Life Below Water

Coral Bleaching Analysis

Smita PRAKAS

My goal is

- to **raise awareness** on coral bleaching incidents over time
- to **analyze trends in marine plastic pollution** over time and across different countries
- to use deep learning techniques for coral images classification to **guide conservation efforts**

What is coral bleaching?

It is the process in which corals **lose their vibrant colors** due to stress, often caused by changes in water temperature or environmental conditions, leading to a decline in their health and vitality.

Hot topic

Le Monde Planète | Comprendre le réchauffement climatique 9 indicateurs de l'urgence climatique

PLANÈTE • OCÉANS

Le réchauffement des océans entraîne un blanchissement massif des coraux dans le monde

L'Agence américaine d'observation océanique et atmosphérique a confirmé qu'« un nouvel événement mondial de blanchissement des coraux » est en cours depuis février 2023, le deuxième épisode en moins de dix ans.

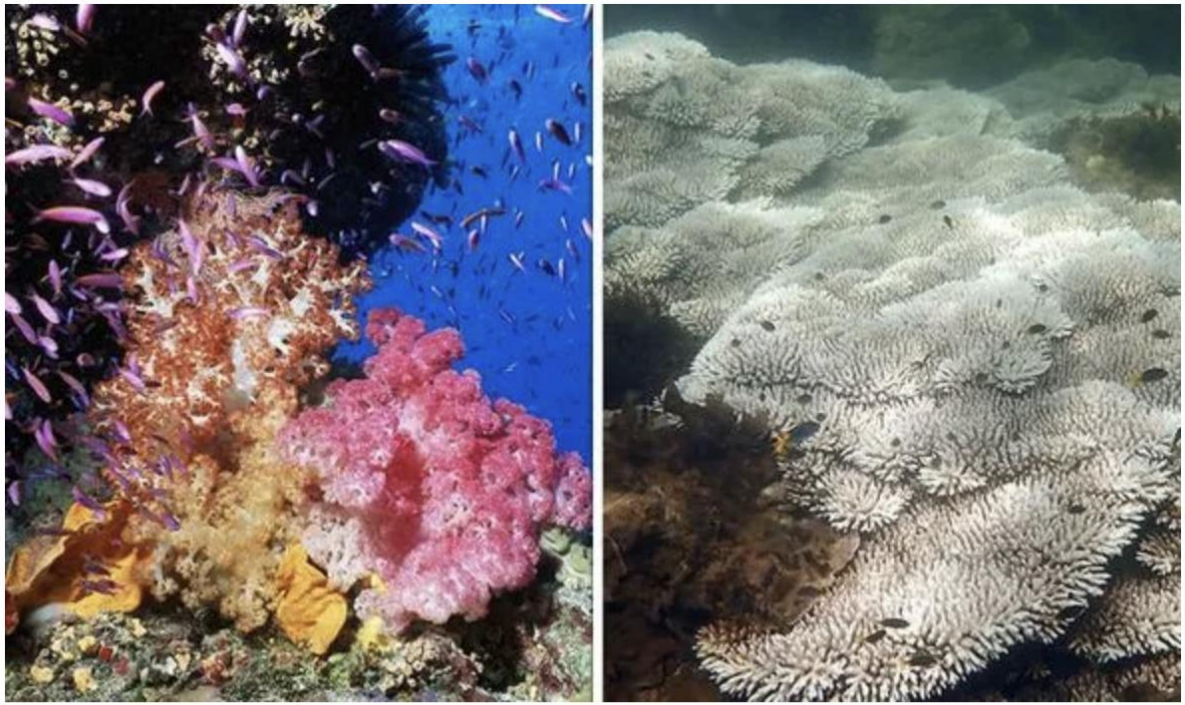
Par Matthieu Goar

Publié le 17 avril 2024 à 05h30, modifié le 17 avril 2024 à 08h50 · Lecture 4 min.

Recent article on Le Monde : Published on April 17th, 2024

Before & after a bleaching event

Before & after a bleaching event



Great Barrier Reef before and after: The coral loses its colour when it's stressed

(Image: PA/Coral Reef Studies)

Why does it matter?

- Coral reefs are vital ecosystems
- Support a vast diversity of marine life
- Provide coastal protection
- Contribute to local economies and livelihoods

Project Management

On Trello :

Research & Data collection

EDA & Cleaning

Visualizations

Flask API

Machine Learning

Data gathering

- **Flat files :** Datasets from Our World in Data and BCO-DMO.org
- **API :** United Nations Statistics Division SDG API
- **Web Scraping :** Oceans & Climate Change | Fox News

EDA and Data Cleaning

- Checking for **data shapes** & **data types**
- Handling **missing values**
- Removing **duplicates**
- Keeping **essential columns**
- **Normalization** & **transformation** of variables

Data Cleaning : overview

Flat files : ‘coral-bleaching-events-per-year.csv’

Initial Shape :
185 rows / 5 columns

After cleaning :
185 rows / 5 columns

	region	code	year	moderate bleaching events (1-30% bleached)	severe bleaching events (>30% bleached)
0	Australasia	AA	1980	3	0
1	Australasia	AA	1981	1	0
2	Australasia	AA	1982	0	0
3	Australasia	AA	1983	3	1
4	Australasia	AA	1984	1	0

Datasets : overview

Flat files : ‘global_bleaching_env.csv’

Initial Shape :
41 361 rows / 62 columns

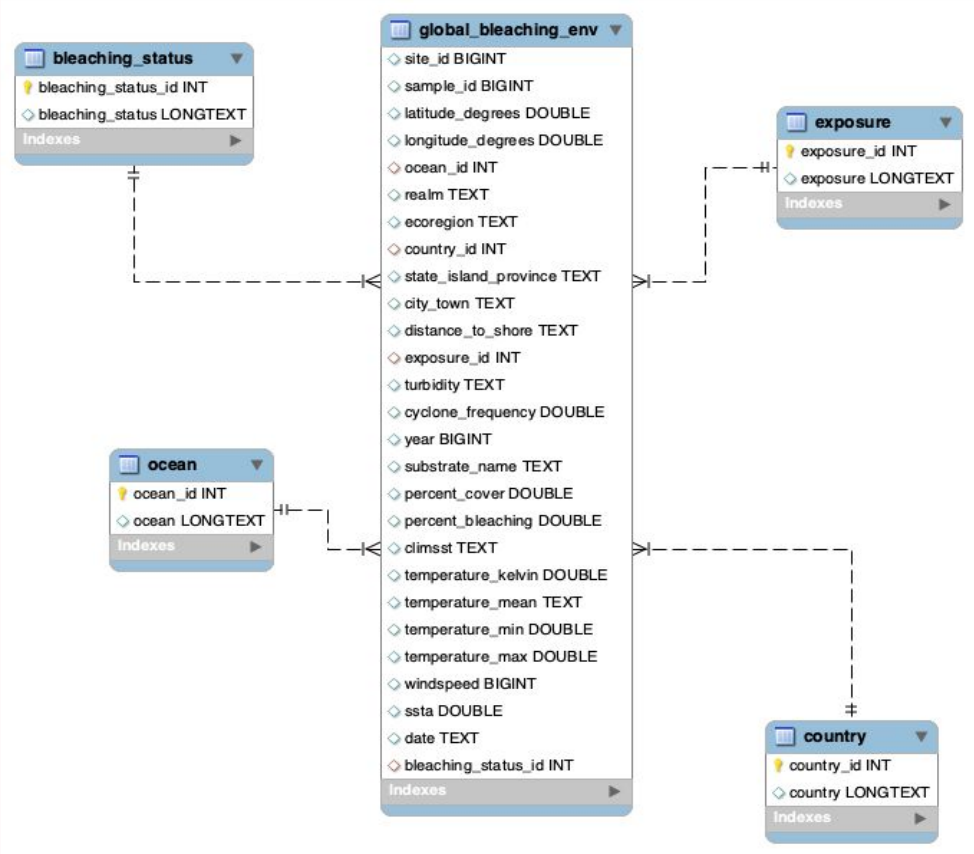
After cleaning :
21 836 rows / 26 columns

	site_id	sample_id	latitude_degrees	longitude_degrees	ocean	realm	ecoregion	country	state_island_province	city_town	...	percent_bleaching
6981	7	10313996	-19.1219	146.8808	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Magnetic Island	...	0.0
6982	7	10313996	-19.1219	146.8808	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Magnetic Island	...	0.0
6989	9	10314035	-18.7979	147.5206	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	...	0.0
6990	9	10314035	-18.7979	147.5206	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	...	0.0
6991	10	10314059	-18.7974	147.5200	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	...	0.0

Database Schema

- Connection **from Python to MySQL**
- Use of **foreign keys** to maintain **relational integrity**
- Use of MySQL for **ERD creation**

Entity Relationship Diagram



Example : MySQL query

```sql
create table sample as
select sample_id, latitude_degrees, longitude_degrees, ocean
, realm, ecoregion, country, state_island_province, city_town
, distance_to_shore, exposure, year, date, turbidity, cyclone_frequency
, sum(case when substrate_name = 'Hard Coral' then percent_cover end) as hard_coral_percent_cover
, sum(case when substrate_name = 'Nutrient Indicator Algae' then percent_cover end) as nutrient_indicator_algae_percent_cover
, sum(case when substrate_name = 'Fleshy Seaweed' then percent_cover end) as fleshy_seaweed_percent_cover
, sum(case when substrate_name = 'Hard Coral' then percent_bleaching end) as hard_coral_percent_bleaching
, sum(case when substrate_name = 'Nutrient Indicator Algae' then percent_bleaching end) as nutrient_indicator_algae_percent_bleaching
, sum(case when substrate_name = 'Fleshy Seaweed' then percent_bleaching end) as fleshy_seaweed_percent_bleaching
from global_bleaching_env b
left join ocean o on b.ocean_id = o.ocean_id
left join country c on c.country_id = b.country_id
left join exposure e on e.exposure_id = b.exposure_id
left join bleaching_status bs on bs.bleaching_status_id = b.bleaching_status_id
group by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15;
```

Overview of the 'sample' table :

sample_id	latitude_degre...	longitude_degre...	ocean	realm	ecoregion	country	state_island_provin...	city_town	distance_to_sh...	exposure
10313996	-19.1219	146.8808	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Magnetic Island	126.24	Sometimes
10314035	-18.7979	147.5206	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	57425.12	Sheltered
10314059	-18.7974	147.52	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	57467.04	Sheltered
10314060	-18.796	147.5213	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	57656.01	Sheltered
10314022	-18.6872	147.0968	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Great Barrier Reef Marine Park	43199.05	Sheltered
10313871	-18.6689	146.507	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Palm Islands	528.77	Sheltered
10313872	-18.6689	146.507	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Palm Islands	528.77	Sheltered
10313873	-18.6689	146.507	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Palm Islands	528.77	Sheltered
10313874	-18.6689	146.507	Pacific	Central Indo-Pacific	Central and northern Great Barrier Reef	Australia	Queensland	Palm Islands	528.77	Sheltered

Flask API

- Ressources exposed from ‘sample’ table on MySQL
- **4 endpoints :**

  /samples

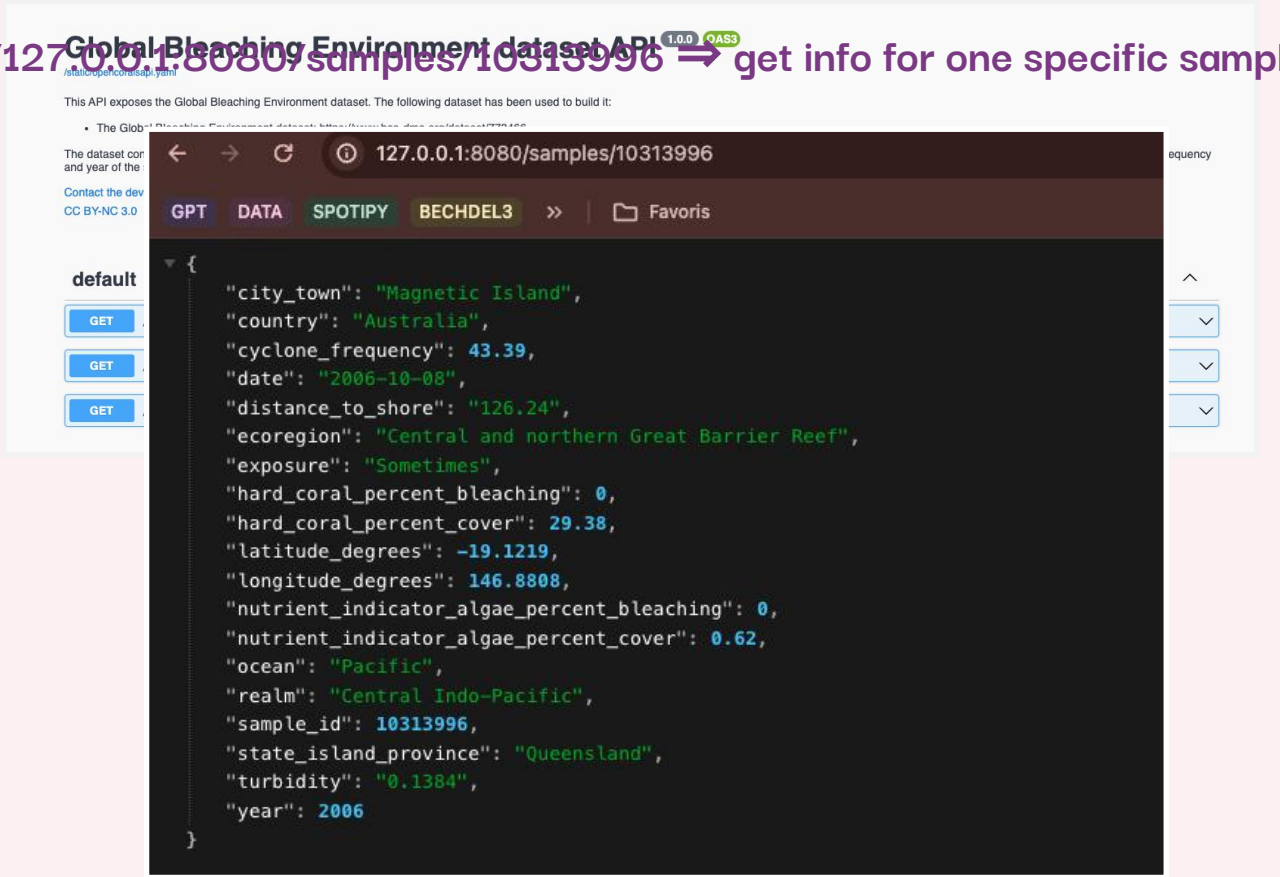
  /samples/<sample_ID>

  /samples/year/<year>

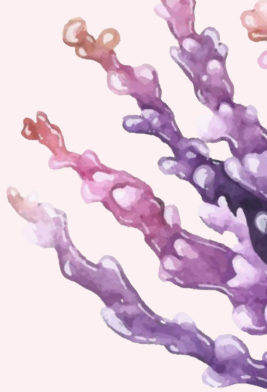
  /get_sample_ids
- **API Documentation** with Swagger UI

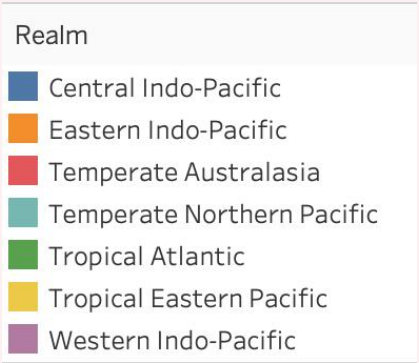
Example Usage

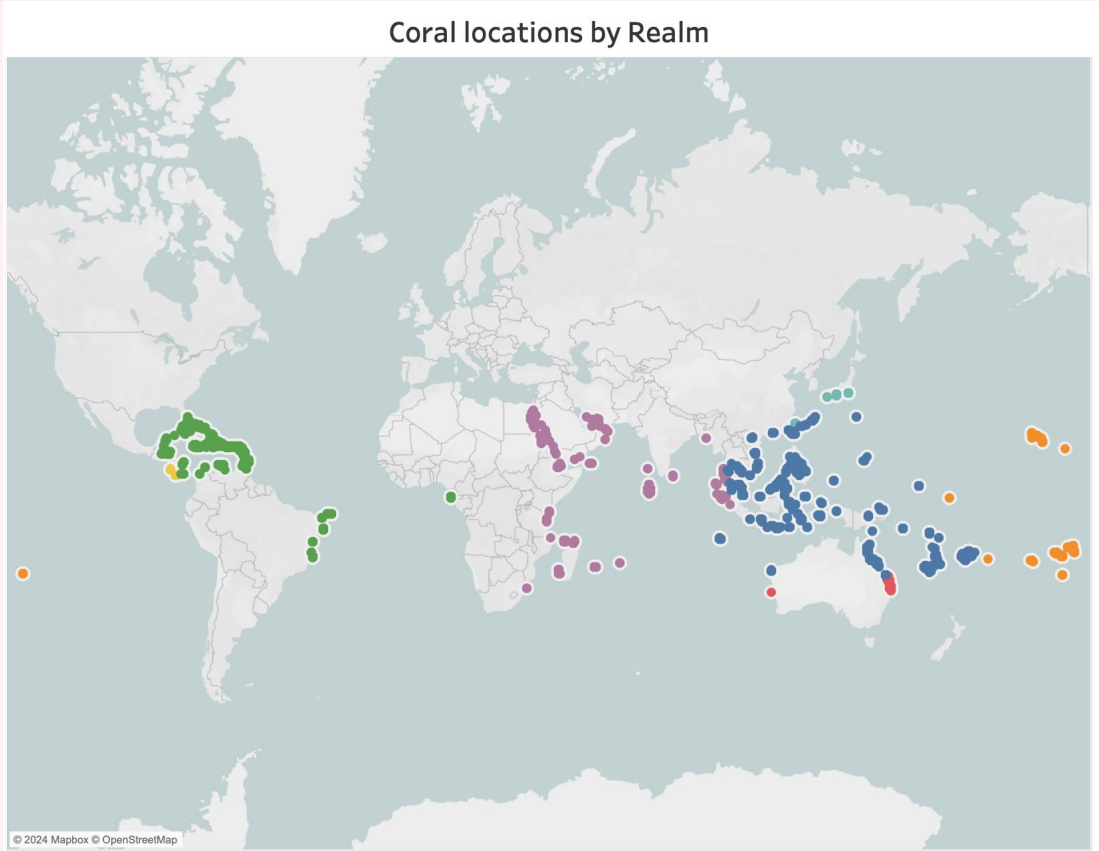
http://127.0.0.1:8080/samples/10313996 ⇒ get info for one specific sample_id



Visualizations

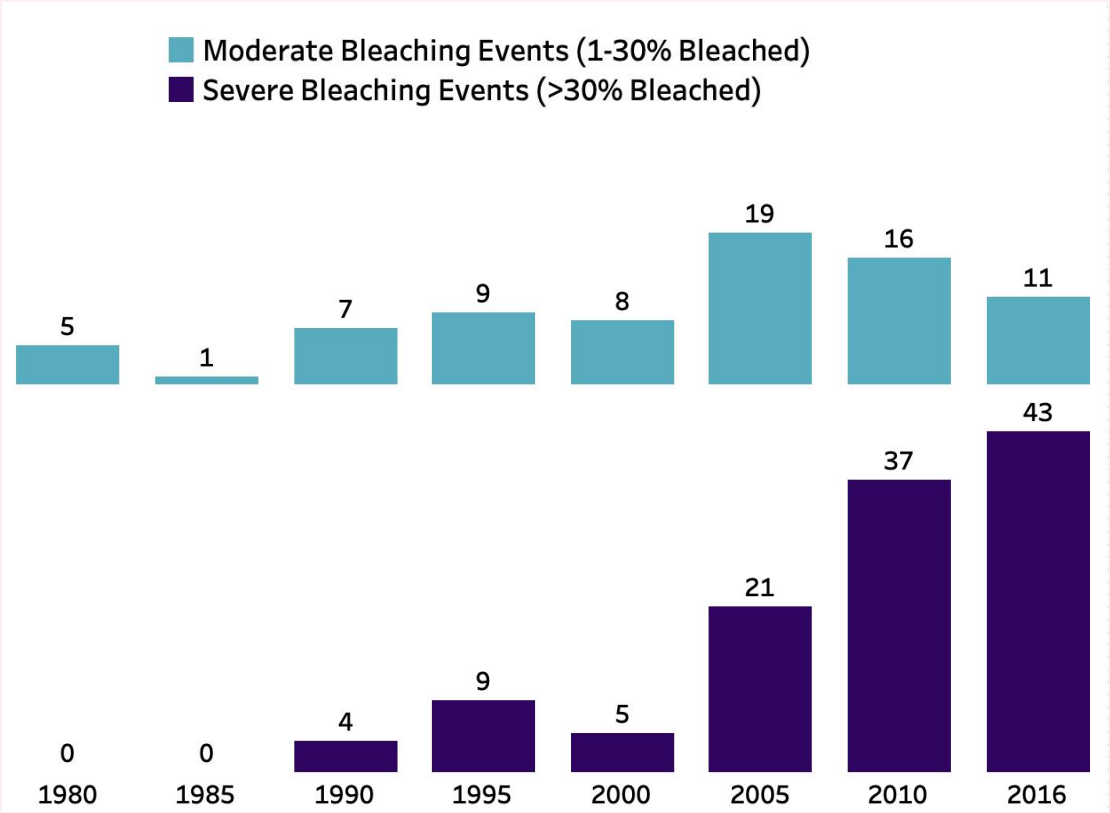






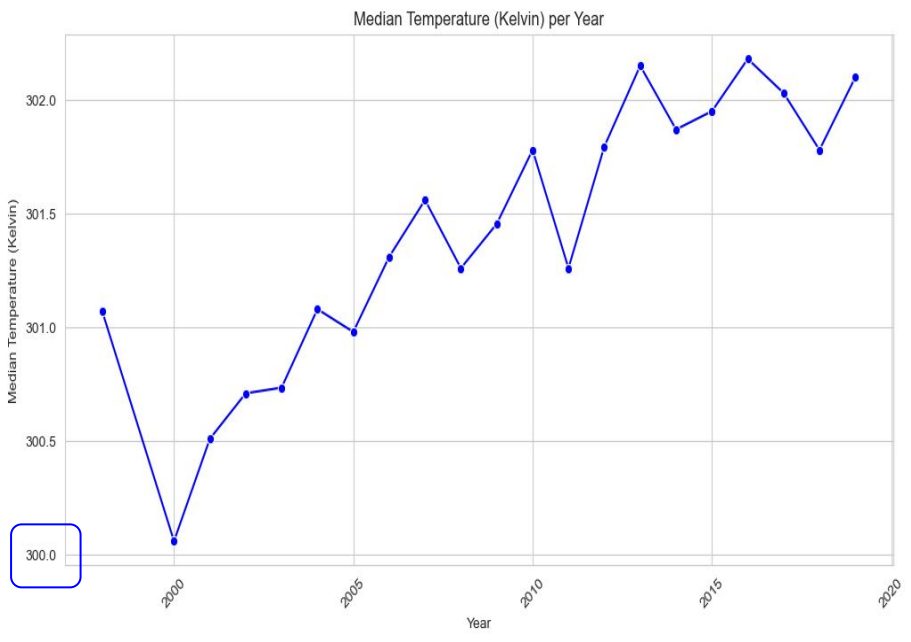
Visualizations

Distribution of bleaching events over the last 40 years, Worldwide

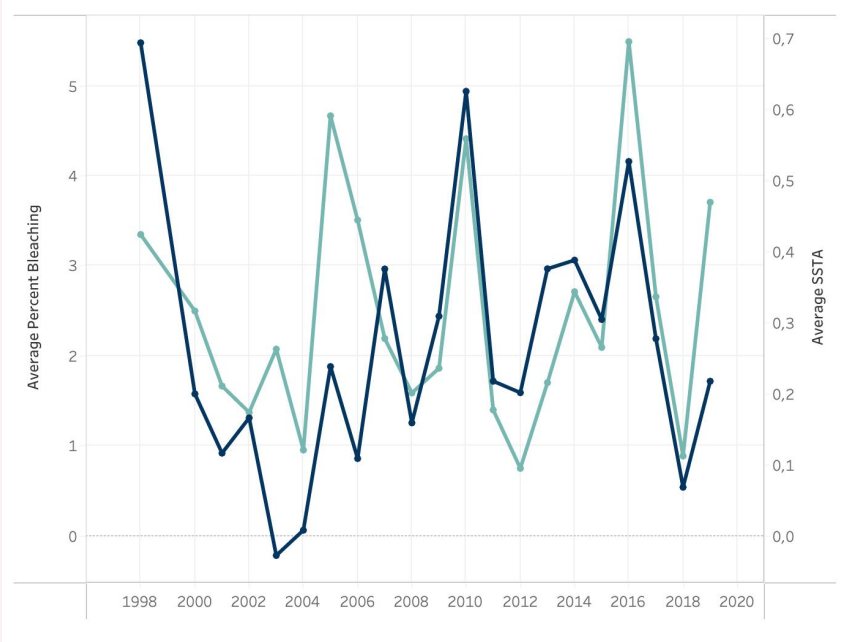


Visualizations

Temperature curve over the last 20 years



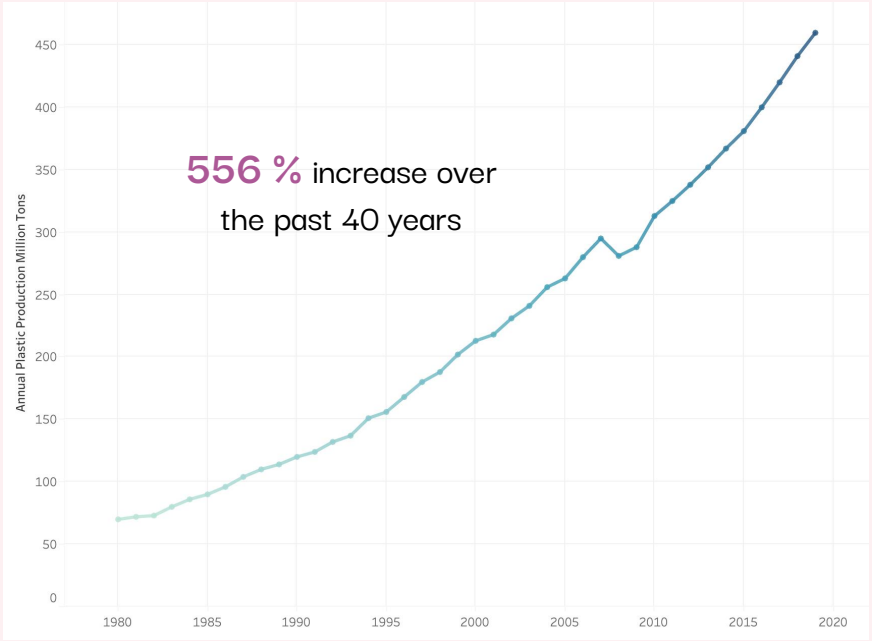
Average Percent Bleaching vs Average SSTA



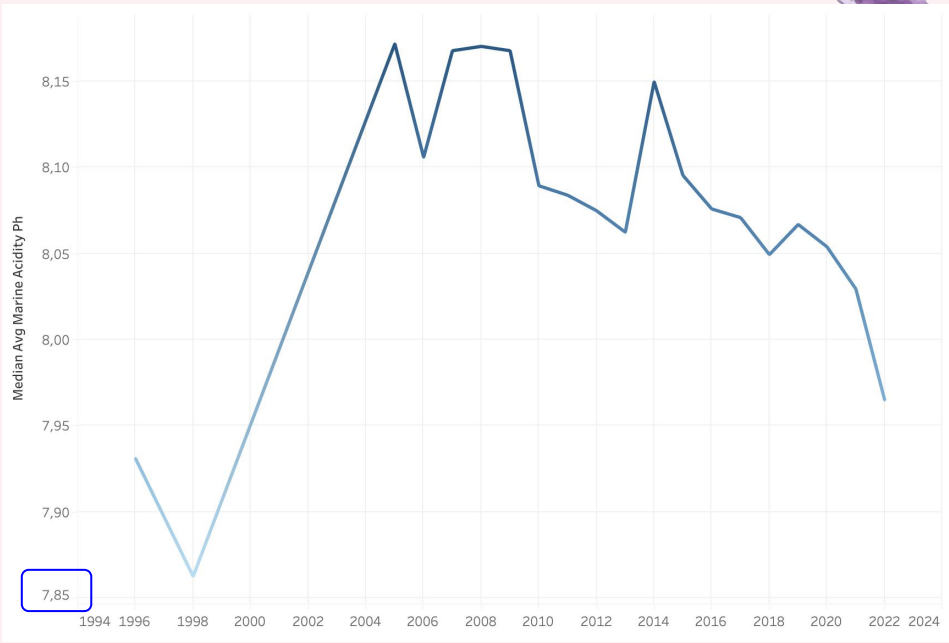
SSTA : Sea Surface Temperature Anomaly

Visualizations

Global Plastic Production over the last 40 years

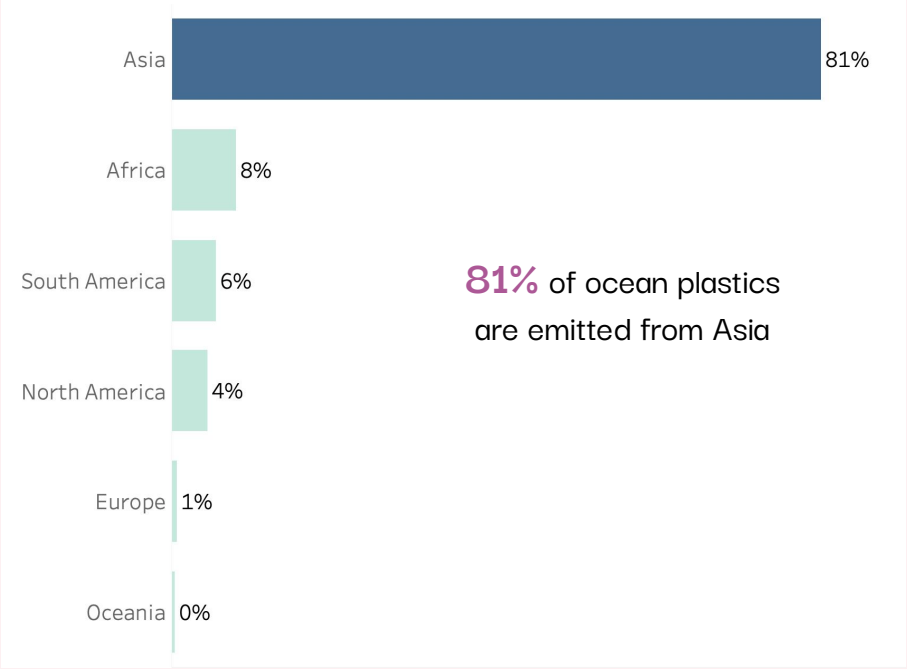


Marine acidy (pH) curve over the last 30 years



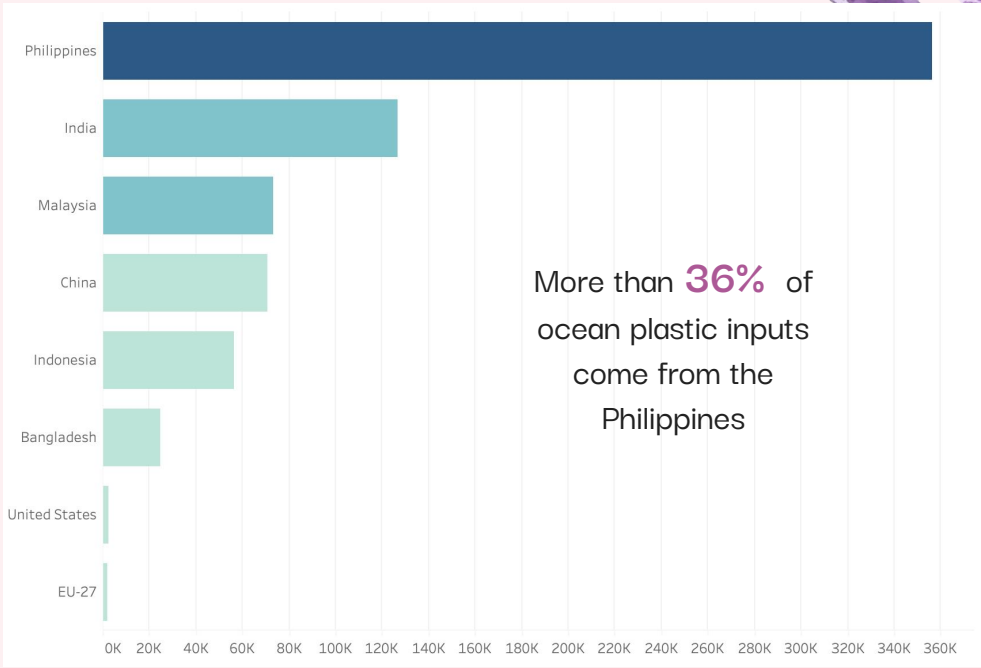
Visualizations

Share of global plastic waste emitted to the ocean, 2019



81% of ocean plastics are emitted from Asia

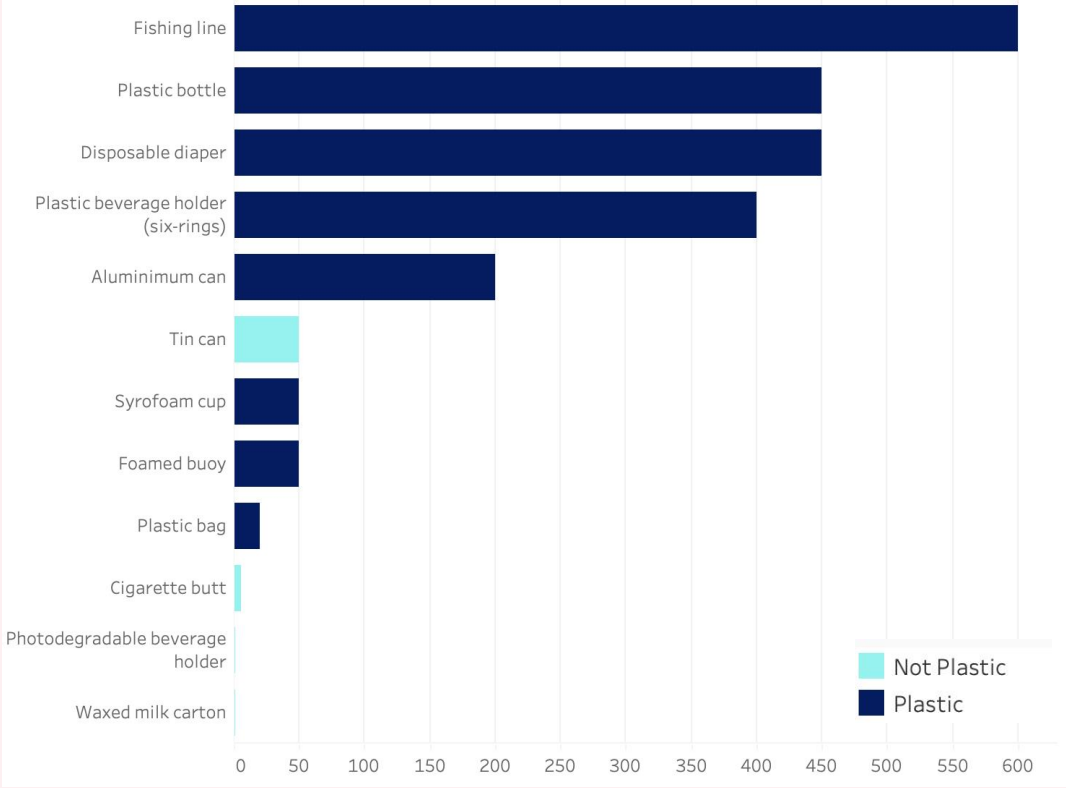
Plastic wasted emitted to the ocean, 2019



More than **36%** of ocean plastic inputs come from the Philippines

Visualizations

Average estimated decomposition times (in years) of typical marine debris items



How can we help corals?

Accurately identifying and monitoring the health of coral reefs for effective conservation and restoration efforts

However..

- manual identification of bleached and healthy corals can be a **time-consuming** and **labor-intensive process**
- would require **expert knowledge** & continuous monitoring given **the vast expanse of coral reefs** and the limited resources available for conservation efforts.

VGG19 Model

- VGG19 can be leveraged to **classify images of corals** into different categories based on their health status, such as healthy corals, bleached corals, or diseased corals.

- By fine-tuning the pretrained VGG19 model on a dataset of coral images, it can learn to distinguish between different **types of coral conditions** with high accuracy.

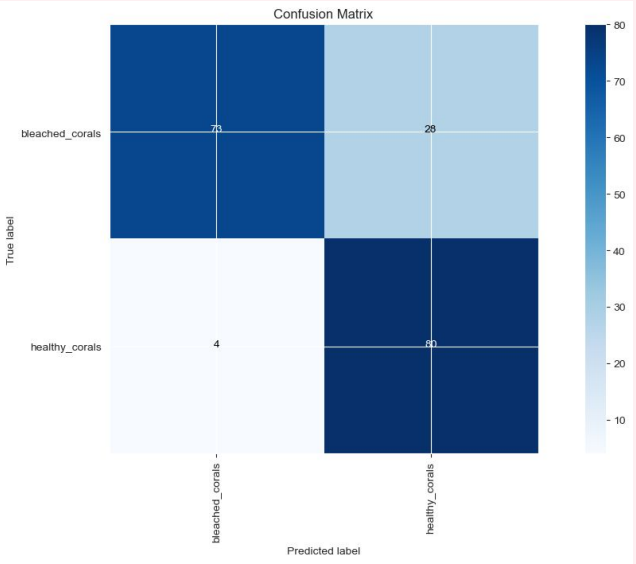
Main Results

Classification report

```
y_test = list(test_df.Label)
print(classification_report(y_test, pred))
```

	precision	recall	f1-score	support
bleached_corals	0.95	0.72	0.82	101
healthy_corals	0.74	0.95	0.83	84
accuracy			0.83	185
macro avg	0.84	0.84	0.83	185
weighted avg	0.85	0.83	0.83	185

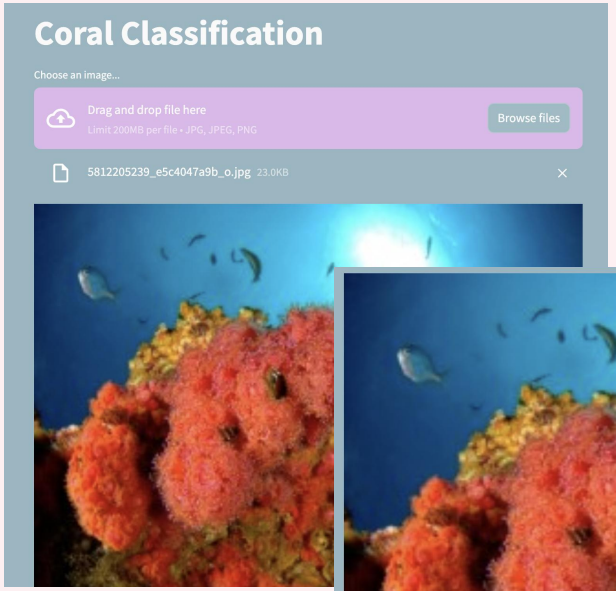
Confusion matrix



Challenges & Next Steps

- Data collection around my topic
- Time management
- Technical challenges using the pre trained model

Demo



http://localhost:8501/

Thank You :)

Do you have any questions? Let's talk !