Other Outputs Assessment Notes

1. **Please enter your name**

   Answer: *The respondent's name (hopefully!).*

2. **The public perception of released statistics is important.**

   Answer: *True.*

   The research community is reliant on the public to be data subjects, as well as support the use of sensitive data. As such it is important to consider how a statistic might be perceived.

3. **What are the general SDC principles?**

   Answer: *Is the statistic based on a sufficient number of observations?*

   Would it take an unreasonable amount of effort to work back?

   Would I be happy for my data to be presented in this way?

   The number of observations is a crucial part of an SDC check and will be considered for practically every statistic. SDC is not a 100% effective solution but we hope that the techniques mean that it is highly unlikely that someone would be able to work back to the underlying data. If an individual is very motivated and has access to significant resources then it may be impossible to completely protect the data. This would not be a realistic scenario though and it would seriously impact on the ability of researchers to work if we aimed for completely safe outputs. It is also important to remember when we are working with sensitive data that it belongs to a person, and would we want our data to be used by another researcher in the same way. If you feel you wouldn't, that is often a good sign that you are not handling the data in the right way.

4. **What are the SDC considerations relating to test statistics?**

   Answer: *Should be based on a sufficient number of observations.*

   Correlations should not be exactly 1 or -1.

   Test statistics are usually considered safe but there are a few considerations. As always, they should be based on a sufficient number of observations and correlations between variables should not be exactly 1 or -1. The latter case is normally indicative of some underlying issue with the generation of the statistic anyway.

5. **The box plot below presents the political attitudes of a number of survey respondents.**

   The respondents were grouped by age and the distribution of responses for each age group have been plotted.

You have been asked to review the output.

What are the issues with the plot?

Answer: *There are outliers, and these individual observations may be identifiable.*

The plot is not clearly labelled.

We do not know how many observations the plots are based on.

The outliers likely refer to individual observations which could be a disclosure issue. We can clearly see that one observation in the first group received a score of 5 on the scale. This may not be an issue however if the variable is derived or otherwise does not exist outside of the data, but the plot is not clearly labelled so we cannot assess this. We also do not know how many respondents are in each group.

6. **Presenting point maps can be problematic from an SDC perspective.**

How could we adjust the map presented below to reduce the disclosure risk?

Answer: *The respondents could mention that this issue with point maps is that they reveal very precise geographic information about the observation. It is often better to generate a choropleth or heatmap which shows the distribution of observations in an area or other useful statistics but it does not give such a precise view of any single observation.*