Your name: _____

Sheet 1 – simple formulae and references

a. Use a formula in Excel to add the three numbers together. Write the sum here: _____
b. Use a formula in Excel to multiply the three numbers. Write the product here: _____

Sheet 2 – built-in functions

Use Excel *functions* to calculate the following values for the long list of numbers:

a. Sum: _____
b. Mean: _____
c. Standard deviation: _____
d. Smallest value (minimum): _____
e. Biggest value (maximum): _____

Sheet 3 – dragging formulae down, sorting

Here I have put data I collected from Hong Kong Cantonese speakers about how concrete (具體) they think certain English words are. The survey had 36 words, and 5 people did the survey; they rated the words from 1 (very abstract) to 6 (very concrete). I want you to figure out what the most concrete and least concrete words are in the list, and also show me a screenshot of what the sheet looks like after you have done this. To do this, you will need to first find the average concreteness (across 5 raters) for each word, and then sort the data in order of concreteness.[1]

a. Most concrete word: _____
b. Least concrete word: _____
c. Screenshot of your data sheet after doing this work (insert below or attach):

[1] There are alternative ways to find the most and least concrete words without sorting the data, but they might be harder. You can try if you want, but it's not necessary.

Sheet 4 – absolute (static) references

These data are similar to the data from Sheet 3. This time, the same 5 raters rated how familiar each word is to them, where 7 is the most familiar and 1 is the least familiar.[2] I again want you to calculate the most and least familiar words; this time, however, there is an extra step to do. I suspect that some raters gave overall higher familiarity ratings (across all words) and some raters gave overall lower familiarity ratings; this might be because some raters have better English proficiency or went to different schools where they started learning English earlier, or maybe they just understood the question differently. I think we should adjust for these differences: for example, if Rater 1 rated most of the words as having a familiarity of "5" but rather 4 rated most words of having a familiarity of "3", then if they both rated the word *fantasy* as "4" then this actually has different meanings: Rater 1 is saying they think the word is less familiar than average, and Rater 4 is saying they think the word is more familiar than average! One way we can make this adjustment is to go in and, for every person's rating for every word, subtract that person's average rating. Then, for example, Rater 1's rating of "4" for *fantasy*, minus Rater 1's average rating of 5, would turn into "-1"; Rater 4's rating of "4" for *fantasy*, minus their average rating of 3, would turn into "1". To help with this, in the top row (cells B1:F1) I have calculated each rater's average; by now you should be able to understand how I did this. What you need to do is first adjust every rating, in cells G3:K38, and then get the averages of those adjusted ratings for each word, in cells L3:L38.

a. Most familiar word: _____
b. Least familiar word: _____
c. Screenshot of your data sheet after doing this work (insert below or attach):

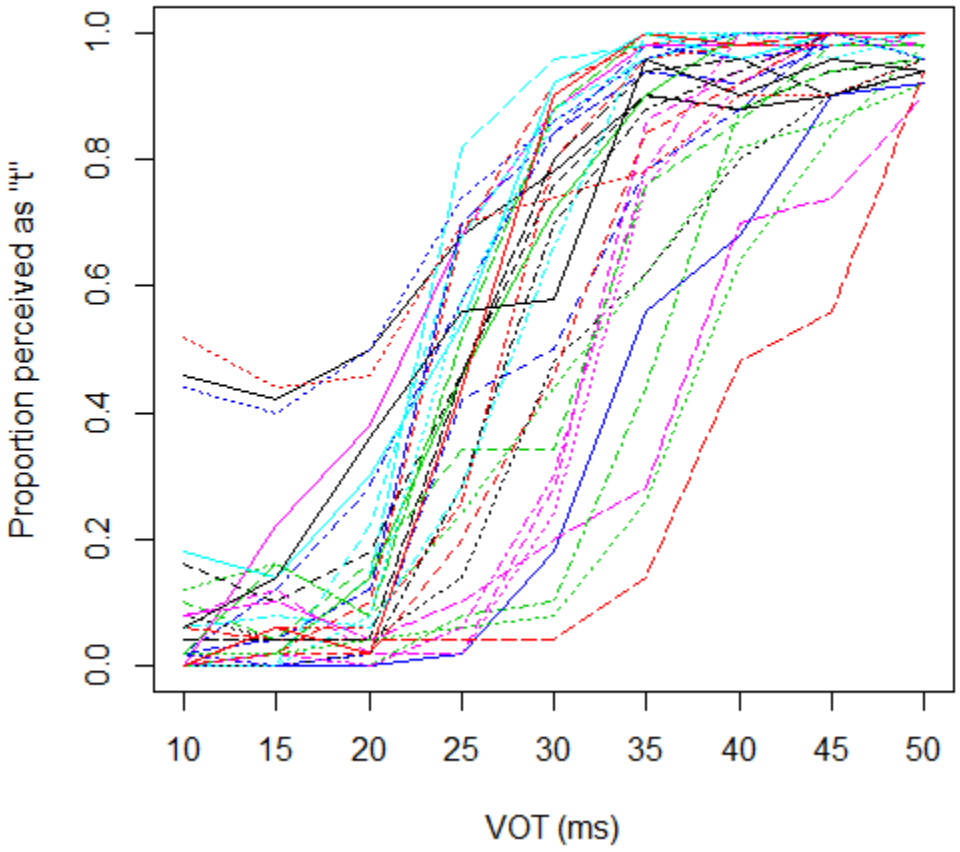
Sheet 5 – graphing

Here are data from a categorical perception experiment. Each participant heard Mandarin words that started with either /t/ or /d/; they differed only in Voice Onset Time (VOT). As you may have learned in other classes, sounds with a higher Voice Onset Time tend to sound more like aspirated sounds (like /p, t, k/), whereas sounds with a lower Voice Onset Time tend to sound more like unaspirated sounds (like /b, d, g/). Each participant heard ten of each of these sounds (e.g., each participant heard the sound with 10 ms of VOT ten times, the sound with 15 ms of VOT ten times, etc.) and had to judge whether they heard /t/ or /d/; the data

[2] Technically, what they actually rated was how old they think they were when they learned the word. This is closely related to familiarity: words that we learned early in life, like *dog*, feel more familiar than words we learned later in life, like *comprehensibility*.

shown here tells us what percentage of the time each participant perceived each VOT as sounding like /t/.

I want you to make two graphs of the data. The first should show all participants' data on top of each other (this is sometimes called a "spaghetti plot", because all the lines look like a bunch of spaghetti). It should look a little bit like my graph below. It won't look exactly the same (because I made this graph in R, rather than Excel), but make sure it has VOT labelled at the bottom like mine does, and make sure the lines are in the right places; other details like colour and line style aren't so important.



The second graph should be similar to this, but with only one line: the *average* across all subjects.

Paste your graphs in below or attach them separately.

a. All-subjects graph:

b. Average graph: