

# Pandas-Lab

## **Question 1.**

For each of the following questions you will use a dataset containing information on global shark attacks.

### **Attribute Information:**

The attributes recorded in the dataset are as follows:

0. Case Number
1. Date
2. Year
3. Type
4. Country
5. Area
6. Location
7. Activity
8. Name
9. Sex
10. Age
11. Injury
12. Fatal (Y/N)
13. Time
14. Species
15. Investigator or Source

## **Question 1.**

(i)

What location globally has the highest number of shark attacks.

New Smyrna Beach, Volusia County 157
-----------------------------------------

(ii)

Read the shark attack dataset into a Pandas Dataframe.

Determine the six countries that have experienced the highest number of shark attacks.

USA	2116
AUSTRALIA	1279
SOUTH AFRICA	565
PAPUA NEW GUINEA	133
NEW ZEALAND	125
BRAZIL	102

(iii)

Modify your code to print out the six countries that have experienced the highest number of fatal shark attacks.

AUSTRALIA	341
USA	246
SOUTH AFRICA	137
PAPUA NEW GUINEA	56
MEXICO	43
BRAZIL	40

,

(iv)

Based on the data in the Activity column are you more likely to be attacked by a shark if you are “Surfing” or “Scuba Diving”.

Numbers of attack when Surfing	731
Numbers of attack when Scuba Diving	48

## Question 2.

(i)

Determine from the dataset what percentage of all recorded shark attacks were fatal.

Percentage of attacks that are fatal:
26.4530424408

(ii)

For each individual country print out the percentage of fatal shark attacks (number of fatal shark attacks expressed as a percentage of the total number of shark attacks). Some countries have recorded 0 fatal and non-fatal attacks. You will need to take this into account in your code.

```
Percentage of attacks that are fatal in USA : 11.7872544322
Percentage of attacks that are fatal in AUSTRALIA : 27.4115755627
Percentage of attacks that are fatal in NEW CALEDONIA : 37.2549019608
Percentage of attacks that are fatal in REUNION : 47.3684210526
Percentage of attacks that are fatal in BAHAMAS : 12.2448979592
Percentage of attacks that are fatal in SPAIN : 23.6842105263
Percentage of attacks that are fatal in CHINA : 50.0
Percentage of attacks that are fatal in JAPAN : 43.75
Percentage of attacks that are fatal in COLUMBIA : 22.2222222222
.....
```

### Question 3.

(i)

In this question we are interested in looking at the number of recorded shark attacks over time for a specific country. Write a function called *calculateYearlyAttacks* that will take in a valid country name as a parameter and the attack dataframe. It should print out the number of recorded shark attacks for the country for every year from 1925 to 2015. The following is a sample output when the function is called and passed “AUSTRALIA” as the country.

```
Number of attacks in AUSTRALIA during 1925 : 4
Number of attacks in AUSTRALIA during 1926 : 5
Number of attacks in AUSTRALIA during 1927 : 11
Number of attacks in AUSTRALIA during 1928 : 9
Number of attacks in AUSTRALIA during 1929 : 24
Number of attacks in AUSTRALIA during 1930 : 10
Number of attacks in AUSTRALIA during 1931 : 9
Number of attacks in AUSTRALIA during 1932 : 10
Number of attacks in AUSTRALIA during 1933 : 12
Number of attacks in AUSTRALIA during 1934 : 16
Number of attacks in AUSTRALIA during 1935 : 17
Number of attacks in AUSTRALIA during 1936 : 21
Number of attacks in AUSTRALIA during 1937 : 18
Number of attacks in AUSTRALIA during 1938 : 9
Number of attacks in AUSTRALIA during 1939 : 9
Number of attacks in AUSTRALIA during 1940 : 10
Number of attacks in AUSTRALIA during 1941 : 4
Number of attacks in AUSTRALIA during 1942 : 8
Number of attacks in AUSTRALIA during 1943 : 2
.....
```