Shark Attacks Marketing opportunities

Objectives

Analyze the data of **Global Shark Attacks** and convince investors to set up an Ocean Guardian franchise in North America and South America.

OCEAN GUARDIAN

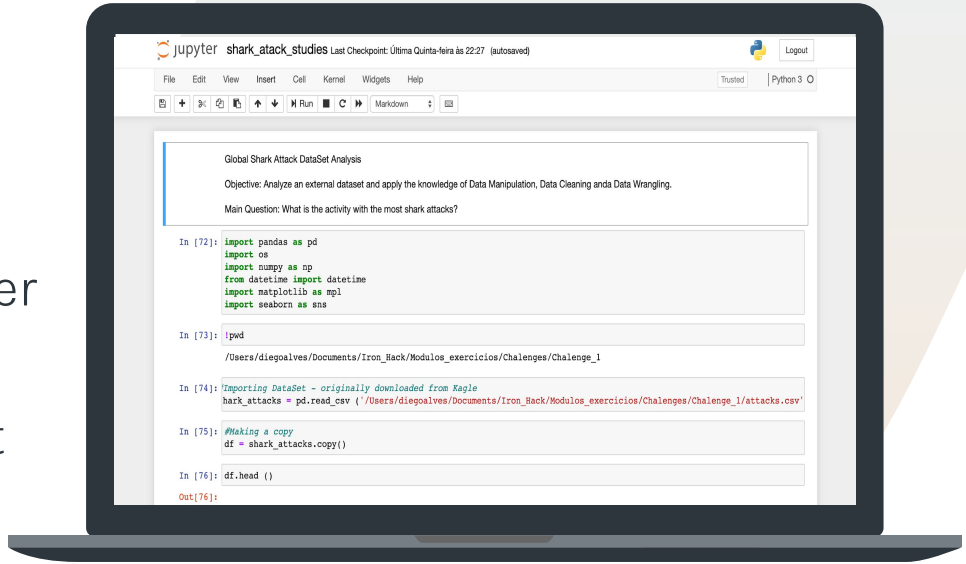
1

[Off Topic]

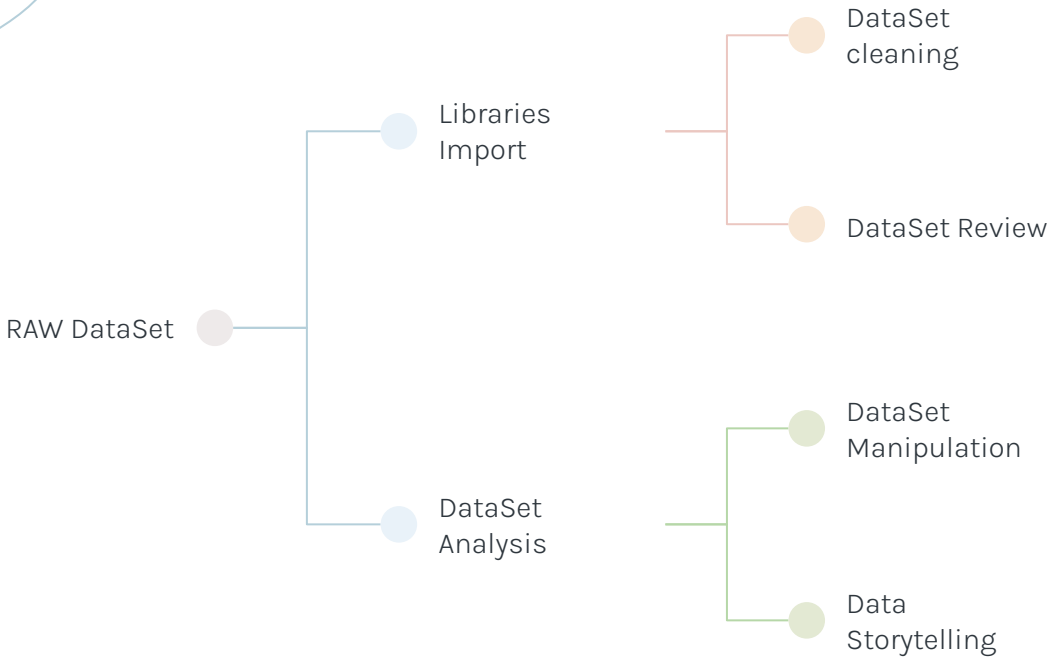
Objective: Analyze an external data set and apply the knowledge of **data manipulation**, **data cleaning** and **data organization.**

Data project

In this stage, the Dataset was prepared so that, after cleaning and organizing the data, it could support the arguments for the proposed purpose.



Workflow used in this project



Global Shark Attack DataSet Analysis

Objective: Analyze an external dataset and apply the knowledge of Data Manipulation, Data Cleaning anda Data Wrangling.

Main Question: What is the activity with the most shark attacks?

In [72]:
```
import pandas as pd
import os
import numpy as np
from datetime import datetime
import matplotlib as mpl
import seaborn as sns
```

In [73]:
```
!pwd
```

```
/Users/diegoalves/Documents/Iron_Hack/Modulos_exercicios/Chalenges/Chalenge_1
```

In [74]:
```
#Importing DataSet - originally downloaded from Kagle
hark_attacks = pd.read_csv ('/Users/diegoalves/Documents/Iron_Hack/Modulos_exercicios/Chalenges/Chalenge_1/attacks.csv'
```

In [75]:
```
#Making a copy
df = shark_attacks.copy()
```

Out[78]:
```
Index(['Case Number', 'Date', 'Year', 'Type', 'Country', 'Area', 'Location',
       'Activity', 'Name', 'Sex ', 'Age', 'Injury', 'Fatal (Y/N)', 'Time',
       'Species ', 'Investigator or Source', 'pdf', 'href formula', 'href',
       'Case Number.1', 'Case Number.2', 'original order', 'Unnamed: 22',
       'Unnamed: 23'],
      dtype='object')
```

In [79]:
```
#Declaring function for clean columns
def clear_columns (df):
    x = []
    for columns in df.columns:
        columns = columns.strip().lower()
        columns = columns.replace ('.', ' ')
        columns = columns.replace (' ', '_')
        x.append (columns)

    df.columns = x
    return df.columns
```

In [80]:
```
clear_columns (df)
```

Out[80]:
```
Index(['case_number', 'date', 'year', 'type', 'country', 'area', 'location',
       'activity', 'name', 'sex', 'age', 'injury', 'fatal_(y/n)', 'time',
       'species', 'investigator_or_source', 'pdf', 'href_formula', 'href',
       'case_number_1', 'case_number_2', 'original_order', 'unnamed:_22',
       'unnamed:_23'],
      dtype='object')
```

```
In [91]: df_droped = df.drop (columns = ['unnamed:_22' ,'unnamed:_23'], inplace = True)
```

```
In [92]: df.head ()
```

Out[92]:

	case_number	date	year	type	country	area	location	activity	name	sex	...	fatal_(y/n)	time	species	investigator_or_sc
0	2018.06.25	25-Jun-2018	2018.0	Boating	USA	California	Oceanside, San Diego County	Paddling	Julie Wolfe	F	...	N	18h00	White shark	R. Collier, (
1	2018.06.18	18-Jun-2018	2018.0	Unprovoked	USA	Georgia	St. Simon Island, Glynn County	Standing	Adyson McNeely	F	...	N	14h00 -15h00	NaN	K.McM TrackingSharks
2	2018.06.09	09-Jun-2018	2018.0	Invalid	USA	Hawaii	Habush, Oahu	Surfing	John Denges	M	...	N	07h45	NaN	K.McM TrackingSharks
3	2018.06.08	08-Jun-2018	2018.0	Unprovoked	AUSTRALIA	New South Wales	Arrawarra Headland	Surfing	male	M	...	N	NaN	2 m shark	B. Myatt, (
4	2018.06.04	04-Jun-2018	2018.0	Provoked	MEXICO	Colima	La Ticla	Free diving	Gustavo Ramos	M	...	N	NaN	Tiger shark, 3m	A .K

5 rows × 22 columns

In [53]:
```
df['activity'].value_counts()
```

Out[53]:
```
Surfing                                             965
Swimming                                            845
Fishing                                             418
Spearfishing                                        324
Bathing                                             154
                                                   ...
Diving (Hookah)                                       1
Gaffing netted shark                                  1
Swimming /  Whale Watching                            1
The steamships Thingvalla and Geiser collided         1
Lobstering                                            1
Name: activity, Length: 1420, dtype: int64
```

In [54]:
```
df['activity']
```

Out[54]:
```
0           Paddling
1           Standing
2            Surfing
3            Surfing
4        Free diving
            ...
6296        Swimming
6297          Diving
6298    Pearl diving
6299        Swimming
6301        Swimming
Name: activity, Length: 5874, dtype: object
```


DF['ACTIVITY'].VALUE_COUNTS(
).NLARGEST(12).INDEX
imgflip.com

In [55]:
```
top_eight = df['activity'].value_counts().nlargest(8).index
```

In [56]:
```
top_eight
```

Out[56]:
```
Index(['Surfing', 'Swimming', 'Fishing', 'Spearfishing', 'Bathing', 'Wading',
       'Diving', 'Standing'],
      dtype='object')
```

In [57]:
```
top_ten = df['activity'].value_counts().nlargest(10).index
```

In [58]:
```
top_ten
```

Out[58]:
```
Index(['Surfing', 'Swimming', 'Fishing', 'Spearfishing', 'Bathing', 'Wading',
       'Diving', 'Standing', 'Snorkeling', 'Scuba diving'],
      dtype='object')
```

In [59]:
```
top_twelve = df['activity'].value_counts().nlargest(12).index
```

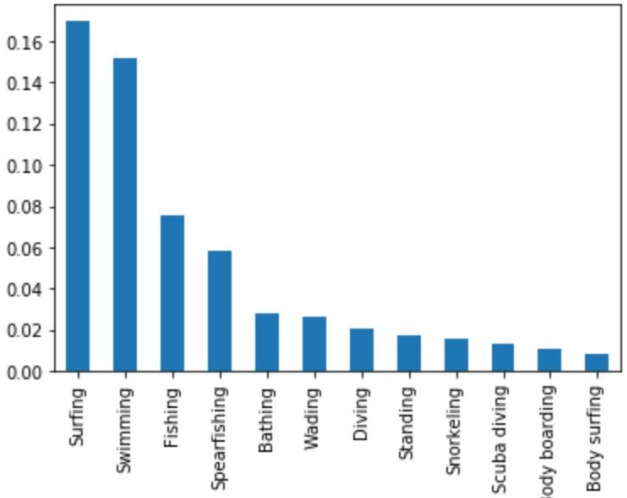
In [60]:
```
top_twelve
```

Out[60]:
```
Index(['Surfing', 'Swimming', 'Fishing', 'Spearfishing', 'Bathing', 'Wading',
       'Diving', 'Standing', 'Snorkeling', 'Scuba diving', 'Body boarding',
       'Body surfing'],
      dtype='object')
```

In [37]:
```
import matplotlib as mpl
```

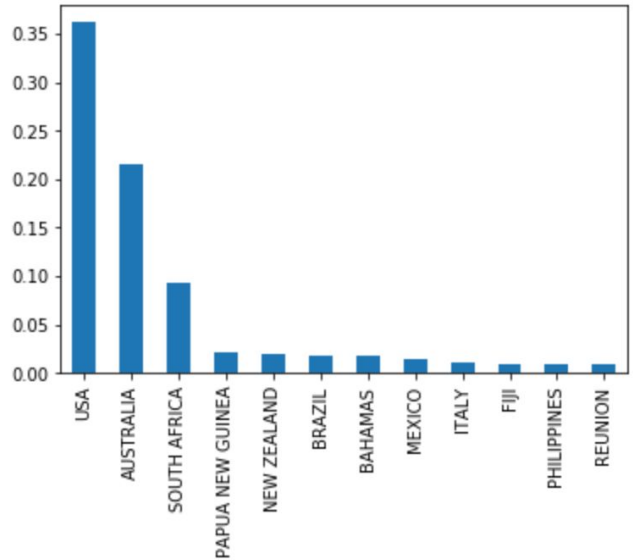
In [39]:
```
df['activity'].value_counts(normalize=True).head(12).plot(kind='bar')
```

Out[39]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x122637b90>
```



In [48]:
```
df['country'].value_counts(normalize=True).head(12).plot(kind='bar')
```

Out[48]: <matplotlib.axes._subplots.AxesSubplot at 0x11ee68650>



2

Business opportunities

Objective: convince investors to set up an Ocean Guardian franchise in North America and South America.

Surfers in the world

23.000.000

It is the estimated total of surfers in the world

3.000.000

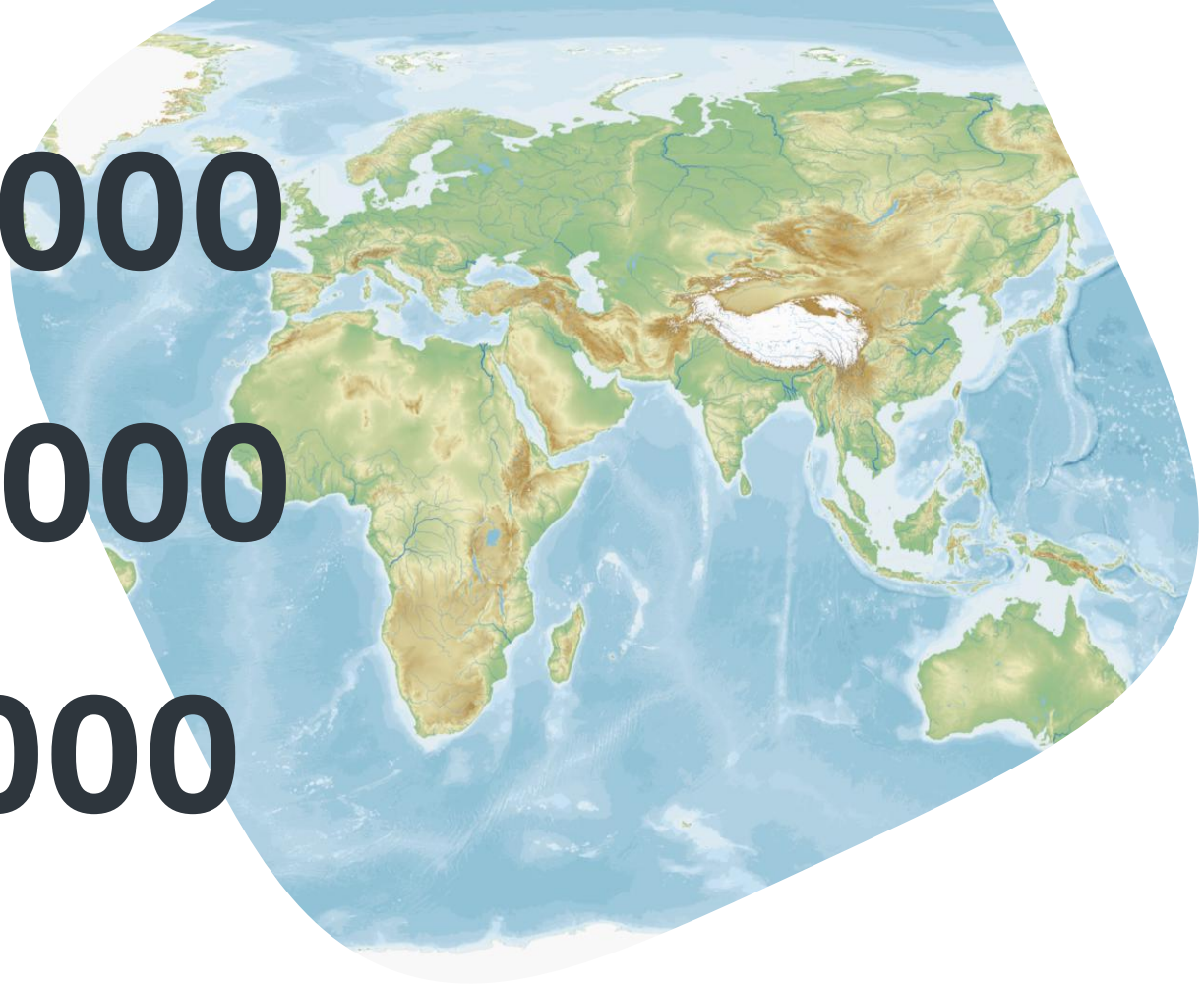
Brazil

2.800.000

E.U.A

1.700.000

Australia



13.500.000

Surfers in the American Continent

"Stop selling. Start helping" – Zig Ziglar

Our Solution


NOT ACTIVATED
TERRA AUSTRALIS

Thanks!

Any questions?

You can find me at diegoalvesteo@gmail.com