# Title - Programming Exercise 1: Tableau Public


## Student name - Dishi Jain

## Student ID - 30759307

## Tutorial Number - 24

## Tutor Name - Shirin

# Table Of Contents

## 1. Data Wrangling

The given data was not suitable to be loaded into Tableau. Hence, by looking at the attributes I visualised that rows could represent the “Year” grouped by “Sitenames” and the columns would represent the “Corals”. I used the Kutools tool to reformat my data in Excel. I selected the data points for one coral type at once from datapoint for 2017 site 1 till 2010 site8 (example for soft corals, I selected from cell D3 till cell K10 from the given data) then clicked on the **Kutools** tab then under **Range** I selected **Transform Range** and then the option **Range to single column**. Finally I chose the cell location at which I wanted my resulting data. I manually arranged the cells for the columns Sitename, Longitude, Latitude and Year. Hence for each coral I reformatted my data to finally look like the image attached below.

When loading the data in Tableau, as seen from the image above, the data points were coming in decimal points instead of percentage values. Hence, I changed the default properties of the Measure ( i.e. soft coral, hard coral, sea pens, sea fans and blue coral) by using the drop down from the Measures section in Tableau for each coral type. Inside Default Properties I selected Number Format and then selected Percentages.

Figure 1: Data as seen in Tableau after reformatting in Excel

| Sheet6 Sitename | Sheet6 Long | Sheet6 Lat | Sheet6 Year | Sheet6 Softcorals | Sheet6 Seafans | Sheet6 Blue Corals | Sheet6 Hard Corals | Sheet6 Sea Pens |
|---|---|---|---|---|---|---|---|---|
| site01 | 143.51500 | -11.8430 | 2017 | 0.838700 | 0.473200 | 0.682300 | 0.44320 | 0.157800 |
| site01 | 143.51500 | -11.8430 | 2016 | 0.802100 | 0.471200 | 0.656100 | 0.41780 | 0.126500 |
| site01 | 143.51500 | -11.8430 | 2015 | 0.753400 | 0.468700 | 0.621300 | 0.40990 | 0.124500 |
| site01 | 143.51500 | -11.8430 | 2014 | 0.749900 | 0.413400 | 0.589800 | 0.40780 | 0.107600 |
| site01 | 143.51500 | -11.8430 | 2013 | 0.577000 | 0.408800 | 0.519800 | 0.39980 | 0.087600 |
| site01 | 143.51500 | -11.8430 | 2012 | 0.564300 | 0.385600 | 0.503700 | 0.20450 | 0.085100 |
| site01 | 143.51500 | -11.8430 | 2011 | 0.554300 | 0.391200 | 0.402500 | 0.19340 | 0.081200 |
| site01 | 143.51500 | -11.8430 | 2010 | 0.562900 | 0.373400 | 0.397800 | *null* | 0.053300 |
| site02 | 147.89800 | 18.9370 | 2017 | 0.212300 | 0.563200 | 0.232100 | 0.34890 | 0.801300 |
| site02 | 147.89800 | 18.9370 | 2016 | 0.192300 | 0.543200 | 0.213400 | 0.30120 | 0.701200 |
| site02 | 147.89800 | 18.9370 | 2015 | 0.172100 | 0.502100 | 0.187800 | 0.25320 | 0.308900 |
| site02 | 147.89800 | 18.9370 | 2014 | 0.157800 | 0.485000 | 0.177900 | 0.24890 | 0.294800 |
| site02 | 147.89800 | 18.9370 | 2013 | 0.148000 | 0.463900 | 0.173700 | 0.24770 | 0.289000 |
| site02 | 147.89800 | 18.9370 | 2012 | *null* | 0.459800 | *null* | 0.24820 | 0.286100 |
| site02 | 147.89800 | 18.9370 | 2011 | *null* | *null* | *null* | *null* | *null* |
| site02 | 147.89800 | 18.9370 | 2010 | *null* | *null* | *null* | *null* | *null* |

## 2. Data Exploration

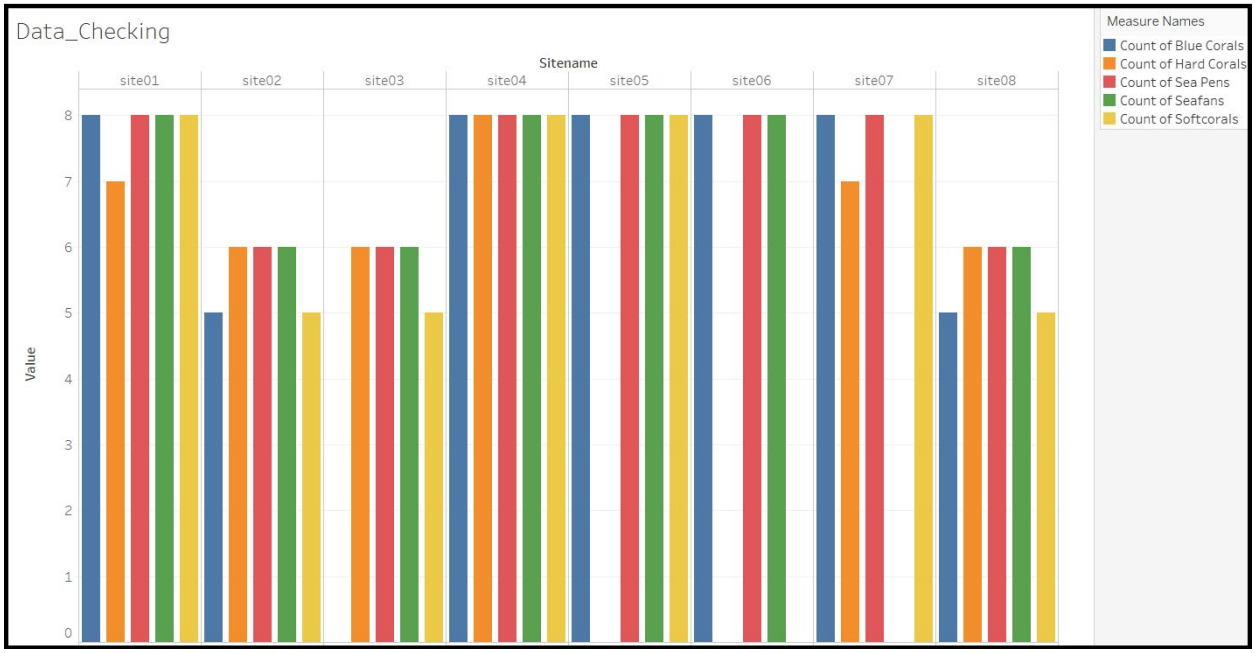
**Error 1 - Finding Missing/Null values**

For each site there should be exactly 8 values for each kind of coral representing the eight years from 2010-2017. However from the below graph in Tableau we can clearly see that there are many missing

values or nulls. Example for site01 all the corals have a count of 8 except for hard coral which has 7 values.

**Fixing the error -** Where the count is completely Null for the entire coral we can fill the values with 0% as bleaching. Example for blue corals in site03, there are 8 null values and hence we can fill the values with 0% i.e. for blue corals in site03 for each year (2010-2017) bleaching is 0%. On the other hand, where the count is greater than 0 but less than 8, we can fill the null values with the average of the remaining values(for a particular site). Hence for blue corals in site 02 and 08, for hard corals in site 01,02,03,07 and 08, for sea pens in site 02,03 and 08, for sea fans in site 02,03 and 08, and for soft corals in site 02,03 and 08 we can fill the Null values with the averages.

Figure 2: Count of values in dataset for corals for each site

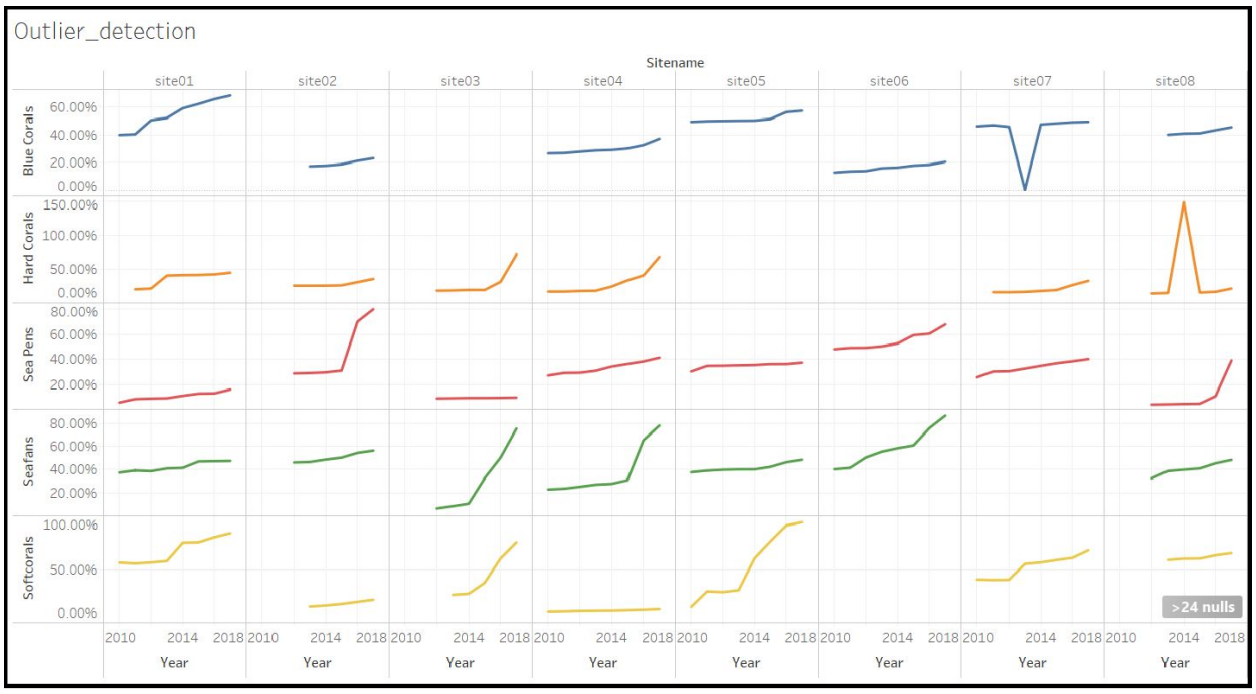



**Error 2- Outlier detection**

By plotting a line graph using dimensions Sitename and Year for the measure as the coral bleaching values, there is a detection of an outlier. For blue corals in site07, we can see a sudden dip in the value for the year 2013. For the rest of the years the bleaching value is almost similar but in 2013 the value falls steeply as seen in the graph. Similarly for had coral in sit08, we see a huge increase in the bleaching value for the year 2014.

**Fixing the error -** Outliers can be fixed by replacing the outlier values by the mean of the remaining values for that particular coral and sitename. Hence, a better analysis and visualisation can be done by fixing such errors.

Figure 3: Outlier detection by plotting Sitename and Year against coral values

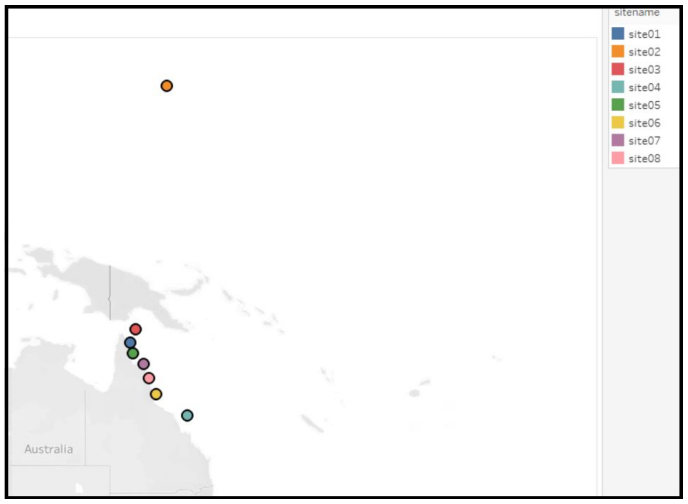



**Error 3 - Defect in latitude value for site02**

By plotting the latitude and longitude of the sites on Tableau, it can be seen that site02 is present far away from the remaining sites. The rest of the sites are closely located around the border of Australia. Hence while looking at the latitude and longitude value for site02, it was observed that just for the site02 the latitude value is a positive number. For the rest of the sites it's a negative number. Hence this is a data entry error which needs to be fixed.

**Fixing the error -** Manually fixing the error is the solution for this error. The value has been changed from 18.937 to -18.937. After fixing the error the site02 now lies in the correct position(as can be seen in the answer of Question 2)

Figure 4: Error in the location of site02

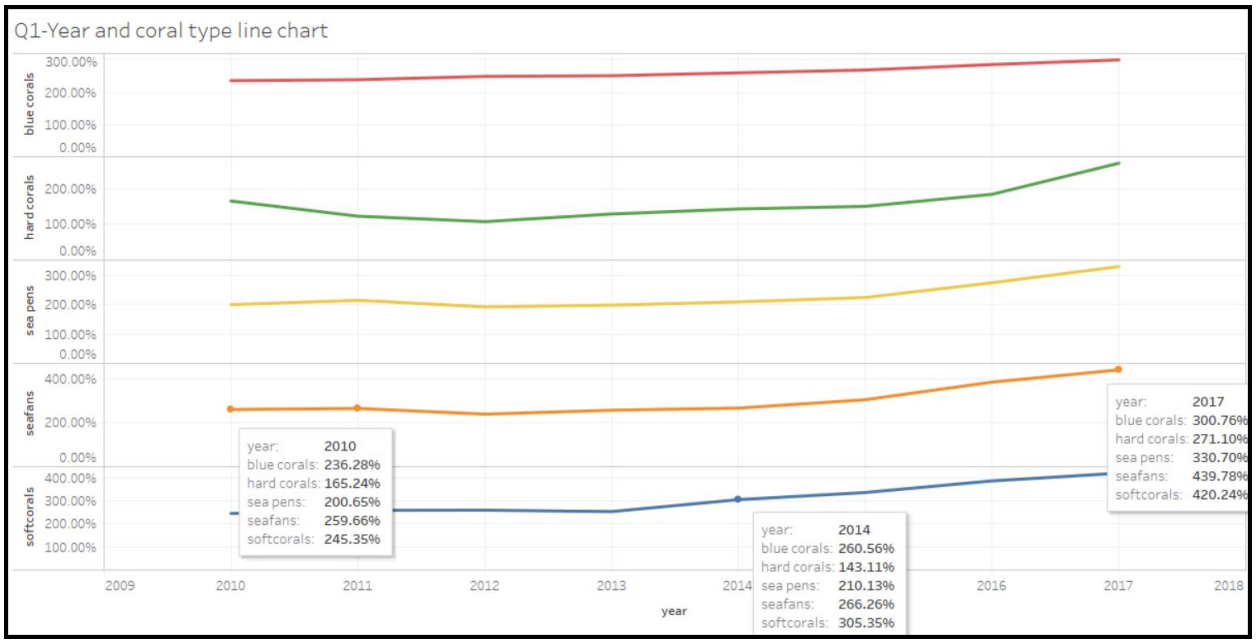


## 3. Answering Questions

Q1) In which years and for which kinds of coral is bleaching the worst?

A1) For this question, I have plotted a line chart with Year as column values and the Sum of the coral types individually year wise as the Row values. In the resulting chart, we can clearly depict in which years and for which kinds of coral the bleaching is at the maximum value hence the worst. The exact percentage of bleaching for 2010 can be seen by placing the cursor near any point in 2010 and same for the rest of the years too. Hence from the chart we can see that for 2010, 2011, 2013 and 2017 Sea Fans coral shows the maximum percentage of bleaching hence this coral was in the worst condition during these years. For the years 2012,2014,2015 and 2016 Soft Corals incurred the worst bleaching percentage.

Figure 5: Line chart showing bleaching percentage for coral types in different Years




Q2) How does the location of the site affect bleaching of the different kinds of corals?

A2) For this question, a scatter plot is generated with Columns as Sitename and Measurename(i.e. coral types) and the Rows as Measure Values(taken as the average of bleaching values for each coral site wise). For site01 the bleaching percentage of the different kinds of coral is usually high except for sea pens coral. It generates the highest percentage of bleaching for hard coral, blue coral and soft coral as compared to all other sites. For site02 blue and soft corals incur relatively less bleaching and on the other hand sea pens and sea fans are closely related with a high value of bleaching. Site03 shows the least level of bleaching for sea pens and sea fans when compared to other sites with 8.98% and 30.49% respectively. Also, this site shows no record for bleaching percentages for blue corals. In the sites 01,02 and 03, most of the data points are spread far away from each other but in site04, most of the data points are closely placed except for soft corals, hence depicting that site04 shows similar trends in bleaching irrespective of the kind of coral with an exception for soft corals. In site04, soft corals experience the least bleaching with a value of 11.04%. Site05 shows no record for bleaching of hard corals but shows high bleaching values for the rest of the corals. Site06 represents the maximum bleaching for sea pens and sea fans, hence this site is least suitable for these kinds of corals. It however is the most suitable for blue

corals as blue corals bleach only by 16.22% on this site. This site however shows no record for hard and soft corals. Site07 shows no record of soft coral bleaching points. This site and Site08 show similar data for hard and blue corals and very diverging data for sea pans and sea pens(hence an opposite effect). These sites show relatively high bleaching values for soft corals and hence making themselves unsuitable for soft corals.

Figure 6: Scatter Plot depicting Sites and Bleaching values for different kinds of corals

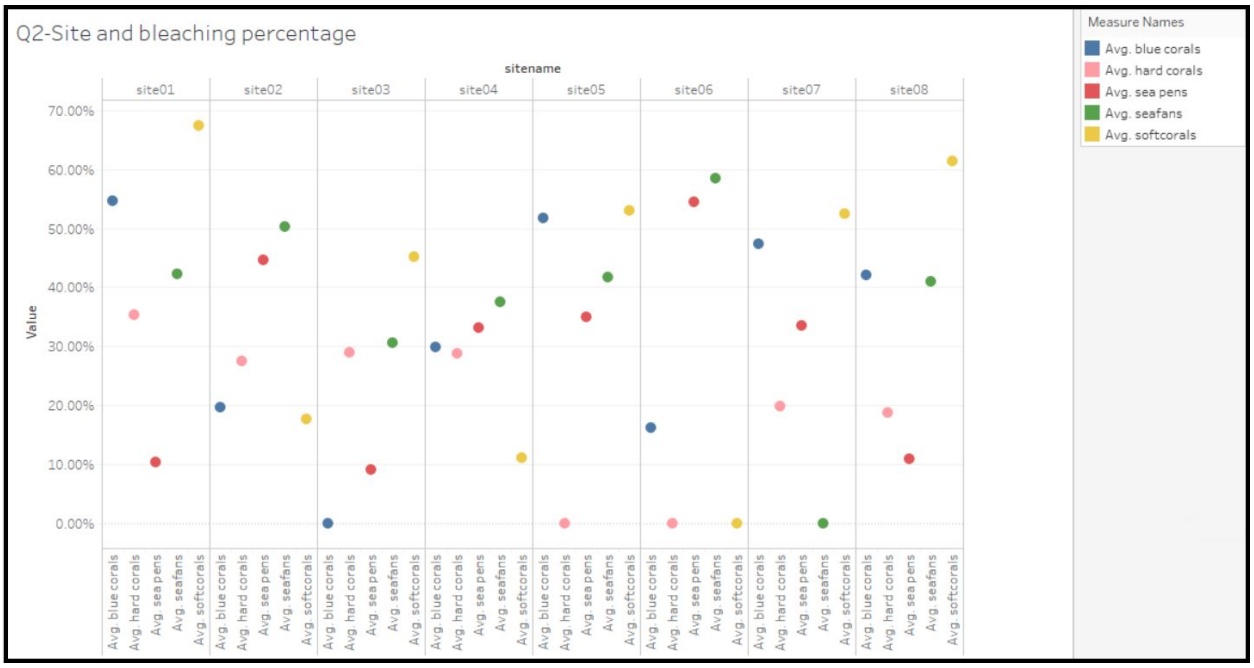



Figure 7: Location of the sites