

Title - Programming Exercise 1: Tableau Public

Student name - Sarthak Sareen

Student ID - 30761182

Tutorial Number - 24

Tutor Name - Shirin

Table Of Contents

Table Of Contents.....	1
Data Wrangling.....	2
Data Exploration.....	3
Answering Question.....	5

DATA WRANGLING

Reformatting of the dataset includes certain steps.

- 1) I have manually arranged the site name ,longitude,latitude and year columns whereas the name of the corals are as in columns.
- 2) Then I selected the data of soft corals (D3:K10) and pasted into another sheet Then I named the selected data as soft corals from the name box at the left upper corner in excel. After that clicked on the adjacent column(1st row) used the formula **=INDEX(softcoral,INT((ROW()-ROW(\$Z\$1))/COLUMNS(softcoral))+1,MOD(ROW()-ROW(\$Z\$1),COLUMNS(softcoral))+1)**

Index () takes the arguments of the data ,rows,columns and retrieves the specified cell value. In the formula above it is taking the soft coral as a data , INT((ROW()-ROW(\$Z\$1))/COLUMNS(softcoral))+1 is saving the rows and MOD(ROW()-ROW(\$Z\$1),COLUMNS(softcoral))+1 is saving the columns.

ROW()-ROW(\$Z\$1) - this function is only counting +1 for every row down a column.

/COLUMNS(softcoral))+1 - this function divides the number of columns in the data +1.

All this is wrapped in the INT () which will give the whole numbers and this gives the INDEX() row number.

MOD(ROW()-ROW(\$Z\$1),COLUMNS(softcoral))+1 - this gives the variable of the column number of the index function. MOD give sthe remainder of the counting of rows and the number of the coumns of the data +1.

This all gives such as - INDEX(softcoral,1,2),INDEX(softcoral,2,1) etc.

Fig 1 - Reformatting

	A	B	C	D	E	F	G	H	I	J
1	83.870%	80.210%	75.340%	74.990%	57.700%	56.430%	55.430%	56.290%		83.87%
2	21.230%	19.230%	17.210%	15.780%	14.800%					
3	75.340%	60.230%	37.210%	26.890%	25.890%					
4	12.450%	11.780%	11.340%	10.980%	10.890%	10.670%	10.230%	10.010%		
5	94.230%	91.230%	76.230%	60.230%	30.230%	28.450%	29.130%	14.450%		
6										
7	67.890%	60.780%	58.760%	56.450%	55.120%	40.120%	39.980%	40.340%		
8	65.230%	63.210%	60.120%	60.030%	58.890%					
9										
10										
11										

- 4) Drag down the formula until you get the REF error. This REf error indicated that all the values have been retrieved into one column.

The cells with no data is giving a value 0 whereas we cannot change the data in the dataset. So I am using if else condition.

=IF(conditon="", "",condition)

Using this as a formula with IF and Else condition -

=IF(INDEX(softcoral,INT((ROW()-ROW(\$Z\$1))/COLUMNS(softcoral))+1,MOD(ROW()-ROW(\$Z\$1),COLUMNS(softcoral))+1)="", "",INDEX(softcoral,INT((ROW()-ROW(\$Z\$1))/COLUMNS(softcoral))+1,MOD(ROW()-ROW(\$Z\$1),COLUMNS(softcoral))+1)).

After this formatting all the values to percentage by using format cells.

Fig 2- Data in Tableau

Abc Data Name	📍 Data Long	📍 Data LAT	# Data Year	# Data Soft Coral	# Data Sea Fan	# Data Blue Corals	# Data Hard Coral	# Data 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

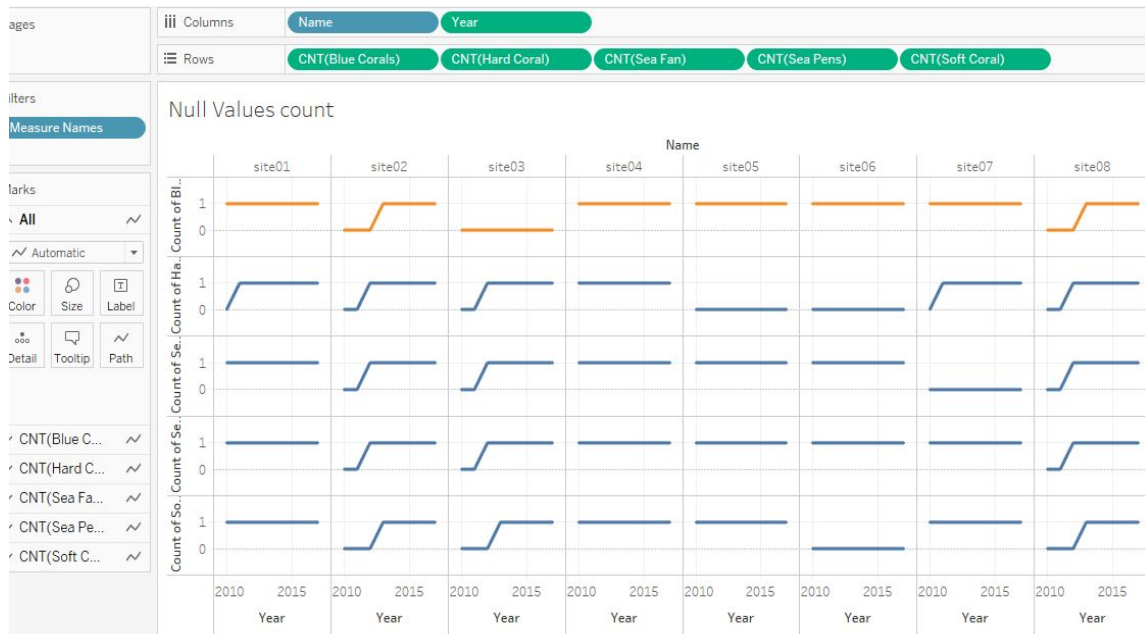
Data Exploration

Error 1 - missing/null values

According to the visualization we can see that in site02 for blue corals the count of the coral bleaching measure is 0 for 2010 and 2011 years. With the help of these we can find the null values of all sites and coral bleaching measure values where the count is 0.

Correction - this null/missing values can be corrected by taking an average of all the values of the specific coral and specific site. If the site does not have any data for a specific coral then 0% for all the years.

Fig 3 finding null/missing values



Error 2 - outliers

Fig 4 outliers

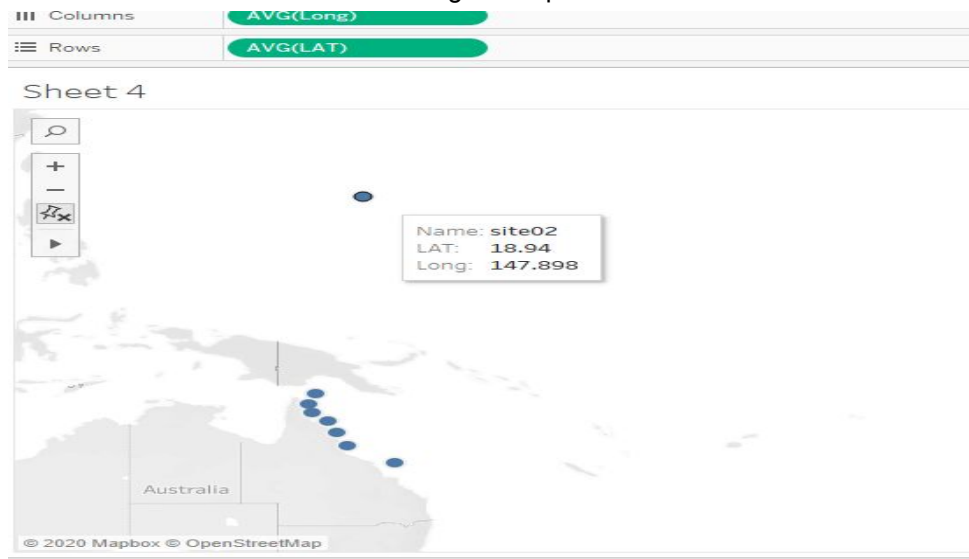


As we can see with the help of the visualization site07 for blue corals in the year 2013 there is a sudden decrease in the value which is a outlier and in the site 08 for hard corals in the year 2014 the value of coral bleaching shows a drastic increase which is also a outlier.

Correction- this outliers can be corrected with taking an average of the all the values of the specific coral and specific site.

Error 3 - latitude of site 02

Fig - 5 map



After examining all the longitude and latitude of other sites the latitude measurement is negative whereas the latitude measurement of the site 02 is positive that's why it is very far from the other sites in the visualization.

Correction-

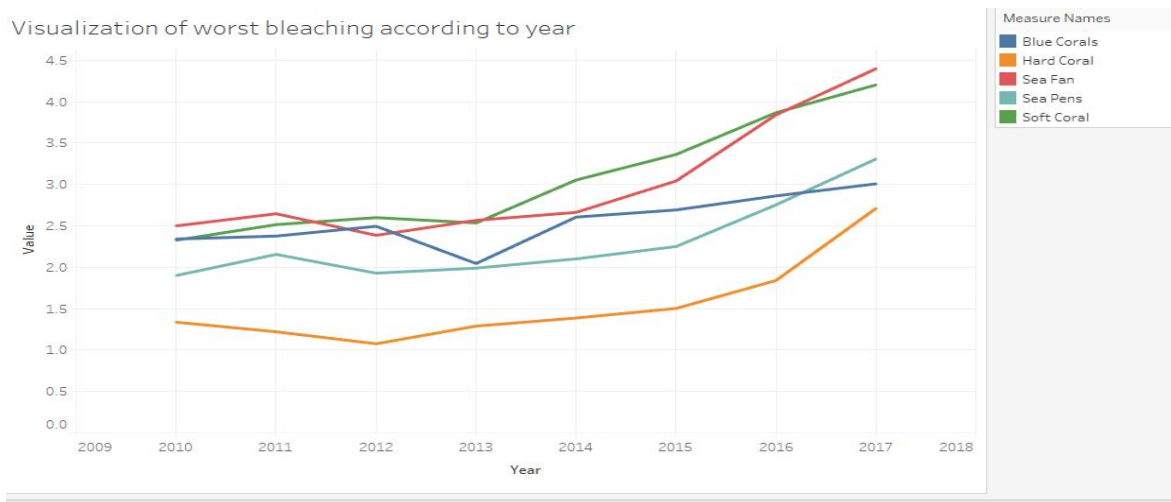
Changing the latitude measure of the site01 to positive.

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

According to the question we have to find the coral in every year which has the worst bleaching.

I have inserting Years in the column section and Measure Values in the rows sections. This gives the visualization of a line chart which depicts all the sum of different corals according to its year

Fig 06 -Worst coral bleaching according to year.



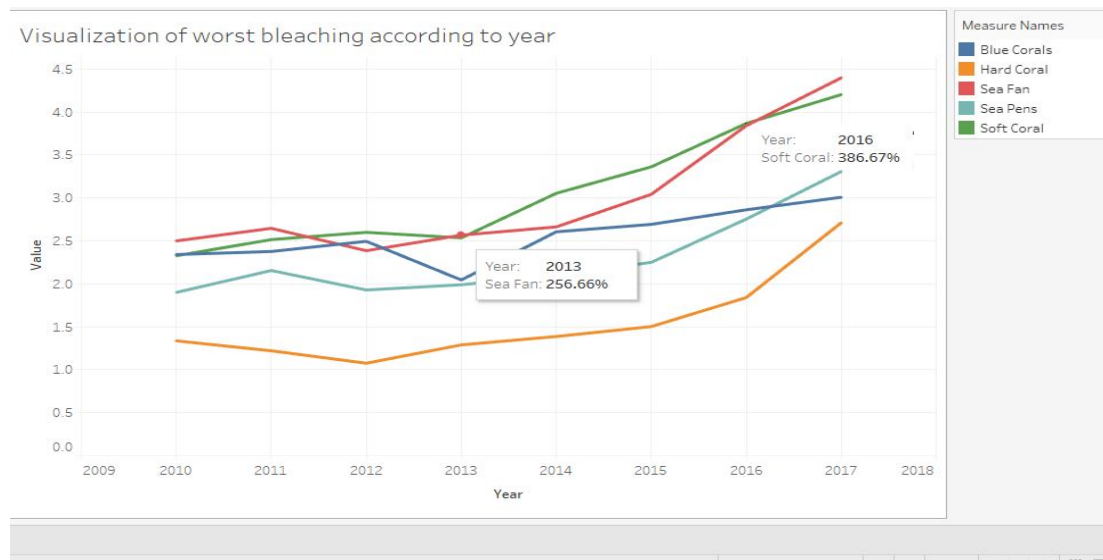
According to the visualization we can depict worst bleaching in different years-

2010 - Seafan

2011 - Seafan

2012 - Soft coral

Fig 7 Worst coral bleaching according to year



2013- sea fan

2014- soft coral

2015- soft coral

2016- soft coral

2017- sea fan

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

According to the question in this we have to find out the bleaching effect in different sites according to different kinds of corals. I have added Name and measure names in the column section and Measure values in the row section.

Fig 8 Site Location on map

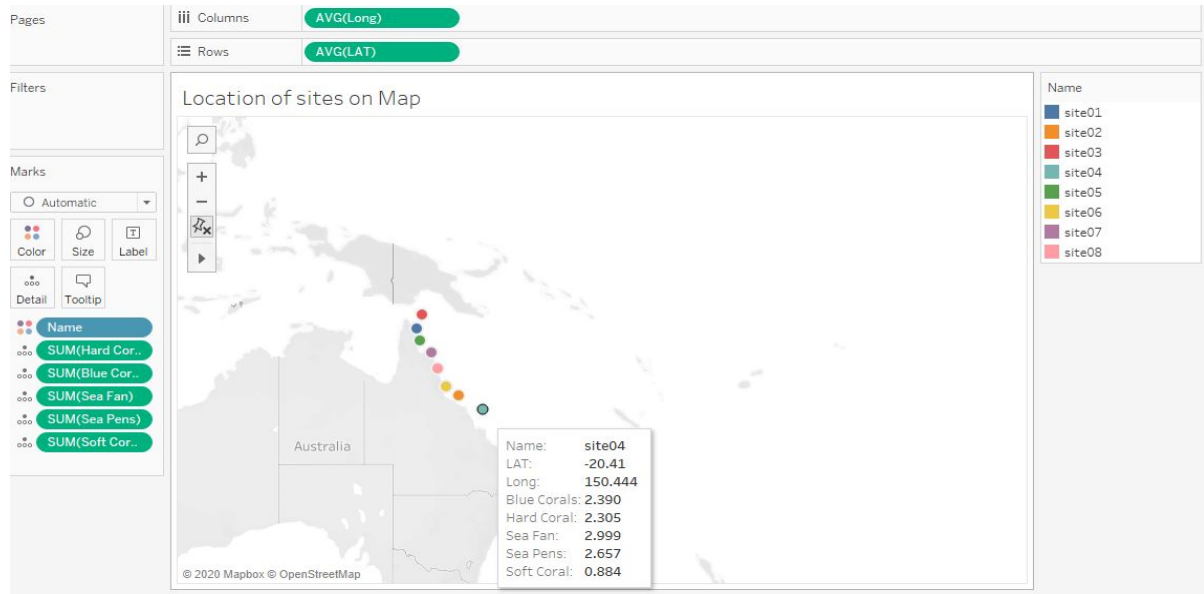
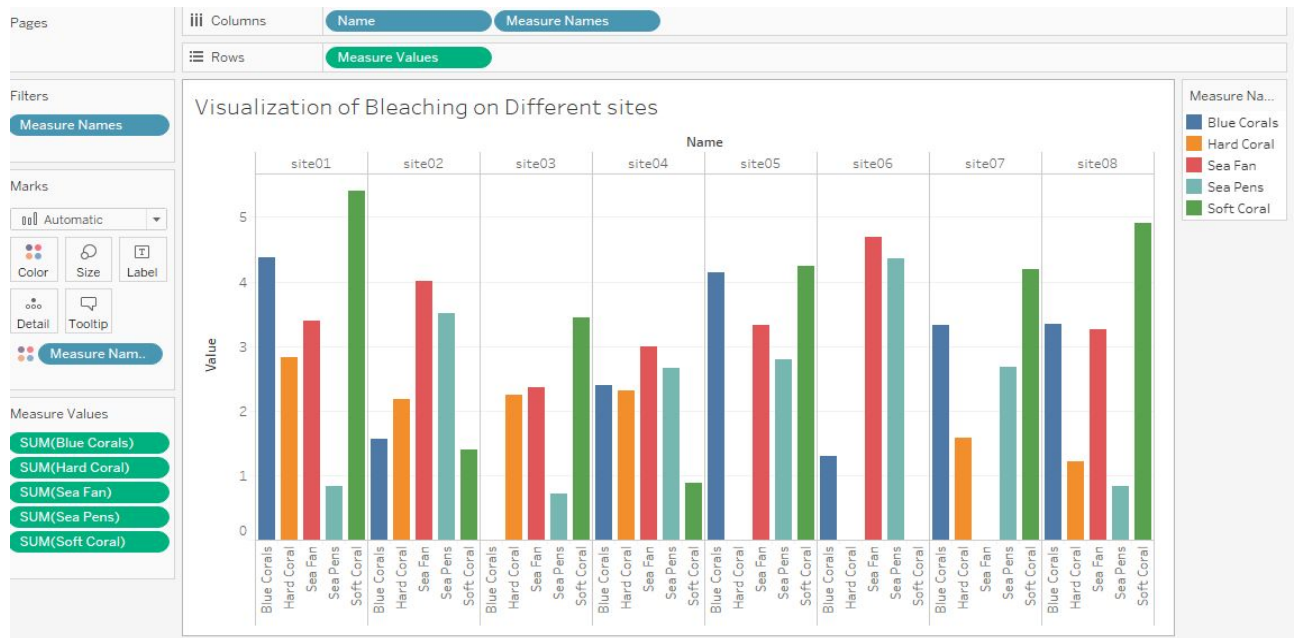


Fig 9 - Visualization on leaching of corals on different sites



According to the visualization above we can depict the trends of bleaching in different corals in different site' locations. Soft Coral in site01 ,site 03 ,site 05, site 07 and site 08 shows the worst bleaching compared to other corals whereas in site 02 and site 04 shows the least bleaching in soft coral. In site 02 and site 06 sea fan shows the worst bleaching effect. In site 01 ,site 03 and site 08 sea pens shows the least bleaching effect whereas in site 02 , site 04 and site 06 the sea pens show the second worst bleaching measure. Blue corals in site 01, site 06 site 07and site 08 also show the second worst bleaching compared to other corals.In site 06 and site 07 there are two least worst bleaching of corals i. e hard, soft corals and sea fan respectively. In site 03 Blue corals shows the least worst bleaching effect compared to other corals. This explains the different trends of bleaching of corals of different kinds and different locations of sites.