ASSIGNMENT 2 FRONT SHEET

Qualification	**BTEC Level 5 HND Diploma in Computing**		
Unit number and title	Unit 14: Business Intelligence		
Submission date	6/29/2021	**Date Received 1st submission**	
Re-submission Date		**Date Received 2nd submission**	
Student Name	Thatsadaphone Inthatpakdy	**Student ID**	GCD191062
Class	Team 2: (Class GCD0805) Mr. Vo Ngoc Thanh Mr. Nguyen Hoang Mr. Thatsadaphone Inthapakdy	**Assessor name**	Tran Trong Minh
Student declaration I certify that the assignment submission is entirely my own work and I fully understand the consequences of plagiarism. I understand that making a false declaration is a form of malpractice.			

Student's signature	Thatsadaphone Inthapakdy

Grading grid

P3	P4	P5	P6	M3	M4	D3	D4

❒ Summative Feedback:		❒ Resubmission Feedback:
Grade:	**Assessor Signature:**	**Date:**
IV Signature:		

UNIVERSITY of GREENWICH

Alliance with FPT Education

6/27/2021

Design Tableau & Build BI Tool

Plastic Pollution & Trading Dashboard with Program Python code for BI Tool

Mr. Vo Ngoc Thanh
Mr. Nguyen Hoang Lam
Mr. Thatsadaphone Inthapakdy
THE OCEAN CLEAN UP

Contents

Group's Work Reporting on designing BI Tools

LO3 Demonstrate the use of business intelligence tools and technologies

Design and install a small application based on selected business intelligence tools to demonstrate the application of business intelligence to the given dataset. Application (though it's small and demonstrated) should be user-friendly and geared toward high-end users. Demo the application and collect comments from users

P3 Determine, with examples, what business intelligence is and the tools and techniques associated with it.

1. General about BI

a. Define Business Intelligence

BI is a process that integrates the technology that businesses use to control the huge amount of data coming from different sources and exploit that data source to enable them to make decisions more effectively in their business activities. BI is present throughout businesses such as supermarkets, banks, telecommunications, that are all places to collect and handle huge amounts of data.

Therefore, BI has a very high applicability when the data source of the enterprise will increase with the time of operation.

b. Real examples of how to apply BI on business:

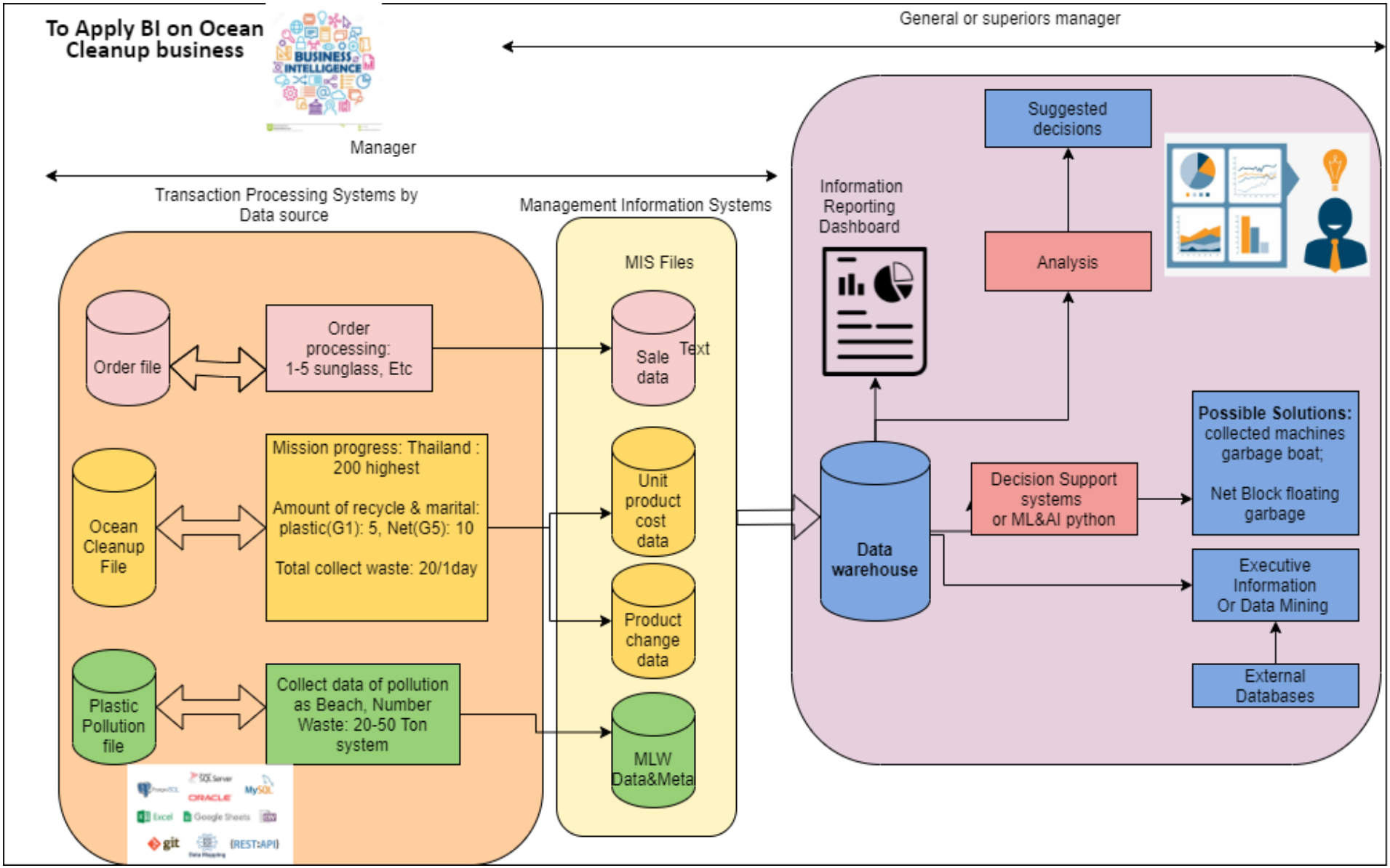


Figure 1 apply BI on business ocean cleanup and design tool architecture on decision levels

We will define used Techniques based on Real example above

To collect raw data in Transaction processing system with raw details by step taken according from ASM1's Business process such as

- **Order file.**
- **Ocean cleanup file.**
- **Plastic Pollution file.**

By collect those raw data will be transacted to become managed Information by MIS convert to those raw data to clean, organize into label as Excel file as well as:

- **Sale data**

In the IS's BI, Online site's Sale database of MIS included in the product database can be report as:

- Add a Product Line order
- Change a Product Price

In query, many products sold due to site a month could be addition of product line or change a product price. Also, ML is used to calculate prediction trending of future product with fix price.

- **Unite product cost & change data.**

In product cost & change database of MIS could be executed

- Mission progress
- Amount of recycle & Materials
- Total collection waste per day

In query could be calculated by ML as "regression amount of decreased plastic waste 15% per months and reuse Material of plastic recycle into industry" to change if more 5 years, how much pollution will be decreased more and revenue of reselling plastic waste.

- **MLW_Data&Meta data.**

In MLW Data-Meta database of MIS could be executed:

- Targeted country's Beach detail and observation.
- Each type & specific of Pollution categories that are collection by Ocean cleanup

In query could be known which Beach demographic is holding Top amount of highest plastic pollution. however, using ML to read meta detail in Dataset which understating then apply regression method will search about conclusion total waste that each beach's holding most with select Top highest. Also, in each beach can specify waste into category such as Plastic, wood, Net and etc. which Ocean cleanup can select country's beach that have highest plastic cate to assign Mission to.

2. BI techniques:

a. Collection & Analysis techniques: cleansing, labeling

- **Data mining:**
  - provides a concise summary of the given data collection
  - Exploiting, analyzing and identifying models for databases or trend forecasts.

Example: E-commerce field collects a lot of historical data as well as current data. Data mining applications are examining that data with its predictive analysis algorithms and offering best-selling options to customers. From many different ways, e-commerce sites will use data mining tools. Many e-commerce companies use Data Mining and Business Intelligence to provide cross-selling and selling through their websites as amazon.

- **Predictive analytics**
  - Predict probabilities and trends.
  - Predict the value of a data item attribute using different statistical.
  - Orient the best process for certain situations.
  - Optimization & simulation.

Example: Send marketing campaigns to customers who are more likely to buy. For businesses with only $ 5,000 budget for increased marketing and three million customers, it is obvious that it is impossible to extend the 10% discount for each customer. Predictive and smart business analysis can help predict customers who have the highest probability of purchasing your product, then send coupons to those people to optimize revenue.

- **Classify**

- Identify the class of the data item.

Example: In the case of black and white images, the intensity of each pixel is used as one of the measurements. In color images, each pixel provides three intensity measurements in three different colors - red, green and blue (RGB).

- **Association, correlation**
  - Identify the relationship between attributes.

Example: For example, an electric utility can produce less energy on a light day based on the relationship between electricity demand and weather. In this example, there is a causal relationship, because inclement weather makes people use more electricity to heat or cool; however, statistical dependence is not sufficient to prove the presence of such a causal relationship.

- **Online Analytical Processing (OLAP)**
  - Allow analysis of different sizes of data.

Example: An e-commerce company wants to compare its sales figures in February and March and also wants to see a wise sales area, then a wise, wise state of time and ultimately wise.

To achieve this, a system is needed to insert data from different OLTP databases into the Data warehouse and apply the ETL process. Later, OLAP developers will take data from OLAP systems and create different types of reports and charts based on business requirements.

- **Model Visualization**
  - Events are made possible by charts, storylines, charts and other visual means.

Example: Languages in the world allow exploring popular languages, see which languages are most spoken and see the languages used throughout the world. This is a great visual storytelling: take an in-depth topic and break it in an easy-to understand way

3. We have selected BI tools for Our dataset BI project from reported ASM1:

a. Programming tools as regression and machine learning for Python

- **Code IDE**

There are many IDE to code Python such as Spyder IDE, Jupyter-notebooK IDE, PyCharm IDE and Visual Code Etc. however, we will use whatever familiar with.

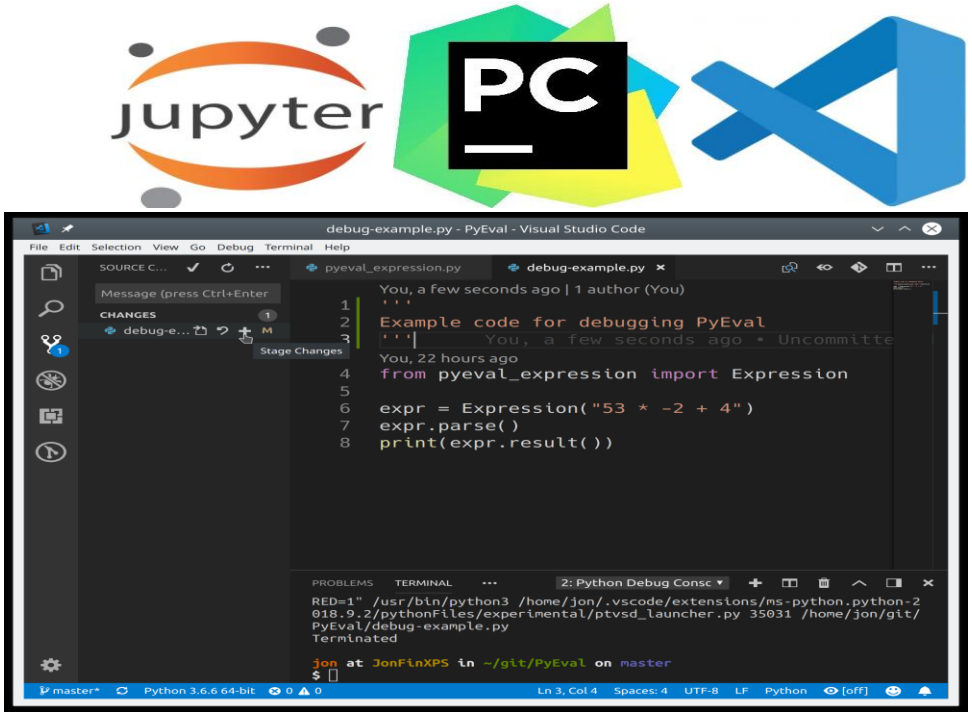


Figure 2 Programming code tool IDE

Select wether IDE is flexible: configure and arrange the user interface to support a wide range of workflows in data science and scientific computing .

Therorefore, that allows us as beginner to create and share documents that contain live code, equations, visualizations and narrative text. Uses include: data cleaning and transformation, numerical simulation, statistical modeling, data visualization, machine learning, and regression searches for relationships among variables.

- **Pyhon Programing language**

Figure 3 apply BI tool woth Best Python programming language for large Data data science

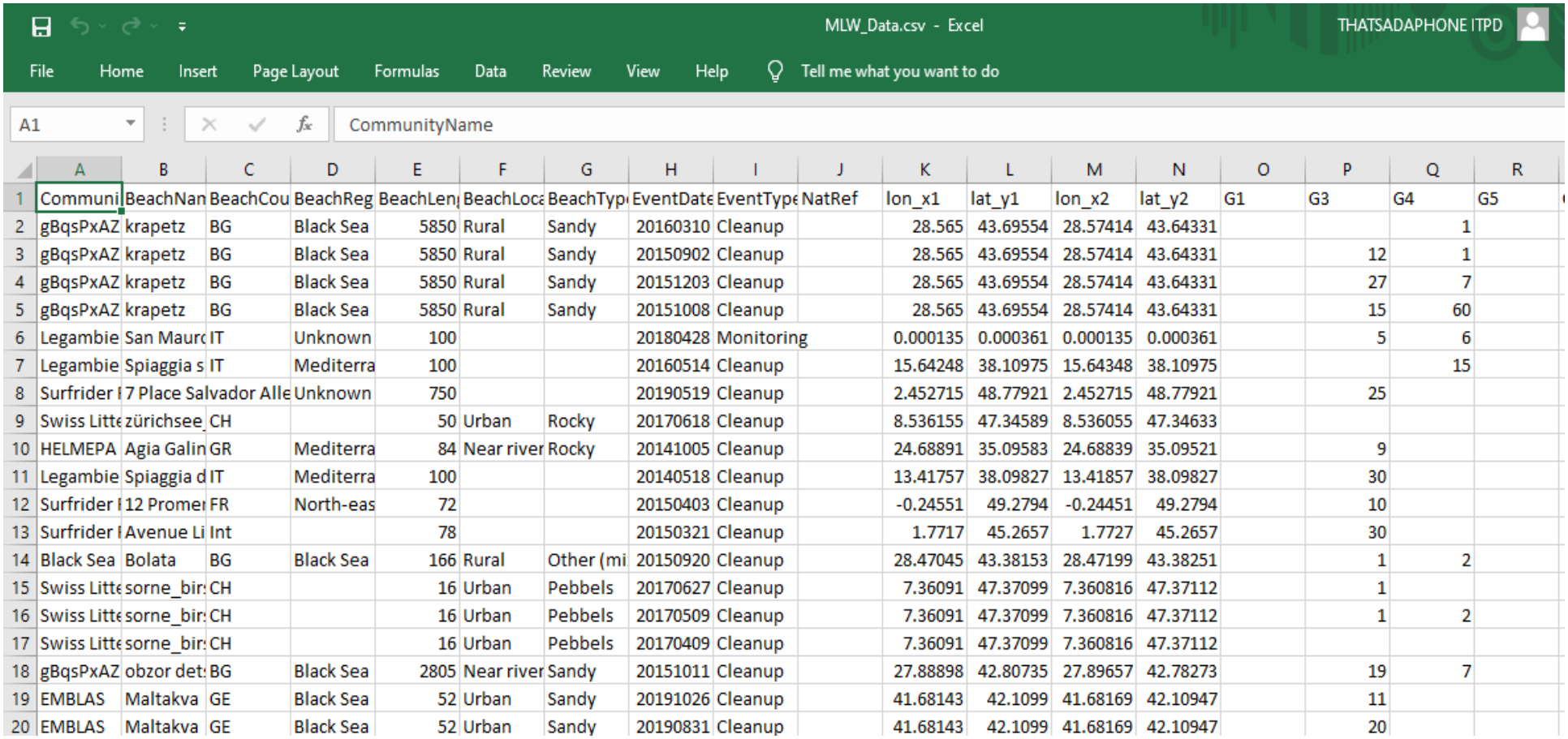
For Python programming language, we prefer to use Python to code Machine Learning in BI's Dataset which able to use working with a variety of modules, which allows it to start up and soon realize that everything is an object that has a namespace itself. Also, to give the program structure while keeping it clean and simple as why Python excels at introspection, which comes from the object nature of Python. Also, regression searches for relationships among variables.

Therefore, Python is most suable for data science as well as ML, Deep learning and AI with powerful libraries. We should use python because data structure of python is very similar of Human language with Easy-to-learn.

b. Database / data warehouse tools by Microsoft Excel

- **Spreadsheets**: Often use excel, open-source spreadsheets. Website based spreadsheet

Example: Below is the spreadsheets indicating plastic pollution on Beach Reviews:



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Communi	BeachNan	BeachCou	BeachReg	BeachLen	BeachLoca	BeachTyp	EventDate	EventType	NatRef	lon_x1	lat_y1	lon_x2	lat_y2	G1	G3	G4	G5
2	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20160310	Cleanup		28.565	43.69554	28.57414	43.64331			1	
3	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20150902	Cleanup		28.565	43.69554	28.57414	43.64331		12	1	
4	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20151203	Cleanup		28.565	43.69554	28.57414	43.64331		27	7	
5	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20151008	Cleanup		28.565	43.69554	28.57414	43.64331		15	60	
6	Legambie	San Maurc	IT	Unknown	100			20180428	Monitoring		0.000135	0.000361	0.000135	0.000361		5	6	
7	Legambie	Spiaggia s	IT	Mediterra	100			20160514	Cleanup		15.64248	38.10975	15.64348	38.10975			15	
8	Surfrider f	7 Place Salvador Alle		Unknown	750			20190519	Cleanup		2.452715	48.77921	2.452715	48.77921		25		
9	Swiss Litte	zürichsee_	CH		50	Urban	Rocky	20170618	Cleanup		8.536155	47.34589	8.536055	47.34633				
10	HELMEPA	Agia Galin	GR	Mediterra	84	Near river	Rocky	20141005	Cleanup		24.68891	35.09583	24.68839	35.09521		9		
11	Legambie	Spiaggia d	IT	Mediterra	100			20140518	Cleanup		13.41757	38.09827	13.41857	38.09827		30		
12	Surfrider f	12 Promer	FR	North-eas	72			20150403	Cleanup		-0.24551	49.2794	-0.24451	49.2794		10		
13	Surfrider f	Avenue Li	Int		78			20150321	Cleanup		1.7717	45.2657	1.7727	45.2657		30		
14	Black Sea	Bolata	BG	Black Sea	166	Rural	Other (mi	20150920	Cleanup		28.47045	43.38153	28.47199	43.38251		1	2	
15	Swiss Litte	sorne_bir	CH		16	Urban	Pebbels	20170627	Cleanup		7.36091	47.37099	7.360816	47.37112		1		
16	Swiss Litte	sorne_bir	CH		16	Urban	Pebbels	20170509	Cleanup		7.36091	47.37099	7.360816	47.37112		1	2	
17	Swiss Litte	sorne_bir	CH		16	Urban	Pebbels	20170409	Cleanup		7.36091	47.37099	7.360816	47.37112				
18	gBqsPxAZ	obzor det	BG	Black Sea	2805	Near river	Sandy	20151011	Cleanup		27.88898	42.80735	27.89657	42.78273		19	7	
19	EMBLAS	Maltakva	GE	Black Sea	52	Urban	Sandy	20191026	Cleanup		41.68143	42.1099	41.68169	42.10947		11		
20	EMBLAS	Maltakva	GE	Black Sea	52	Urban	Sandy	20190831	Cleanup		41.68143	42.1099	41.68169	42.10947		20		

Figure 4 Example plastic pollution dataset of MLW_data.csv

- **OLAP tools**: OLAP is online analytical processing. Help users interactively analyze data from different sources in multidimensional mode based on the user's perspective. The OLAP functions are drilled deep, drilling on them is consolidation, exchange, synthesis, and division.

Example:

Datawarehouse is an OLAP system:

- A company can compare its mobile phone sales in September with sales in October, then compare those results with another location that can be stored in the censorship database.
- Amazon analyzes customer purchases to offer a personalized home page with products that may interest their customers.

c. Data visualization tools as report, queries, dashboards by Tableau

- **Digital dashboard**: Use for real-time reporting by graph of the current state that the user wants to see.

Example: Investor relations dashboard provides invaluable data focusing on the company's investors. Help ensure the health and consistency of these relationships include Return on Assets, Return on Equity and P / E Ratio as well as extremely important stock prices, this digital panel software provides all important investment information in one central location.

Figure 5 Digital dashboard

- **Data Visualization Software**: Help users visually analyze data sets to create customer visualization for meaningful insight. The images help users to distinguish and share with the company's department.

Example: This is a graph that visually shows the location of the Covid19 virus is being reported and previous transmissions discovered around the world.

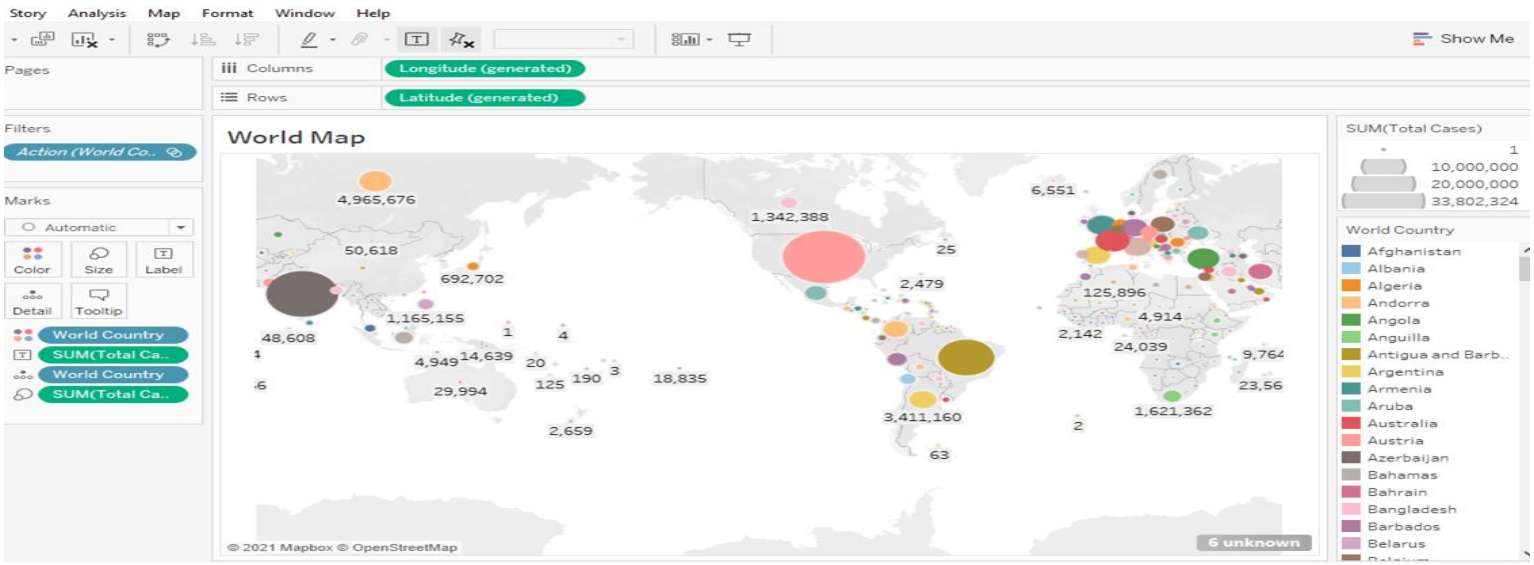


Figure 6 Example designed Dashboard of breakout convid19

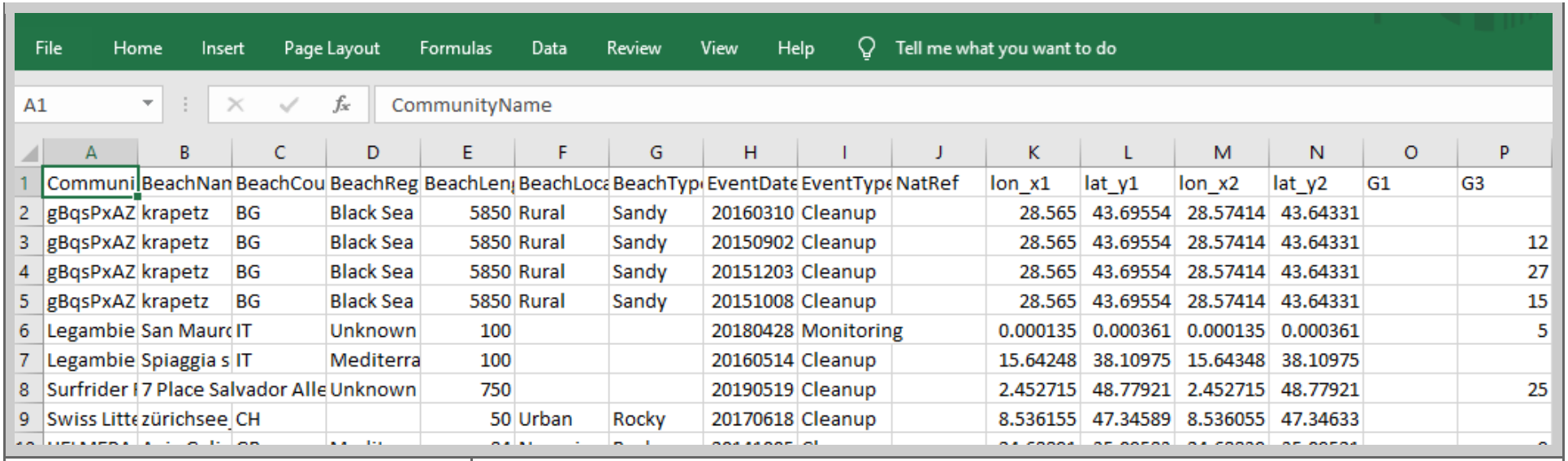
P4 Design BI tool with application UI to perform a specific task to support problem-solving or decision-making at an advanced level.

Team Working on Tableau software of BI Tool

1. Explain dataset

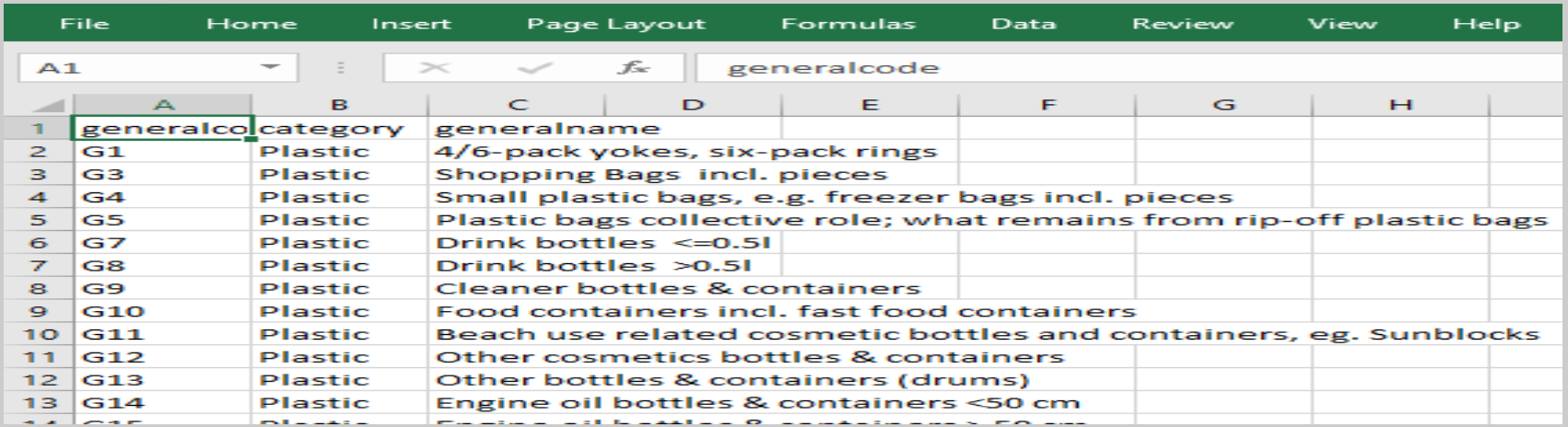
Global plastic pollution Dataset's Columns List	Explain Dataset and Type
MLW Data	**This dataset is used store Targeted Beach's Detail information and sort Each rubbish code of categories into columns.**

1	Community Name	Identity every country that have beach. Ex: Surfrider Foundation Europe, Marnoba
2	Beach Name	Display the available dirty beach. Ex: Krapetz, Maltakva…
3	Beach Country code	Beach with country code. Ex: BG, IT, GR, CH…
4	Beach Regional Sea	Location sea that Beach stay with. Ex: Black Sea, Mediterranean Sea…
5	Beach Length m	Each beach scales. Ex: 5850, 50 ,100, 750…
6	Beach Location	Each Beach Location Type. Ex: Rural, Urban, Near river….
7	Beach Type	A kind of Beach. Ex: Rocky, Sandy, Pebbles….
8	Event Date	Activated time of Event. Ex: 20150902, 20151008….
9	Event Type	Activity on Beach. Ex: Cleanup or Monitoring
10	Nat Ref	Employ the system works. EX: RO000002, __none__ ….
11	G1-G200	Each General code of rubbish category which collect number amount of specific waste. EX: 1, 3, 22, 50, 100



A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Communi	BeachNan	BeachCou	BeachReg	BeachLen	BeachLoc	BeachTyp	EventDate	EventType	NatRef	lon_x1	lat_y1	lon_x2	lat_y2	G1	G3
2	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20160310	Cleanup		28.565	43.69554	28.57414	43.64331		
3	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20150902	Cleanup		28.565	43.69554	28.57414	43.64331		12
4	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20151203	Cleanup		28.565	43.69554	28.57414	43.64331		27
5	gBqsPxAZ	krapetz	BG	Black Sea	5850	Rural	Sandy	20151008	Cleanup		28.565	43.69554	28.57414	43.64331		15
6	Legambie	San Maurc	IT	Unknown	100			20180428	Monitoring		0.000135	0.000361	0.000135	0.000361		5
7	Legambie	Spiaggia s	IT	Mediterra	100			20160514	Cleanup		15.64248	38.10975	15.64348	38.10975		
8	Surfrider I	7 Place Salvador Alle		Unknown	750			20190519	Cleanup		2.452715	48.77921	2.452715	48.77921		25
9	Swiss Litte	zürichsee_	CH		50	Urban	Rocky	20170618	Cleanup		8.536155	47.34589	8.536055	47.34633		

MLW Meta		**This Dataset is used to keep Category detail which identify common kind and specific of waste as Plastic, wood, net etc.**
12	General code	ID code to identity garbage with collecting rated number on each country's Beach. EX: G3 : 12 , G8: 31, G9 : 50…
13	Category_ garbage	Sort Common kind of garbage category in order. Ex: Plastic, Rubber, Paper and Cloth….
14	General name	Focus on the specific of each garbage. Ex: Buckets, Fenders, Net, CD, Package.



generalcode	category	generalname
G1	Plastic	4/6-pack yokes, six-pack rings
G3	Plastic	Shopping Bags incl. pieces
G4	Plastic	Small plastic bags, e.g. freezer bags incl. pieces
G5	Plastic	Plastic bags collective role; what remains from rip-off plastic bags
G7	Plastic	Drink bottles <=0.5l
G8	Plastic	Drink bottles >0.5l
G9	Plastic	Cleaner bottles & containers
G10	Plastic	Food containers incl. fast food containers
G11	Plastic	Beach use related cosmetic bottles and containers, eg. Sunblocks
G12	Plastic	Other cosmetics bottles & containers
G13	Plastic	Other bottles & containers (drums)
G14	Plastic	Engine oil bottles & containers <50 cm

Plastic connectivity table		**This Dataset is used to store information trading of rubbish between counties that Export and Import. We use this information to trade our recycle plastic with marketing.**
15	Country from	ID country that exports plastic. Ex: Thailand, Peru, China, Vietnam, India, Australia…
16	Country to	Tell Trading with other countries by Plastic to the destination country that imports it. Ex: Australia, New- Zealand, New Caledonia, Fiji, Vanuatu….
17	Total number particles released	Store total export quantity by Ton. Ex: 390636, 6122, 2337, 9525….
18	Average daily particles inside EEZ	Store import average 24/7. Ex: 110.91, 6.7444, 12.592 ….

	A	B	C	D
1	Country_from	Country_to	Total_number_particles_released	Average_daily_particles_inside_EEZ
2	Indonesia	Australia	390636	110.91
3	Fiji	Australia	6122	12.592
4	Vanuatu	Australia	2337	6.7444
5	Papua New Guinea	Australia	9525	4.8078
6	Philippines	Australia	215974	4.1962
7	Vietnam	Australia	176511	3.8257
8	East Timor	Australia	3390	3.2787
9	South Africa	Australia	100943	2.5649
10	China	Australia	752074	2.0412
11	Malaysia	Australia	106756	1.6821
12	Solomon Islands	Australia	1943	1.6407

Table 1 Explain our BI Project's main 3 dataset

2. Dataset's Pre-process steps in Machine Learning:

Data preprocessing is a data mining technique that involves transforming raw data into an understandable format. The real-world data are generally incomplete: lacking attribute values, lacking certain attributes of interest, or containing only aggregate data. Noisy: containing errors or outliers. Inconsistent: containing discrepancies in codes or names.

Therefore, in our BI's ML project, we must use data preprocessing on our Plastic pollution dataset which is a proven method of resolving such issues. Also, in another name of Data processing can be use as SDLC of ML project to application.

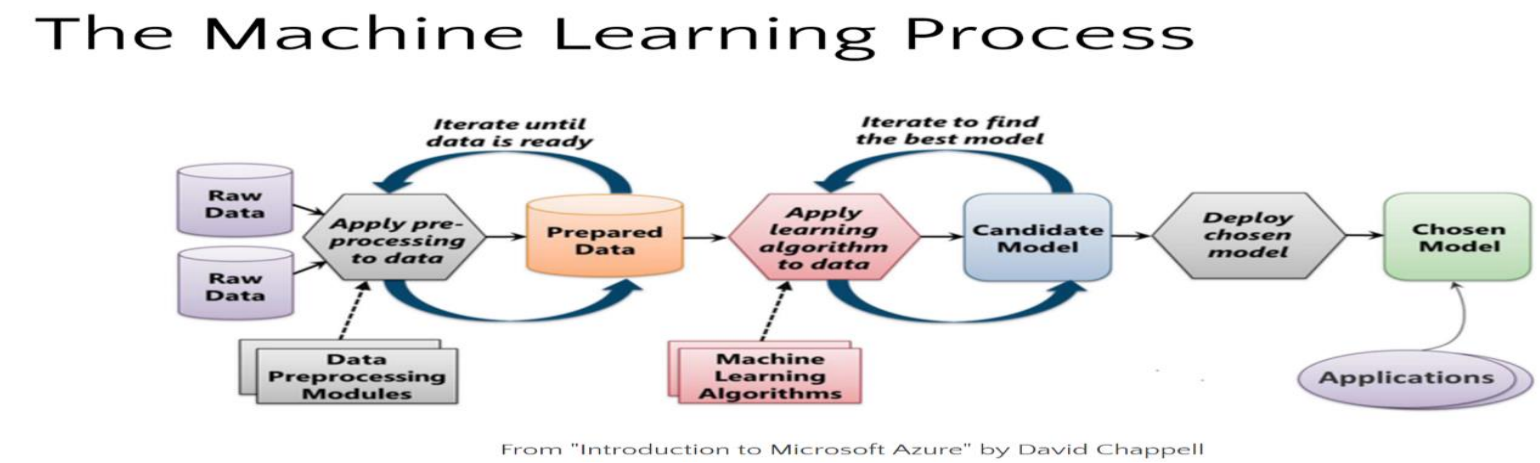
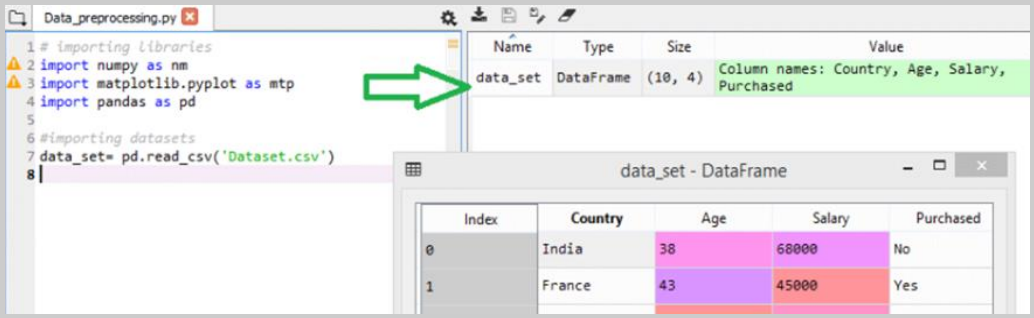


Figure 7 the ML process (Mohit S, Oct,2020)

Pre-process Steps in ML Application Interface (Gayal.K 2020)		Explain	Condition to apply on our Plastic pollution dataset (Status)
1	**Acquire the dataset**	To build and develop Machine Learning models, you must first acquire the relevant dataset. This dataset will be comprised of data gathered from multiple and disparate sources which are then combined in a proper format to form a dataset. Dataset formats differ according to use cases.	We are able to create a dataset by collecting data via different source. Once the dataset is ready, we must put it in a CSV or XLSX file formats. (**Apply to**)
2	**Import all the crucial libraries**	Python is the most used and preferred library by Data Scientists. The predefined Python libraries can perform specific data preprocessing jobs. The three core Python libraries used for this data preprocessing in Machine Learning are: **NumPy**: the fundamental package for scientific calculation in Python which is used for inserting any type of mathematical operation in the code. Also, to add large multidimensional arrays and matrices in your code.	▪ **NumPy as np for using mathematical method.** (**Apply to**) ▪ **Pandas as pd for analysis with managing the dataset**. (**Apply to**)

		Pandas: Data manipulation and analysis which is used for importing and managing the datasets. It packs in high-performance, easy-to-use data structures and data analysis tools for Python.
		
		Matplotlib: Python 2D plotting library which used to plot any type of charts in Python. Also, to deliver publication-quality figures in numerous hard copy formats and interactive environments across platforms:
		
		`# importing libraries`
		`import numpy as nm`
		`import matplotlib.pyplot as mtp`
		`import pandas as pd`	

3	**Import the dataset**	to import the dataset/s that we have gathered for the ML project at hand. But before import, we must set the current directory as the working directory in IDE  extract the independent variables, use “iloc[ ]” function of the Pandas library: `x= data_set.iloc[:,:-1].values`	(Apply to) **Create path and files name to read as set directory**

		[['India' 38.0 68000.0] ['France' 43.0 45000.0] ['Germany' 30.0 54000.0] ['France' 48.0 65000.0] ['Germany' 40.0 nan]	
4	**Identifying and handling the missing values**	pivotally to identify and correctly handle the missing values, failing to do this, we might draw inaccurate and faulty conclusions and inferences from the data. Needless to say, this will hamper our ML project. two ways to handle missing data: - **Deleting a particular row** – In this method, to remove a specific row that has a null value for a feature or a particular column where more than 75% of the values are missing. - **Calculating the mean** – This method is useful for features having numeric data like age, salary, year, etc. by this, we can calculate the mean, median, or mode of a particular feature or column or row that contains a missing value and replace the result for the missing value.	(**Apply to**) **Deleting a particular row, the Dataset in MLW_Data.csv have too much null value or not has unusable value. So, we will delete some row and select only Generalcode that is included on plastic category** (**Apply to**) **Calculating the mean for additional miss-value as Total rubbish and average value of mode.**
5	**Encoding the categorical data**	Categorical data refers to the information that has specific categories within the dataset. In the dataset cited above, there are two categorical variables such as country and purchased. ML models are based on mathematical equations. So, able to understand that keeping the categorical data in the equation will cause certain issues since only need numbers in the equations. array([[2, 38.0, 68000.0], [0, 43.0, 45000.0], [1, 30.0, 54000.0],	(**Not Apply to**) the country and beach column won't cause problems because Beach's pollution detail and categories are separated into two files, if uses category, would use as related database with primary key generalcode with G1-213. So, we don't need to convert it into numerical values.

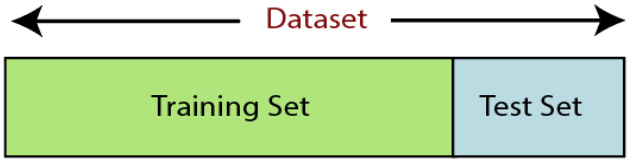
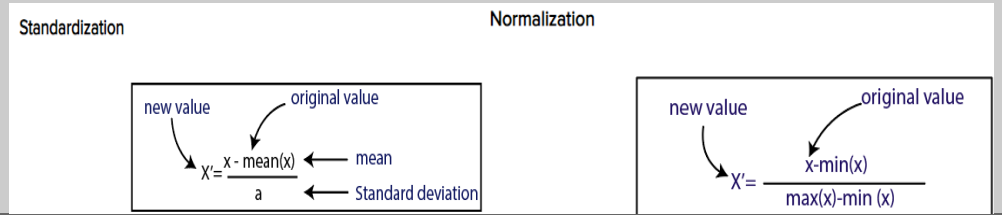
6	**Splitting the dataset**	Every dataset for Machine Learning model must be split into two separate sets – training set and test set. 	(**Apply to**) **Split Global plastic pollution dataset into two sets as MLW_Data(Beach detail set) and MLW_Mata(category Set) for avoid case problem as mess-data and over-labeling.**
7	**Feature scaling**	A method to standardize the independent variables of a dataset within a specific range. In other words, feature scaling limits the range of variables so that you can compare them on common grounds. 	(**Apply to**) **we must combine and modify each calculation into one Method to predict solution as well as each define function by programing code.**

Table 2 apply process Steps of building python BI or programing BI tool on ML application into our Dataset

3. Design BI tool, application UI to perform a specific task for decision-making at an advanced level. (By Draw.io, UI of Reporting's Interface: Sheet, Story and Dashboards)

We are using Datawarehouse Architecture to be based Steps to design BI Tools to do a specific task as below:

- **Data gathering and Data cleaning/standardization sages**

Excel Collect: a raw data will be stored in files by integration of data mining with cleaned and labeling in appropriate Table of Excel file or database.

- **Analysis and Reporting sages**

Tableau Software: to act as Datawarehouse which able to connect many multiple Data source as Excel files into a one place database. Also, this capable software has many functional tools which include analyze, Reporting, Dashboard result to react Data data-driven decision and Embedded in Application.

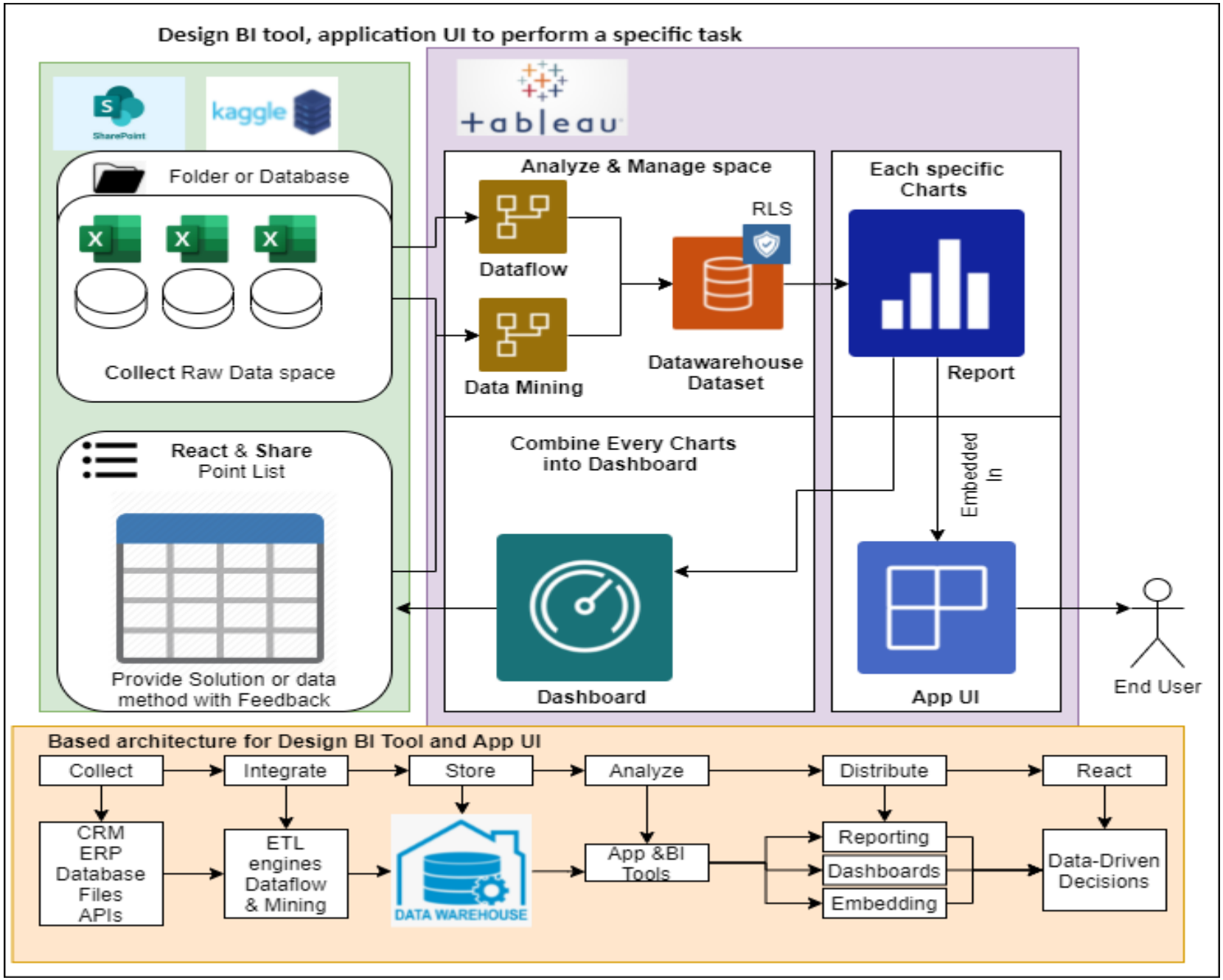

Design BI tool, application UI to perform a specific task
SharePoint
kaggle
tableau
Folder or Database
Collect Raw Data space
React & Share Point List
Provide Solution or data method with Feedback
Analyze & Manage space
Dataflow
Data Mining
RLS
Datawarehouse Dataset
Combine Every Charts into Dashboard
Dashboard
Each specific Charts
Report
Embedded In
App UI
End User
Based architecture for Design BI Tool and App UI
Collect
Integrate
Store
Analyze
Distribute
React
CRM
ERP
Database
Files
APIs
ETL engines Dataflow & Mining
DATA WAREHOUSE
App &BI Tools
Reporting
Dashboards
Embedding
Data-Driven Decisions

Figure 8 Design Taken software include by BI tool architecture to develop our dataset Dashboard.

Base on our Global-plastic Polluion dataset, we clould divide into Three Specific tasks of different chart. By each one will contain different value for specific reports. Also, we have design each chart interact that related to our dataset.

This Bar report is used to compare between specific name and value number. We use bar chart beacus our datasets have specific Beach_Name with Total coculation of general_code rubbish which contain value number in every beach; general_code_name with Categories_specific and Country_From with Total and Average of Im-Explort.

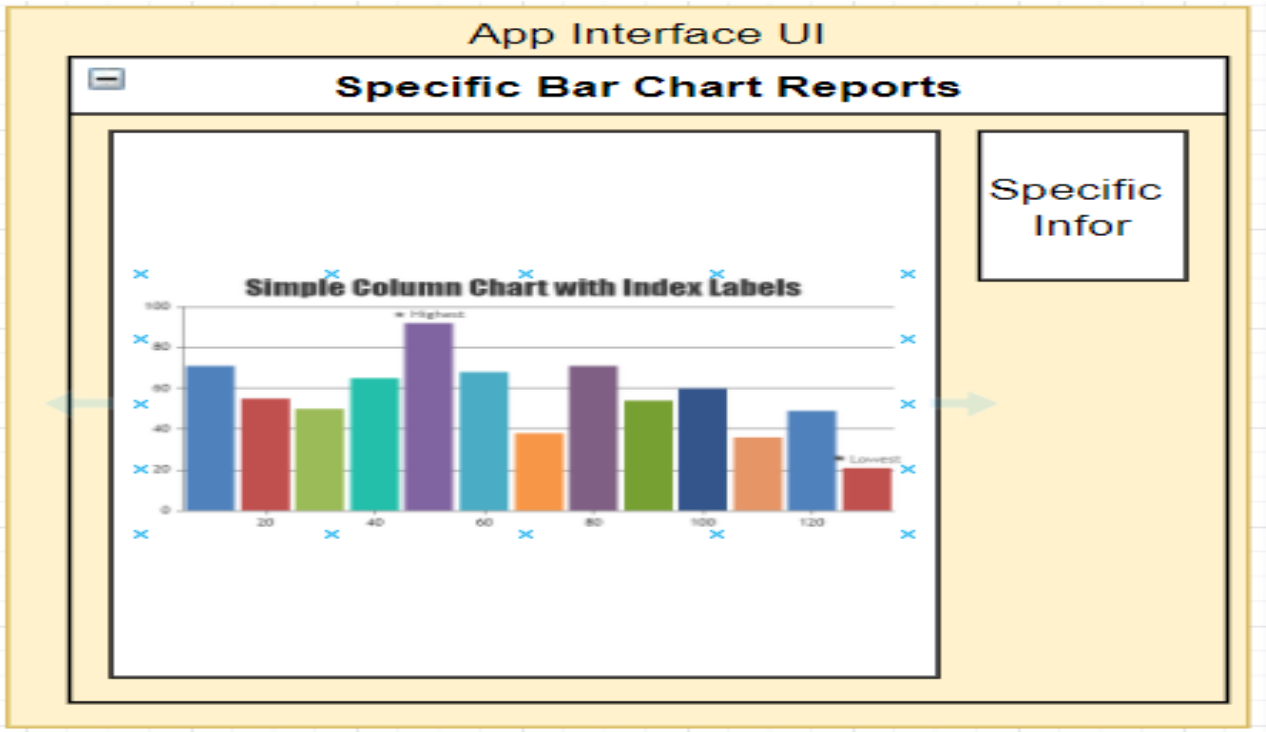


Figure 9 design overall bar charts UI reports

Here our Dataset Bar Chart reports of Global plastic pollution:

- **Bar chart reports**
  1) **Top 10 Beach with highest Rubbish chart**

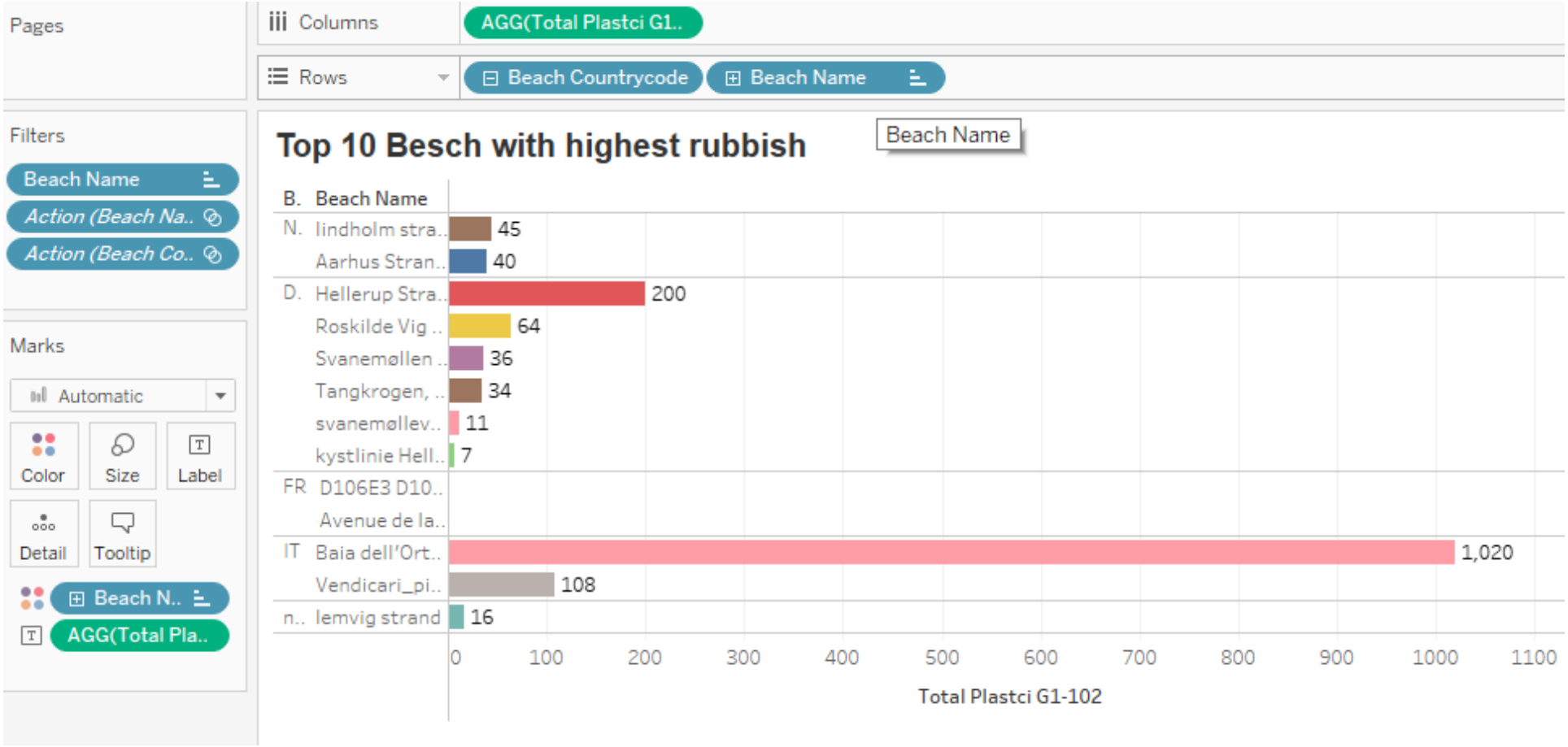


Figure 10 specific bar chart report of Top 10 beach with highest rubbish

Total Plastci G1-102 for each Beach Name broken down by Beach Countrycode. Color shows details about Beach Name. The marks are labeled by Total Plastci G1-102. The data is filtered on Action (Beach Name) and Action (Beach Countrycode,Beach Name). The Action (Beach Countrycode,Beach Name) filter keeps 1,642 members. The view is filtered on Beach Name.

Summary Top 10 Beach with highest Rubbish Chart report	
Count:	13
AGG(Total Plastci G1-102)	
Sum:	1,581
Average:	143.73
Minimum:	7
Maximum:	1,020
Median:	40.00

2) Fix pollution Next 5 years

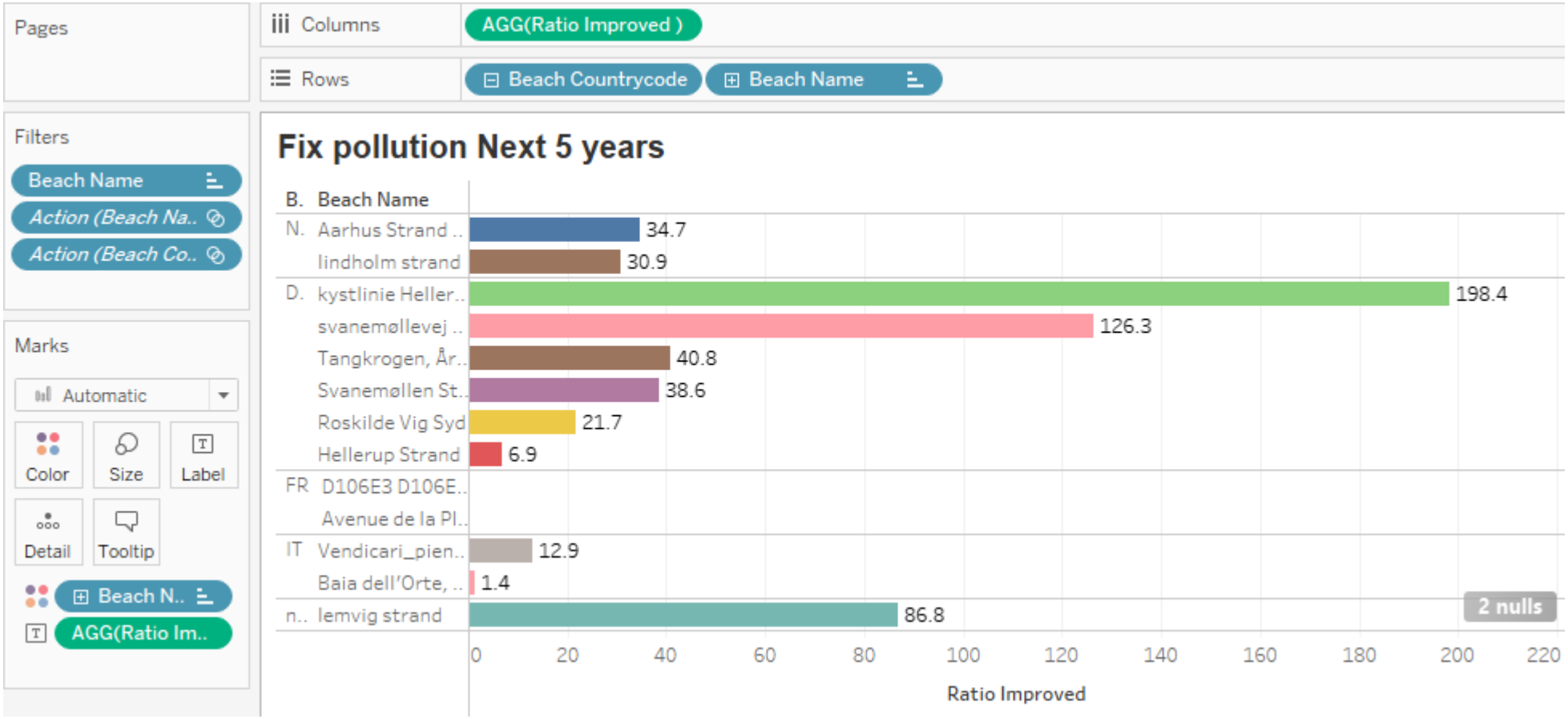


Figure 11 specific bar chart report of Fix pollution Next 5 years

Ratio Improved for each Beach Name broken down by Beach Countrycode. Color shows details about Beach Name. The marks are labeled by Ratio Improved. The data is filtered on Action (Beach Name) and Action (Beach Countrycode,Beach Name). The Action (Beach Countrycode,Beach Name) filter keeps 1,642 members. The view is filtered on Beach Name.

Summary Fix pollution Next 5 years
Count:11
AGG(Ratio Improved)

Calculation method: ((100/(([Total Plastic G1-102]*30)*12))/0.001)*5	
Sum:	**599.4**
Average:	**54.5**
Minimum:	**1.4**
Maximum:	**198.4**
Median:	**34.7**

3) Countries Trading Import and Amount Export

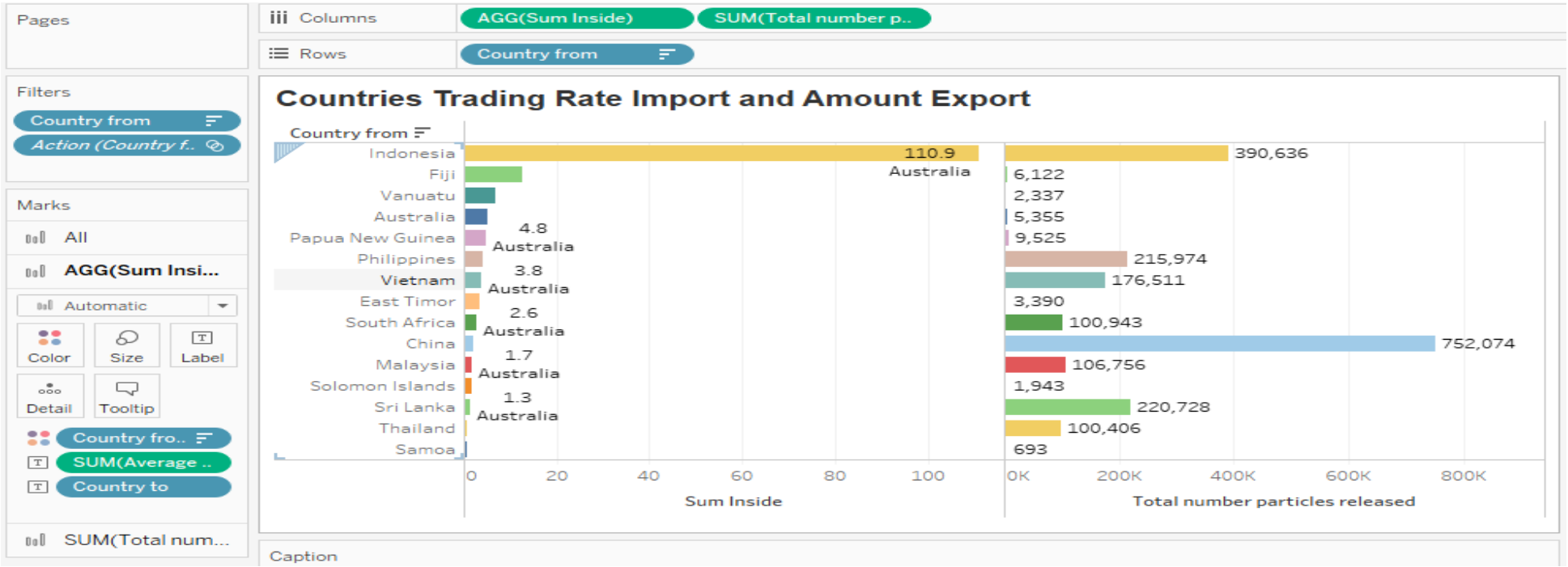


Figure 12 specific bar chart report of countries Trading rate Import and Amount Export

Sum Inside and sum of Total number particles released for each Country from. Color shows details about Country from. For pane Sum Inside: The marks are labeled by sum of Average daily particles inside EEZ and Country to. For pane Sum of Total number particles released: The marks are labeled by sum of Total number particles released. The data is filtered on Action (Country from), which keeps 30 members. The view is filtered on Country from, which keeps 15 of 30 members.

Countries Trading Rate Import and Amount Export	
SUM(Total number particles released)	
Sum:	2,093,393
Average:	139,559.53
Minimum:	693
Maximum:	752,074
Median:	100,406.00
AGG(Sum Inside)	
Sum:	161.9
Average:	8.5
Minimum:	0.0
Maximum:	110.9
Median:	2.0

- **Bubble chart reports**

This Bubble report is used to compare and measure the ratio between specific name and average number. our datasets have specific Country_Frome with Average daily inside & Total release of Im-Export.

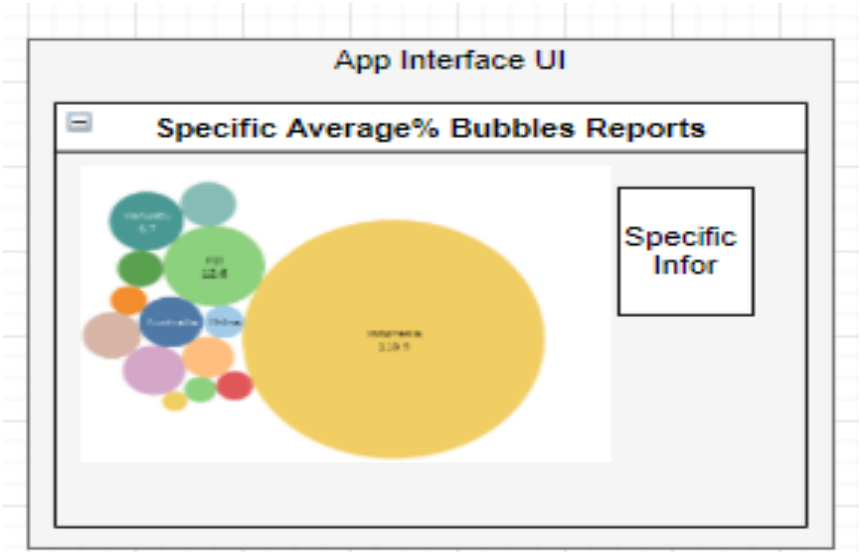


Figure 13 design over bubble chart UI report

4) Countries Trading Import Bubble chart:

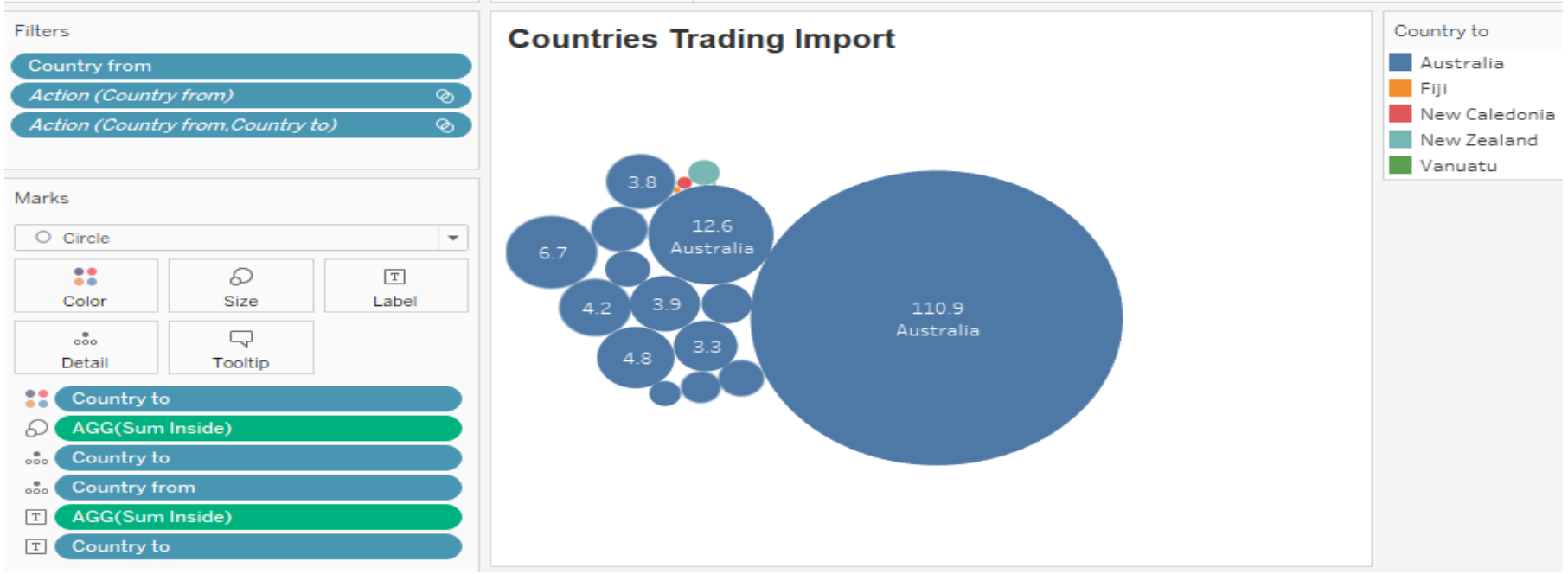


Figure 14 specific bubble chart report of Countries Trading Import

Sum Inside, Country to and Country from. Color shows details about Country to. Size shows Sum Inside. The marks are labeled by Sum Inside, Country to and Country from. Details are shown for Country to. The data is filtered on Action (Country from), which keeps 30 members. The view is filtered on Country from, which keeps 14 of 30 members.

Summary Countries Trading Import	
Count:14 SUM(Sum Inside)	
Sum:	161.3
Average:	11.5
Maximum:	110.9
Median:	3.6
Standard deviation:	28.8

- **Mapping chart reports**

This Mapping report is used to match a mapping of both between specific name and average & Value number. our datasets have specific Country_From with average_Im-Export and Beach_Name with Total_rubbish value number in every country.

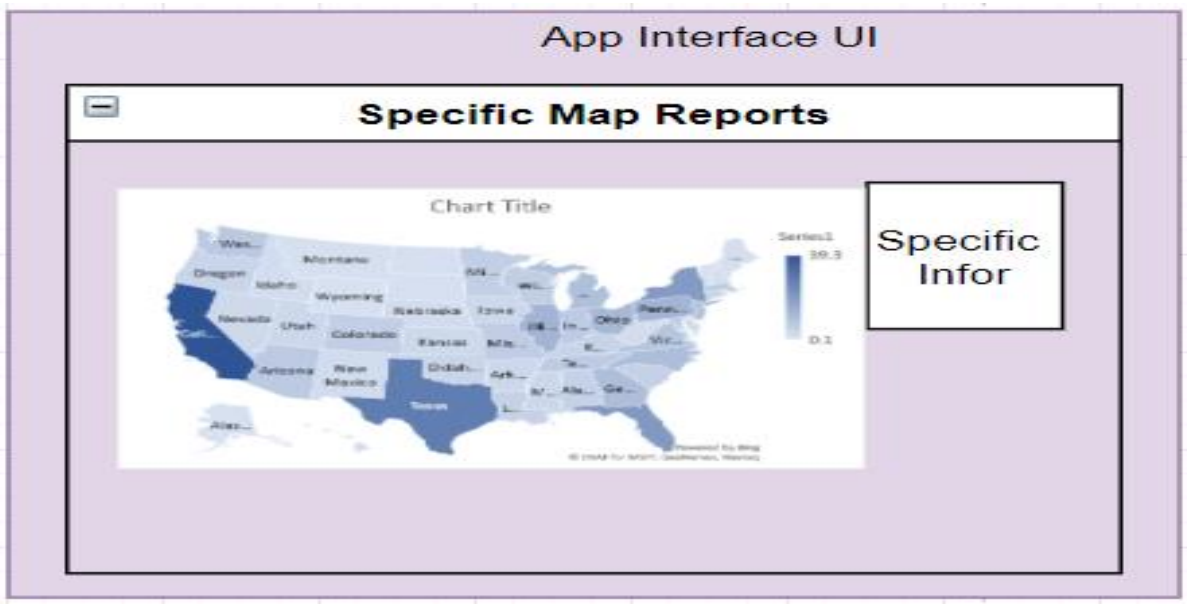


Figure 15 design overall mapping chart UI reports

5) Beach's Plastic Pollution:

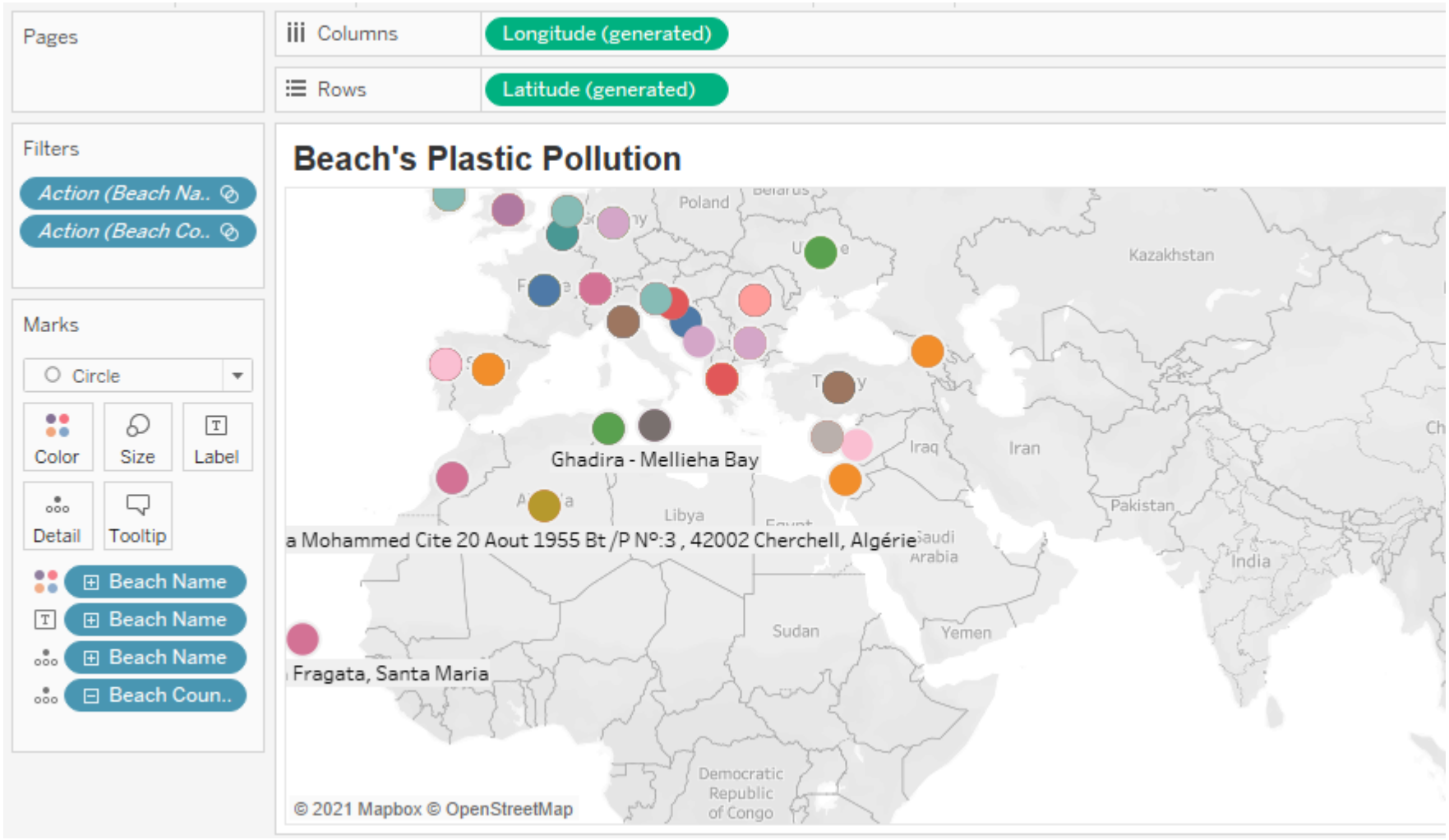


Figure 16 specific Map chart report of beach's Plastic pollution

Map based on Longitude (generated) and Latitude (generated). Color shows details about Beach Name. The marks are labeled by Beach Name. Details are shown for Beach Name and Beach Countrycode. The data is filtered on Action (Beach Name) and Action (Beach Countrycode,Beach Name). The Action (Beach Countrycode,Beach Name).

6) Countries Trading Export:

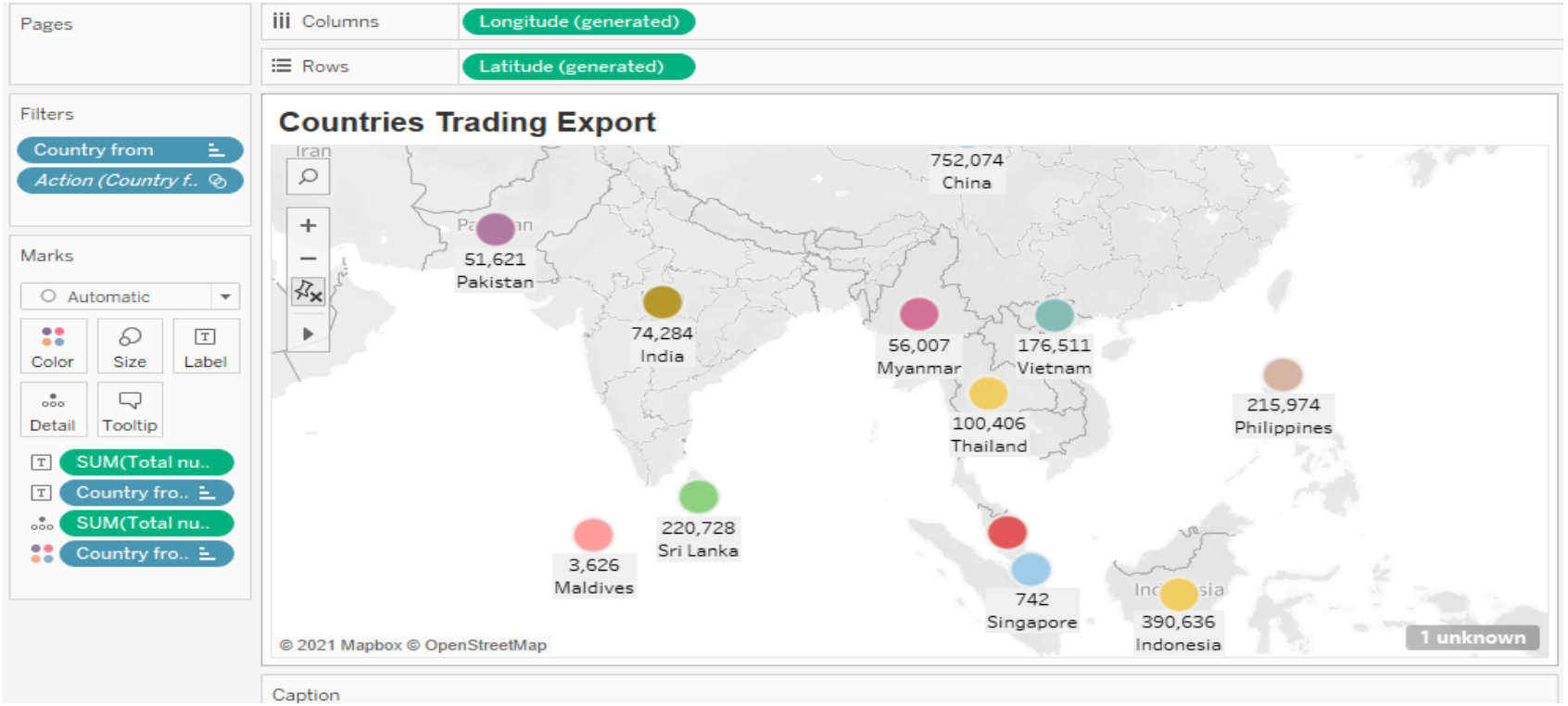


Figure 17 specific Map chart report of countries trading Export

Map based on Longitude (generated) and Latitude (generated). Color shows details about Beach Name. The marks are labeled by Beach Name. Details are shown for Beach Name and Beach Countrycode. The data is filtered on Action (Beach Name) and Action (Beach Countrycode,Beach Name). The Action (Beach Countrycode,Beach Name) filter keeps 1,642 member.

M3 Customize the design to ensure that it is user friendly and has a functional interface (Tableau IS).

After done on designing specific tasks of each chart reporting above, we must combine them into concussion report as Dashboard base on our BI research's Purpose of dataset and also, by a result of Dashboard, we cloud identify target problem & reason to driven decision-making process with suggestion Feedback.

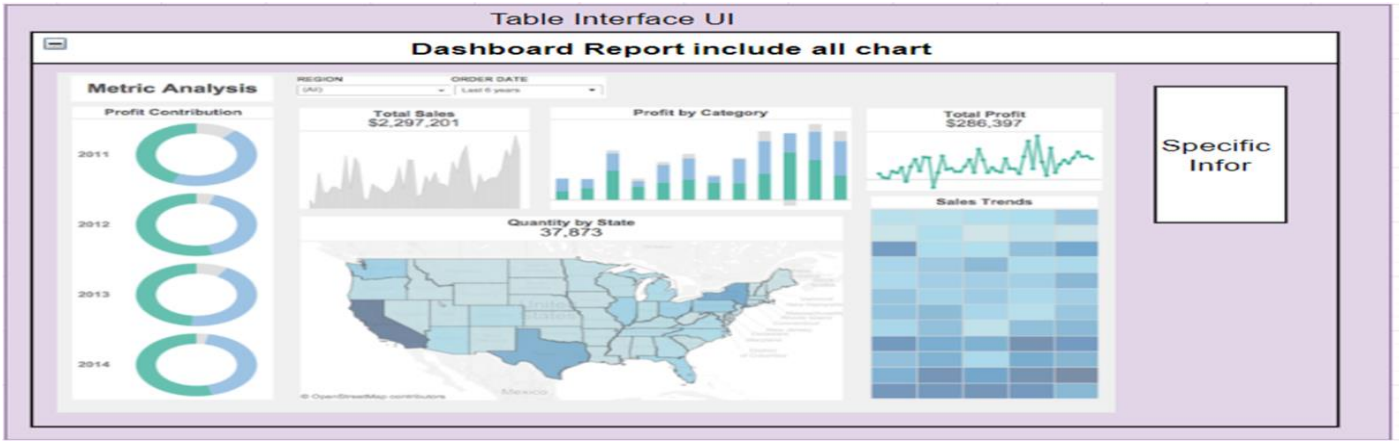
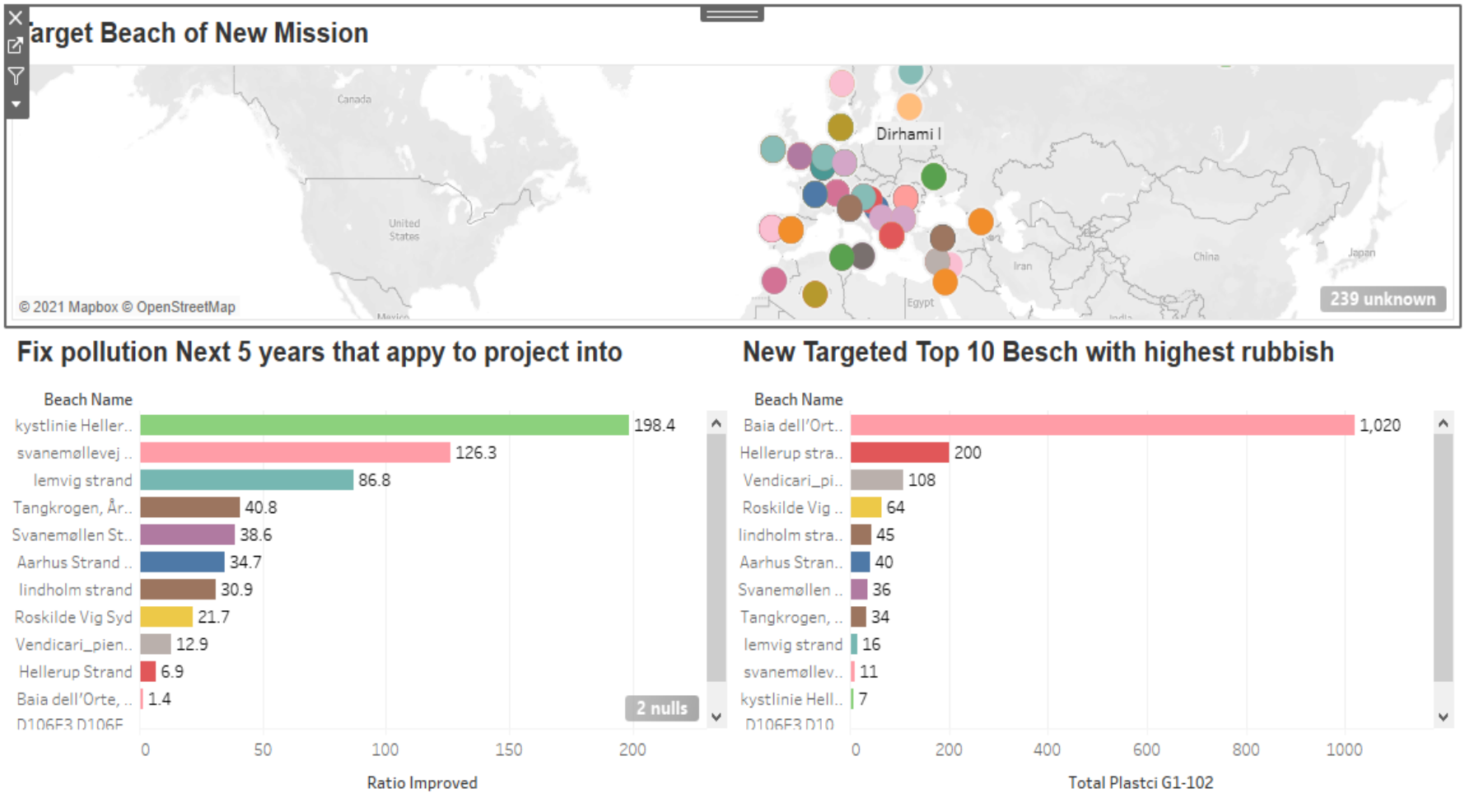


Figure 18 Design overall dashboard

Research BI purpose for Business improvement (possibilities)	Description Research for Business by BI
Supply to new consumption industry (High)	Look for a new industry to provide them with the benefit of Waste for their factory industry as an electric power generator by heat with burning garbage.
Run project with country that've high plastic garbage (High – Medium)	Install our business process with new countries that need our help to clean up.
New revenue with partnership country (Medium)	Make connections with local representative companies of those countries as shareholders that can provide support.
New phototype of production (Low)	Figure out or search for a new trend for new production which is made by recycling plastic.
Assign New missions Cleanup (EXTREME)	Apply and deploy our business process to new places with the community.
Increase new job and labor on economy chart (High)	Provide new opportunities to local people and build consciousness on the thought of a clean environment.

Figure 19 our Business's precious purpose on building BI Tool and research BI dataset

Designed dashboards
Frist Dashboard



Dashboard1
source Link: **https://public.tableau.com/app/profile/bee2839/viz/Competitor_16249584721700/Dashboard1**

Purpose of dashboard 1.

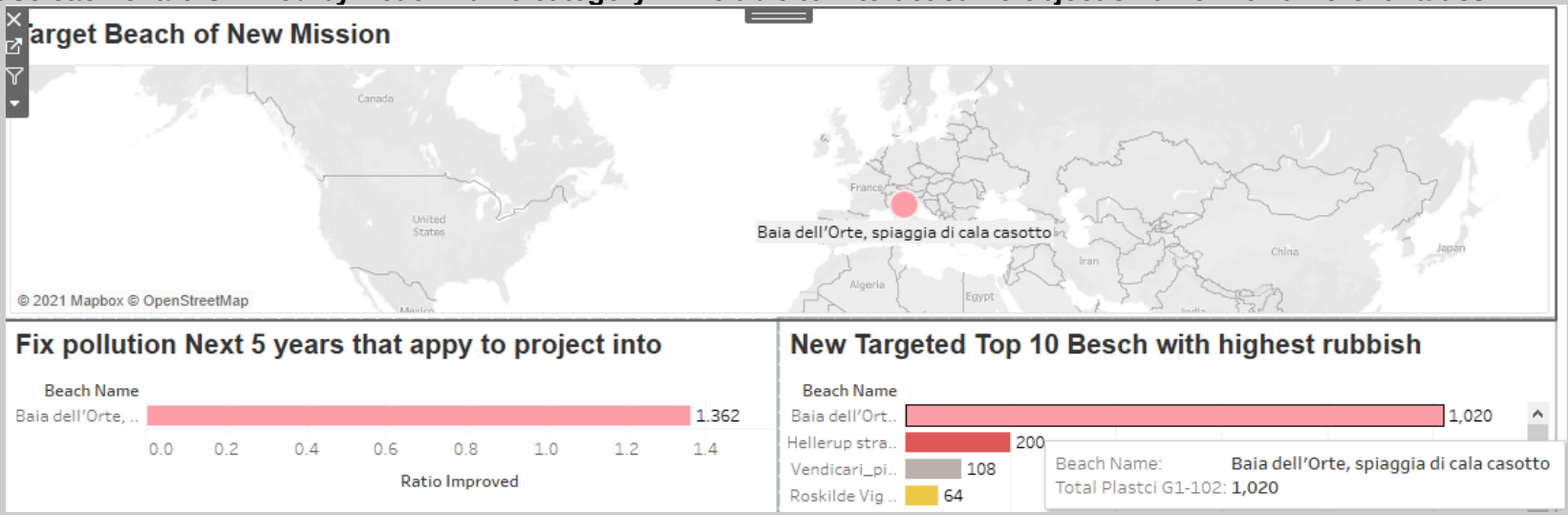
This dashboard is purposely crated for broad observation on amplification behavior of plastic pollution. By first dashboard, we can do tracking and locate beach on mapping with analyze amount of plastic rubbish. Also, analysis calculation is formatted formular to support decision-making

process which be based on present analysis amount Top 10 Beach while compared with apply project into by average amount within 5 years that pollution will be decreased.

- **Dashboard contains charts**
  - **Purpose of chart**

1) **Target beach of new mission: it's attribute carry Beach name to locate beach on countries**
2) **New target top 10 Beachs: include beach name with calculation sum of plastic rubbish category to conclude highest plastic amount on countries.**
3) **Fix pollution Next 5 years: contain Beach name with prediction formular to calculate future average a year by converting daily amount x30 into x12 months of a year which divide for year average.**

- **Interaction on dashboards**
  - **Result: to add action on dashboard by hover object in some chart that will be highlight a same object in another chart too because each chart is linked by Beach name category while able to interact same object's name with different values.**

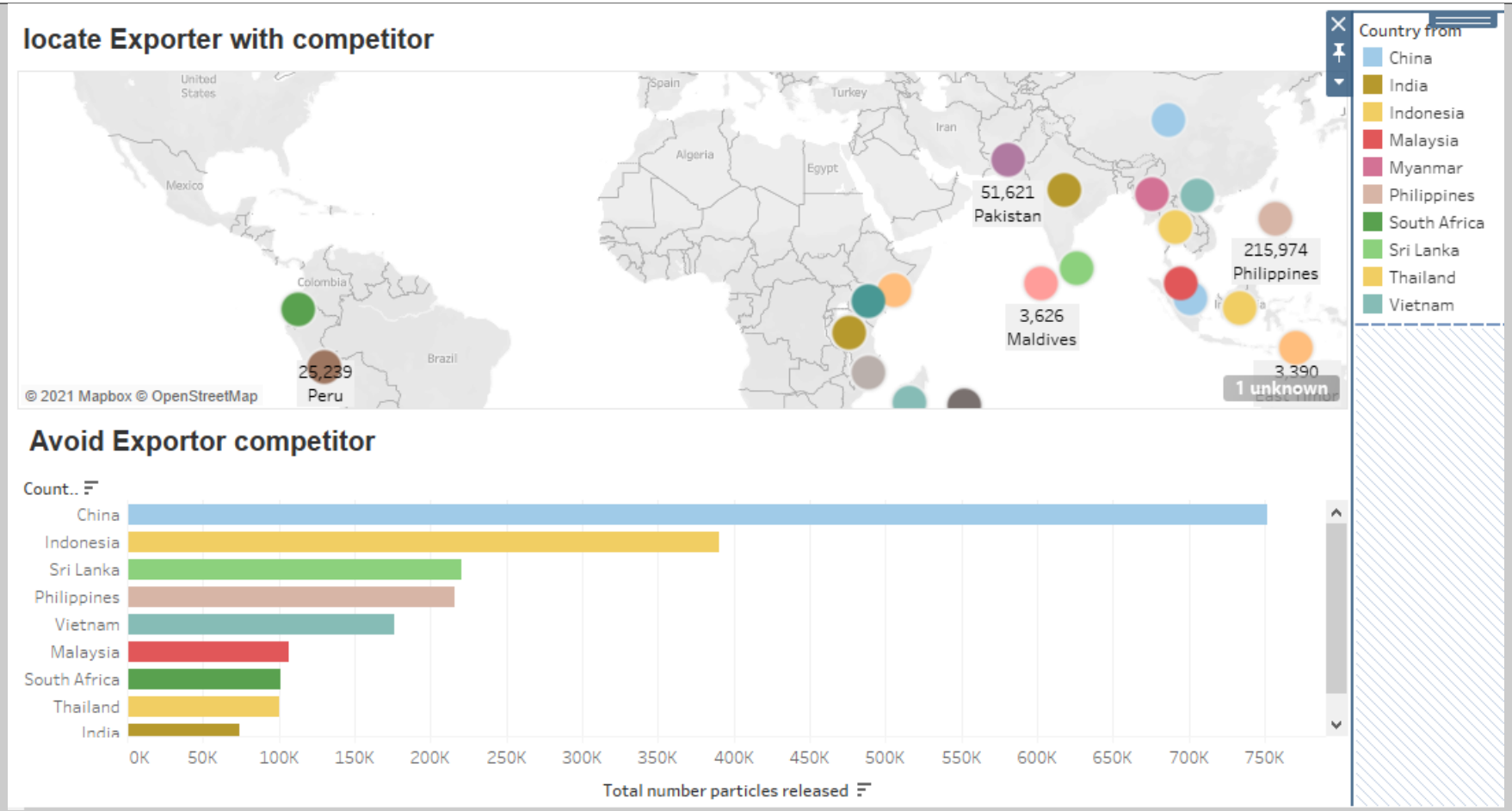


  - **Conclusion: Therefore, we are able select new target mission of highest beach based as 1.020 ton of Bala dell's ort while is compare on possible prediction future amount to be fixed within 5 years by 1.362.**

  - **Suggestion: However, we should apply more Machine Learning's algorithm to analyze effective condition of Data such as "IF obj THEN ELASE obj END" instead of formula.**

Second Dashboard

Purpose of Dashboard 2



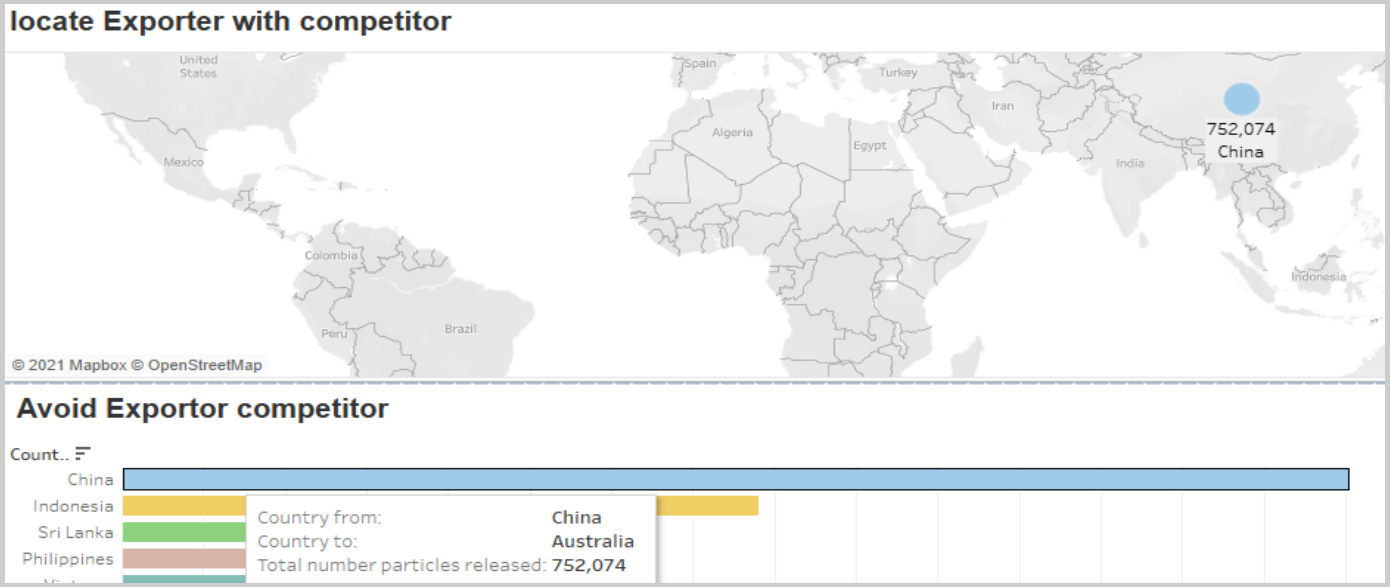
Dashboard2
source Link: https://public.tableau.com/app/profile/bee2839/viz/Competitor_16249584721700/Dashboard2

- **Each dashboard contains charts**
  - **Purpose of chart**

1) **Locate Export with competitor:** Include both country From-To 's detail and value of Total released of plastic which observe whose country exports the most plastic and also, that country might have chance to be who pollute largest of Plastic rubbish while might run a same project of plastic recycle as theft compactor to our business process. **Ex as I saw real thing: Thailand has own large recycle factory industry while also most importer of global waste.**

2) **Avoid Export:** Include both country From-To 's detail with sets calculation's Values as Total_number_relaese able to calculate each country's amount of plastic competitor export

- **Interaction on dashboards**
  - **Result: to add action on dashboard by hover object in some chart that will be highlight a same object in another chart too because each chart is linked by Country From category while able to interact same object's name with different values.**

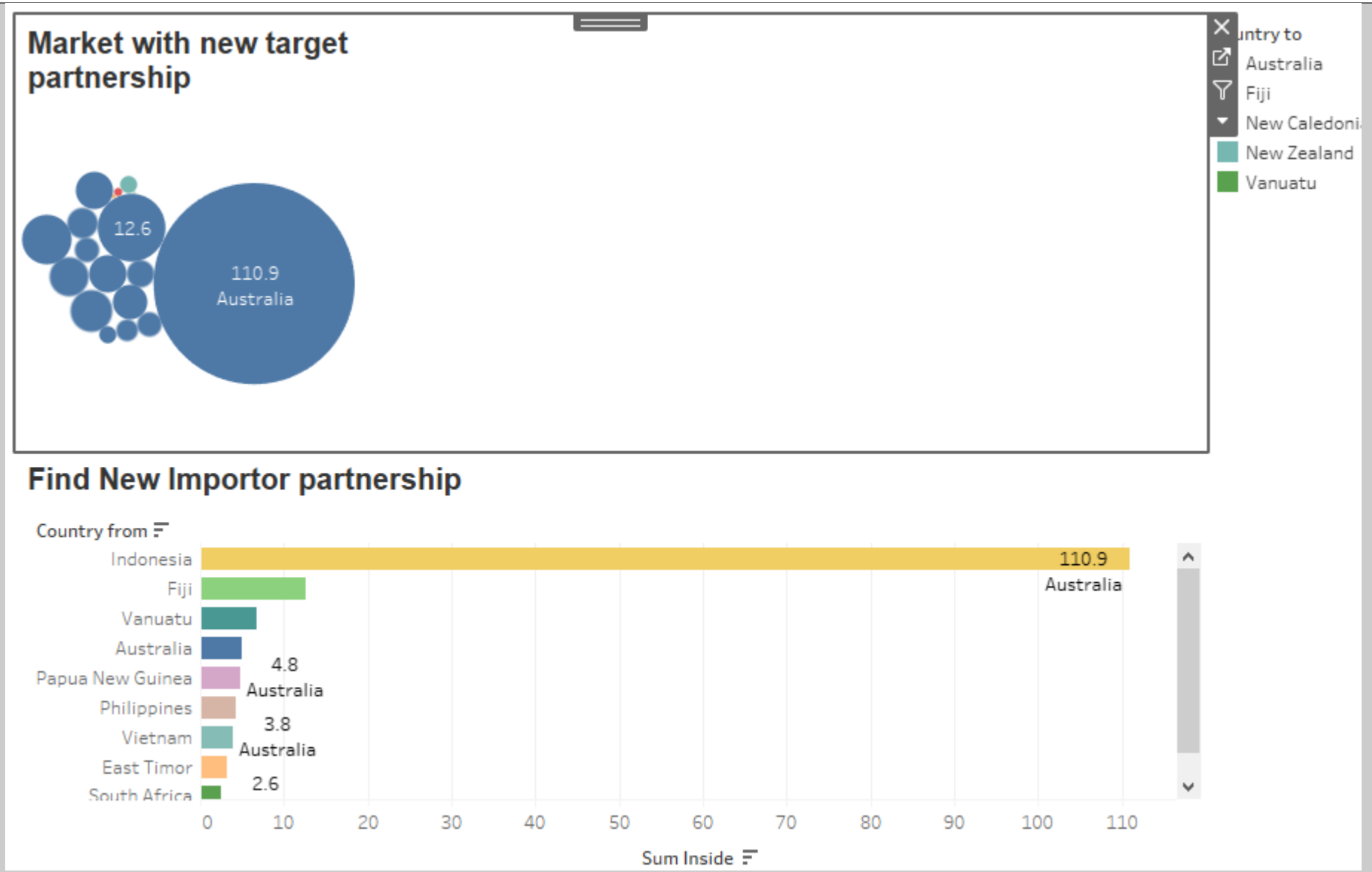


  - **Conclusion: Therefore, this second Dashboard help us to analyze and estimate value amount of each country who might be our new target mission.**

  - **Suggestion: in my opinion, we decide to avoid China as the biggest competitor as 752,074 beach China has many companies work collecting rubbish which difficult to compete with.**

Third Dashboard

Purpose of Dashboard 2


Market with new target partnership
12.6
110.9
Australia
untry to
Australia
Fiji
New Caledoni
New Zealand
Vanuatu
Find New Importor partnership
Country from
Indonesia
Fiji
Vanuatu
Australia
Papua New Guinea
Philippines
Vietnam
East Timor
South Africa
110.9
Australia
4.8
Australia
3.8
Australia
2.6
0
10
20
30
40
50
60
70
80
90
100
110
Sum Inside

Dashboard3

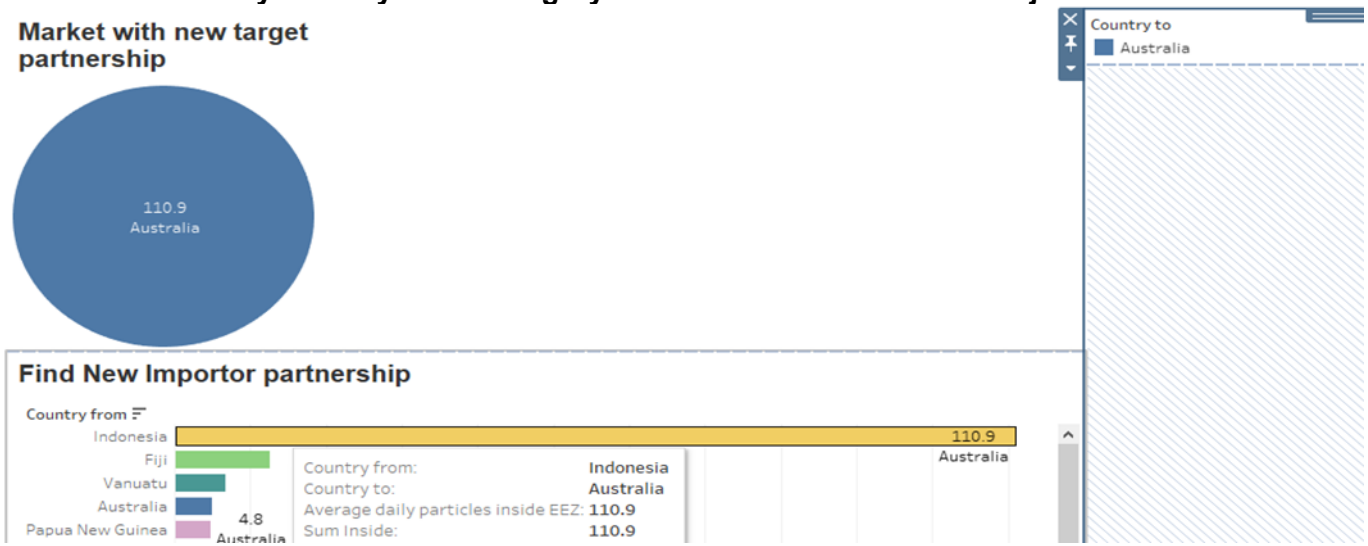
source Link: https://public.tableau.com/app/profile/bee2839/viz/PlastictradingcommunitysDashboard2/Dashboard4

- **Each dashboard contains charts**
  - **Purpose of chart**

  1) Market with new target partnership: Include both country From-To 's detail and value sum of Average Inside_EEZ, this chart's used to search most importer as well as Buyer to make new partnership for purchase contract trading from Ocean cleanup's recycled plastic.

  2) Countries Trading Import and Amount Export: Include both country From-To 's detail with sets calculation's Values as Sum inside able to calculate each country's amount of plastic Export & import. Also, remind whose countries are Seller and buyer.

- **Interaction on dashboards**
  - **Result: to add action on dashboard by hover object in some chart that will be highlight a same object in another chart too because each chart is linked by Country From category while able to interact same object's name with different values.**



  - **Conclusion: Therefore, this Dashboard help us to analyze benefic of recycles plastic with new revenue partnership who deal with resell plastic production as 110.9 ton/day of Australia which import from so many counties while second importer with Fiji.**

- **Suggestion: in my opinion, we should make friendly partnership with Australia by huger ratio of bubble chart that Australia is biggest importer from most countries. Also, we cloud add more calculation of recycle plastic's Price per Ton that country can make profit, also some more ML to predict an opportunity chance of Plastic trending in new counties.**

Figure 20 designed combination of reports charts into different two specific dashboards dataset as plastic pollution and trading dashboard

D3 Provide a critical review of the design in terms of how it meets a specific user or business requirement and identify what customization has been integrated into the design.

Viewers

Give feedback / comments

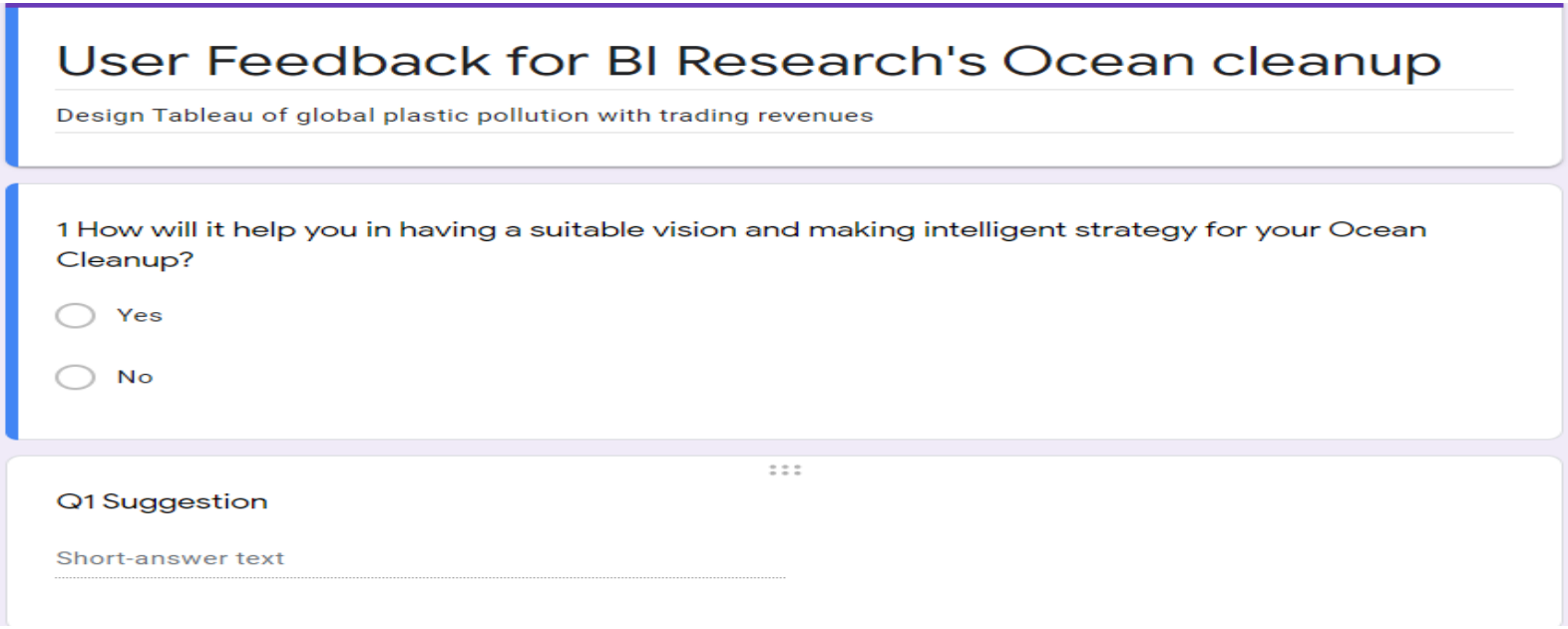
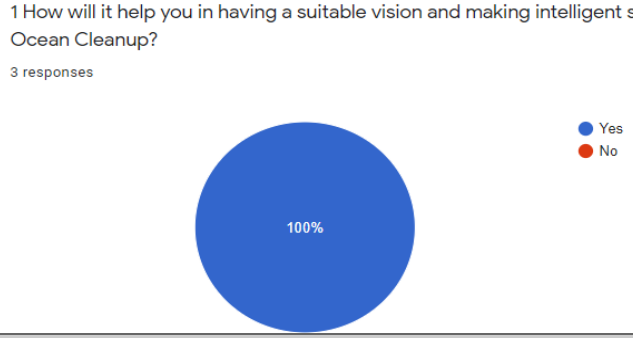
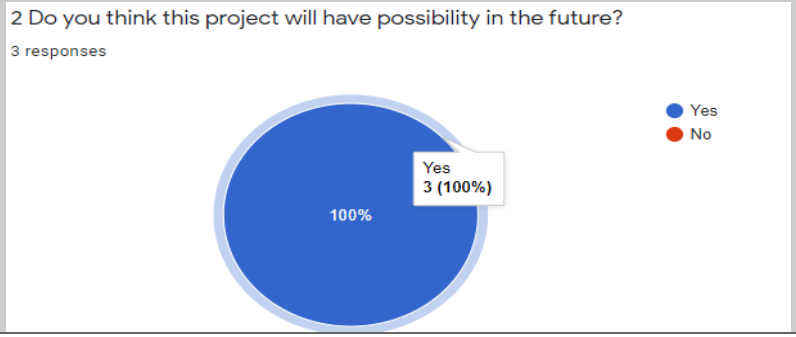
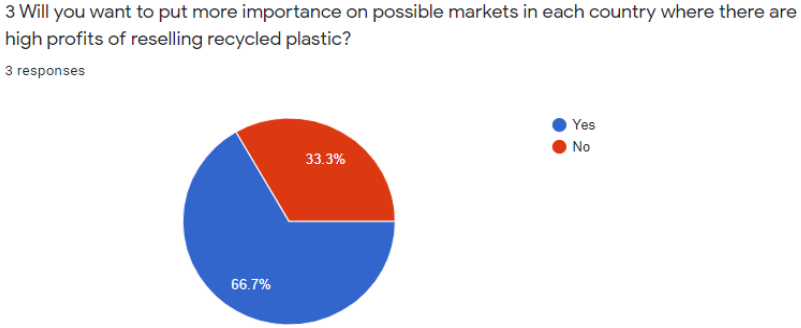
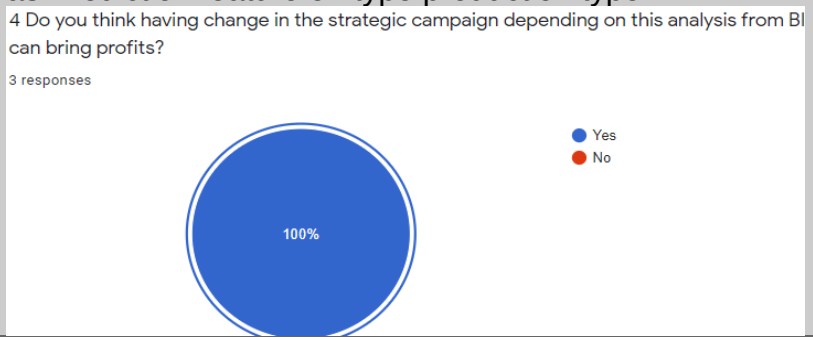
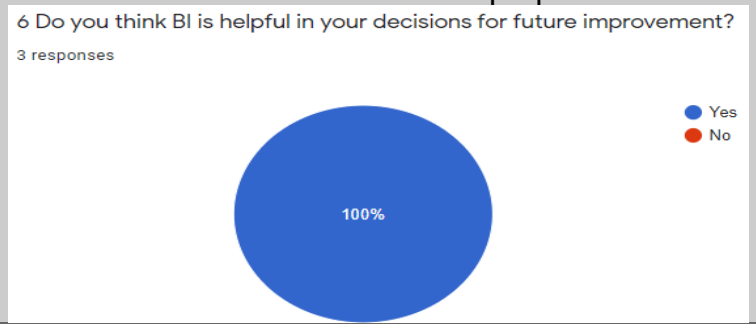


Figure 21 designed User feedback form of Tableau's dashboards for collect user satisfaction with future improvement comments by google form

Feedback Form of Idea business to Global plastic pollution Based on our demonstration		
Question	**Repose Review (Good & Bad points)**	**Recommendation (Need to improve & be add for future)**
Based on statistics and important information from data that BI tools analyzed and presented analytical findings in reports of summary of dashboards, charts and maps, (Q) How will it help you in having a suitable vision and making intelligent strategy for your Ocean Cleanup?	**(Yes),** which able select new target mission of highest beach based on possible prediction future amount and to analyze benefic of recycles plastic and estimate competitor and new revenue partnership who will deal with reselling plastic in long-term. 	(**Apply**) Prediction on how pollution rate grows next 5 years
(Q)Do you think this project will have possibility in the future?	**(Yes),** by BI Result, as we know Global-plastic pollution is grown ever year while is overload garbage theft to countries but well, it keeps up high demand to eliminates garbage with offer great price by government. However, Garbage is not garbage for throwing, but it can still be a profit treasure as recycle them into useful material. 	(**Comment**) Very impressive result, but couldn't find any trend production type

Based on the dashboard, it illustrates the strength of our project. (Q)Will you want to put more importance on possible markets in each country where there are high profits of reselling recycled plastic?	**(Yes),** Our Business strength is based on modern Ocean cleanup technology which none has developed yet with this project's profile focus on ocean service fee with local government and reselling recycled material to directly factory while is a sable market 	(**Comment**) Our company Ocean Cleanup's Priority stability market place
(Q)Do you think having change in the strategic campaign depending on this analysis from BI can bring profits?	(**Middle agree**), currently version of Tableau still lack potential on Machine leaning of developed functionality as Prediction feature on type production type. 	(**Apply**) Provide more dataset of production trend
(Q)What can we improve in our project in the future?	(**Yes**), for additional feature ML and Algorithm	• (**Apply**) Add site production's dataset or connect site's database for new market productions phototype • (**Apply**) Fix Security of user personal data

	5 What can we improve in our project in the future? (Guys, pls provide suggestion how apply more feature, new dataset, security and ML with tell how in short) 2 responses Add site production's dataset for new product's phototye Fix Secureity of user personal data	
(Q)Do you think BI is helpful in your decisions for future improvement?	(Yes), this BI research's Dashboard can capably successful almost cover all of our purpose. 	(Apply) for Open new path project of marketing by AI_Marketing of BI research which benefit cash-back to advertiser and marketer.

Personal Reporting

LO4 Discuss the impact of business intelligence tools and technologies for effective decision-making purposes and the legal/regulatory context in which they are used

P5 Discuss how business intelligence tools can contribute to effective decision-making.

1) Present Our point of view about how BI tools can contribute to effective decision-making

In my opinion, we through BI is able to helps extract crucial facts from an amount of unstructured data and transform them into actionable information that enables companies to make informed strategic decisions, improving operational efficiency and business productivity. This actionable information provides crucial insights that reveal the underlying currents of customer behaviour:

Our point of view about business intelligence tools	
Improve Business Productivity	Businesses can channel their vital resources and workforce to foster business productivity as they need not divert their resources to BI gathering work which is handled by the BI team such as results in cost saving, time saving and efficient reporting supporting improved business productivity. The BI gathering team can pull up vital information from customer interactions and present it in a manner that is easy to understand, communicate and execute. BI gathering teams ensure that detailed reports are provided to companies ensuring that all the crucial information is presented in a well-documented and presentable manner so that information can be used to efficiently manage the business.
Crucial Information Easily Accessed	BI provides crucial information to companies, improving their ability to make quick decisions and generating a competitive advantage.
Good Return on Investment (ROI)	BI ensures stronger ROI as companies can significantly reduce costs, enhance revenue, improve margin, drive cost. Also, BI ensure qualitative ROI such as improvements in business and process efficiency, employee productivity, better strategic decision making, customer satisfaction and much more. Therefore, we're relying on BI can reduce operational costs due to improved operational efficiency and optimized business processes.
Informed Decision Making (Our built python BI Tool)	Reporting based on accurate and timely information helps companies measure the performance of their processes. BI tool helps companies make informed decisions on strategic issues by providing crucial information on current and historical performance of the company along with future trends, expected demands, customer behaviour etc. By BI teams ensure that the company receives real-time advanced reports to ensure that the company can efficiently utilize the information at hand to better manage the business.

Table 3 point of view to BI tool on effective decision-making (Iyoob J.. 2012)

Therefore, according above, BI Tool can support companies with rich data resources which can contribute us achieve our business goals and targets by guiding timely strategic decisions. Interactions with customers in the form of voice calls, chat interactions, emails etc

Also, BI Tool can be thoroughly analysed by BI gathering teams to shed light on aspects such as customer preferences; likes and dislikes of customers; technical difficulties faced by customers; customer reactions to promotions; and the online shopping customer experience on a website. This analysis can lead to improving conversion rates and much more.

2) Example Using our own built Python on analysis dataset of Plastic pollution while calculate method with predicted solution as well as effective decision-making

- **Coding Python sample on BI Toll with Visual Code IDE**

➢ **Designing our own Program ML on BI by Python language:**

Table 4 building own BI Tool's program Python

Building own python program's BI tool Snippet Code (Team Working)
Source code Here:

```python
import numpy as np
from numpy.lib import average
import pandas as pd
from pandas.core.frame import DataFrame
import matplotlib.pyplot as plt


def openCSVFile(path):
    file = pd.read_csv(path)
    pd.set_option('display.max_columns', 4)
    return file


def getCVSLines(path):
    with open(path) as f:
        for i, line in enumerate(f):
            print("line {0} = {1}".format(i, line.split()))


def filtCSVCollumn(columns, group_by_name, top, comparation_param, src_file):
    data = pd.DataFrame(src_file[columns])
    if(top != 0):
        dataFrame = data[columns].groupby(group_by_name).sum()
        result = dataFrame.sort_values(
            comparation_param, ascending=False).head(top)
    else:
        if(group_by_name != ''):
            result = data[columns].groupby(
                group_by_name, as_index=False).sum()
        else:
            result = data[columns]
    return result


def predict(years, file):
    data = DataFrame(file)
    total_average = data["Average_daily_particles_inside_EEZ"].sum()
    predict_result = ((total_average * 30)*12) * years
    difference = predict_result - predict_result/years
    return predict_result, difference

```

```
42
43  def incomeTotalPlastic(file, country_name, current_value):
44      data = DataFrame(file)
45      country_total = data.loc[data["Country_from"] == country_name]
46      total_each_country = country_total["Total_number_particles_released"]
47      result = int((total_each_country * current_value)*30*12)
48      return result
49
50
51  def drawBarChart(data):
52      x = np.arange(len(data["Country_from"]))
53      y1 = data["Total_number_particles_released"]
54      y2 = data["Average_daily_particles_inside_EEZ"]
55      width = 0.4
56      # plt.rcParams["figure.figsize"] = (30, 8)
57      # plot data in grouped manner of bar type
58      plt.barh(x-0.4, y1, width)
59      plt.barh(x, y2, width, color='orange')
60      plt.yticks(x-0.2, data['Country_from'])
61      plt.ylabel("Countries")
62      plt.xlabel("Amount")
63      plt.legend(["Total number particles released",
64                  "Average daily particles inside EEZ"])
65      plt.show()
66
67
68  def main():
69      # Path of the Dataset file.
70      path = 'D:\Python Code\ASM\Resource\dataset.csv'
71      # Open dataset.csv
72      file = openCSVFile(path)
73      print("[OPENING CVS FILE]\n\n", file, '\n\n')
74
75      # Display file Lines
76      print("[RUNNING GET CVS FOLLOWING LINES]\n =>{}\n\n".format(getCVSLines(path)))
77
78      # Define Select column: Country from, Average daily inside
79      fill_result_of_average = filtCSVCollumn(
80          columns=["Country_from", "Average_daily_particles_inside_EEZ"],
81          top=0,
```

```
        group_by_name='',
        comparation_param='',
        src_file=file
    )
    print("[FILL RESULT] => [Country from, Average daily inside]\n\n{}\n\n".format(
        fill_result_of_average))

    # Define Select column: Country from, Total number release
    fill_result_of_total = filtCSVCollumn(
        columns=["Country_from", "Total_number_particles_released"],
        top=0,
        comparation_param='',
        group_by_name='',
        src_file=file,
    )
    print("[FILL RESULT] => [Country from, Total number release]\n\n{}\n\n".format(
        fill_result_of_total))

    # Define function "Sum(Average and Total) of each country per day"  (Total sum)
    fill_result = filtCSVCollumn(
        columns=["Country_from", "Average_daily_particles_inside_EEZ",
                 "Total_number_particles_released"],
        group_by_name='Country_from',
        top=0,
        comparation_param='',
        src_file=file,
    )
    print("[FILL RESULT] => [Sum(Average or Total) of each country]\n\n{}\n\n", fill_result)

    # Select Top 10 of Max(total Sum ) of all country
    fill_result_top = filtCSVCollumn(
        columns=["Country_from", "Total_number_particles_released"],
        group_by_name="Country_from",
        top=10,
        comparation_param='Total_number_particles_released',
        src_file=file
    )
    print('[Filtered results] => Select Top 10 of Max(total Sum ) of all country\n\n{}\n\n'.format(
        fill_result_top))
```

```
121
122     # Average of Total number particles released
123     data = DataFrame(file)
124     average_total = data["Total_number_particles_released"].mean()
125     print("Average of Total number particles released: ",
126           round(average_total, 2), '\n\n')
127     # Input how many(year)
128     year = int(input("Enter the number of year you wanna predict: "))
129     predict_result, diff = predict(years=year, file=file)
130     print("[PREDICT RESULT] - In next {} year(s), average will be {} and increases {}"
131           .format(year, round(predict_result, 2), round(diff, 2)))
132
133     country_name = input("Which country you wanna know?\n>_ ")
134     result_income = incomeTotalPlastic(
135         file=fill_result_of_total, country_name=country_name, current_value=60)
136     print(result_income)
137     print("[RESULT VALUE] => The income of {} by sell Plastic recycle are ${}".format(
138         country_name, result_income))
139
140     drawBarChart(fill_result)
141
142
143 main()
144
```

Link file Here:

Our Team2's-built source code of Python BI tool
By anyone who have link can access file's code

Link Here: https://colab.research.google.com/drive/1zT4XgHmyV_xgzBRjrkF-3_wOBa2RxEIh?usp=sharing

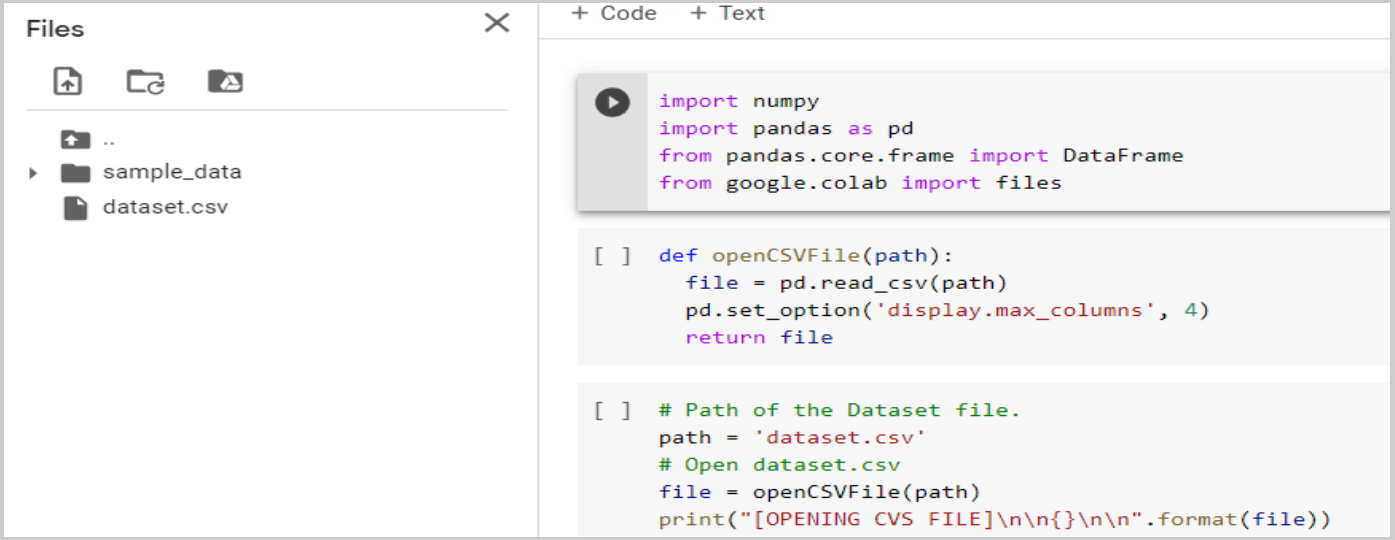
Run first statements in order

first of all, Upload dataset file into Files of Google Colab:

Statement 1

```
1  import numpy as np
2  from numpy.lib import average
3  import pandas as pd
4  from pandas.core.frame import DataFrame
5  import matplotlib.pyplot as plt
6
```

Above are the libraries and objects. This will be used throughout our Python program. As above we can see some familiar libraries like Pandas for data analysis.



use dataset.csv in Files

Statement 2

As below is the function to open and read CSV files, this function is to support reading the entire file according to the given path and then returning a file object.

```
8  def openCSVFile(path):
9      file = pd.read_csv(path)
10     pd.set_option('display.max_columns', 4)
11     return file
12
```

To use the above function, I have written below the command block. Here I have created a main function and contains all the calls of the main program. Below is the command block to read the CSV file from the given path and print it to the screen.

```
68 def main():
69     # Path of the Dataset file.
70     path = 'D:\Python Code\ASM\Resource\dataset.csv'
71     # Open dataset.csv
72     file = openCSVFile(path)
73     print("[OPENING CVS FILE]\n\n", file, '\n\n')
74
```

Output:

```
[OPENING CVS FILE]

         Country_from     Country_to  Total_number_particles_released  \
0           Indonesia      Australia                           390636
1                Fiji      Australia                             6122
2             Vanuatu      Australia                             2337
3   Papua New Guinea      Australia                             9525
4         Philippines      Australia                           215974
5             Vietnam      Australia                           176511
6          East Timor      Australia                             3390
7        South Africa      Australia                           100943
8               China      Australia                           752074
9            Malaysia      Australia                           106756
10    Solomon Islands      Australia                             1943
11          Sri Lanka      Australia                           220728
12           Thailand      Australia                           100406
13              Samoa      Australia                              693
14              India      Australia                            74284
15            Somalia      Australia                            11258
16           Maldives      Australia                             3626
17           Tanzania      Australia                             4878
18     Comoro Islands      Australia                             6292
19        New Zealand      Australia                              742
20             Myanmar      Australia                            56007
21          Mozambique      Australia                             5381
22                Peru      Australia                            25239
23           Mauritius      Australia                             7031
24           Singapore      Australia                              742
25            Pakistan      Australia                            51621
26               Kenya      Australia                             2084
27          Madagascar      Australia                             3443
28             Ecuador      Australia                            10999
29           Australia      Australia                             1071
30           Australia    New Zealand                             1071
31           Australia  New Caledonia                             1071
32           Australia           Fiji                             1071
33           Australia        Vanuatu                             1071

    Average daily particles inside EEZ
```

	Average_daily_particles_inside_EEZ
0	110.910000
1	12.592000
2	6.744400
3	4.807800
4	4.196200
5	3.825700
6	3.278700
7	2.564900
8	2.041200
9	1.682100
10	1.640700
11	1.285200
12	0.771730
13	0.606090
14	0.146220
15	0.118410
16	0.073751
17	0.033881
18	0.031828
19	0.024384
20	0.014203
21	0.013518
22	0.008812
23	0.007957
24	0.006588
25	0.005989
26	0.005647
27	0.002310
28	0.001369
29	3.930000
30	0.800000
31	0.180000
32	0.040000
33	0.010000

Statement 3

Below is the function to call to get the rows of data in the CVS file that we have given the link as shown below. After reading a line, that line will be printed immediately on the console screen

```
14  def getCVSLines(path):
15      with open(path) as f:
16          for i, line in enumerate(f):
17              print("line {0} = {1}".format(i, line.split()))
18
```

Call the function to proceed to call the getCVSLines function to list the data lines of the file.

```
75      # Display file Lines
76      print("[RUNNING GET CVS FOLLOWING LINES]\n =>{}\n\n".format(getCVSLines(path)))
77
```

Output:

```
line 0 = ['Country_from,Country_to,Total_number_particles_released,Average_daily_particles_inside_EEZ']
line 1 = ['Indonesia,Australia,390636,110.91']
line 2 = ['Fiji,Australia,6122,12.592']
line 3 = ['Vanuatu,Australia,2337,6.7444']
line 4 = ['Papua', 'New', 'Guinea,Australia,9525,4.8078']
line 5 = ['Philippines,Australia,215974,4.1962']
line 6 = ['Vietnam,Australia,176511,3.8257']
line 7 = ['East', 'Timor,Australia,3390,3.2787']
line 8 = ['South', 'Africa,Australia,100943,2.5649']
line 9 = ['China,Australia,752074,2.0412']
line 10 = ['Malaysia,Australia,106756,1.6821']
line 11 = ['Solomon', 'Islands,Australia,1943,1.6407']
line 12 = ['Sri', 'Lanka,Australia,220728,1.2852']
line 13 = ['Thailand,Australia,100406,0.77173']
line 14 = ['Samoa,Australia,693,0.60609']
line 15 = ['India,Australia,74284,0.14622']
line 16 = ['Somalia,Australia,11258,0.11841']
line 17 = ['Maldives,Australia,3626,0.073751']
line 18 = ['Tanzania,Australia,4878,0.033881']
line 19 = ['Comoro', 'Islands,Australia,6292,0.031828']
line 20 = ['New', 'Zealand,Australia,742,0.024384']
line 21 = ['Myanmar,Australia,56007,0.014203']
line 22 = ['Mozambique,Australia,5381,0.013518']
line 23 = ['Peru,Australia,25239,0.0088125']
line 24 = ['Mauritius,Australia,7031,0.0079569']
line 25 = ['Singapore,Australia,742,0.006588']
line 26 = ['Pakistan,Australia,51621,0.005989']
line 27 = ['Kenya,Australia,2084,0.0056468']
line 28 = ['Madagascar,Australia,3443,0.0023101']
line 29 = ['Ecuador,Australia,10999,0.0013689']
line 30 = ['Australia,Australia,1071,3.93']
line 31 = ['Australia,New', 'Zealand,1071,0.8']
line 32 = ['Australia,New', 'Caledonia,1071,0.18']
line 33 = ['Australia,Fiji,1071,0.04']
line 34 = ['Australia,Vanuatu,1071,0.01']
[RUNNING GET CVS FOLLOWING LINES]
```

Statement 4

The function below is created by us to serve 3 main problems: Grouping data of duplicate countries in the data table. The second task is to perform the top of the highest or lowest countries. And finally, to do the ordering of the data sheet

```
def filtCSVCollumn(columns, group_by_name, top, comparation_param, src_file):
    data = pd.DataFrame(src_file[columns])
    if(top != 0):
        dataFrame = data[columns].groupby(group_by_name).sum()
        result = dataFrame.sort_values(
            comparation_param, ascending=False).head(top)
    else:
        if(group_by_name != ''):
            result = data[columns].groupby(
                group_by_name, as_index=False).sum()
        else:
            result = data[columns]
    return result
```

Below is the call statement placed inside the main function. This command is called to produce the required filtered columns where the required columns are **Country_from** and **Average_daily_particles_inside_EEZ**

```
    # Define Select column: Country from, Average daily inside
    fill_result_of_average = filtCSVCollumn(
        columns=["Country_from", "Average_daily_particles_inside_EEZ"],
        top=0,
        group_by_name='',
        comparation_param='',
        src_file=file
    )
    print("[FILL RESULT] => [Country from, Average daily inside]\n\n{}\n\n".format(
        fill_result_of_average))

```

Next is the code block below, also in main. To call the function filtCSVCollumn() to get two columns Country_from and Total_number_particles_released

```
89      # Define Select column: Country from, Total number release
90      fill_result_of_total = filtCSVCollumn(
91          columns=["Country_from", "Total_number_particles_released"],
92          top=0,
93          comparation_param='',
94          group_by_name='',
95          src_file=file,
96      )
97      print("[FILL RESULT] => [Country from, Total number release]\n\n{}\n\n".format(
98          fill_result_of_total))
```

Output:

```
[FILL RESULT] => [Country from, Average daily inside]

         Country_from  Average_daily_particles_inside_EEZ
0           Indonesia                          110.910000
1                Fiji                           12.592000
2             Vanuatu                            6.744400
3   Papua New Guinea                             4.807800
4         Philippines                            4.196200
5             Vietnam                            3.825700
6          East Timor                            3.278700
7        South Africa                            2.564900
8               China                            2.041200
9            Malaysia                            1.682100
10   Solomon Islands                             1.640700
11         Sri Lanka                             1.285200
12          Thailand                             0.771730
13             Samoa                             0.606090
14             India                             0.146220
15           Somalia                             0.118410
16          Maldives                             0.073751
17          Tanzania                             0.033881
18    Comoro Islands                             0.031828
19        New Zealand                            0.024384
20            Myanmar                            0.014203
21         Mozambique                            0.013518
22               Peru                            0.008812
23          Mauritius                            0.007957
24          Singapore                            0.006588
25           Pakistan                            0.005989
26              Kenya                            0.005647
27         Madagascar                            0.002310
28            Ecuador                            0.001369
29          Australia                            3.930000
30          Australia                            0.800000
31          Australia                            0.180000
32          Australia                            0.040000
33          Australia                            0.010000
```

```
[FILL RESULT] => [Country from, Total number release]

         Country_from  Total_number_particles_released
0           Indonesia                           390636
1                Fiji                             6122
2             Vanuatu                             2337
3   Papua New Guinea                             9525
4         Philippines                           215974
5             Vietnam                           176511
6          East Timor                             3390
7        South Africa                           100943
8               China                           752074
9            Malaysia                           106756
10   Solomon Islands                             1943
11         Sri Lanka                           220728
12          Thailand                           100406
13             Samoa                              693
14             India                            74284
15           Somalia                            11258
16          Maldives                             3626
17          Tanzania                             4878
18    Comoro Islands                             6292
19       New Zealand                              742
20           Myanmar                            56007
21        Mozambique                             5381
22              Peru                            25239
23         Mauritius                             7031
24         Singapore                              742
25          Pakistan                            51621
26             Kenya                             2084
27        Madagascar                             3443
28           Ecuador                            10999
29         Australia                             1071
30         Australia                             1071
31         Australia                             1071
32         Australia                             1071
33         Australia                             1071
```

Statement 5

This command block is intended to perform the total summation of all data lines with duplicate country names in the given data set. At the same time, display the resulting value on the screen

```python
# Define function "Sum(Average and Total) of each country per day"  (Total sum)
fill_result = filtCSVCollumn(
    columns=["Country_from", "Average_daily_particles_inside_EEZ",
             "Total_number_particles_released"],
    group_by_name='Country_from',
    top=0,
    comparation_param='',
    src_file=file,
)
print("[FILL RESULT] => [Sum(Average or Total) of each country]\n\n{}\n\n", fill_result)

```

Output:

```
[FILL RESULT] => [Sum(Average or Total) of each country]

{}

          Country_from  Average_daily_particles_inside_EEZ  \
0            Australia                           4.960000
1                China                           2.041200
2      Comoro Islands                           0.031828
3           East Timor                           3.278700
4              Ecuador                           0.001369
5                 Fiji                          12.592000
6                India                           0.146220
7            Indonesia                         110.910000
8                Kenya                           0.005647
9           Madagascar                           0.002310
10            Malaysia                           1.682100
11            Maldives                           0.073751
12           Mauritius                           0.007957
13          Mozambique                           0.013518
14             Myanmar                           0.014203
15         New Zealand                           0.024384
16            Pakistan                           0.005989
17  Papua New Guinea                           4.807800
18                Peru                           0.008812
19         Philippines                           4.196200
20               Samoa                           0.606090
21           Singapore                           0.006588
22     Solomon Islands                           1.640700
23             Somalia                           0.118410
24        South Africa                           2.564900
25           Sri Lanka                           1.285200
26            Tanzania                           0.033881
27            Thailand                           0.771730
28             Vanuatu                           6.744400
29             Vietnam                           3.825700

    Total_number_particles_released
0                              5355
1                            752074
2                              6292
```

	Total_number_particles_released
0	5355
1	752074
2	6292
3	3390
4	10999
5	6122
6	74284
7	390636
8	2084
9	3443
10	106756
11	3626
12	7031
13	5381
14	56007
15	742
16	51621
17	9525
18	25239
19	215974
20	693
21	742
22	1943
23	11258
24	100943
25	220728
26	4878
27	100406
28	2337
29	176511

Statement 6

Below is the command block below to perform data filtering and get the top 10 countries with the highest Total number of particles released

```
111     # Select Top 10 of Max(total Sum ) of all country
112     fill_result_top = filtCSVCollumn(
113         columns=["Country_from", "Total_number_particles_released"],
114         group_by_name="Country_from",
115         top=10,
116         comparation_param='Total_number_particles_released',
117         src_file=file
118     )
119     print('[Filtered results] => Select Top 10 of Max(total Sum ) of all country\n\n{}\n\n'.format(
120         fill_result_top))
121
```

Output:

```
[Filtered results] => Select Top 10 of Max(total Sum ) of all country

              Total_number_particles_released
Country_from
China                                  752074
Indonesia                              390636
Sri Lanka                              220728
Philippines                            215974
Vietnam                                176511
Malaysia                               106756
South Africa                           100943
Thailand                               100406
India                                   74284
Myanmar                                 56007
```

Statement 7

Below is the block of code to calculate the average Total number of particles released

```
122     # Average of Total number particles released
123     data = DataFrame(file)
124     average_total = data["Total_number_particles_released"].mean()
125     print("Average of Total number particles released: ",
126           round(average_total, 2), '\n\n')
```

Output:

```
Average of Total number particles released:  69324.12
```

Statement 8

The function below allows us to predict what the average daily particles inside EEZ will be for a certain upcoming period of time.

```
35  def predict(years, file):
36      data = DataFrame(file)
37      total_average = data["Average_daily_particles_inside_EEZ"].sum()
38      predict_result = ((total_average * 30)*12) * years
39      difference = predict_result - predict_result/years
40      return predict_result, difference
41
```

```
127     # Input how many(year)
128     year = int(input("Enter the number of year you wanna predict: "))
129     predict_result, diff = predict(years=year, file=file)
130     print("[PREDIC (variable) year: int  year(s), average will be {} and increases {}"
131           .format(year, round(predict_result, 2), round(diff, 2)))
132
```

Output:

```
Enter the number of year you wanna predict: 5
[PREDICT RESULT] - In next 5 year(s), average will be 292322.86 and increases 233858.29
```

Statement 9

```python
def incomeTotalPlastic(file, country_name, current_value):
    data = DataFrame(file)
    country_total = data.loc[data["Country_from"] == country_name]
    total_each_country = country_total["Total_number_particles_released"]
    result = int((total_each_country * current_value)*30*12)
    return result

```

```python
    country_name = input("Which country you wanna know?\n>_ ")
    result_income = incomeTotalPlastic(
        file=fill_result_of_total, country_name=country_name, current_value=60)
    print(result_income)
```

Output:

```
Which country you wanna know?
>_ Vietnam
3812637600
[RESULT VALUE] => The income of Vietnam by sell Plastic recycle are $3812637600
```

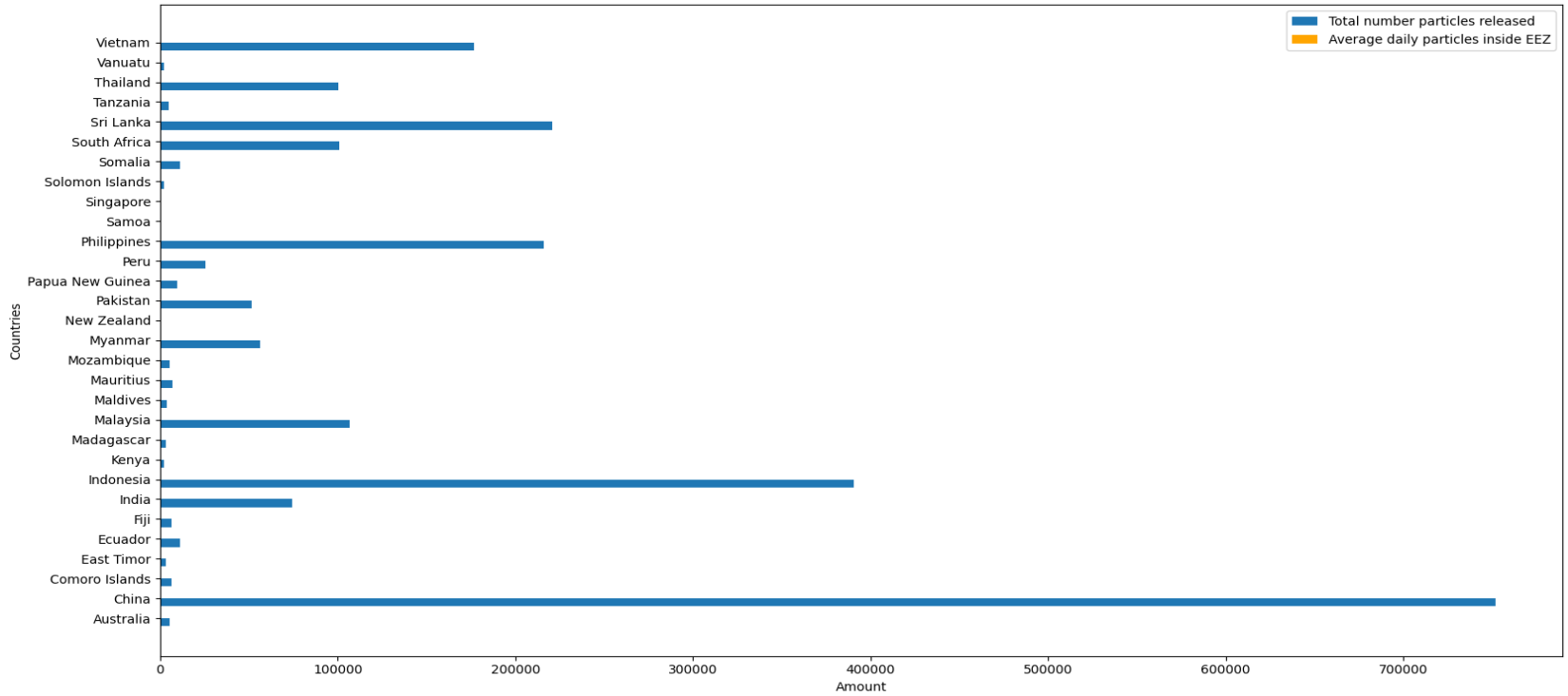
Statement 10

```python
def drawBarChart(data):
    x = np.arange(len(data["Country_from"]))
    y1 = data["Total_number_particles_released"]
    y2 = data["Average_daily_particles_inside_EEZ"]
    width = 0.4
    # plt.rcParams["figure.figsize"] = (30, 8)
    # plot data in grouped manner of bar type
    plt.barh(x-0.4, y1, width)
    plt.barh(x, y2, width, color='orange')
    plt.yticks(x-0.2, data['Country_from'])
    plt.ylabel("Countries")
    plt.xlabel("Amount")
    plt.legend(["Total number particles released",
                "Average daily particles inside EEZ"])
    plt.show()


```

```python
drawBarChart(fill_result)

```

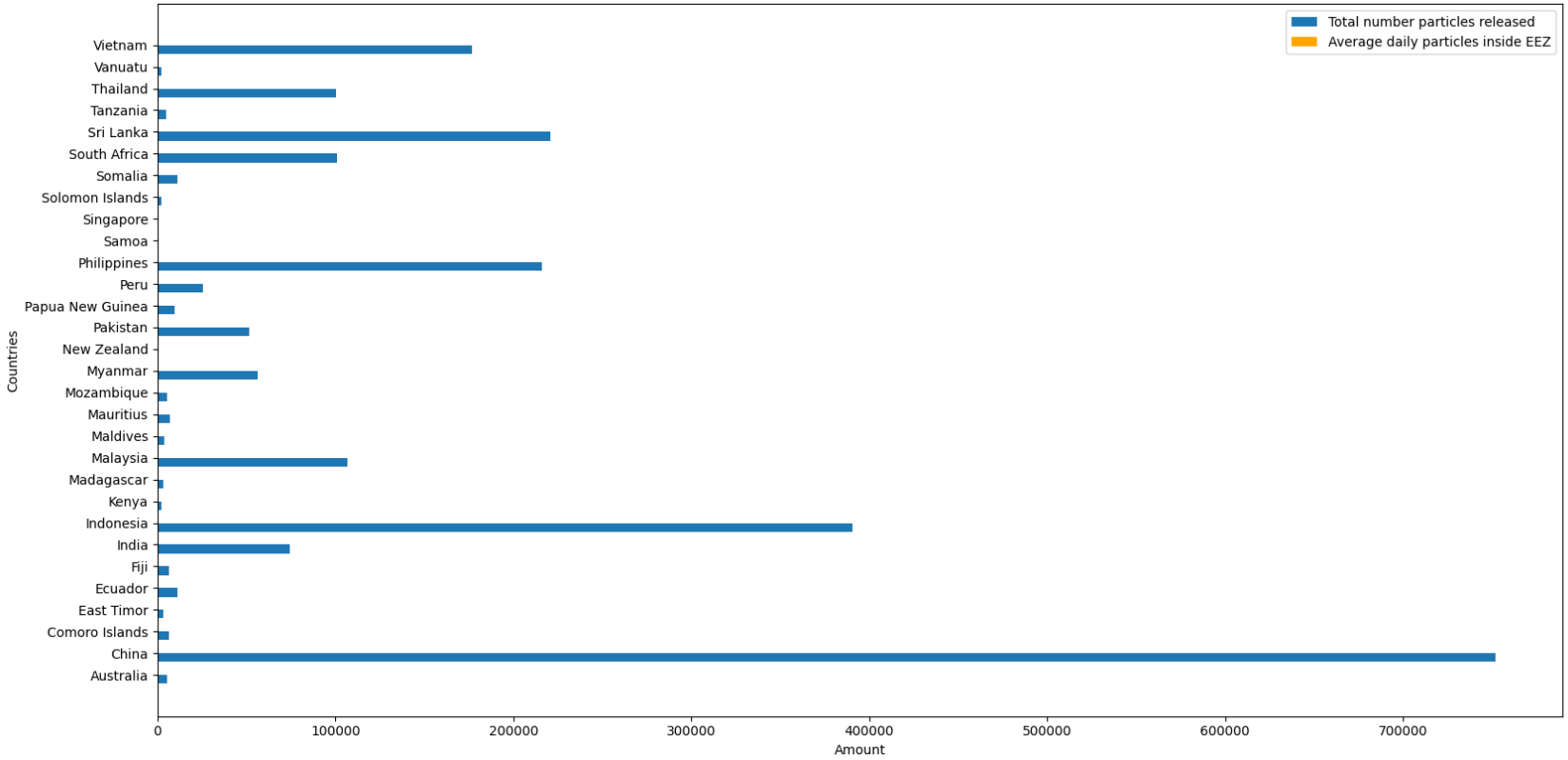
Output:

On this chart we have not seen the Average daily particle inside the EZZ value, which is almost invisible in the chart. Why so ? In essence, this range of values still appears on the graph. However, because the two-valued correlation between Total number particle released and Average daily particle inside EZZ is too large, the value of **x Axis** must be pushed to a higher value to be able to represent the Total number particle released value, making the Average daily particle inside EZZ barely represented on the chart.



Concussion of effective our Built BI Tool to decision-making (we have built our own BI tool by Python code programming language for more advance of effective decision-making of ML which program tool's analysis and report**)**

As we can also see above, the BI analysis tool is capable of analyzing and processing data according to user requirements. Accordingly, users can rely on the data analyzed and diagnosed by the tool to provide future reasonable methods for their projects and businesses. Let's take an example to better see its use,



[Filtered results] => Select Top 10 of Max(total Sum) of all country

Country_from	Total_number_particles_released
China	752074
Indonesia	390636
Sri Lanka	220728
Philippines	215974
Vietnam	176511
Malaysia	106756
South Africa	100943
Thailand	100406
India	74284
Myanmar	56007

Based on the chart and the result of the data analysis above, we can completely know where the amount of waste is the most, and where is the most reasonable and potential place for us to apply the project to. Typically, China, Indonesia, the Philippines, Sri Lanka and Vietnam are the countries with the largest amount of plastic waste generated. This shows that these places are the places that need the earliest and most drastic recycling projects to prevent pollution effectively.

As another example, as we have also seen above, we have a function to diagnose the amount of waste in 5 years if our project does not apply. (Note this is diagnostic reference data. and assuming the amount of garbage grows constant over the years). In the figure below I will diagnose the hypothetical situation of the amount of waste after 5 years.

```
Enter the number of year you wanna predict: 5
[PREDICT RESULT] - In next 5 year(s), average will be 292322.86 and increases 233858.29
```

It can be seen that the decision tool is very effective in making decisions for the development and operation of our business. It provides us. Provides metrics and diagnostics, making it easier for us to track and decide what to do next with our Business

P6 The legal issues involved in the secure User's data exploitation of business intelligence Tools

BI has always been an important resource for law firms and one that primarily serves the roles of library and analysis within the firm. Historically, through packaged reports, dashboard-based information displays, and data analysis, BI has been used to capture a moment in time and put out static information on different subjects such as customers, markets, and lawyers. BI's current role varies from displaying market statistics, making decisions, analyzing past outcomes, and tracking goals; however, the task is shifting as law firms ' and business needs become more complex.

As we do e-commerce business, e-commerce is changing rapidly and companies are trying to adapt to those changes. If anyone is tourism business owner, he should bear in mind the legal aspects of e-commerce when deciding to sell his goods or services online. The Legal issues involved in the secure exploitation of business intelligence tools has been explained below:

- **Maintain privacy of clients:**

Client's confidentiality has revolutionized the physical world which expect to protect information with increasing dependency on online portals and access to devices due to the confidential information of each individual shouldn't be shared to anyone.

- **Privacy:**

Most User have their personal data distributed all over the digital globe as items considered to be safe may be unintended outlets, such as email or private accounts. Software behaviors of the majority of workers.

Privacy has evolving legal consequences but principled concerns are still present as well as doesn't user even know how their accounts are being monitored as below:

- The tech world points out
- Privacy issues can quickly turn into a slippery slope
- Slowly eroding a person's right to privacy altogether.

- **Digital Ownership:**

Digital media has allowed information to flow freer than ever before that comes with a legal and reaction against this exchange of ideas. Ownership be defined in the digital realm as things can be copied and shared very quickly online which makes intelligent property difficult to manage. Legal concepts like copyright failed to keep up with modern days.

Therefore, most Industries in the music and film industry have constrained for tougher legal rights of intellectual property while other movements have aimed to provide tougher opportunities to the exchange of ideas on the digital domain.

- **Data Gathering:**

At some point, we know our online life is being monitored. However, In the name of some national security such as the USA have approved laws authorizing government action to track private citizens. Such steps rekindled a controversy as what & why of information should be obtained which companies need to decide specific information have to gather.

Therefore, we solve this problem by a constant query which information we are monitoring user while users have the right to know as our marketing survey form only that how to collect & uses their information on survey base on database.

- **Security Liability:**

Problem: “Security liability coverage protects against risks associated with the failure to protect electronic data containing others’ private information, the inadvertent transmission of a computer virus, the inability of authorized users to access your website or computer network, and failure to notify individuals as required by any security breach notification law that applies.”

According to (CITATION, Sec 19) Security problems, in order to secure sensitive information and valuable properties which is a much more complex protection mechanism for data networks to be electronic. This increased protection comes with increased supervision.

However. Both security devices bear inherent threats, which means that what rights can be forfeited is a question of. In the end, IT professionals need to balance risk with independence to build a security that is both effective and ethical.

M4 Conduct research to identify specific examples of organizations that have used business intelligence tools to enhance or improve operations.

❖ Big data and data visualization to raise awareness about the scope of the problem

So, to apply data science into this as well as the rapid development of key technologies involving big data and machine learning is opening up new possibilities for measuring and tracking Earth’s resources. It also allows for better coordination between researchers to share and analyses key data on pollution.

Using data analysis and data visualization, environmental scientists and organizations can gather, clean and analyses existing data on plastic pollution and use data visualization tools in order to gain better insights and raise awareness about the causes, effects and the scope of the issue.

In addition, scientists could train machine learning models to identify the key variables and make predictions that could be used in building long-term strategies to reduce ocean plastic.

Here are some examples of some companies and our organizations ocean cleanup that have used business intelligence tools to enhance or improve their business operations:

1) Our organization Ocean Cleanup (N.Koleva, June,2019):

- **Problem**: The unprecedented and ever-growing amount of plastic pollution is causing serious damage to the world's ecosystems. Of all of them, the ocean and its lifeforms are suffering the most. Every year, over 8 tons of plastic are dumped into the ocean and by 2025, this number is expected to double. Studies show that one million seabirds and 100,000 marine mammals die every year from plastic ingestion, with a growing number of species facing a risk of extinction.

Figure 22 Golbal-plastic pollution

- **Solution**:

The Ocean Cleanup is a non-profit engineering environmental organization based in the Netherlands, that develops technology to extract plastic pollution from the oceans. The approach involves placing cleanup systems in ocean gyres to scoop up marine debris as the system is pushed by wind and current.

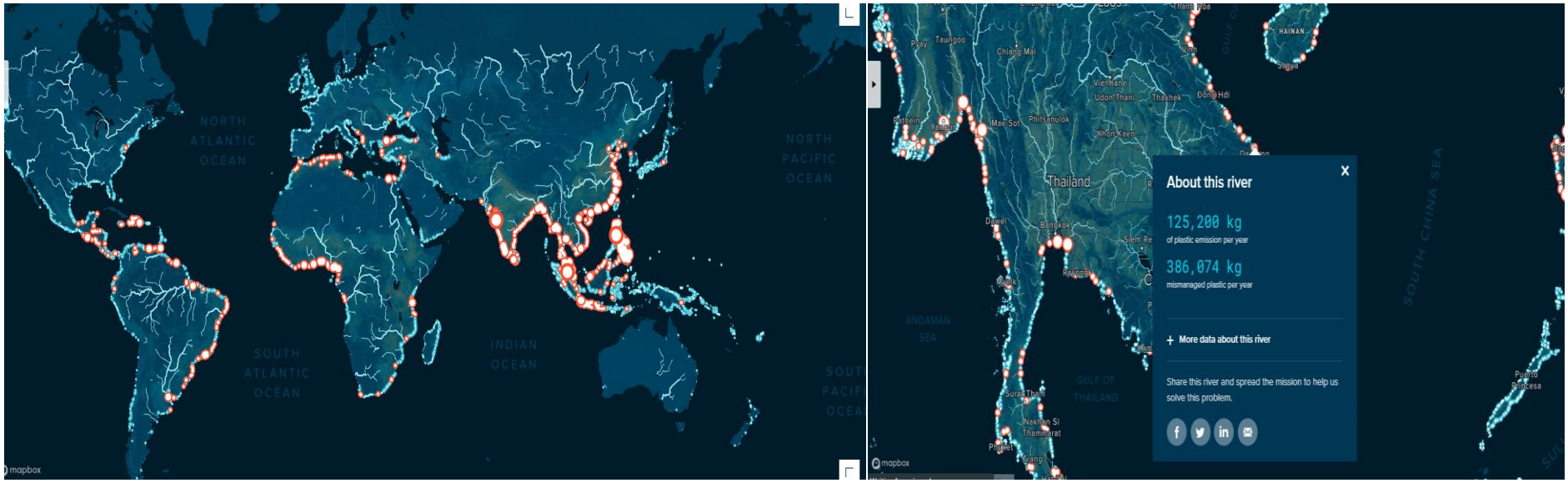


Figure 23 The Price Tag OF Plastic Pollution & The 1000 Top Polluting Rivers (Ocean cleanup, 2021)

Therefore, to solve the problem, The organization also conducts scientific research into oceanic plastic pollution including two expeditions to the North Pacific Gyre, collecting huge amounts of data, and publicizing multiple scientific papers which The Ocean Cleanup uses satellite imaging and machine learning to help clean up and capture the 5 trillion pieces of plastic trash they have observed in the world's "ocean garbage patches." They estimate that within 5 years they could collect 50% of the ocean waste.

2) Provide our conduct specific examples from our team's Business research result of BI Tools by Tableau and Python on Ocean Cleanup

By our team's business also conducts scientific research into oceanic plastic pollution including two result below wit tableau and Python:

Tableau Dashboard result.

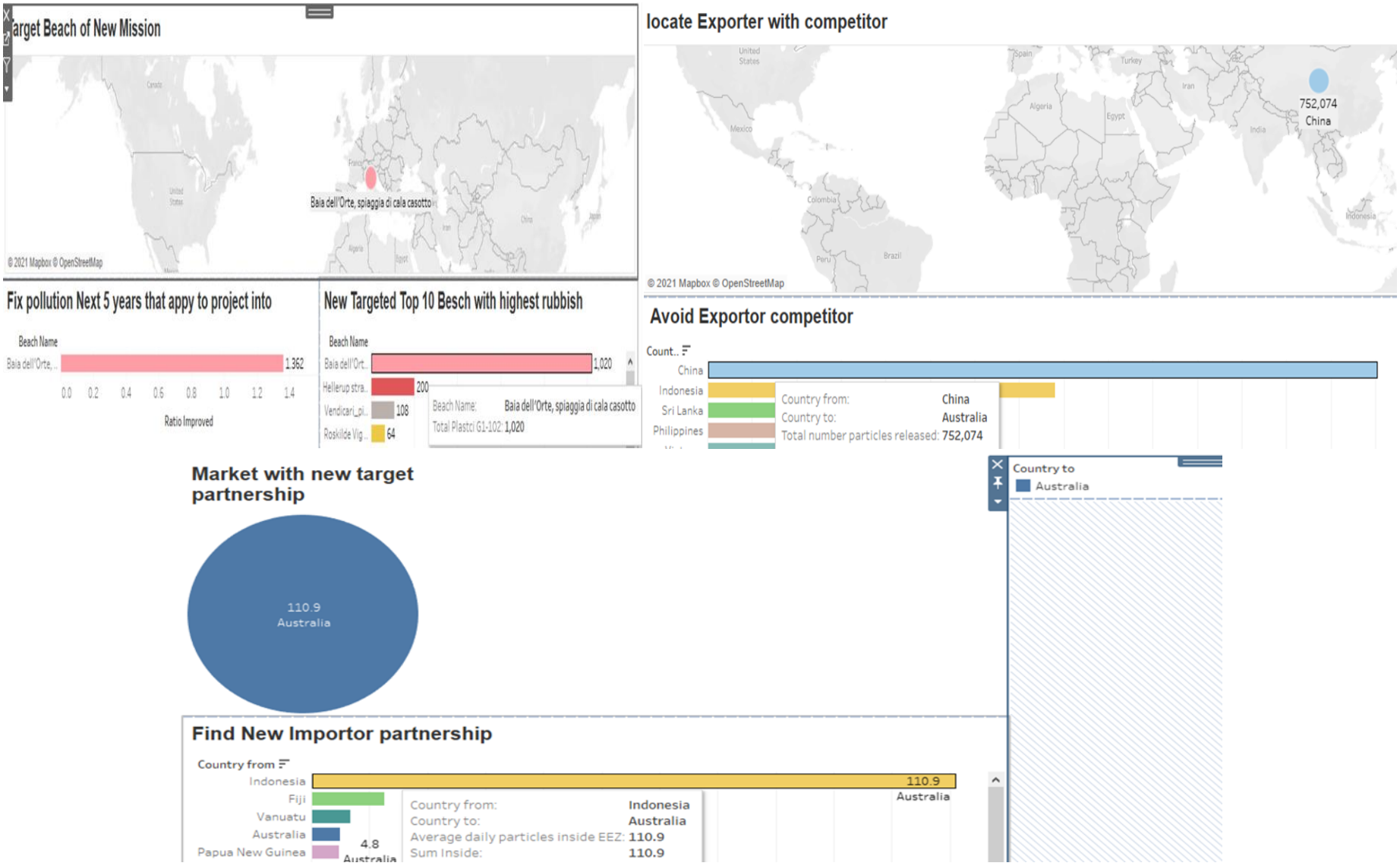


Figure 24 conducted specific example of tableau

Using tableau to visualize huge amounts of data, and spreading multiple scientific charts which we use dashboards and apply to help calculate future amount trillion of plastic trash such as

- **Identify new Mission Ocean cleanup**

they have observed in the world's "Target beach of new mission" we estimate that within 5 years they could decrease possible average amount 198%/5 years of the ocean waste, if we apply Ocean cleanup project into new target as Baia Dall ort and Krystina Heller Beach to be new mission.

- **Identify Competitor and New Partnership of market**

Another result of trading plastic recycle production after cleanup project, we visualize counties trading data with estimated charts into dashboards and observe trading trend of plastic trash in the world's "Locate Exporter with Competitor" and "New Target Partnership" we are able to estimate that within those counties importer as Australia have import highest average of plastic into country as 120.9 ton from Indonesia while also the one is most importing with top 10 exporters, by this observation Market for partnership, Australia is the most potential contracted retailer or buyer to make trading with our project production of plastic waste.

Built python result.

```python
def predict(years, file):
    data = DataFrame(file)
    total_average = data["Average_daily_particles_inside_EEZ"].sum()
    predict_result = ((total_average * 30)*12) * years
    difference = predict_result - predict_result/years
    return predict_result, difference

```

```python
def incomeTotalPlastic(file, country_name, current_value):
    data = DataFrame(file)
    country_total = data.loc[data["Country_from"] == country_name]
    total_each_country = country_total["Total_number_particles_released"]
    result = int((total_each_country * current_value)*30*12)
    return result

```

```python
# Input how many(year)
year = int(input("Enter the number of year you wanna predict: "))
predict_result, diff = predict(years=year, file=file)
print("[PREDIC (variable) year: int year(s), average will be {} and increases {}"
      .format(year, round(predict_result, 2), round(diff, 2)))

```

```python
country_name = input("Which country you wanna know?\n>_ ")
result_income = incomeTotalPlastic(
    file=fill_result_of_total, country_name=country_name, current_value=60)
print(result_income)
```

```
Enter the number of year you wanna predict: 5
[PREDICT RESULT] - In next 5 year(s), average will be 292322.86 and increases 233858.29
```

```
Which country you wanna know?
>_ Vietnam
3812637600
[RESULT VALUE] => The income of Vietnam by sell Plastic recycle are $3812637600
```

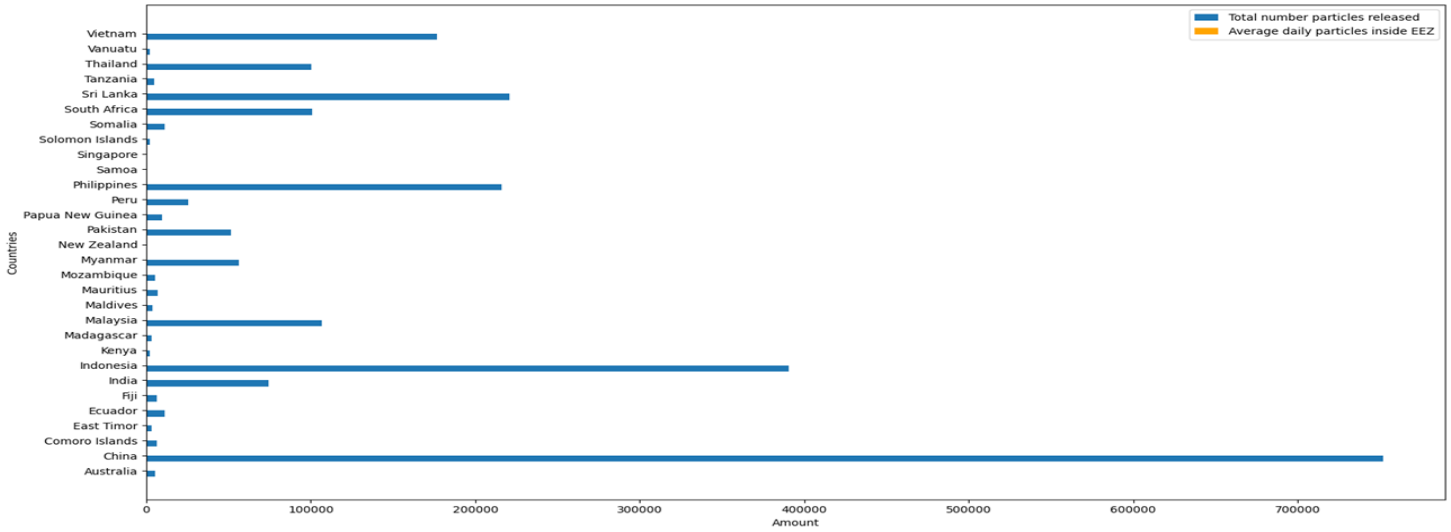


Figure 25 conducted specific example of python

Try to Using python to visualize data, and spreading multiple function and method which apply ML or DL to help regression future trading of plastic trash,

- **Predict possible of future Increased pollution.**

The firstly function “Predict years”, we estimate that input within 5 years, we could predict possible total average of increased pollution within 1- 5 years and different In 1 years of the ocean waste.

- **Predict daily income into annual income of a year.**

The secondly function “income total plastic” to estimate selected country as Vietnam who have release plastic waste or export which could make income buy selling plastic recycle with amount of 3.812 Million USD by 60$ USD / Ton.

- **Visualization data into image chart.**

The last, we visualize those function’s data in big data for decision-makers over perspective or overview by evaluate the result, How to do? By we generate country_from as X and Total_number_release as Y into bar-chat which show plot which country can make most income depend on how mush or highest they have released.

D4 Evaluate organization's BI usage to extend target audience and more competitive within the market, taking security legislation into consideration.

Overview of Our research's BI usage to improve our team business on Ocean Cleanup:

While data science can be of great help with analyzing the current impact of plastic waste and predicting future development, one of the biggest problems is what we don't know. Properly documenting data about plastic pollution can be challenging, given the rate at which it's growing and how little we know about its long-term effects, but with increased computing power and AI algorithms, data scientists can better understand natural systems and ocean pollution patterns to optimize interventions.

Moreover, the democratization of AI and data science technologies means that emerging environmental actors can now produce valuable new knowledge on the matter.

However, making our business process more accurate and automated. The fact, must understand our organization to conducts scientific research into oceanic plastic pollution including collecting huge amounts of data, publicizing multiple scientific papers which uses satellite imaging with ML to help capture the pieces of plastic trash that have observed in the world.

Also, our organization uses some platform to maximize the use of our data to become more efficient and effective. Our Team2 Data Science supports the organization with logistics, image detection, including identifying plastic in drone images and barrier identification, marketing operations as well as day-to-day use.

While innovation and technology can't magically solve complex environmental and social issues such as ocean plastic pollution, democratizing data science and encouraging the use of AI for good can empower organizations and individuals to make a difference.

Therefore, taking Legal issues into consideration that are involved in the secure exploitation of business intelligence tools is important if a company wants to implement BI tools for enhancement or improvement their business. Many methods and ideas of using BI tools and techniques to extend the target audience of the business and make them more competitive within the market, taking security legislation into consideration.

Here are some ways that the organizations could use BI to extend their target audience and make them more competitive within the market, taking security legislation into consideration according to (Mathew, A. 2019):

Table 5 own opinion on Extend 10 Usage 's way lists of business intelligence Tools

Extend 10 Usage 's way lists of business intelligence Tools in our Ocean Cleanup

Target Audience
To understand what drives revenue for Ocean cleanup's business:
While BI can be a very broad term, to me it is knowing exactly what is going to drive revenue for the organization's business. To know that is to know who the organization's end user is, what influences their decisions, and how they consume their information. In today's marketplace, there are thousands of outlets pushing information, and not everyone can consume everything. BI is knowing where and how media outlets and through articles, videos, ads, podcast, etc. organization's ideal customer consumes their information, what is going to drive them to purchase and subscribe, and executing a plan to take advantage of that knowledge.
More Competitive Within the Market
By personalizing the sales strategy and anticipate objections:
BI is understanding more about the other business that the organization are trying to work with in any capacity. Generally, sales teams research this information to be prepared for objections specific to a company's situation. It's particularly useful, if the company that the organizations are trying to sell to is cutting budgets, having an increased personnel turnover, showing decreased sales numbers, about to merge, or outsourcing with competitors at lower prices.
To avoid bottlenecks and problems in a fragile market:
BI is critical to both executives and owners equally as it provides insights into the true health of an organization that are not readily available from a profit and loss statement. There is a plethora of BI applications available on the market today that integrate with organizations accounting software. Including forecasting to avoid tailbacks and provide perfect timing for decision making. An easy-to read dashboard that is reviewed to help organizations be more informed and avoid problems in a very delicate market.
By tracking inventory and capitalize on trends:
BI or analytics has become critical in retail given the fast pace at which consumers expect orders to be delivered to their doorsteps for example: Retailers and brands need to know where their inventory is located and how well it's selling. Analytics provides this insight to buyers and executives so they can capitalize on buying trends.
To discover the motivates of **consumer**'s **behavior:**
The numbers that drive BI because the way the executive suite sees the business is not exactly how customers see the business, and social media snippets aren't enough to fill the gap. To truly understand and satisfy customers' needs and correct problems, they must first understand why a phenomenon is happening, rather than go with a gut feeling. Adding qualitative analytics to their dashboard alongside the quantitative data can give them that big picture because it's adding the 'why' to the 'what.' But they're formulating BI strategy to tell them what &why user are doing thing.
To help building organization's own brand in a competitive, complex retail environment:
In the context of consumer-packaged goods (CPG) marketing, good BI is the data that enables marketers to plan and react faster, and more efficiently, throughout the decision-making process. This data allows insights professionals to find creative ways of growing the business through initiatives like cross-promotion strategies with complementing product categories.
P**owering up productivity:**
BI has the potential to release inefficiency bottlenecks, refine existing business processes, automate routine tasks and bring new levels of organization and arranging to work. The efficiency and productivity can be including more responsive customer service, better use sale's time, and closer measurement of product development cycles and marketing campaigns which is also evident at a more senior level thanks to

automated reporting and dashboards. User often work remotely that Cloud CRM solution has halved the number of calls back to the office significantly boosted productivity.

Security such as Maintain privacy of clients, Privacy, Digital Ownership, Data Gathering and Security Liability

- **Tightener up Data Accuracy and Compliance:**

Without addressing the issues regarding the implementation of business intelligence, the business is at risk through poor decision making based on inaccurate data and from increasingly strong data compliance regulations. So not only does holding data in separate siloes make it almost impossible to achieve the degree view of the organization's customers which risks the very practical matter of data accuracy and consistency, which will have a negative impact on all areas of their business. It is indeed underlying integrity of their data in any BI project but we also see improved data governance as a strong motivating factor behind such investments.

However, data helps to improve transparency and expose inaccuracies and gaps that will lead to wasted marketing spend, not to mention potential brand damage caused by sending insensitive communications. Data protection regulations around the world are increasingly tightening the rules around the capture, storage and convention of personal data. The new laws include requirements to keep data accurate and up-to-date, to demonstrate grounds for processing data and formulate a clear privacy policy for improved transparency. If the company has poorly managed records, the employees are high questionable to be in a position to meet these requirements with risk falling of the regulator as well as the company's customers.

Decision-making

- **To determine the Return on Investment of organization's marketing strategy:**

In a market saturated with apps, social media platforms, analytics tools, and pay-per-click campaigns, business intelligence is crucial in helping small businesses figure out if the marketing strategy they've invested in is producing ROI. BI can translate into analytics reports where businesses can base decisions on solid research, data, and facts, rather than intuition, assumption, or gut instinct.

- **By helping the organization in making smart decision making:**

As a business manager or owner, it's vital to have a firm hold on what the organization's data is telling as known as information doesn't equal to intelligence as the case that information is siloed in different parts of their business. The goal of a BI creativity is to convert the information into structured, analyzable insight.

In other words, BI that can inform strategic decision-making across base on organization's customers' own experiences which clearly having up date, data-driven intelligence not to better business decisions but will eventually contribute to greater financial routine. The technical support to intelligent decision-making is a single, centralized source which collected data on all of business activities and customer interactions. Customer Relationship Management (CRM) solutions typically play an invaluable role here.

Therefore, a well-implemented CRM acts between teams can enable organizations to run reports that deliver a key of business metrics on productivity, staff performance, product preferences, sales cycles, customer behavior, core customers, revenues and market trends all ready for analysis by the management team. It's then possible to quickly pinpoint performance hotspots and to flag up where processes could be emulated in other parts of the business, or where adjustments need to be made. Any resulting decisions will be based on hard facts rather than guesswork or assumptions.

References

Mohit S, (Oct,2020): Data Preprocessing: 6 Necessary Steps for Data Scientists from https://hackernoon.com/what-steps-should-one-take-while-doing-data-preprocessing-502c993e1caa

Goyal. K.(Jan, 2020): Data Preprocessing in Machine Learning: 7 Easy Steps To Follow From https://www.upgrad.com/blog/data-preprocessing-in-machine-learning/

Iyoob. J (Sep. 2012): BUSINESS INTELLIGENCE HELPS IN STRATEGIC DECISION MAKING from https://www.etechgs.com/blog/business-intelligence-helps-strategic-decision-making/

N.Koleva (June.2019): Solving the Ocean Plastic Pollution Problem With Data from https://blog.dataiku.com/solving-the-ocean-plastic-pollution-problem-with-data

Tableau's solution: Coca-Cola Bottling Company empowers the enterprise with Tableau Mobile dashboards to drive bottom line from https://www.tableau.com/solutions/customer/eliminating-reporting-bottleneck-ccbc

Tableau's solution: Des Moines Public School District improves intervention programs with predictive analytics form https://www.tableau.com/solutions/customer/des-moines-public-school-district-improves-intervention-programs-predictive

A.Mathew (Jun, 2016): Key Benefits of Using Business Intelligence Tools from https://www.liveadmins.com/blog/4-key-benefits-of-using-business-intelligence-tools/

Ocean Cleanup(2021): The Price Tag OF Plastic Pollution & The 1000 Top Polluting Rivers from https://theoceancleanup.com/sources/ & https://theoceancleanup.com/the-price-tag-of-plastic-pollution/