I think I followed the majority in deciding to analyse a tool that I had used before, in my case, Tableau. Having said that, I haven't used it a lot, and I decided to try to analyse some geospatial data, which I hadn't done before.

I didn't have a dataset in mind to analyse, but was able to find a file on the Internet that contained a map of Australian postcodes. Opening the file in Tableau was straightforward, as it includes "Spatial file" in a list of common file types. This was handy, as I've never really used spatial data before, and the zip file I downloaded actually contained five files, with file tyes that I wasn't familiar with: .cpg, .dbf, .prj, .shp and .shx.

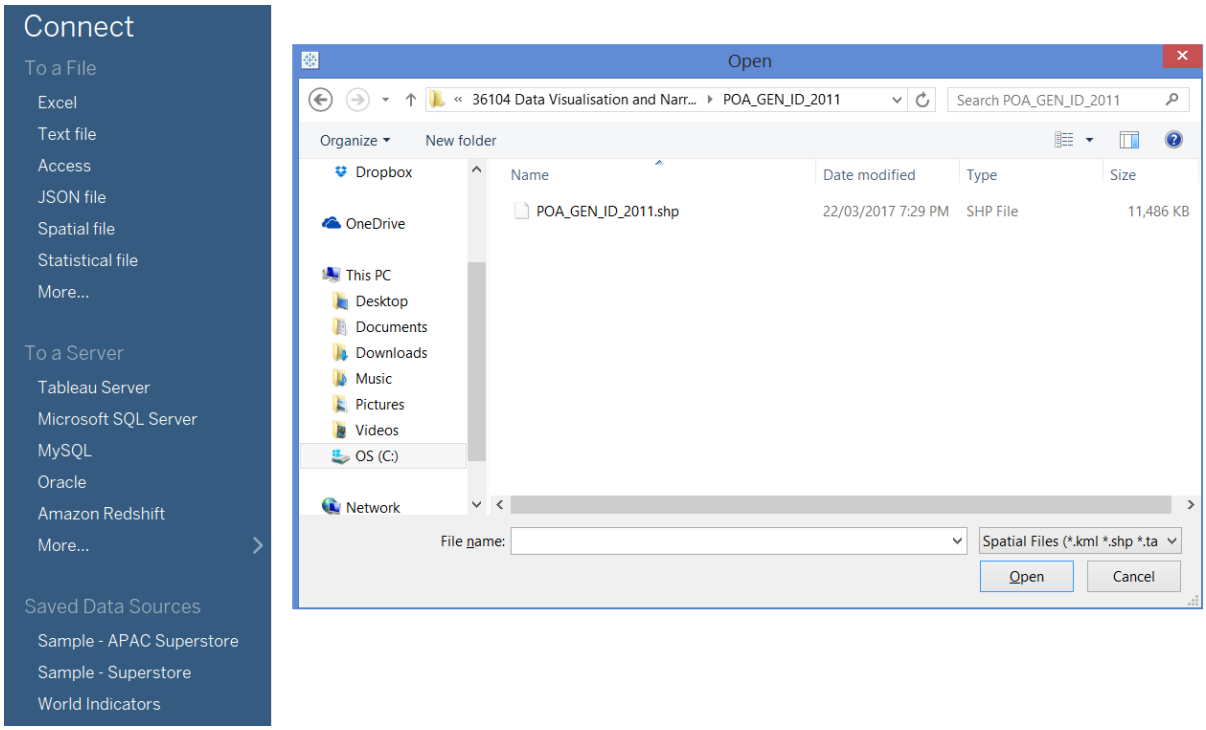


Figure 1: Opening a source file

When Tableau opens a file, it shows the data and makes a guess as to what type it is. In this case, there were four columns, which it classified as numeric, string, string and geospatial. You can easily change the data type if Tableau gets it wrong, or you want numeric data to be used as a string, for example.

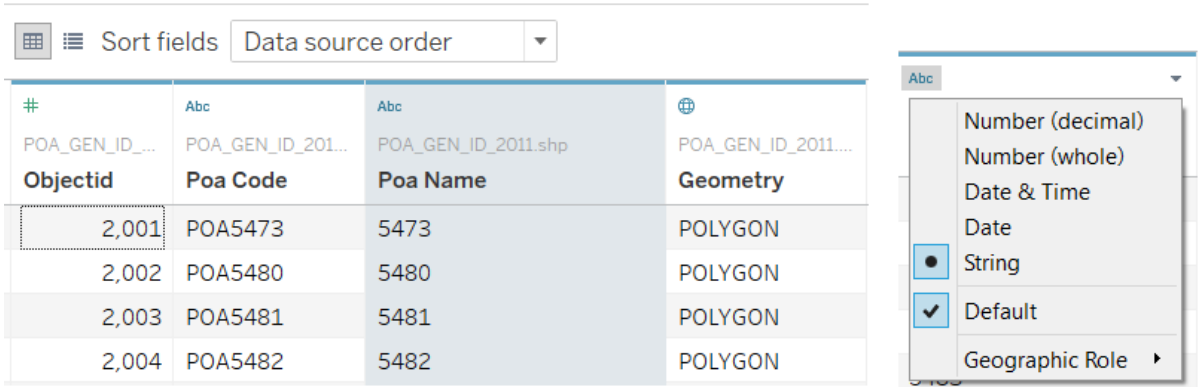


Figure 2: Configuring data types

Once you have the data read in, you then open a worksheet to start using it. Again, Tableau performs an analysis of the data and separates it into Dimensions and Measures. Once again, you can change this manually, though that may depend on the type of data. Tableau may also create some additional default measures, such as the number of records in the dataset.

Tableau allows you to quickly get started by suggesting graphs for the data that you're interested in. You just select the dimensions and measures that you want (on the left in Figure 3 below), and it highlights the appropriate graph types in a legend (top right). You can also hover your mouse over a graph type and it tells you what combination of measures and dimensions are required for that type of graph (bottom right).

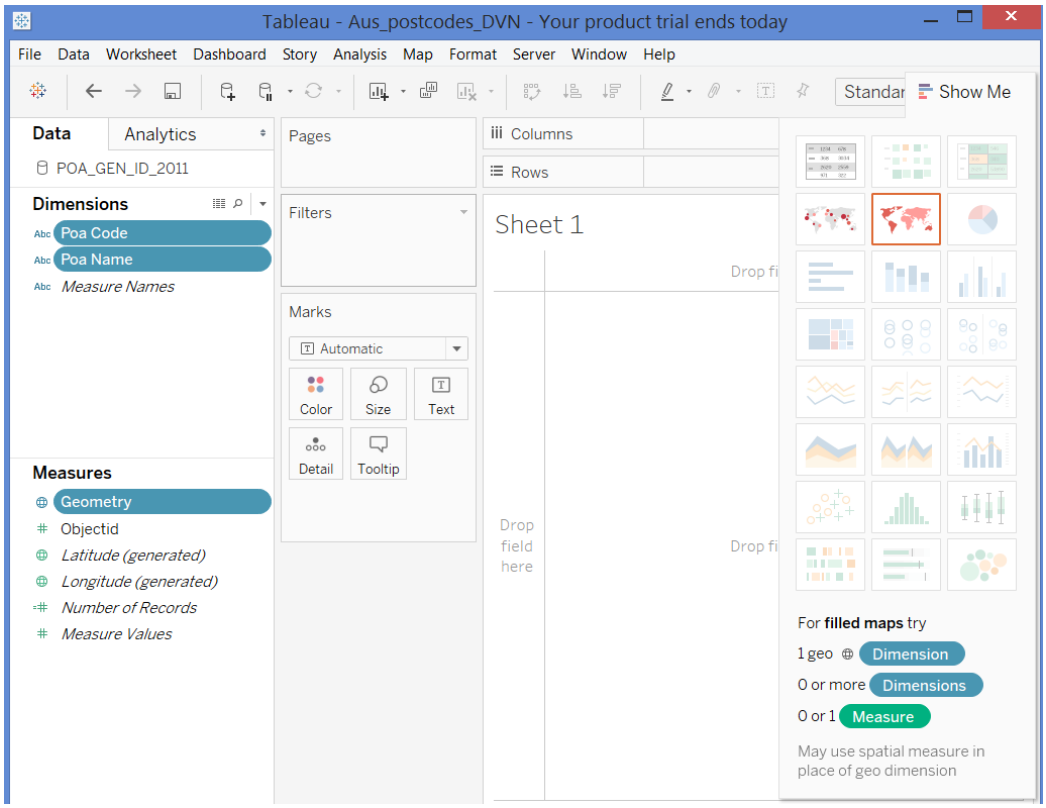


Figure 3: Choosing data and a graph type

This allowed me to create a rather plain map showing the postcode boundaries in Australia. I was easily able to add some detail by dragging the “Poa Code”and “Poa Name” fields to the Detail box in the Marks area in the middle of the sheet (note that the Detail box only appeared in Marks after a map type of graph was chosen).

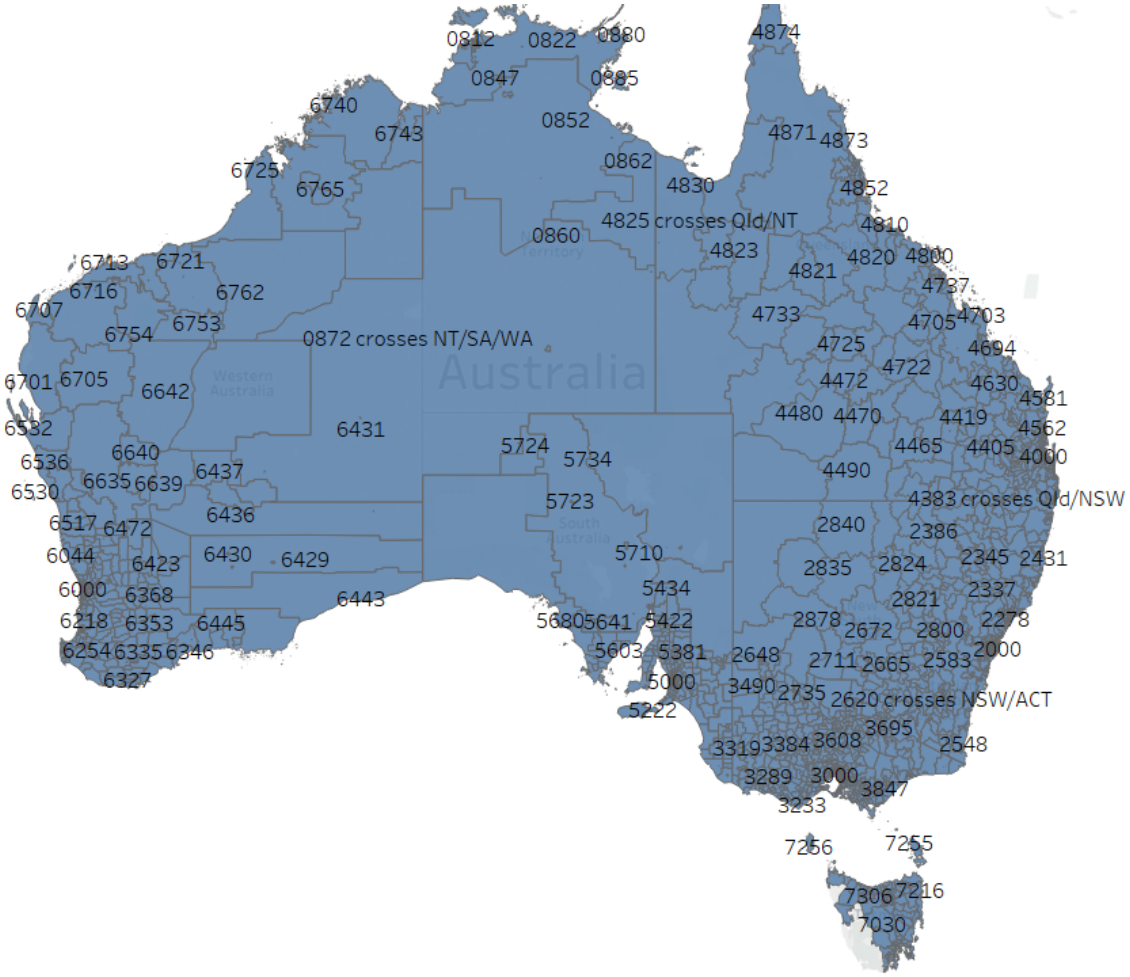


Figure 4: A basic map with labels

To try to add some more interesting information to the map and to explore more of Tableau’s capabilities, I went to the ABS (Australian Bureau of Statistics) website and downloaded some more data on postcodes, including area, suburb names and state. In Tableau, I went back to the Data Source tab and chose to add a data source, then opened one of the files I had downloaded. Tableau automatically prompts you to create a join for the files. I chose the two columns I wanted to join, but suspected something was wrong as there was no data showing in the pane below, as there had been in Figure 2 above.

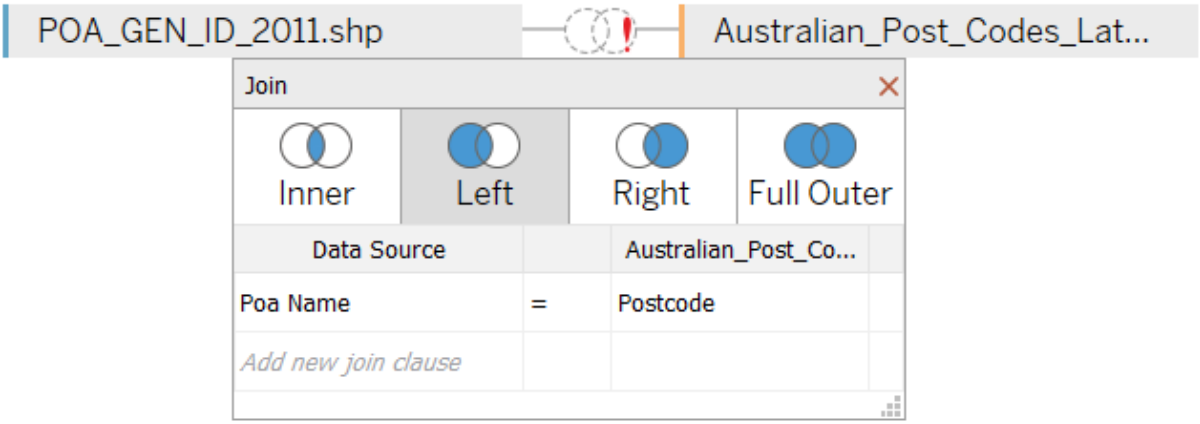


Figure 5: Creating a table join

On returning to the worksheet tab, an error appeared, telling me that the columns chosen to be joined must have the same data type.

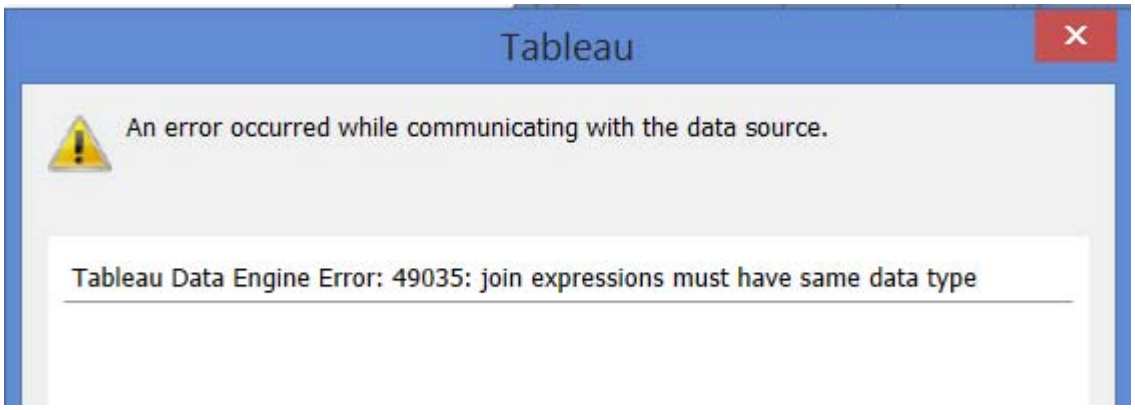


Figure 6: Data join error

I returned to the Data Source tab and saw that the Postcode column from the second file was recognised as a number (Tableau actually assigns it a sort of dual role as a number, but also recognises it as geographic data and assigns it another role as a postcode). When this was changed to a string, Tableau was able to complete the join.

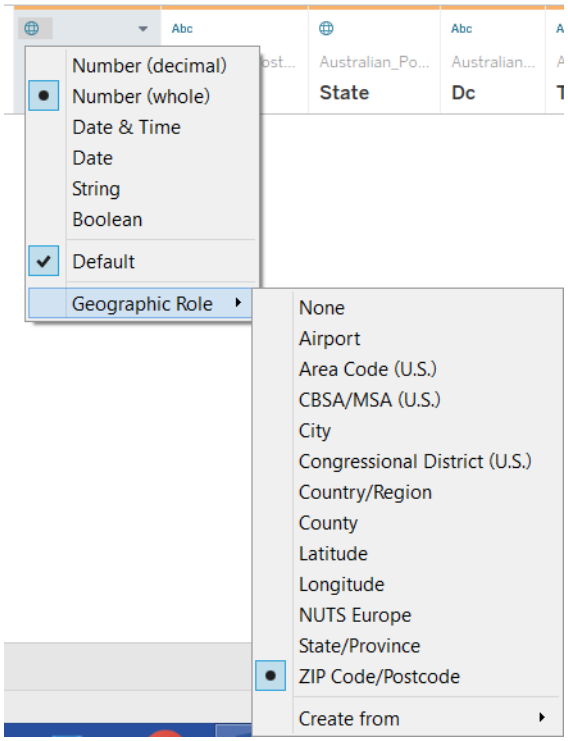


Figure 7: Postcode's geographic role

I added suburb names to the postcode numbers as a label on the graph, and then coloured each postcode according to area, by dragging the area measure to the Color box in the Marks area. The map was too small to see much detail, so I chose to make the State dimension available as a filter, and limited it to Queensland only. I then decided to add the latitude as a second filter to enable me to filter out Brisbane and the rest of the south. Not happy with the default colour scheme, I was able to choose another by clicking on the Color box. Map labels for North, Central and Southern QLD were added.

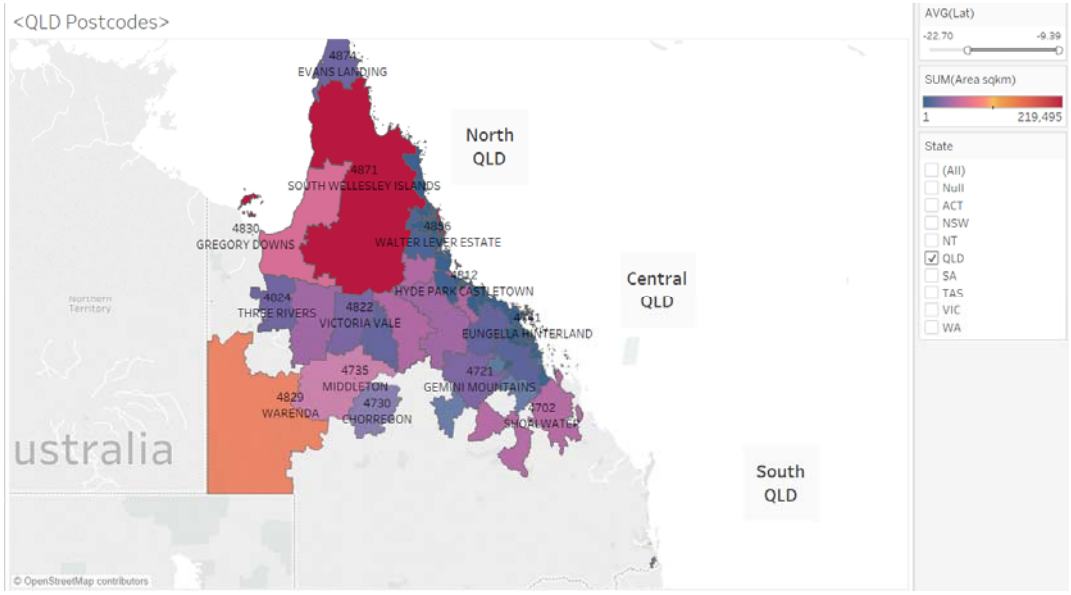


Figure 8: North and Central QLD postcodes with area

Deciding that I wanted a similar map of northern Australian postcodes for WA and NT, I was able to copy the worksheet to two new sheets and change the filters and titles, so that it took only a matter of minutes. I was then able to create a dashboard and copy the three sheets to it, so that I could see all three maps at once.

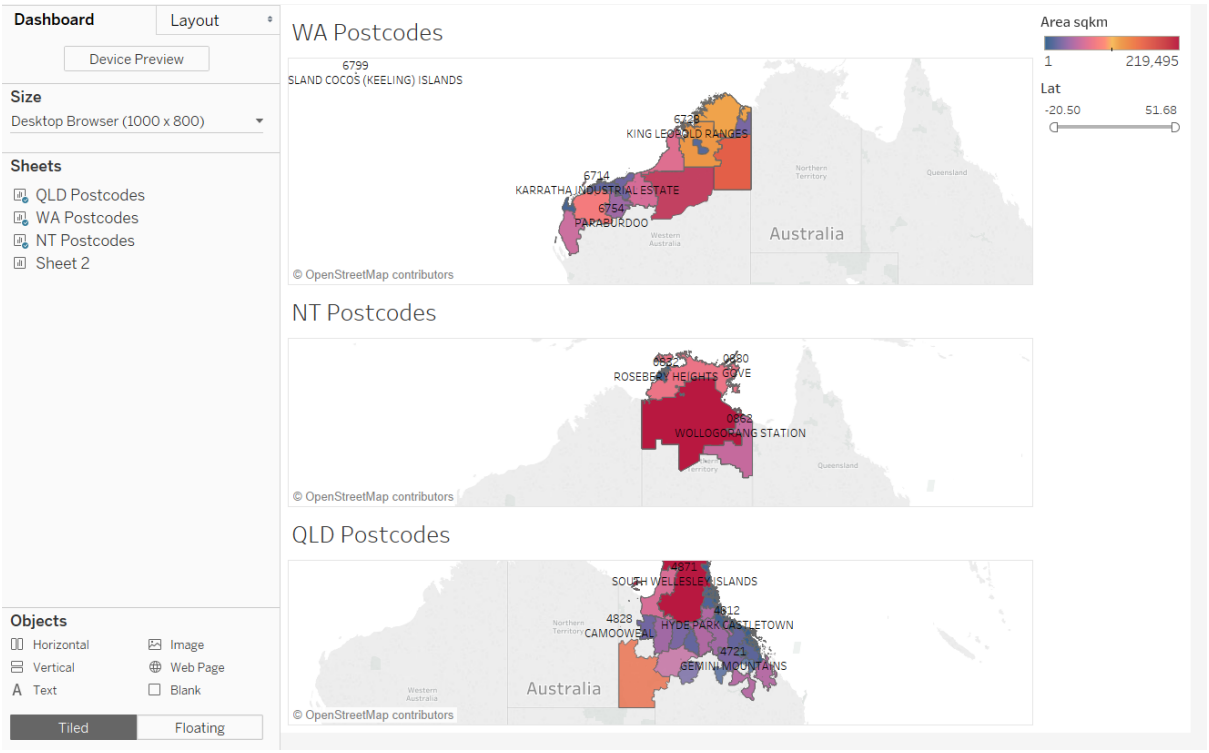


Figure 9: Dashboard showing maps of northern Australian postcodes in WA, NT and QLD

At this point, I felt I had done enough with the maps and went on to try creating graphs from some of the other data columns (see below for some examples). While the fact that Tableau does not make you explicitly spell out each step to create graphs is its strength, it can sometimes be difficult to get it to do exactly what you want. I found that Tableau would sometimes move a field from rows to columns or vice versa when I changed the graph type. It would often default to using the sum of the chosen field, or I would need to change it from a measure to a dimension to get it to do what I wanted. Also, despite its ease of use, a tool like Tableau does not have the same flexibility as you have with a programming language such as R or Python. Any complicated analyses have to do be done outside of Tableau and the results imported as a data source.

However, the difficulties outlined above could well be surmountable with more training and practice. Overall, I think that Tableau is well suited to quickly creating visualisations using clean and well prepared data. It is easy to view the raw data by simply dragging it into a table. You can add new calculated fields inside Tableau using the data imported from your data sources. You can combine information and easily add features such as labels, colour schemes and trend lines, change graph types and make an axis logarithmic. Perhaps best of all, it is easy to create maps using geospatial data with very little training required of the user.

Columns: State

Rows: Suburb

Sheet 2

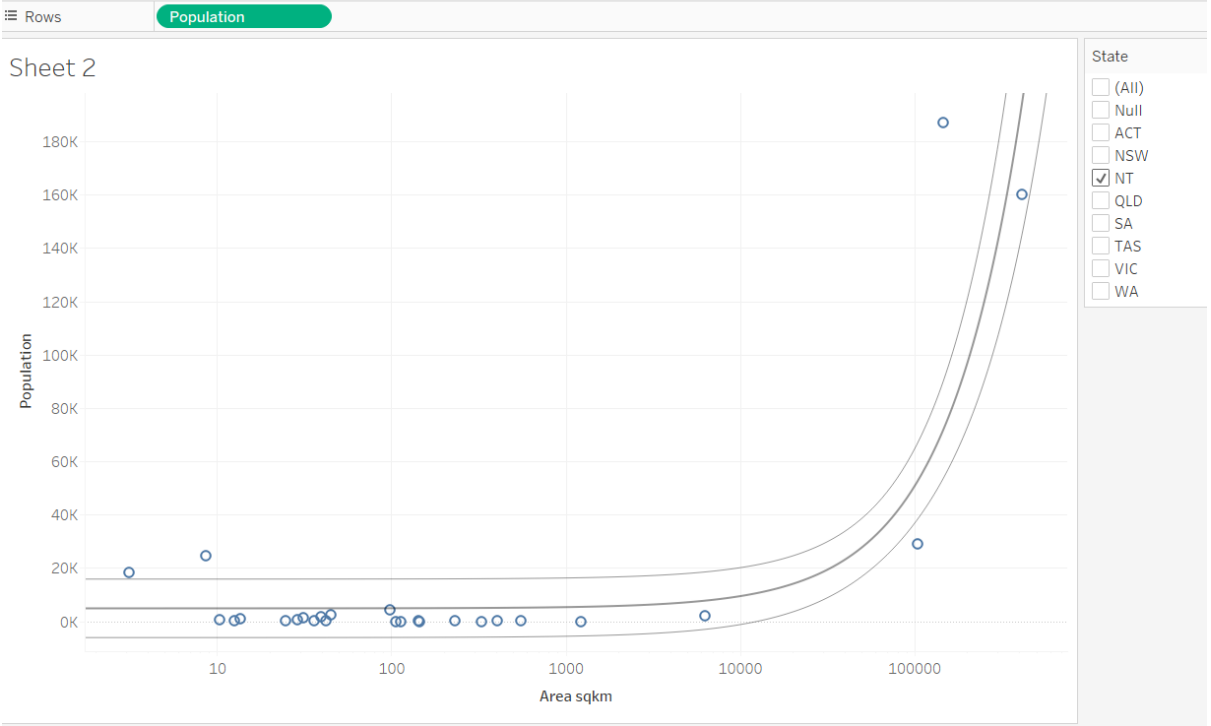
	State								
Suburb	Null	ACT	NSW	NT	QLD	SA	TAS	VIC	WA
Null	1,320,951								
AARONS PASS			5,397						
ABBA RIVER									1,195
ABBEY									1,195
ABBEYARD								1,873	
ABBEYWOOD					777				
ABBOTSBURY			20						
ABBOTSFORD			6		2,091			2	
ABBOTSHAM							511		
ABELS BAY							179		
ABERCORN					1,015				
ABERCROMBIE RIVER			5,079						
ABERDARE			1,511						

State	Area sqkm	Population	Popn/SqKm
Null	1,320,951	305,926	45,081
ACT	10,041	657,996	76,580
NSW	12,658,704	12,659,004	1,115,907
NT	7,839,159	6,149,833	10,058
QLD	39,548,001	27,288,771	253,347
SA	10,453,883	7,457,206	1,096,367
TAS	952,104	1,185,969	99,078
VIC	2,326,651	6,457,027	862,080
WA	14,388,337	7,682,996	398,286

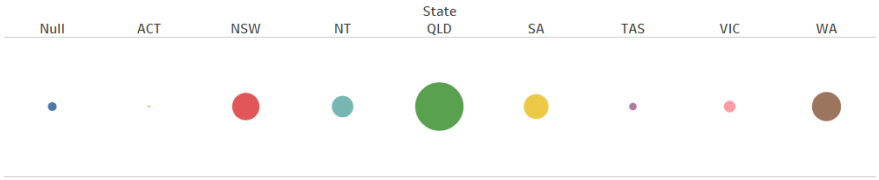
Columns: Area sqkm

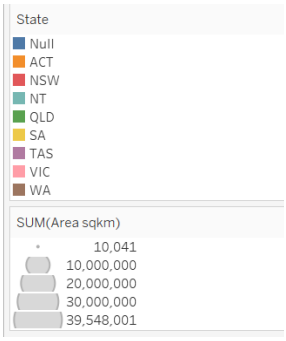
Rows: Population

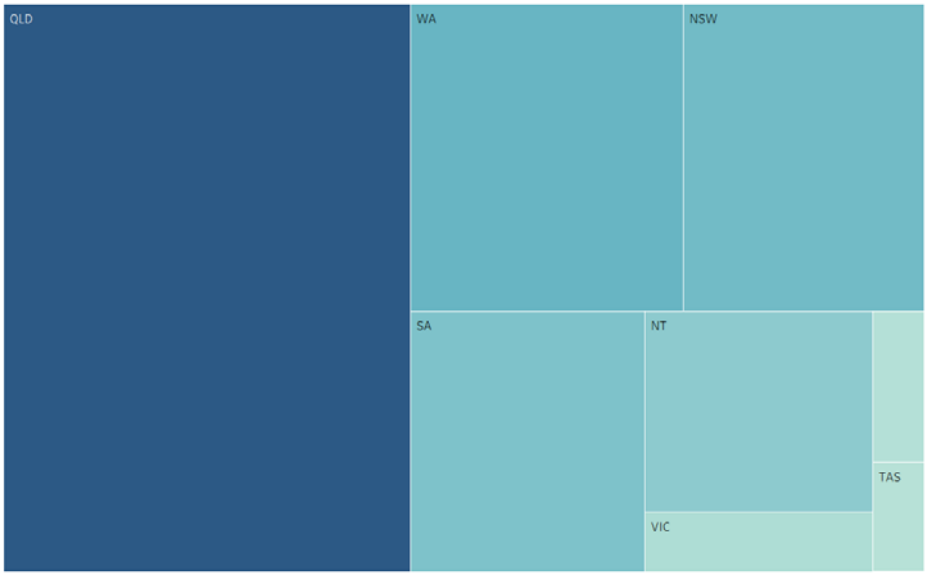
Sheet 2



Sheet 2







Sheet 2