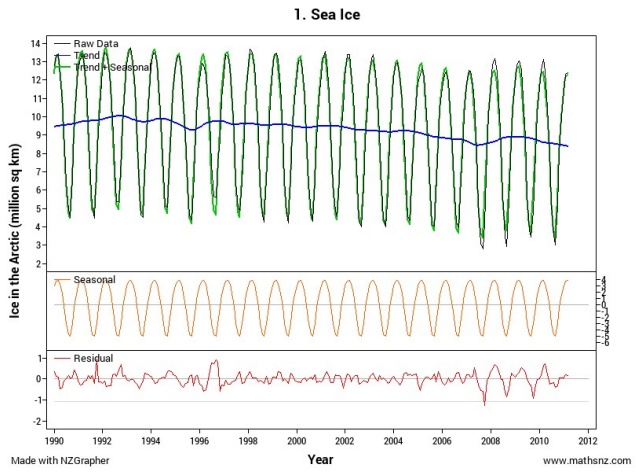
The next step involves a little bit of mathematical calculations. To work out if a point is an outlier tend to use the rule of thumb that the residual is more than 10% of the overall spread. NZGrapher automatically draws these in as light grey lines for you in the residuals section. To easily identify points if you tick the 'Point Labels' button it puts the id of the row next to each point.

To calculate this we use the following calculation:

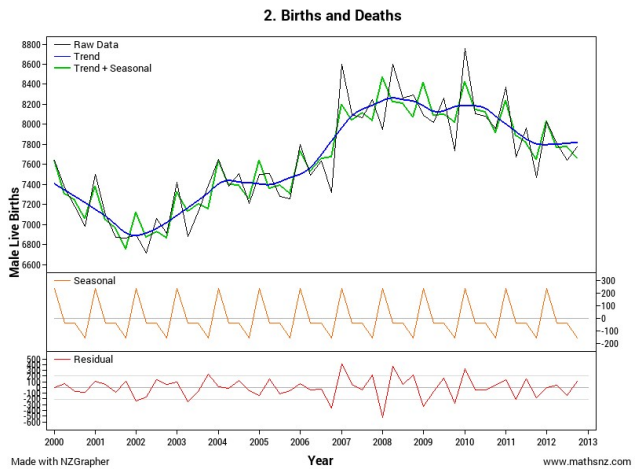
$$\frac{\text{Absolute Highest Value} - \text{Absolute Lowest Value}}{10}$$

If any of the residuals are either larger or smaller than this value we need to comment on them and what might be causing them. You could also be thinking about how big the variation is of the residuals is as a component of the overall variation.

For each of the graphs calculate if there are any outliers and comment on any unusual features. The first two have been done for you.



Absolute Highest Value: 14,
Absolute Lowest Value: 3.
$\frac{14-3}{10} = 1.1$
Looking at the residuals graph there in only one point that is more than 1.1 million square kilometres away from the trend. This occurred in September 2007 and may have been due to an unusually hot summer.



Absolute Highest Value: 8700
Absolute Lowest Value: 6700
$\frac{8700-6700}{10} = 200$
Looking at the residuals graph due to the inconsistency of the data there are a large number of residuals between 2007 and 2010 that are outside the acceptable range. This is during the financial boom so it may be due to people being more willing to have children and therefore not worrying about the timing so much, therefore the normal patterns do not happen.