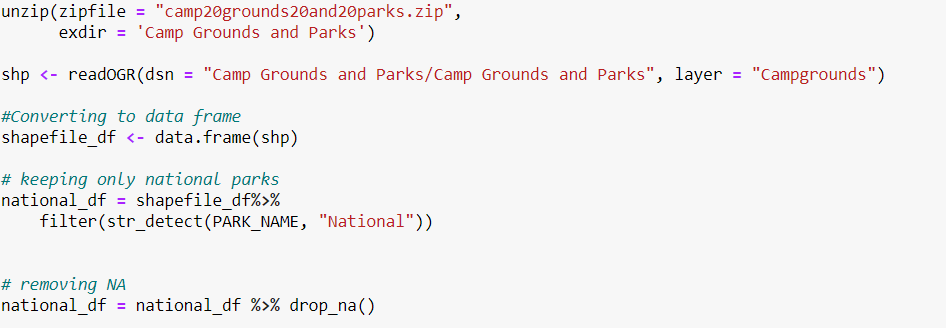
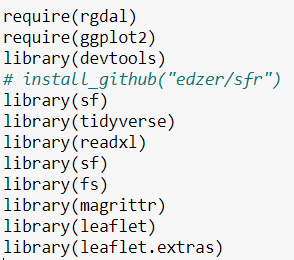
**Open Data List:**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **No.** | **Names** | **Link** | **Description** | **File format** | **Frequency of source updates** | **Frequency of Iteration** | **Granularity** | **Copyrights/licensing details** |
| 1 | Parks Victoria Camp Grounds and Huts (GovHack 2016) | <https://discover.data.vic.gov.au/dataset/parks-victoria-camp-grounds-and-huts> | This dataset shows discrete huts and visitor campgrounds within Parks and Reserves managed by Parks Victoria | Shapefiles | N/A | N/A | shapefile with multiple fields of information | State of Victoria (Department of Environment, Land, Water and Planning  Licence under  [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) |
| 2 | Venomous bites and stings 2017–18 | <https://www.aihw.gov.au/reports/injury/venomous-bites-and-stings-2017-18/data> | Hospitalised due to contact with a venomous animal or plant data | Excel | N/A | - | high granularity | Australia Institute of Health & Welfare  License under Creative Commons BY 3.0 (CC BY 3.0) |
| 3 | Maps of Victoria hospital Location | <http://www.health.vic.gov.au/maps/> | Victoria Hospital | KML | annually | - | detailed fields with common names species and zones | Department of Health  Licence under  [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) |
| 4 | Animal distribution | <https://doi.ala.org.au/doi/87720af4-6772-4159-8dbd-6fdbe76d827f;jsessionid=EA3E4FCB4D0B8CC128CDB685E81E5FCF> | Animal distribution map in AU | CSV | N/A | - | - | Atlas of Living Australia Contact  License under Creative Commons Attribution (Australia) (CC-BY 3.0 (Au)) |
| 5 | FOI - Index Centroid - Vicmap Features of Interest | <https://discover.data.vic.gov.au/dataset/foi-index-centroid-vicmap-features-of-interest> | A dynamic Victorian database of features of interest such as emergency facilities | Shapefiles | - | - | shapefile with multiple fields of information | State of Victoria (Department of Environment, Land, Water and Planning  Licence under  [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) |

## Open Dataset No 1:

The dataset Parks Victoria Campgrounds and Huts (GovHack 2016) is downloaded from <https://discover.data.vic.gov.au/> which is a open data set from the Victorian govt. website. It contains the location data, addresses and other information about various campgrounds in Victoria. The dataset is downloaded as a zip file which contains the shapefile information. This dataset is then wrangled using R and then used for plotting on google maps for the map visualization.

The wrangling (data cleaning) process is explained below:

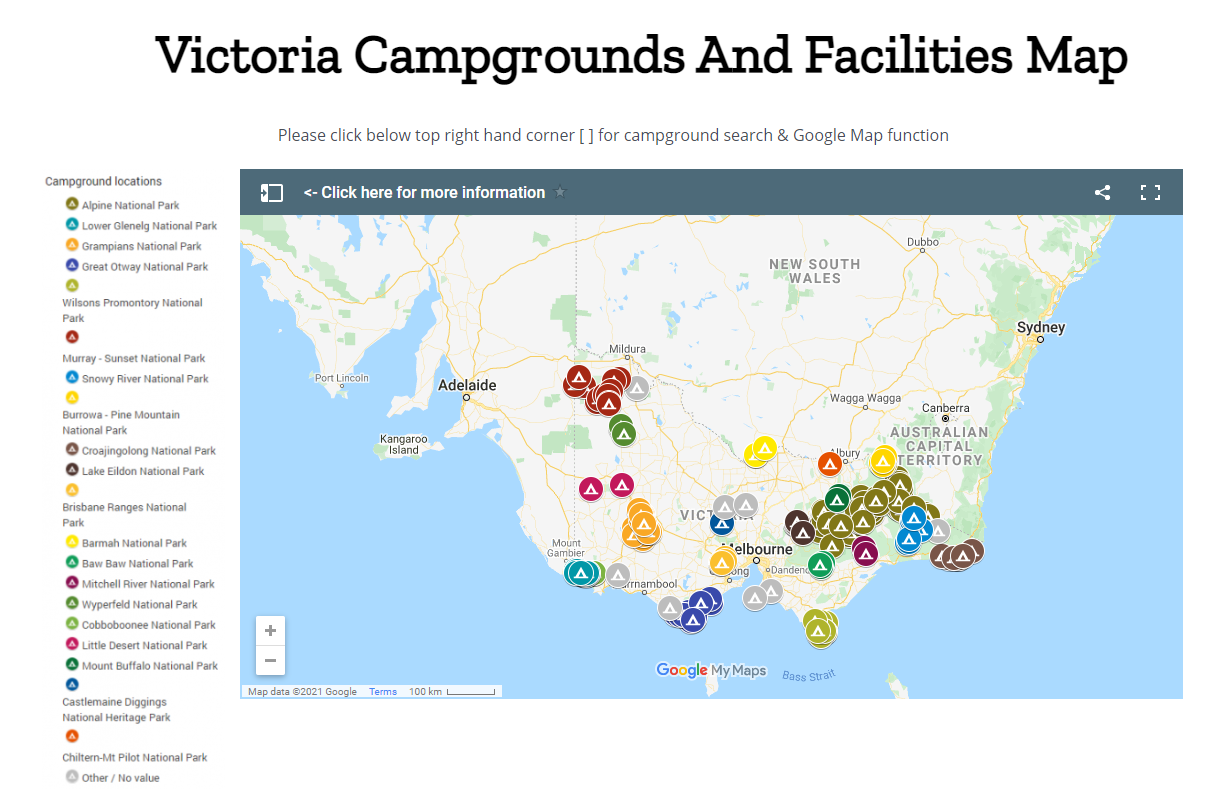


Steps for above:

* After downloading the zip file, unzip it and declare the folder to save the data.
* Use readOGR function to extract the shapefile information
* Convert the shp file to a data frame and filter out only the national parks campgrounds.
* Remove nulls form the data using drop NA.



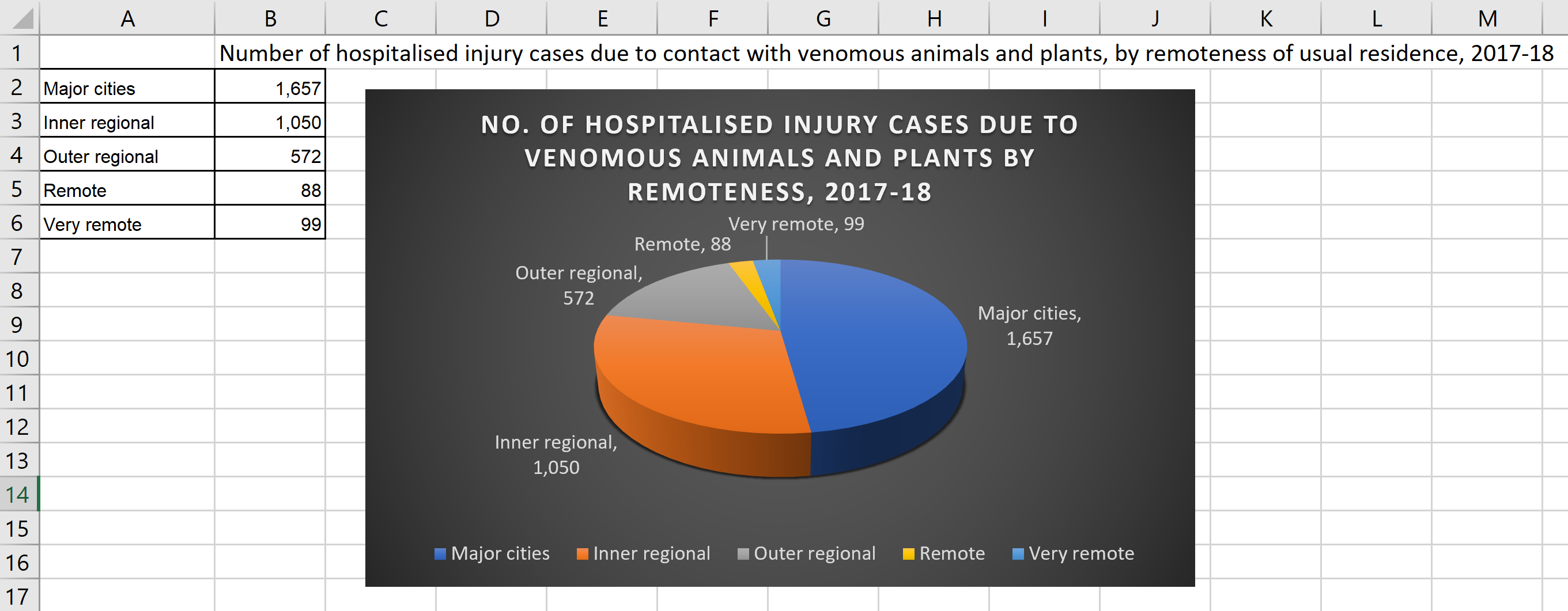
* Keep only the required columns for our purpose.
* Rename a few columns and replace some null fields with ‘NONE’.
* Write the data frame into a CSV which is used in the visualization.

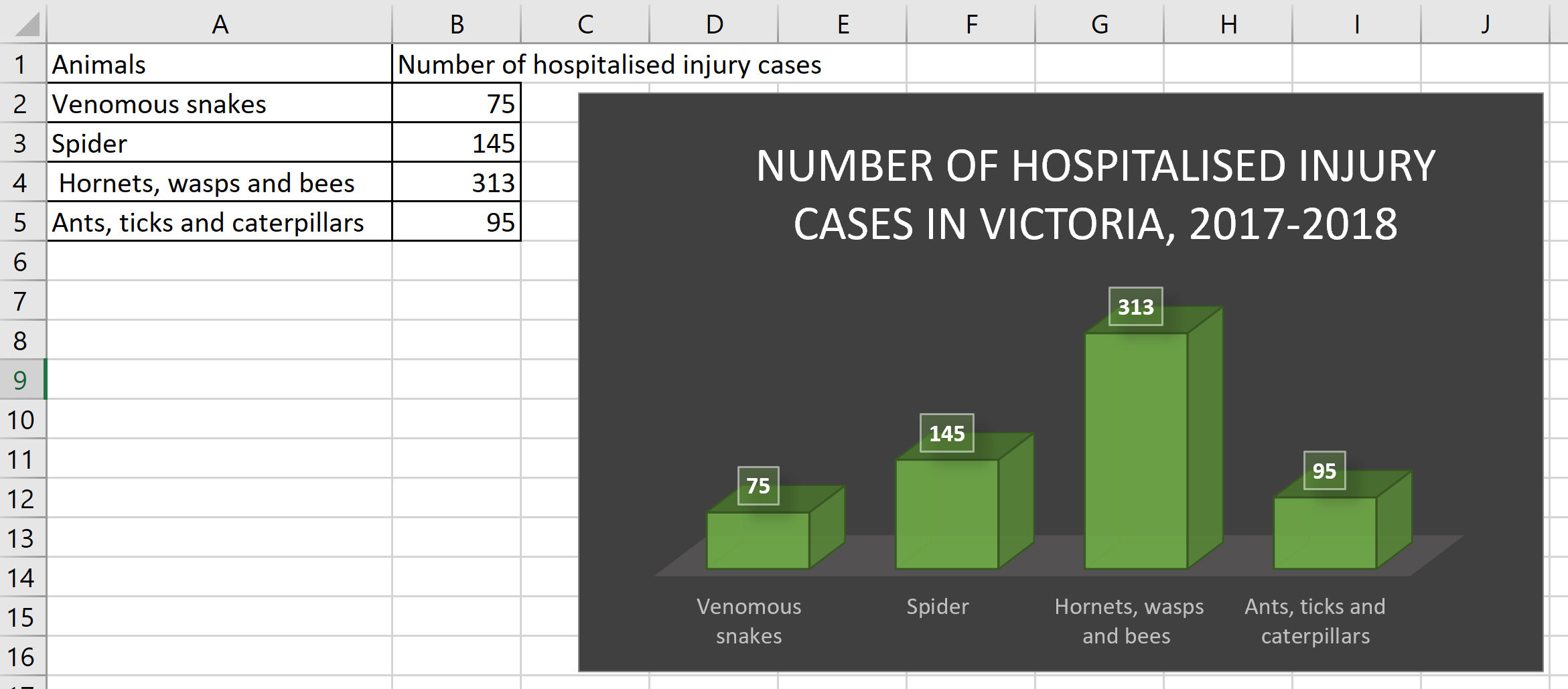


The final csv is used in google Maps for the visualization.

## Open Dataset No 2:

For the dataset - Venomous bites and stings 2017–18, this is a PDF report from Australian Institute of Health and Welfare which released in March 2021. The dataset can be downloaded as Excel file. To make 2 charts, we just need to consolidate the data and generate the chart in Excel as below.





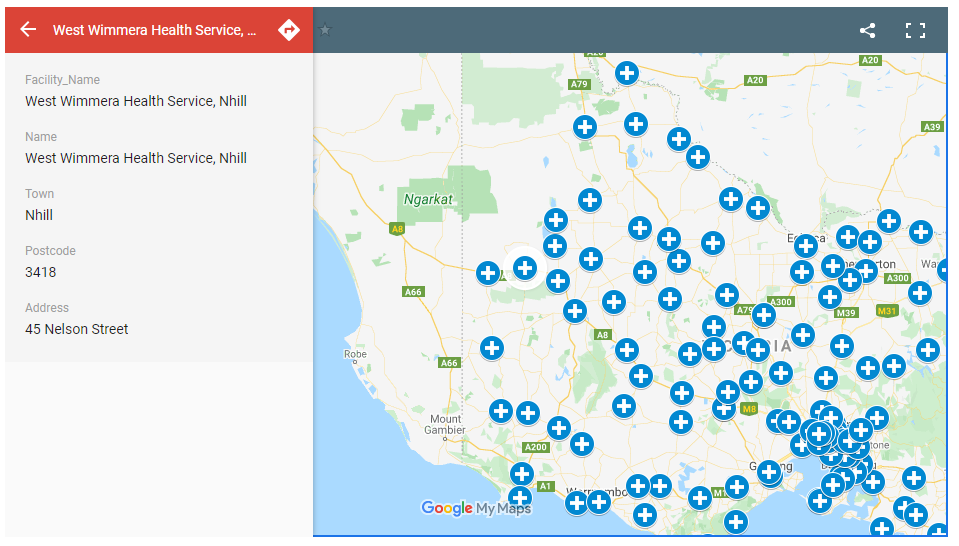
As the report does not mention how often the data and report will be collect but their previous release was more than a decade, in 2008. Therefore, it is reasonable to believe there will not be any near future release. This dataset does not get loaded into our database as we just based on the data to produce 2 charts on the Dangerous Animal Page and part of the input for the interactive video.

## Open Dataset No 3:

This dataset is collected from the Victorian govt. health information website (http://www.health.vic.gov.au/maps/). It contains the locations of hospitals with in Victoria. The file is downloaded as a kmz file and required wrangling to convert to proper format for visualization purposes.



Steps:

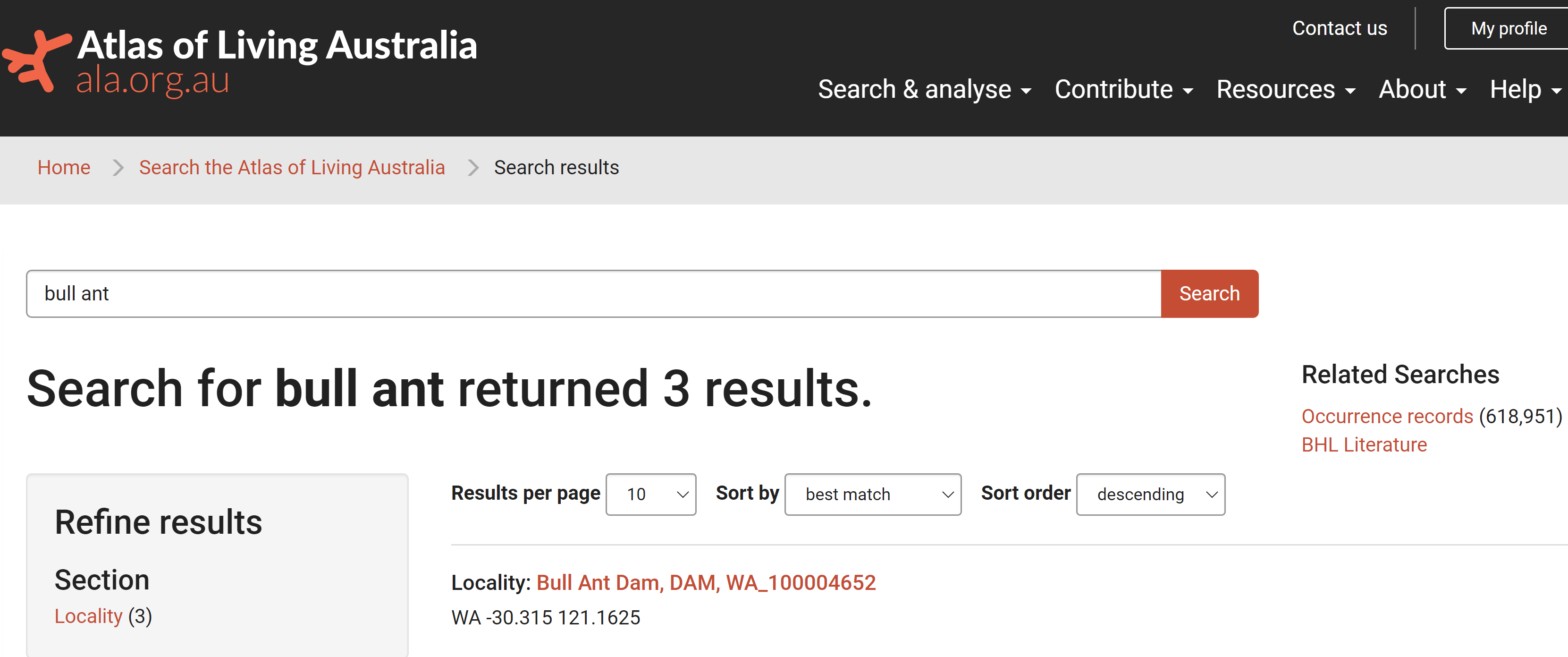
* Renaming the kmz file by adding .zip, then extract to get the doc.kml file.
* Convert the kml file to csv format using the mygeodata.cloud website.
* Unzip the file form the website and read it into R using read.csv function.
* Select the required columns and rename the latitude and longitude column names
* Write the data frame into a csv for use in the visualization.

The final hospital file contains location information along with facility names, town name, postcode and address which is displayed in the visualization.

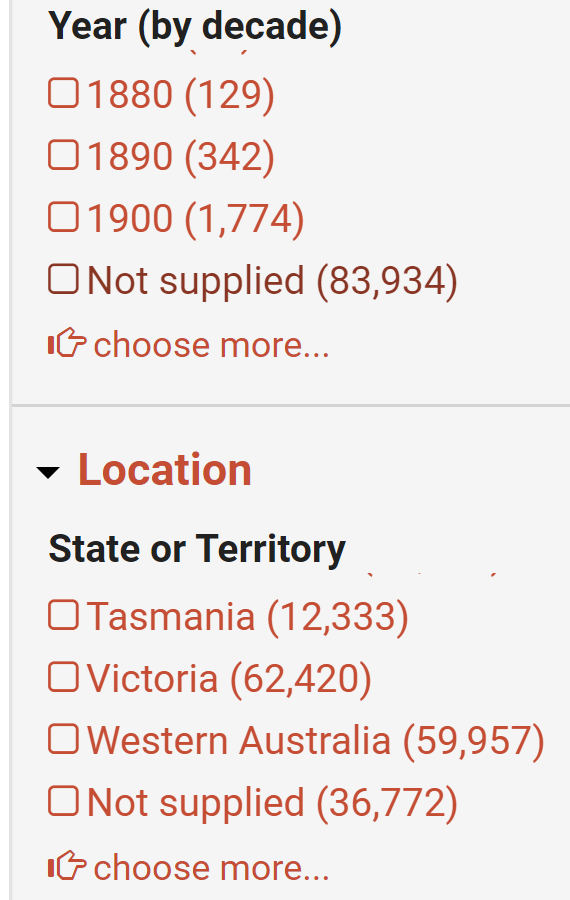
## Open Dataset No 4:

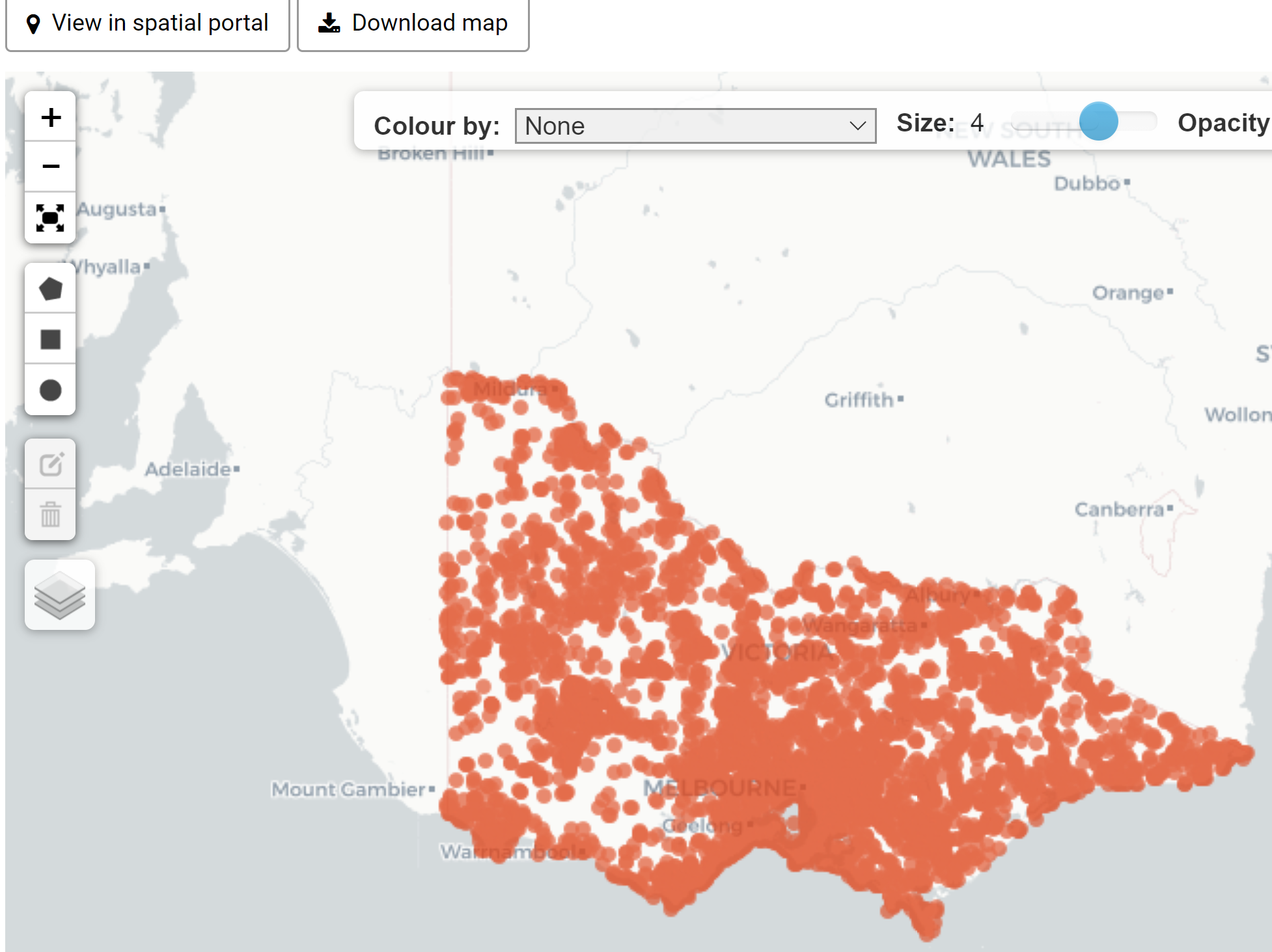
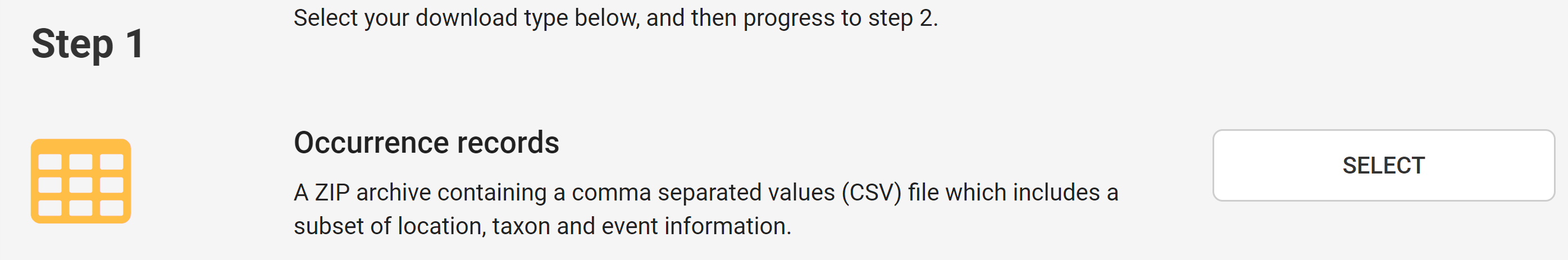
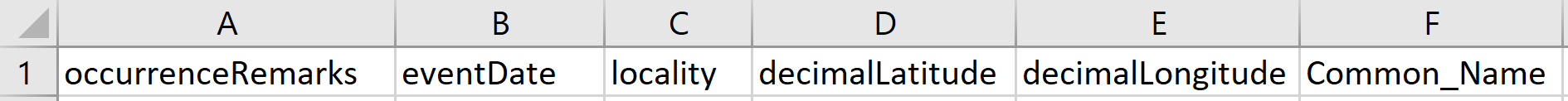
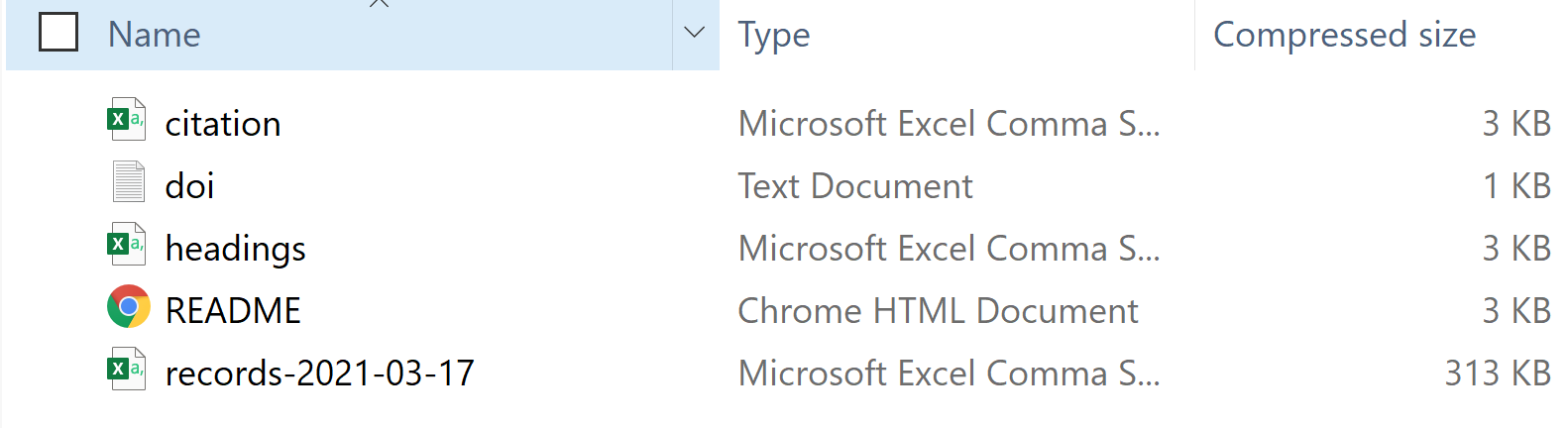
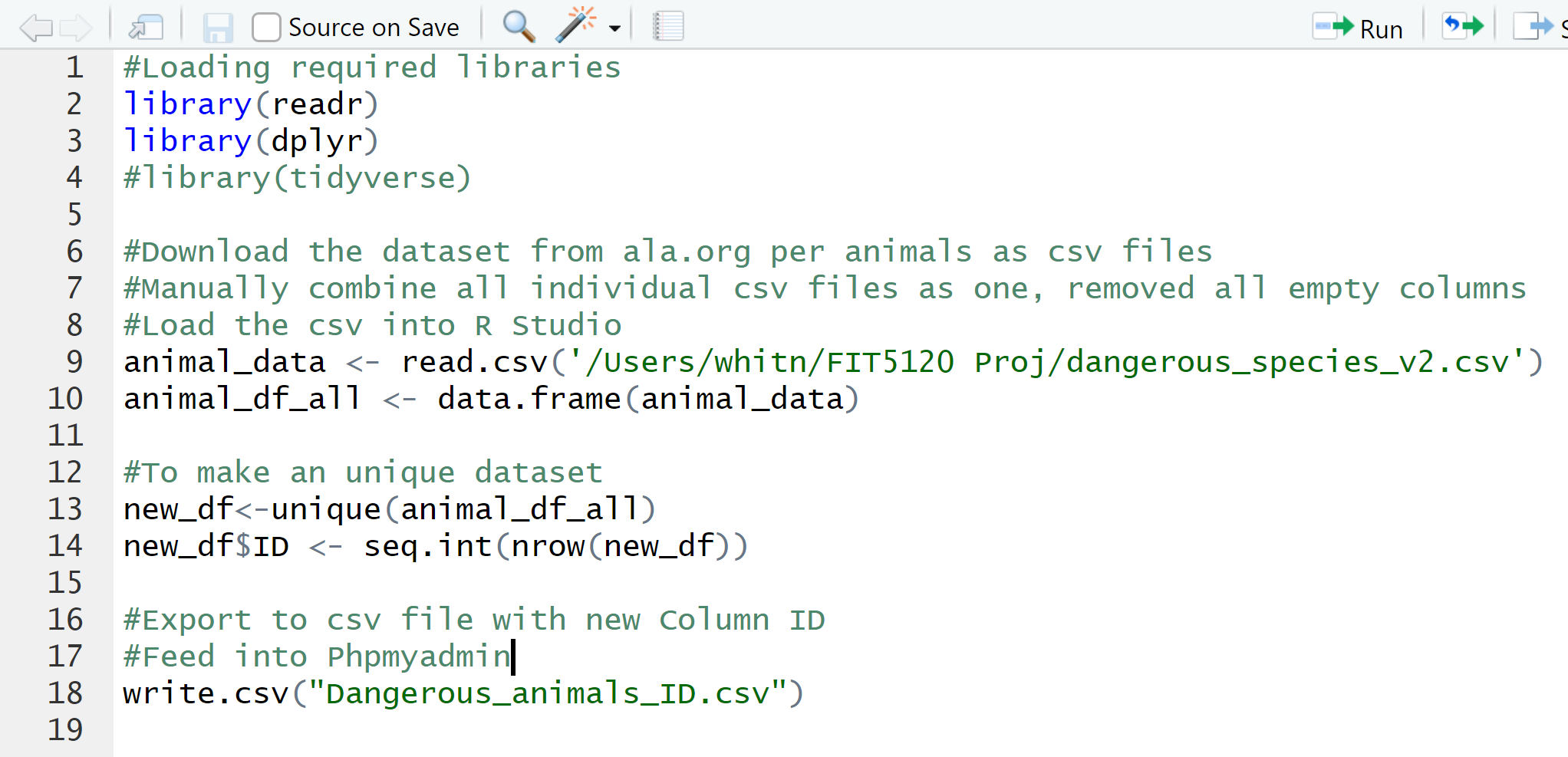
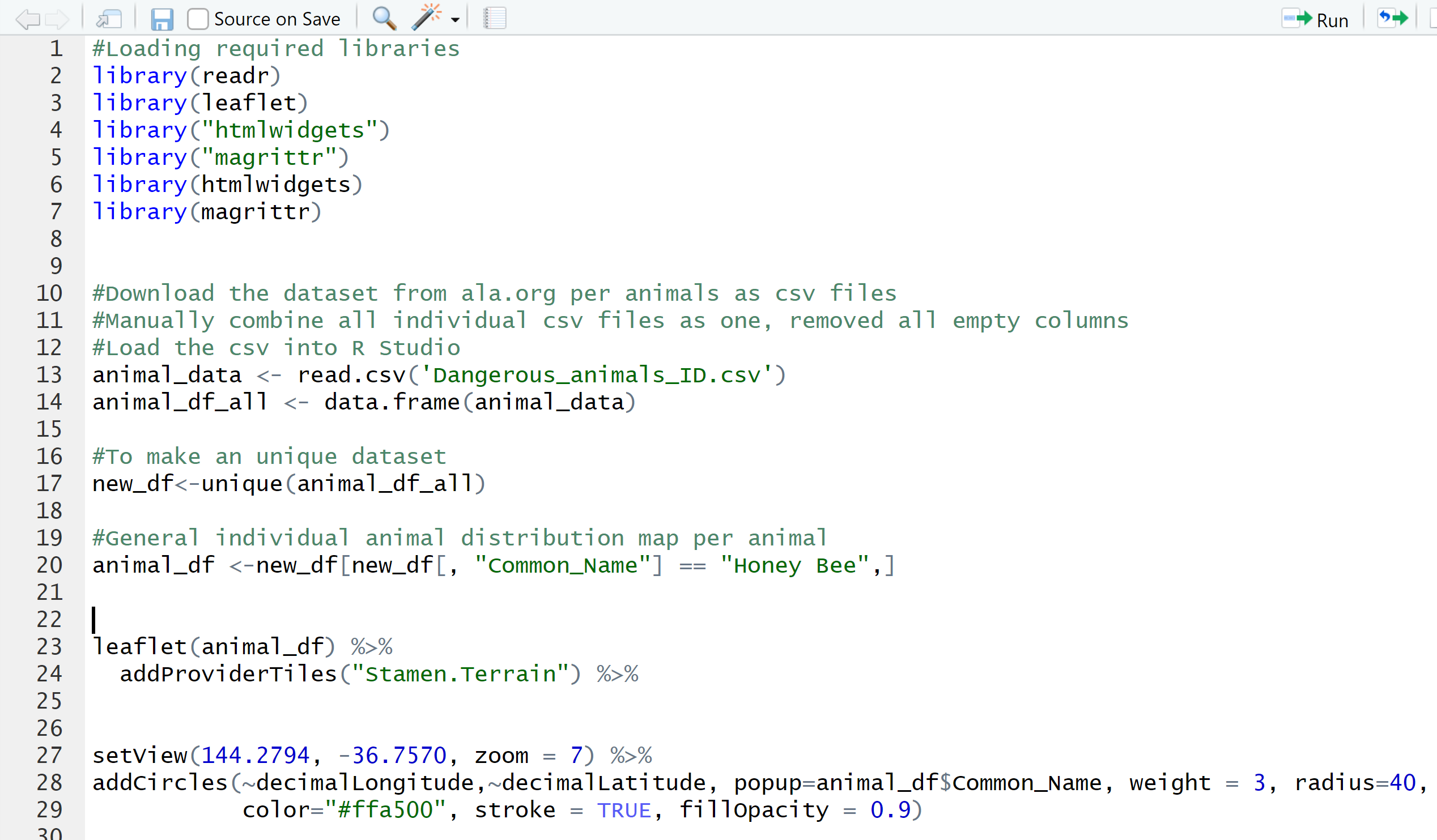
To produce the distribution map for each animal, there are 10 csv files to be downloaded and process. As the data source is based on report of their occurrence which are not a one-off data, would recommend refreshing those distribution maps annually with below steps:

1. Go Atlas of Living Australia <https://www.ala.org.au/>
2. Search for individual animal according to the Dangerous Animal Rating table
3. Click Occurrence records under “Related Searches” (see below)



1. Filter as latest decade or unknow and Victoria to retrieve only latest Victoria data (see below)



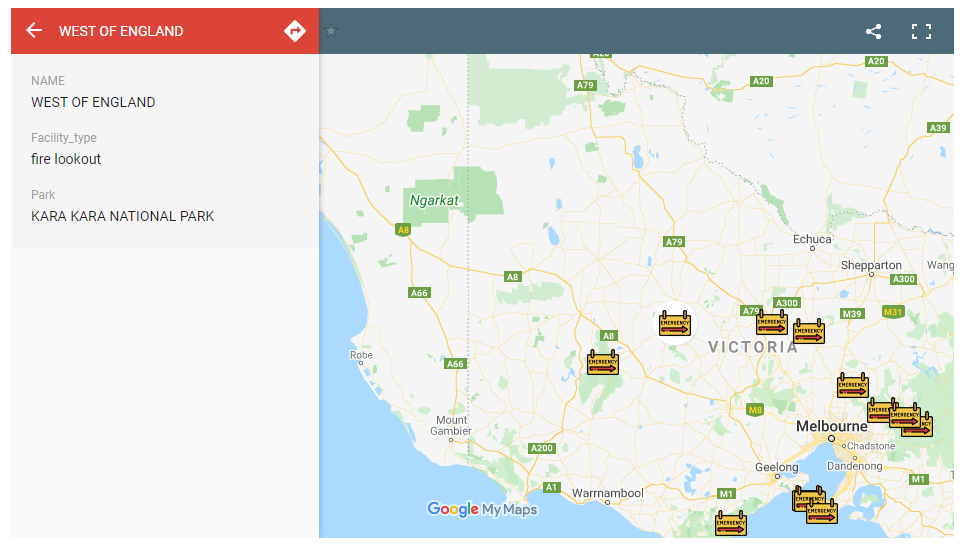
1. Then a map should only show data in Victoria, click the download button. (see below) 
2. Download the zip file and the Exile Data file is with name “records” and the download date. 
3. There are 58 columns, but we only need below 5 columns (A to E) instead. (see below)
4. As we need to do this exercise for 10 animals, we manually added a column F as “Common\_Name” which is the animal’s name.
5. We need to combine all Excel files to one super file and save it as a csv file for further data cleaning and processing in R.
6. We need to delete duplicate rows and then created a new column “ID” as the primary key with unique sequential numbers. Then the updated csv file can be loaded into phpMyAdmin 
7. To make the distribution map by R, we need to retrieve the data as of their individual name for the field “Common\_Name” (see below for Honey Bee) 

Open Dataset No 5:

FOI - Index Centroid - Vic map Features of Interest is downloaded from <https://discover.data.vic.gov.au/> website which is open data from the Victorian government. It contains a dynamic Victorian database of features of interest such as emergency facilities. We extracted the emergency facilities data from here. Same as the other datasets, R was used to clean the data for visualization purposes.

Steps:

* Unzip the Vic map features of interest zip file downloaded from the website.
* Read the shapefile info using readOGR function.
* Convert the shapefile into a data frame and filter out the data where facility type is “emergency facility” and parentname (type of park) is “NATIONAL PARK”.
* Keep the required columns of data and rename some columns for clarity and better understanding of column names.
* Write into a CSV file for the map.



The emergency facilities within Victoria national parks are visualized in the map. The final csv file contains location information as well as name, facility type and name of the park the facility is located.