**4. Data Analysis Scaffolding Sheet – Margin of Error – how accurate is your result?**

If you choose Option 1…. ***Is there a significant difference in [dependent variable] between [this] and [that]?***

Use this *little book of statistics* to give you an idea of how to evaluate the ‘*accuracy*’ of the answer to your RQ.

1. Calculate the standard deviation (s) for both [this] and [that] datasets using the formula =STDEV( : )
2. Calculate the sample size (n) using the formula =COUNT( : )
3. Calculate SEM for both [this] and [that] using the formula =s/SQRT(n)
4. Click on the graph. Click on ‘Chart Design’ tab. Click on ‘Add Chart Element’.
5. Click on Error Bars. Select ‘More Error Bar Options’.
6. Select Custom. Select ‘Specify Value’.
7. Delete all default values in the ‘positive error value’ and ‘negative error value’. Leave curser in empty box.
8. Go back to your spreadsheet and highlight *ALL* SEM data for both [this] *and* [that].
9. Repeat for the ‘negative error value’. Error bars should appear on your graph ☺
10. Email your excel spreadsheet to the teacher.

    

    

If the error bar is small, the accuracy is high. The average (mean) sits in the *middle* of the error bar. The standard error of the mean (SEM) determines the height of the error bar. *Note:* you can also use standard deviation (s) instead of SEM to determine the height of the error bar, but SEM is preferred for population studies (*also note*: SEM error bars are *smaller* than s error bars).

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

If you choose Option 2…. ***Is there a linear relationship between [dependent variable] and [independent variable]?***

Whilst it’s possible to add error bars to every data point on a scatterplot graph, it would require each data point to be a mean value, which requires a lot of sampling and data collection that goes beyond the scope of this subject. If you are interested in learning how, ask your teacher.

Whilst an r value that is close to 0 would suggest the data is inaccurate, there may also be *no* relationship between the two variables (and has nothing to do with accuracy). This is something that would need closer investigation, taking into consideration the experimental design (e.g. limitations of the sampling method), controlled variables, measured variables, as well as how your data compares to what the scientific literature says.