MONASH University

FIT 5147
Data Exploration and Visualization

PROGRAMMING EXERCISE 1: Tableau Public

31187366 - Prashant Hariprasad Jajoria

Introduction

The report gives insights about the % of Coral bleaching for different types of Corals, namely, Hard corals, Sea Pens, Blue corals, Soft corals and Sea fans, for 8 different sites in the Great Barrier Reef for the years 2010 to 2017, by visualizing the data using Tableau Public. This report helps to find the worst affected Corals in the given time period and how the location of site affect bleaching of different types of Corals.

Errors in Data Entry

While going through the data in Excel manually, following possible Data entry errors can be seen:

a. Bleaching % for Hard coral at site 8 for year 2014 was 148.800%, which looks like a possible Data entry error, as this percentage is far more than the pervious years Bleaching percentage. The value of Bleaching % for this outlier data was replaced by the Mean of the Percentage bleaching for Hard coral over the given years.

b. The location of Site 2 was far away from the Australian continent, which seems another possible data entry error. The value of latitude for Site 2 was 18.937 which I made to -18.937. Now the location for site 2 is in the Great Barrier Reef region.

c. The value for % Bleaching for Blue Coral at site 7 in the year 2013 is 0.470%, which is very less than the previous years value of 47.320%. I replaced the value with the mean of % Bleaching for Blue coral. The value now for Site 7 in the year 2013 is now 47.425%.

d. Also, the value for % Bleaching is missing for some years for different corals at various sites. Site 6 does not have Bleaching % for Soft Coral and Hard Coral for all the years. Site 7 does not have percentage values for Sea fans for all the years. Similarly, Site 5 has no values for Hard corals for all 8 years. While visualizing the data, I could see the number of null values for each sites and year. Therefore, I replaced the null values with 0, which will not affect the analysis later in my work.

Following figures shows the visualizations of Data Entry errors:

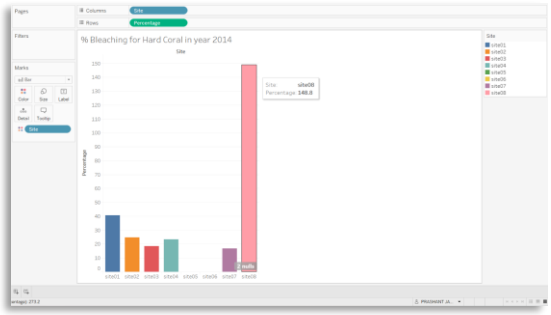

Figure 1: Percentage Bleaching for Hard Coral in 2014 at different sites

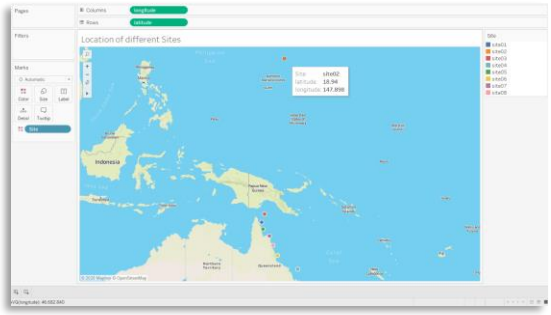

Figure 2: Location of different sites

Figure 3: Percentage Bleaching for Blue Coral in 2013 at different sites

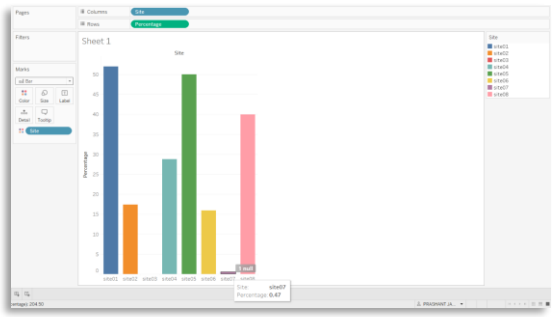

Figure 4: Percentage Bleaching for Hard Corals at different Sites from 2010 to 2017

Data Wrangling

The given raw data is not suitable to be used for visualization in Tableau Public as it is not granular. Before visualizing the data in Tableau Public, it needs to be reformatted. The data needs to be granular meaning it needs to be in its lowest level with each row having information about the name of site, coral type, latitude and longitude, the year of bleaching and percentage of bleaching. To reformat the dataset i.e. data wrangling, I used Python programming language and made use of Pandas Data Analysis Library.

Steps taken to wrangle data

1. Used Spyder IDE for the Data Wrangling process.
2. Imported Pandas Data Analysis library.
3. Used pandas.read_excel() function from Pandas library to read the given excel file.
4. Stored the given dataset in Pandas DataFrame data structure.
5. Used the function pandas.melt() with parameters, id_vars = ['name','latitude','longitude'], var_name = 'Year', value_name = 'Percentage', to make the dataset Granular.
6. Replaced all the Null values with 0 using pandas.DataFrame.fillna() function with parameter inplace=True.
7. Since value of Bleaching percentage for Hard coral at site 8 for year 2014 was an outlier, replaced it with the Mean of the Percentage Bleaching for the corresponding Coral type.
8. Added new Column named 'Coral Type' to dataset with value Soft Coral, Hard Coral, Sea Pens, Blue Coral and Sea Fans for corresponding rows.
9. Exported the reformatted in the form of CSV file by using the pandas.DataFrame.to_csv() function with parameters mode='w', index = True, header=True
10. Imported the structured CSV file in Tableau Public. Renamed the all the Header of CSV file with correct Title.

Dataset after Wrangling

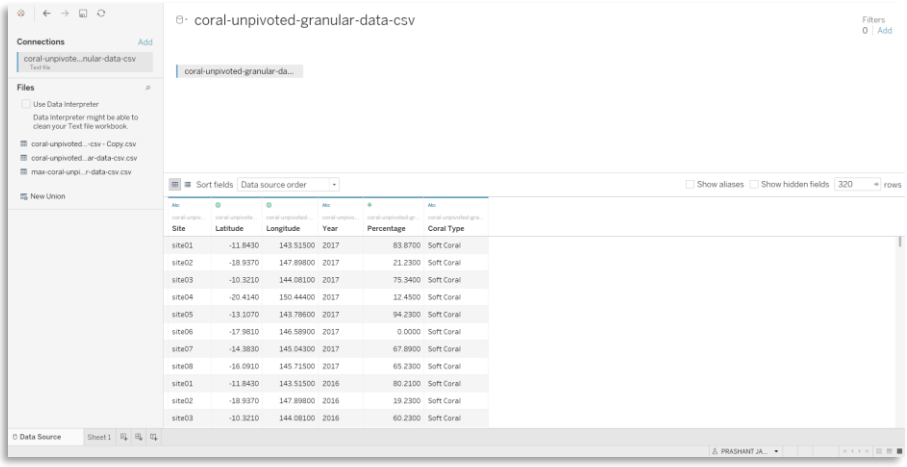

Figure 5: Dataset after formatting and cleaning

Questions and Findings

1. In which years and for which kinds of coral is bleaching the worst?

The bar graph below displays the average percentage bleaching for all the corals in the years from 2010 to 2017.

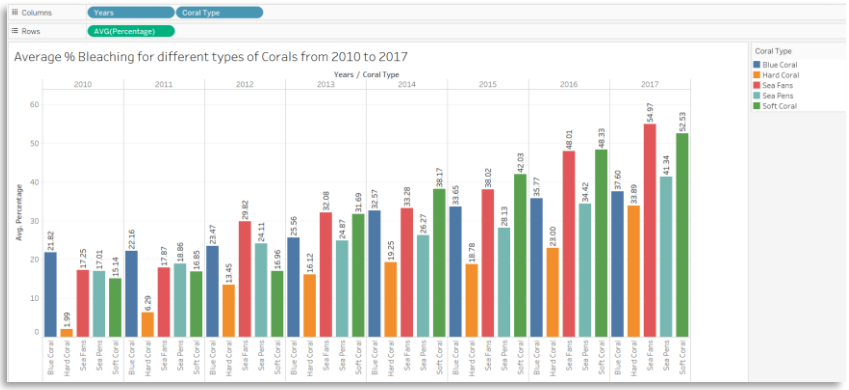

Figure 6: Average % Bleaching for different types of Corals from 2010 to 2017

It can be clearly seen that percentage bleaching for Sea Fans was the highest amongst all the corals in 2017 period with average Bleaching percentage of 54.97%. This was followed by Bleaching of Soft Corals with average percentage bleaching of 52.53% in the same year 2017.

The same can be verified by plotting line graph of corals over the given time period. The line graph helps us see the growth of Bleaching over the years.

Following is the trend observed for various corals:

Blue Coral: Had the minimum Bleaching percentage in 2010 with value of 21.82 and increased gradually over the years and finally attaining the highest value of 37.60%.

Hard Coral: This type of coral was the least affected amongst the given corals over the years. Hard corals saw a steep increase in percentage of Bleaching from 2010 to 2014 and decreased in year 2015. Later it saw a sharp increase in bleaching percentage attaining value of 33.89%.

Sea Fans: The bleaching percentage for Sea Fans kept on increasing steadily from 2010 to 2012 and plateaued from 2012 to 2015 and increasing thereafter and eventually attaining value of 54.97 which was the highest amongst all the corals in the given years.

Sea Pens: This coral had a slow growth in Bleaching percentage from 2010 to 2015 and increased significantly in the coming years with the highest value of 41.34% in 2017.

Soft Coral: The increase of bleaching for this type of coral from 2012 to 2013 is the most eye-catching trend in this graph. The bleaching percentage saw two-fold rise in value with value of 31.69 in 2013.

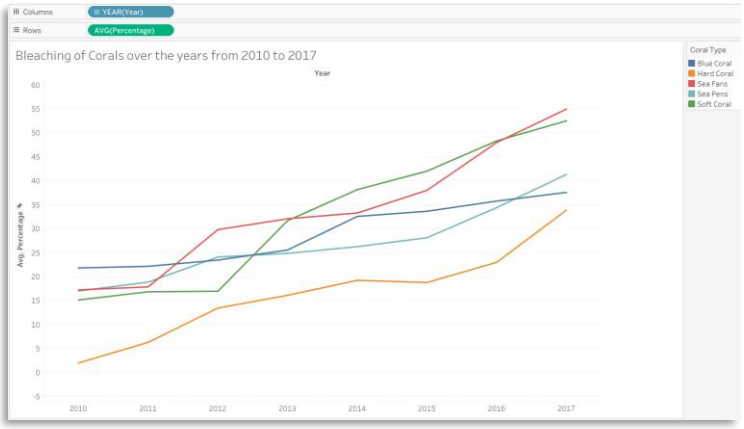

Figure 7: % Bleaching of Corals from 2010 to 2017

2. How does the location of the site affect bleaching of the different kinds of corals?

Blue Coral recorded the maximum average percentage of Bleaching at site 1 with value of 54.67% and minimum of 12.31% at site 2. There was no data for site 3 for this type of coral.

Hard Coral had the maximum average Beaching at site 1 and minimum at site 8 with value of 30.96% and 13.38%.

Sea Fans had the highest average Bleaching percentage value of 58.56% at site 6 and lowest at site 8.

Sea Pens recorded the maximum value of Bleaching percentage at site 6 and minimum at site 8.

For Soft Coral the highest value was 67.53% at site 1 and lowest at site 2 with 11.03%.

Looking at the Bar Graph for values of various Corals at different sites, we can conclude that Site 1 had the most adverse and it affected Blue coral, Hard coral and Soft coral the most.

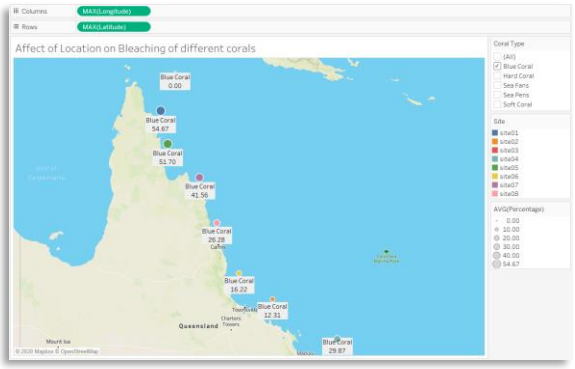

Figure 8: Bleaching of Blue Coral

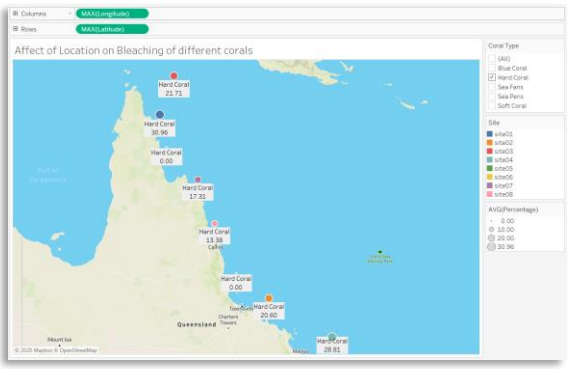

Figure 9: Bleaching of Hard Coral

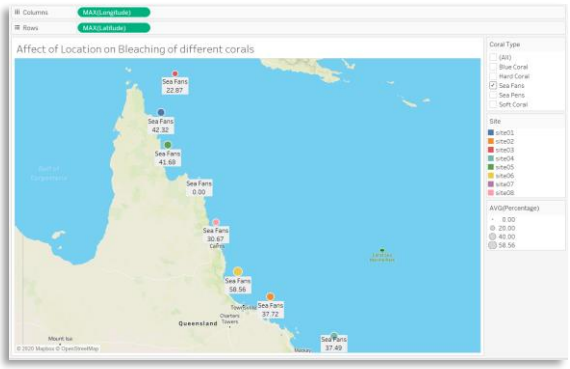

Figure 10: Bleaching of Sea Fans

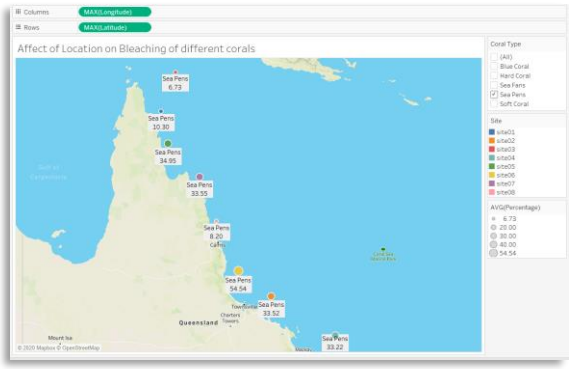

Figure 11: Bleaching of Sea Pens

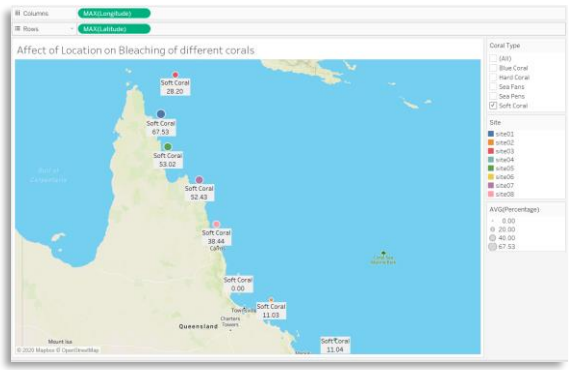

Figure 12: Bleaching of Soft Coral

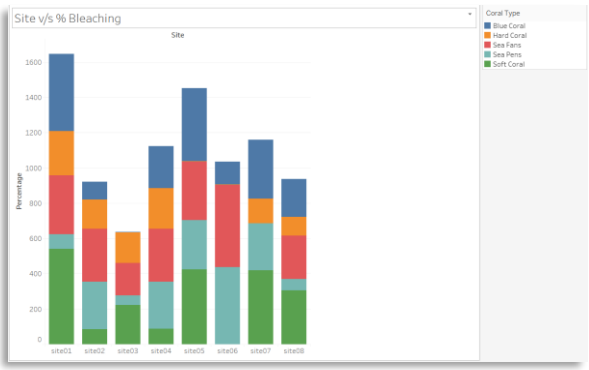

Figure 13: Site v/s % Bleaching