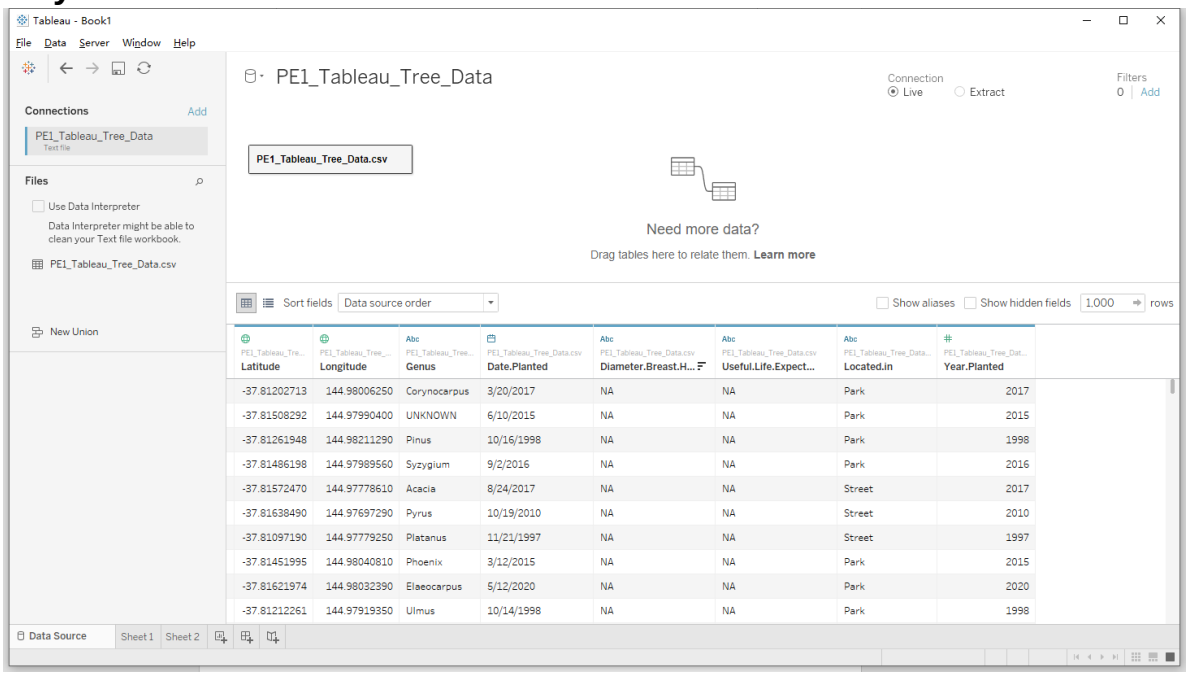
Programming Exercise 1- Tableau Public

Student name: Yue Zhang
Student ID: 30976316
Tutorial Number 2
Tutor Name Benjamin Lee, Xiaojiao Du

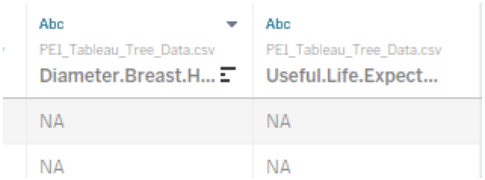
Data Cleaning and Checking

- **Read the data into Tableau and provide an image to show what your data looks like.**

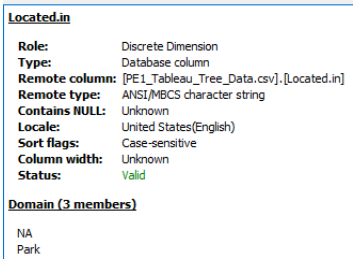


- **Check data for possible errors.**

1. There are NA values in Diameter.Breast.Height and Useful.Life.Expectency.Value.



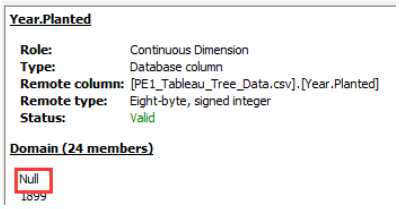
2. There are NA values in located in.



3. There is inconsistency between data planted and year planed, in some rows, the year in date planted is not equal to the year in year planted Value.

	Date.Planted	Diameter.Breast.H...	Useful.Life.Expect...	Located.in	Year.Planted
	1/1/2000	94	60	Park	1900
	1/1/2000	87	5	Park	1900
	1/2/2000	84	30	Park	1900

4. There is a null in year planted.



5. There is a '2061' value in year planted.

	Latitud	Longitu	Genus	Date.Pl	Diamet	Useful.	Located	Year.Pl	ed
57	-37.8153	144.9796	Acacia	2/9/2061	NA	NA	Park	2061	

6. When I doing the exploration process, I found there is one tree that is in sea area according to its latitude and longitude, which means that the latitude and longitude is invalid in this situation.

- **Brief explanation on your process**

For the possible errors founded in previous step:

1. Do not change anything because these two lines are not used in the current data exploration.
2. The same as 1.
3. For the inconsistent values, I found that when the year in Year.Planted is not equal to the year in Date.Planted, the value in Year.Planted is 1900 and 1899, which are not valid values, so I have made the Year.Planted the same as the year in Date.Planted to keep the consistency.

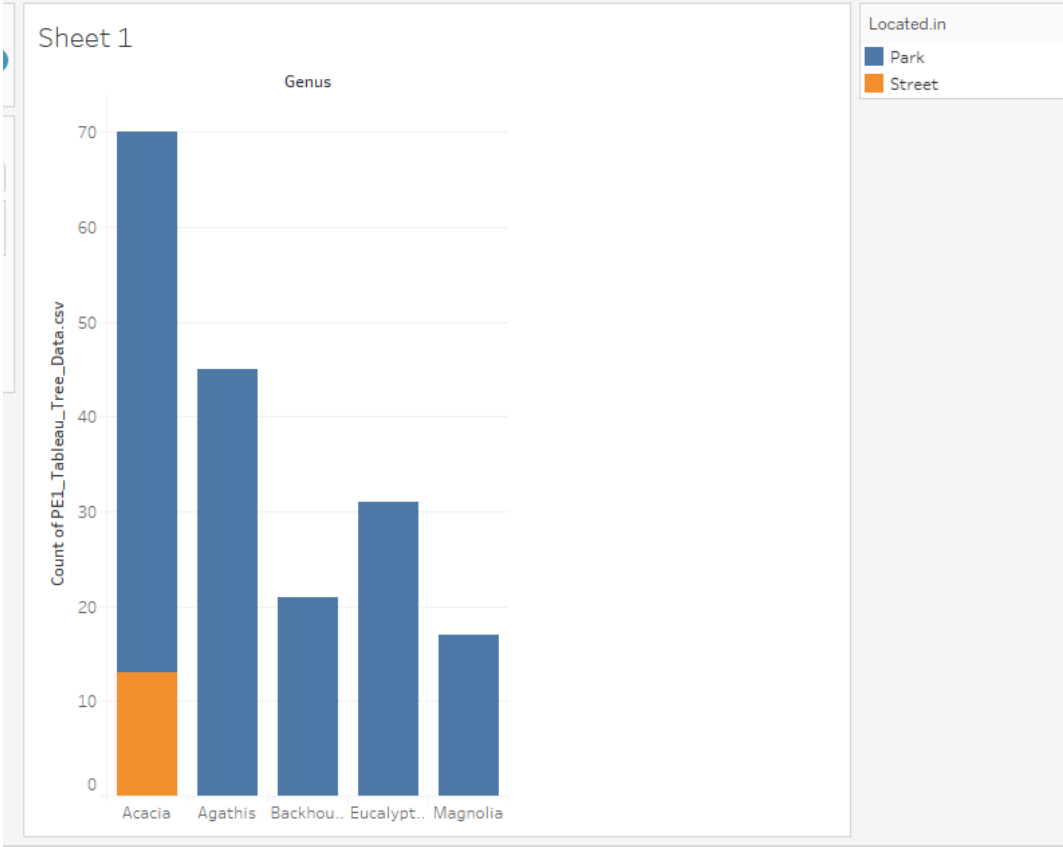
.8135	144.9763	Populus	1/1/2000	87	5	Park	1900
.8157	144.9806	Syzygium	2/1/2000	54	60	Park	1900
.8142	144.9746	Populus	2/1/2000	80	5	Park	1900
.8113	144.9781	Corymbia	30/12/1999	20	60	Street	1899
.8127	144.9778	Corymbia	1/1/2000	30	60	Street	1900
7.814	144.9783	Hesperocy	2/1/2000	34	20	Park	1900

4. For the records with null value, I have checked the data planned, it is also null, but the trees are indeed here (it has valid latitude and longitude), so I do not delete them.

5. For the one record that has 2061 year planned, the date planned is also 2061, it is a unexpected value, so I delete that record.
6. Delete that row.

Answering Questions

- **How are Acacia, Agathis, Backhousia, Eucalyptus, and Magnolia distributed spatially?**



First, regarding weather it is at street or park, as the screenshot shown above, For Acacia, around 13 of this kind of tress are planted at street, and about 57 of them are planted at the park. For the rest of other 4 kinds of tress, they are all planted in the park.
In terms of the spatial distribution, the figure below illustrates the spatial cluster for those 5 kinds of trees. To help illustrate the distribution, I have marked the map into four parks (park A, B, C, D)
For Acacia trees, they are mainly at the southern part of park B and street, with only few of them are at the north-west of park B.
For Agathis trees, they are mainly planted at park B, while few of them are at park A. Particularly, the majority of this kind of tree are planted at the south-east of the park B and near the Acacia.
For Backhousia trees, the most of these tress are at park D, with only around two of them are at park B and closed to the lake.
For Eucalyptus trees, they are mainly planted at park A and park C. Also, they are mainly surrounded at the lake in park A. However, only two Eucalyptus trees are at park B.

For Magnolia trees, only one Magnolia tress is planted at park A, and one is at park D, while they are mainly planted at park B and distributed randomly in that park.
It is interesting to note that park C only has Eucalyptus, and park D almost have one kind of trees which is Backhousia. In addition, there is a variety of tree species are located around the lake in park B.

- **How does the relative number of planted trees by the aforementioned genera (plural for genus) vary in the year 2013, 2014, 2015, 2016, and 2017?**

The figure below demonstrates the number of trees planted of the five kinds of trees that mentioned previously from 2013 to 2017.
In general, the number of Acacia trees planted were the most in this period.
For Acacia trees, it begin at 2013 and went up to around 37 in 2014, then this figure drop to about 14 in 2017.
For Agathis trees, the number of this kind of trees started at 2014 with around 1 and rose from 2014 to 2015 to 4. Furthermore, after 2015, no Agathis tree was planted.
For Backhousia trees, only one has been planted in 2013, and there was no more plant after 2013.
For the number of Eucalyptus trees, it starts at 2 in 2015 and decreased to 1 in 2016, then there was no Eucalyptus planted during this period.

For Magnolia trees, in 2015, initially there was about 1 tree planted in 2015, then the number of this kind of trees grew stably to around 3 in 2017.