Learning Journal- Thomas Kongonis

All time is in 24 hour format

Amendment- 22/8/19

timestamps were not utilised for the first data carpentry lesson

**Data Carpentry 01- formatting data tables in spreadsheets.**

**11/8/19**

**Time was not recorded**

* Diagnosed problems
* Printed spreadsheets to troubleshoot due to headache from looking at screens, this gave me a bit of a break whilst continuing work
* Begun cleaning the Tanzania sheet:
* Initially made copy to begin working.
* Firstly: dwelling table re named all of the values to a constant- implemented camel font whilst retaining underline for the titles.
* Changed key id 3, room to null value as the -99 seems to suggest that a measurement was not taken
* Capitalised all values that did not have a second word for continuity.
* Figured out what the colour coded square meant. This was to suggest that this house had a special room or barn and this accounted for the extra room.
* Decided that it would be more effective to create a new table for barn.
* Implemented this with the values being either ‘Yes’ or ‘No’ and removed colour code and separate value external to the table.
* Encountered problem, could not get black outlines on this new part of table, left this to address if the problem is found again.
* Created new sheet to include the fresh and all-inclusive table.
* Listed Key ids 1-20 and assigned another table to signify country.
* Decided not to merge any cells to make it more exportable and readable.
* Instead decided to add another row titles inc\_Dwelling and signalled Yes for all inputs.
* Copied corrected roof\_type inputs for key\_id 1-10 and added them into new table.
* Proceeded to do the same for wall type floor type and rooms and added inc\_barn table.
* Moved to Tanzania dwelling table for cleaning.
* Proceeded to format in same way as Mozambique.
* Diagnosed cowshed problem that was similar to barn problem, except no colour and asterisk was used instead.
* Implemented same solution- implemented inc\_cowshed in new table.
* Utilised search function and implemented freeze top row to increase ease.
* Solved border problem by searching through format cells menu.
* Set up animal section of the table, rather than merging cells, decided to create an inc\_animals section.
* Decoded awful table in Mozambique section, had difficulty and had to redo after realising that there were some unspecified animals, this was the same for the Tanzania section.
* Added null sections for key id 10 and 20 as they were not measured, used null values because all the ones and zeros confused me, will remove later.
* Moved to plots table.
* Created inc\_plots section for clarity.
* The confusion in this table lied a lot in all the unsuspecting null values, the decision was made to treat these as null values, this included id 6,9 and 10
* Created a constant water use and summer water use value sections. Id 3 only used water in summer.
* As for 7 and 8, their water use was constant, so the decision was made to make their summer water value yes.
* Id 8, in the original table had the water value as one, the decision was made to include a section that states amount of plots watered and this was set to 1, the other ids that had continual and summer watering were set to include all the plots.
* Table was formatted and given lines.
* Upon review a mistake was seen, looked after cows’ section was ignored, this was inserted at the end of the animal section. The dead cow value was signified as yes.
* Animal death section was added to account for this data
* Date collected section added to include all data
* Only available for Tanzania.

**Lesson 02- Metadata**

**14/8/19**

**13:00-14:30**

When reading through the SAFI\_clean table, besides being still slightly confused by spreadsheet data inputs, there appeared to be several things that stood out that i felt would require metadata in a separate document to help clarify. These included:

* The countries each of these towns are in, its possible that this is all the same country and I wasn’t sure if there was continuity from the last spreadsheet looked at. I thought it would be useful to list the country/ies that these towns are in and possibly their proximity to eachother.
* A second thing that stood out to me in both tables was the types of bricks, it would be useful to know what they are comprised of and how well they fulfil their ‘brickness’.
* Thirdly, what exactly connotes a room would be useful. By this I mean is there a minimum or maximum size.
* Fourthly, a big concern I had was: what justifies items\_owned? Is there a monetary amount or a necessity for a large purpose? For example, a shirt could fall under this category so it would be good to know what the lower range is for inclusion.
* Fifth, what constitutes a meal would be necessary metadata. We all have very different notions of what a meal is. Is a full belly necessary or is just staving of hunger enough?
* Sixth, does months lacking food constitute and amount? For example, would one meal a week class as lacking food or having food for the month?

**Scoping exercise:**

**14/8/19**

 15:00

* I made this exercise more difficult for myself by starting with paper, moving to word then finishing with LateX. I was scared of the unknown but in reality LateX was pretty cool and i’ll be using it going forward.
* I begun by utilising the information given to us to create a funnel. I then began to write it all down and transcribe to word as I realised that I did a few things wrong and edited. (misunderstood some of the distinctions between the categories)

15:30

* During this process I had an idea for the final assessments that I noted and added into the exercise.
* After this I moved to Overleaf, after creating my account and selecting a new document I saw the rich text section which I begun to utilise first.

16:00

* I proceeded to write down titles and then I made bullet points and copied and wrote down the exercise in full.

16:15

* After I finished inputting the information I then decided to go to the learn LateX page as it all seemed too simple and I was surprised that I had made a document without breaking it.

16:20

* I then proceeded to read and realise what all the code was for. I decided to play around with some of the code to show that I made an attempt to learn.
* Initially I saw that the title was made from code and it utilised my name as it appeared in my student email. I altered this and moved my name and student number into the author section and kept the \title section as: Scoping Exercise.

16:30

* I then saw that there was code for sections and sub sections which I decided to play with. I took the Jobs, pains, pain relief, gains and gain enabler sections and coded headings in rather than just typing. Then as I continued to read I saw the code for \tableofcontents. I decided to get this working. So I added the code into the start of the document below the title of the document.
* I then proceeded to change the code for the headings to \secton{}.

16:40

* After I updated the document, I realised I had made a mistake, the table of contents was there but the title and author was no longer.

16:45

* To fix this I attempted to take the begin document function and add it above the title.

16:50

* After this didn’t work, I deleted the begin document function and retyped it. this then meant that the end document function was just below so I broke the document and it all disappeared.
* I realised this problem and fixed it but the title still wasn’t there. I referred back to the site and realised I was missing a \maketitle function for some reason. I input that and then it was all fixed.

 17:20

* I then saved the document as a pdf and a tex file, uploaded both of these to a folder in cloudstor and just uploaded the pdf to Ilearn.

**22/8/19 Dates in Spreadsheets:**

**14:30-15:00**

* Begun by copying dates into a new sheet and proceeded to input values
* Inserted values through utilising the codes and got bad results
* Trying to input the days was fine, however inputting the month as =MONTH(11) returned a value of 1. Tested this in the year section and the =YEAR(2016) function returned the result 1905.
* Upon reviewing the lesson again, I realised that the mistake was that it was utilising the number date system and the number 2016 implied 2016 days from December 31st 1899. The solution that I found to this problem was by inputting the date, year and month code and linking the date column to the code for example =Date(A2). After applying this to the entire table I got the correct results.

**Second part:**

**15:00-15:30**

* Followed instructions, the first column was changed to say Nov-17 then changed the year to 2017.
* Realised that the version of excel I was using utilised American dating. As it populated the new data point because there was no year it instantly assumed that it was current year so showed 2019.

**Quality Assurance lesson:**

**22/8/19**

**15:30**

* Followed instructions and input the minimum and maximum numbers for the No\_membrs column to 1 and 30 respectively.
* Created an input message for the Room column that included a min of 1 and a max of 10 as this seems to be fair, any house with more than 10 rooms would probably be a mansion!
* The input message stated that the column must not exceed a max of 10.

**List selection:**

**15:45**

* Followed first example, decided to apply this to the Town column. There was only 3 towns so this made the list option very fitting for this column. I included the correct spelling by following what has already been written and left a meaningful comment that stated my intention and that any new towns added to the data set must be first approved and then added to the dataset through the data validation tools.

**Exporting lesson:**

**16:00**

* Utilised the .csv file type to save the copied SAFI clean spreadsheet.

**Scoping II**

**22/8/19**

**16:15**

* Experimented further with coding in Overleaf.

**16:20**

* Created sections and subsections with the relevant code.
* A difficulty faced however was no matter what I tried to create a space between text without going to a new section, I could not. I went through information on the overleaf site and found a code that could be implemented, this was the \vspace code. I set this to 5mm and achieved success with what I wished to happen. This allowed for me to split up the writing as I wished.

**16:50**

* I then proceeded to apply this to the sections and subsections so that I did not need to create constant new lines of code and over complicate the document.

**23/8/19**

**11:15**

* Finished inputting necessary information and backed up the document ready for upload.

23/8/19

11:54

* Removed a 1500 word portion at the beginning of the document, it defeated the purpose of the learning journal, rather than documenting the process of cleaning the spreadsheet for lesson 1 I rewrote it to be clearer. Yet after reading through the documentation it was clear exactly what I did and why and this is what I chose to retain, the initial documentation not the secondary annotation that was superfluous.