EXPLORING AVERAGES

On the GeoGebra applet you can see the following 11 data points displayed on a numberline. For example, they could be runs scored by cricket players in an innings

1, 4, 6, 8, 9, 10, 12, 14, 17, 22, 44

You can click the boxes at the top of the applet to display the median, mean, range and a box plot.

You can change the value of the data points by dragging them along the number line.

1. Drag point B to the right hand side to increase its value. Explain **how** and **why** this effects…

   a) The median

   ……..

   …….

   b) The Mean

   ……..

   …….

   c) The Range

   ……..

   …….

2. **Click the reset button.** Drag point A to the right to increase its value. As you drag record what happens to each of the statistics, try to think about why this is and fill in the table. The first column is done for you!

	8<A<9	9<A<10	10<A<12	12<A<14	14<A<17	17<A<22	22<A<37	A>44
Median	No effect							
Mean	Increases							
Range	No effect							
Lower Quartile	No effect							
Upper Quartile	No effect							
IQR	No effect							

3. **Click the reset button.** Is it possible to change the value of just one data point without affecting the data's mean? Why is this?

   …….

   …….

4. **Click the reset button.** Is it possible to change the value of two data points so that the there is no effect on the mean, median and range? If so give an example of how this can be done.

    ……

    ……

5. Is it possible to arrange the data points so that the mean and median are the same? If so give an example.

    ……

Sometimes statisticians will exclude some data points from an analysis because they are very different from other data points, possibly indicating an error in the measurements. You can do this by clicking the select outliers button and clicking on data points.

6. **Click the reset button.** Click on select outliers and click on point B. How much does each of the following statistics change by?

    a) The median   ………

    b) The mean   ………

    c) The range   ………

    d) The IQR   ………

7. Which average is most affected by the outlier?   ………………………………..

8. Which measure of spread is most affected by the outlier?   ………………………………..

9. Which average/measure of spread would be best in the following scenarios?

    a) I want an average that will take account of the value of all data points   …………………

    b) I want a measure of spread that I can use to compare the variability within two data sets. I want to

        ignore any extreme values   …………………

    c) I want an average that could be used for qualitative data such as car colour   …………………

    d) I want an average that could be used to compare wages at various companies. I want it to represent the majority of the employees rather than be greatly affected by high-level manager's salaries

        …………………

    e) I want a measure of spread that I can use to compare the gap between the best and worst paid employees at various companies   …………………