Notes about writing answers for Statistics:

1) Discuss briefly the overall shape of the data/graphs separately.
   a) whether the distributions are symmetrical or skewed.
   b) if there are any extreme values or outliers (an outlier is a value a long way from the rest – over a whisker length away at least).
   c) whether the data is smooth or clumped.

2) From this point on try to direct your comments towards **comparisons** between the two measured groups.

   Try to always put **numbers** to back up comments, including the size of differences – you have them so you should use them. Refer directly to the features of the graphs that show what you are talking about.

3) **Compare medians** as your starting point to talk about typical values. You can then bring up the difference in the means (but bear in mind they can be distorted by extreme values).

   Discuss whether the difference in the medians is meaningful. If the medians for each group is outside the Inter-Quartile Range (IQR) for the other, then the difference is unlikely to be a statistical fluke. We say the difference is "significant".

   If a median is well inside the IQR for the other group, then any difference may have occurred as a result of "random sample variability".

4) Compare Inter-Quartile Ranges to discuss the range of typical values, using words like "consistent", "inconsistent", "variability" and "spread of values".

   Compare the Ranges, but take into account the effect of any extreme values.

   The meaning of this variability in the real world can be usefully explained.

5) If there are any other significant differences between the distributions mention them now.

6) **Directly answer the question posed at the start of the assessment**.

7) **Relate your result to the population**. It is vital to not forget that the purpose of sampling is to make some statement (inference) about the real world.

   Don't get too caught up with saying the result "proves" anything. Saying it is "consistent with" or "supports" is far better.

   Then qualify how far you think the results can be taken to represent the real world. Does the result only reliably apply to a portion of the population, or can it reasonably be expected to be true everywhere?

8) Consider improvements, such as breaking down the data in another way to see if there is more information hiding.

   Consider if the data has been gathered in a sound way. Is it "representative" of the population? Does the method have a "bias"?

   Comments about problems and improvements must be **statistical** in nature. If the comment isn't about the data, graphs or statistics, it probably isn't relevant.