Jackfish Bay - Water and Sediment Chemistry 2017

Divya Dharshini (8928823)

Business Analytics, Conestoga College
INFO8076-23F-Sec1-SQL and Data Analysis

Jey Kumaresan

Dec 11, 2023

INTRODUCTION

In 1987, Jackfish Bay, a lake in Ontario, was declared an area of concern due to pollution from a pulp mill in Terrace Bay. This pollution led to a decline in environmental water quality and health. A variety of regulations and action plans were implemented to rejuvenate the lake's ecosystem, particularly the population of White Sucker fish in Jackfish, which had seen a 70% decrease since 1988. In 2017, the Ministry of Environment collected and tested water and sediment samples from Jackfish Bay for a range of elements.

Objective

Conduct a water health analysis in Moberley Bay, Tunnel Bay, Far Field Bay, and Far Far Field Bay. Based on the analysis, identify a location with optimal conditions for fish egg incubation to enhance the population of native species. Provide recommendations to the Ministry of Health.

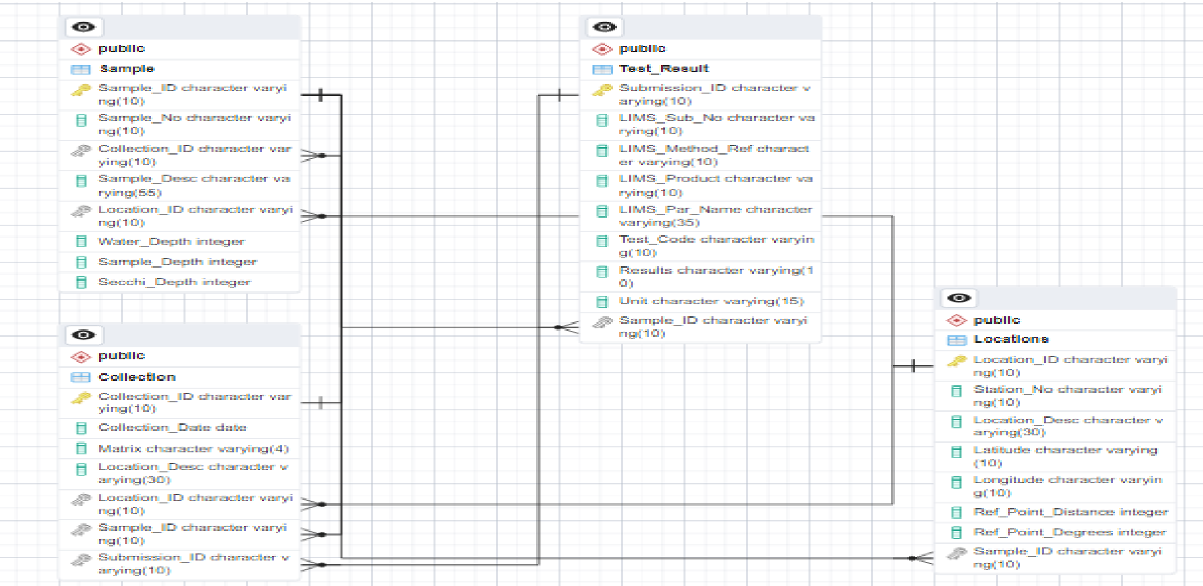
Data Source

Jackfish Bay Area of Concern in Recovery – Water and Sediment Chemistry from Data Ontario
[Jackfish Bay Area of Concern in Recovery – Water and Sediment Chemistry - Dataset - Ontario Data Catalogue]

Data Processing

In the provided Excel sheet, the descriptions of the location and sample have been condensed for improved clarity. The entire dataset is divided into four tables: locations, collection, sample, and test result. A primary key column is incorporated to establish relationships between the tables.

Entity Relationship Diagram



QUERIES, OUTPUTS, AND INSIGHTS

1. Nitrogen analysis

Input:

```
SELECT c.Location_Desc, AVG(CAST(tr.Results AS FLOAT)) as avg_nitrogen
FROM Collection c
JOIN Test_Result tr
ON c.Submission_ID = tr.Submission_ID
WHERE tr.LIMS_Par_Name LIKE '%Nitrogen%'
GROUP BY c.Location_Desc
ORDER BY avg_nitrogen DESC
```

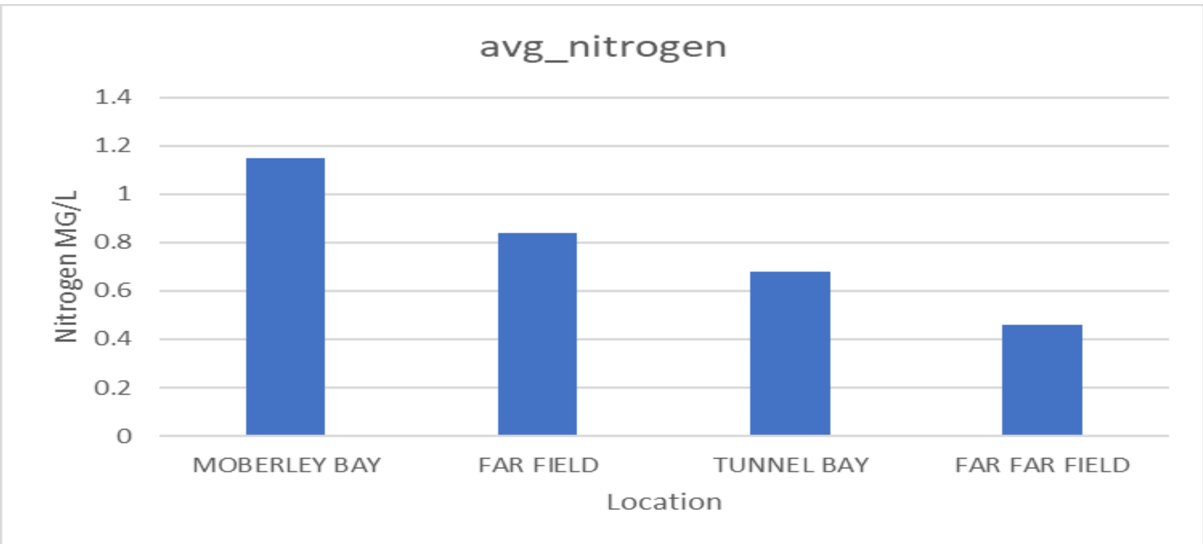
Output:

location_desc	avg_nitrogen
MOBERLEY BAY	1.149129
FAR FIELD	0.8376
TUNNEL BAY	0.6803
FAR FAR FIELD	0.4615

Insights:

The above query shows the locations with the highest amount of Nitrogen in the water. Nitrogen causes overstimulation of the growth of aquatic plants and algae. This is also the most harmful substance affecting aquatic species. Moberley Bay topped the list, while Far Far Field had the lowest. Through this, we can outline the recovery process of the locations for a better rejuvenation of the lake.

Visualization:



2. Heavy metal analysis – Lead

Input:

```
SELECT c.Matrix, c.Location_Desc, AVG(CAST(tr.Results AS FLOAT)) as avg_lead
FROM Collection c
JOIN Test_Result tr ON c.Submission_ID = tr.Submission_ID
WHERE tr.LIMS_Par_Name = 'Lead' AND c.Matrix IN ('WS', 'SE')
GROUP BY c.Matrix, c.Location_Desc
ORDER BY avg_lead ASC;
```

Output:

matrix	location_desc	avg_lead
WS	TUNNEL BAY	0
WS	MOBERLEY BAY	0
WS	FAR FIELD	0.025
WS	FAR FAR FIELD	0.1
SE	MOBERLEY BAY	9.1
SE	FAR FAR FIELD	19.8
SE	FAR FIELD	22.35
SE	TUNNEL BAY	25.4

Insights:

This query shows the amount of lead which was found in their water and sediment test samples with respect to locations. Tunnel Bay sediment samples have the highest lead, and Far Field water samples have the lowest. Lead is harmful to all aquatic species. This can be tackled by controlling the waste from the mill being disposed into the lake.

3. Locations with elements from the CDD family

Input:

```
SELECT c.Location_Desc, COUNT(*) as cdd_count
FROM Collection c
JOIN Test_Result tr ON c.Submission_ID = tr.Submission_ID
WHERE tr.LIMS_Par_Name LIKE '%CDD%'
GROUP BY c.Location_Desc
ORDER BY cdd_count DESC
```

Output:

location_desc	cdd_count
MOBERLEY BAY	28
TUNNEL BAY	14
FAR FIELD	14
FAR FAR FIELD	7

Insights:

The CDD family is one of the highest pollutants of the lake, and this query finds out the count of this pollutant for each location. Moberley Bay topped the list for elements from the CDD family at a 28 CDD count, while the Far Far-field location has a 7 CDD count. This clearly shows the disparity between the pollutant levels in the locations.

4. Average pH by lakes

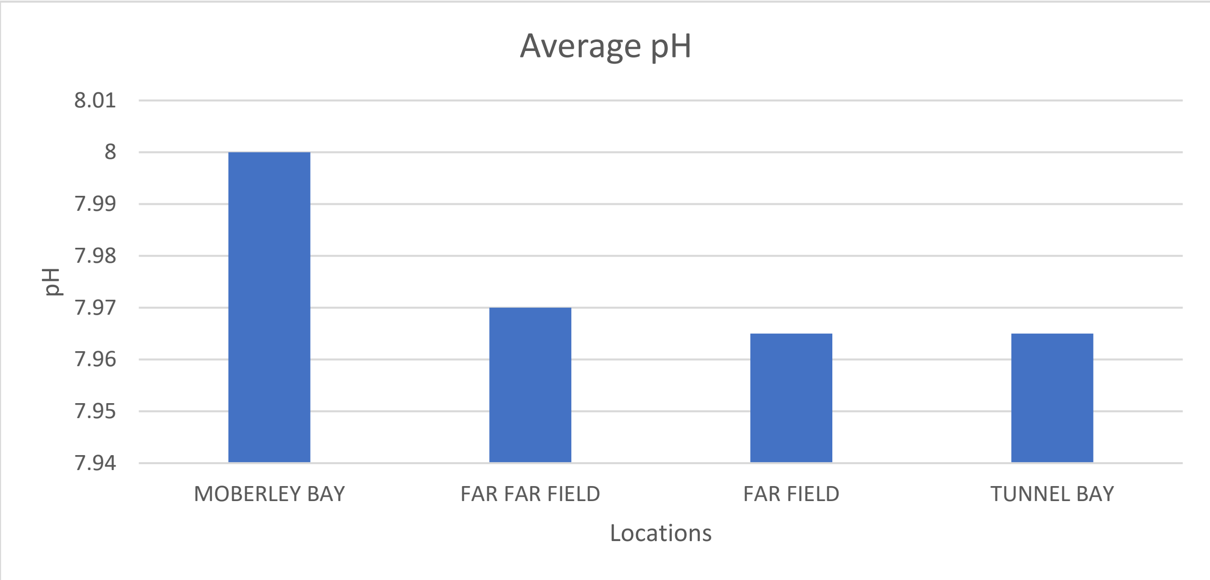
Input:

```
SELECT c.Location_Desc, AVG(CAST(tr.Results AS FLOAT)) as avg_pH
FROM Collection c
JOIN Test_Result tr ON c.Submission_ID = tr.Submission_ID
WHERE tr.LIMS_Par_Name = 'pH'
GROUP BY c.Location_Desc
ORDER BY avg_pH DESC;
```

Output:

location_desc	avg_ph
MOBERLEY BAY	8
FAR FAR FIELD	7.97
FAR FIELD	7.965
TUNNEL BAY	7.965

Visualization:



Insights:

The pH of water inhibits shell growth in aquatic animals. So, tackling the pH level is very important for improving the lake. The pH level of the lake was calculated with respect to the locations from the results derived from their samples.

5. Comparing the Secchi depth of all the lakes

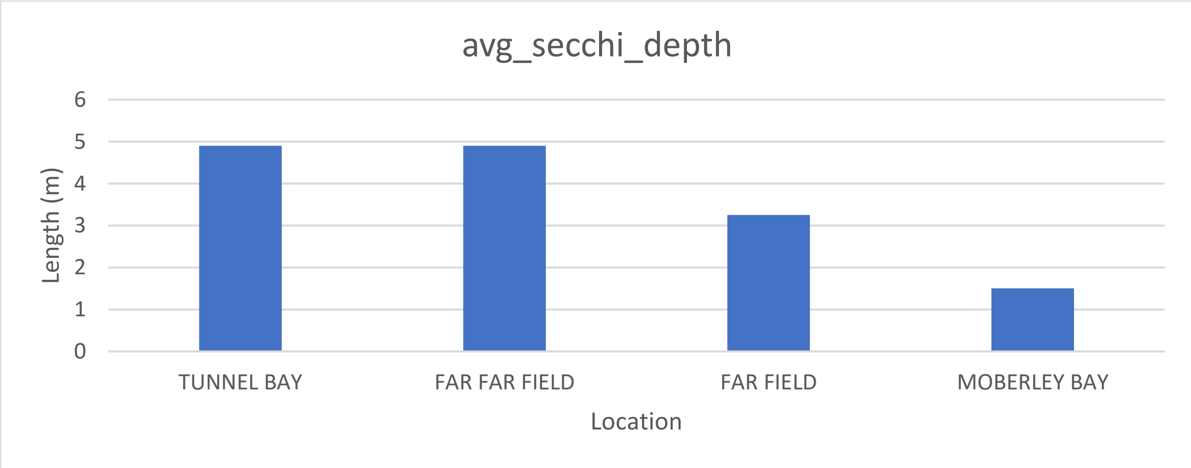
Input:

```
SELECT c.Location_Desc, AVG(s.Secchi_Depth) as avg_secchi_depth
FROM Collection c
JOIN Sample s ON c.Sample_ID = s.Sample_ID
GROUP BY c.Location_Desc
ORDER BY avg_secchi_depth DESC;
```

Output:

location_desc	avg_secchi_depth
TUNNEL BAY	4.9
FAR FAR FIELD	4.9
FAR FIELD	3.25
MOBERLEY BAY	1.5

Visualization:



Insights:

The Secchi depth is related to water clarity and measures how deep light can penetrate the water. The deeper the light penetrates the water, the lesser the growth of algae and eutrophication-related waterbody characteristics, such as chlorophyll concentration, hypolimnetic oxygen depletion rate, and fish yield. This calculates the average Secchi depth to understand the sample details and their contents. Location Tunnel Bay came out to be at 4.9, while Moberley Bay was at the least on the list at 1.5.

6. Comparing dissolved solids and suspended solids in all the locations

Input:

```
SELECT c.Location_Desc,
    tr.LIMS_Par_Name,tr.test_code,
    AVG(CAST(tr.Results AS FLOAT)) as avg_quantity
FROM Collection c
JOIN Test_Result tr ON c.Submission_ID = tr.Submission_ID
WHERE tr.LIMS_Par_Name IN ('Solids -  dissolved', 'Solids -  suspended')
GROUP BY c.Location_Desc, tr.LIMS_Par_Name,tr.test_code
order by c.location_desc
```

Output:

location_desc	lims_par_name	test_code	avg_quantity
FAR FAR FIELD	Solids - dissolved	RSF	69.5
FAR FAR FIELD	Solids - suspended	RSP	0.85
FAR FIELD	Solids - dissolved	RSF	79.25
FAR FIELD	Solids - suspended	RSP	0.725
MOBERLEY BAY	Solids - dissolved	RSF	108
MOBERLEY BAY	Solids - suspended	RSP	0.916667
TUNNEL BAY	Solids - dissolved	RSF	67.5

TUNNEL BAY	Solids - suspended	RSP	0.5

Visualization:

Insights:

The test results and codes were considered by comparing dissolved and suspended solids with respect to each location. There is a contrast difference between the dissolved and suspended levels of the solids in each location. Moberley Bay has the highest, and the bottom-placed location is Tunnel Bay.

7. Percentage of organic pollutant vs inorganic pollutant

Input:

```
WITH ElementSums AS (
    SELECT
        c.Location_Desc,
        SUM(CASE WHEN tr.LIMS_Par_Name LIKE '%CDD%' AND tr.Results ~
E'^\\d+(\\.\\d+)?$'
               THEN CAST(tr.Results AS NUMERIC) ELSE 0 END) AS cdd_sum,
        SUM(CASE WHEN tr.LIMS_Par_Name NOT LIKE '%CDD%' AND tr.Results ~
E'^\\d+(\\.\\d+)?$'
               THEN CAST(tr.Results AS NUMERIC) ELSE 0 END) AS other_sum
    FROM
        Collection c
```

```
    JOIN
        Test_Result tr ON c.Submission_ID = tr.Submission_ID
    GROUP BY
        c.Location_Desc
)

SELECT
    es.Location_Desc,
    es.cdd_sum,
    es.other_sum,
    CASE WHEN (es.cdd_sum + es.other_sum) > 0
         THEN es.cdd_sum * 100.0 / (es.cdd_sum + es.other_sum)
         ELSE 0 END AS cdd_percentage,
    CASE WHEN (es.cdd_sum + es.other_sum) > 0
         THEN es.other_sum * 100.0 / (es.cdd_sum + es.other_sum)
         ELSE 0 END AS other_percentage
FROM
    ElementSums es
ORDER BY
    es.cdd_sum DESC;
```

Output:

location_desc	cdd_sum	other_sum	cdd_percentage	other_percentage
MOBERLEY BAY	663.678	417573.3	0.158685	99.84132
TUNNEL BAY	305.471	140943.7	0.216264	99.78374
FAR FIELD	288.842	272393.3	0.105926	99.89407
FAR FAR FIELD	104.276	94202.46	0.110571	99.88943

Insights:

The organic pollutant is related to chemical pollution, whereas inorganic pollutants arise due to radiant energy, noise and heat. The percentage of organic and inorganic pollutants was calculated to understand the type of contamination each location consists of. Through this, we can devise a proper plan to rejuvenate and improve the quality of the lake.

8. Top 5 elements with high variance between Moberley and Far Far field

Input:

```
WITH ParameterQuantities AS (
    SELECT
        LIMS_Par_Name AS LIMS_Parameter_Name,
        MAX(CASE WHEN L.Location_Desc = 'MOBERLEY BAY' AND TR.Results ~
E'^\\d+\\.?\\d*$' THEN TR.Results::numeric END) AS Quantity_MOBERLEY_BAY,
```

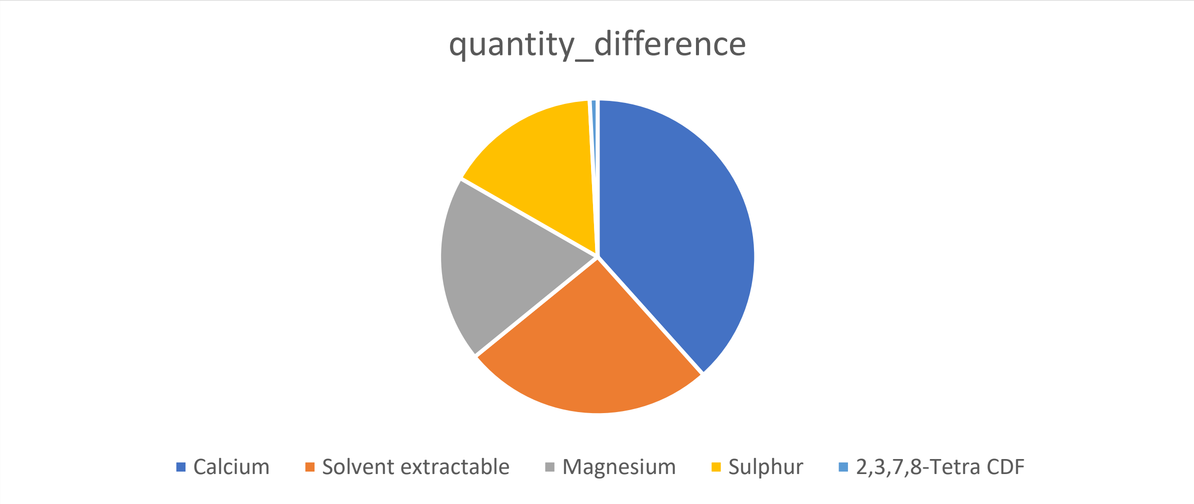
```
        MAX(CASE WHEN L.Location_Desc = 'FAR FAR FIELD' AND TR.Results ~
E'^\\d+\\.?\\d*$' THEN TR.Results::numeric END) AS Quantity_FAR_FAR_FIELD,
        COALESCE(MAX(CASE WHEN L.Location_Desc = 'MOBERLEY BAY' AND
TR.Results ~ E'^\\d+\\.?\\d*$' THEN TR.Results::numeric END), 0) -
        COALESCE(MAX(CASE WHEN L.Location_Desc = 'FAR FAR FIELD' AND
TR.Results ~ E'^\\d+\\.?\\d*$' THEN TR.Results::numeric END), 0) AS
Quantity_Difference
    FROM
        Test_Result TR
    JOIN
        Sample S ON TR.Sample_ID = S.Sample_ID
    JOIN
        Collection C ON S.Collection_ID = C.Collection_ID
    JOIN
        Locations L ON C.Location_ID = L.Location_ID
    WHERE
        L.Location_Desc IN ('MOBERLEY BAY', 'FAR FAR FIELD')
    GROUP BY
        LIMS_Par_Name
)

SELECT
    LIMS_Parameter_Name,
    Quantity_MOBERLEY_BAY,
    Quantity_FAR_FAR_FIELD,
    Quantity_Difference,
    RANK() OVER (ORDER BY Quantity_Difference DESC) AS Rank
FROM
    ParameterQuantities;
```

Output:

lims_parameter_name	quantity_moberley_bay	quantity_far_far_field	quantity_difference	rank
Calcium	13700	5350	8350	1
Solvent extractable	5700	100	5600	2
Magnesium	10300	6130	4170	3
Sulphur	3660	200	3460	4
2,3,7,8-Tetra CDF	179	2.98	176.02	5

Visualization:



Insights:

The top 5 elements with high variance between Moberley Bay and Far Far Field were calculated to compare the test results. Calcium topped the list and was found more in Moberley Bay, whereas 2,3,7,8-Tetra CDF was the least in the top 5 and, again, found more in Moberley Bay as compared to Far Far field.

9. Location which has the highest amount of Phosphorus

Input:

```
SELECT c.Location_Desc, AVG(CAST(tr.Results AS FLOAT)) as avg_phosphorus
FROM Collection c
JOIN Test_Result tr ON c.Submission_ID = tr.Submission_ID
WHERE tr.LIMS_Par_Name LIKE '%Phosphorus%'
GROUP BY c.Location_Desc
ORDER BY avg_phosphorus DESC
```

Output:

location_desc	avg_phosphorus
TUNNEL BAY	226.80151
FAR FIELD	211.36272
MOBERLEY BAY	205.4465484
FAR FAR FIELD	149.62149

Insights:

Phosphorus increases the growth of algae and large aquatic plants, which can result in decreased levels of dissolved oxygen. In phosphorus levels, Tunnel Bay topped the list with the highest amount of phosphorus, and the Far Far field had the least amount of phosphorus.

10. Average water depth vs average Secchi depth

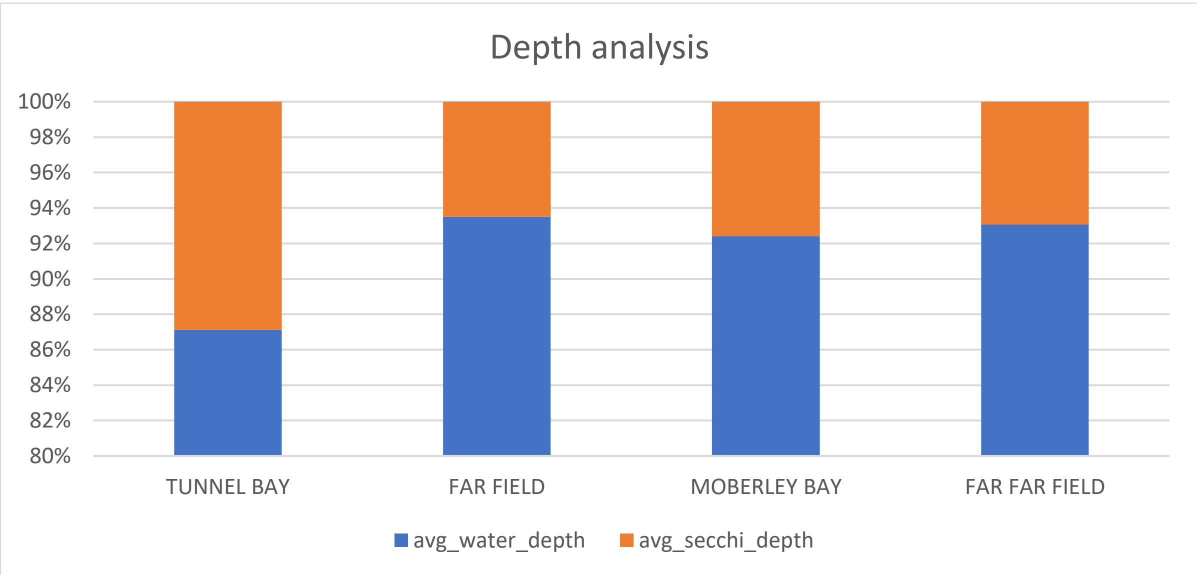
Input:

```
SELECT
   Location_Desc,
   AVG(Water_Depth) as avg_water_depth,
   AVG(Secchi_Depth) as avg_secchi_depth
FROM Sample
JOIN
   Locations ON Sample.Location_ID = Locations.Location_ID
GROUP BY Location_Desc;
```

Output:

location_desc	avg_water_depth	avg_secchi_depth
TUNNEL BAY	33.16034483	4.9
FAR FIELD	46.70074627	3.25
MOBERLEY BAY	18.28832808	1.5
FAR FAR FIELD	65.93551402	4.9

Visualization:



Insights:

The average water and average Secchi depth at each location were compared with the given details from the dataset. The higher the water level, the Higher the Secchi depth will be. This results in more growth of algae and the contamination of the lake. These results show that both the averages are not the same for any location.

SUMMARY

The analysis indicates that Moberley Bay has the highest nitrogen concentration at 1.149129 MG/L, which is below the safe limit of 1.2 MG/L. Despite this, it surpasses in CDD family elements, signifying greater water contamination. Conversely, Tunnel Bay exhibits safer levels in all aspects. Consequently, introducing white sucker eggs in Moberley Bay is not advisable, as the conditions do not promise a favourable environment for the revival of the white sucker population in this context.

Way Forward

- The top priority will be implementing long-term monitoring for Jackfish Bay while it recovers naturally before the delisting criteria are met.
- We must continue assessing fish health, water, and sediment quality
- continue enforcing regulations to ensure the pulp mill complies with the federal and provincial regulatory requirements.

Outlook

Jackfish Bay has seen significant improvement and progress toward restoration since its designation as an AOC and will remain an AOC in recovery until monitoring indicates that the remaining beneficial uses have been restored. Under the Canada-Ontario Agreement on Great Lakes Water Quality and Ecosystem Health, we will work with the province of Ontario to continue assessing and monitoring results for environmental recovery. Once restoration has been confirmed, Jackfish Bay will be removed from the list of AOCs.

REFERENCE

- Blok, A. (2021, July 12). A long road to recovery for Lake Superior fish. Environmental Health News
A long road to recovery for Lake Superior fish - EHN

- “Jackfish Bay Area of Concern in Recovery – Water and Sediment Chemistry.” Ontario Data Catalogue. Retrieved from
https://data.ontario.ca/dataset/jackfish-bay-area-of-concern-in-recovery-water-and-sediment-chemistry/resource/0b2495bc-8522-4ef9-9c87-fa627f7cabe9

- “Jackfish Bay – Remedial Action Plan (North Shore).” InfoSuperior. Retrieved from
https://rap.infosuperior.com/jackfish-bay/ apa format reference