Analytical Programming

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 162

(ii)

Read the shark attack dataset into a Pandas Dataframe.

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

```
USA              2160
AUSTRALIA        1303
SOUTH AFRICA      571
PAPUA NEW GUINEA  133
NEW ZEALAND       126
BRAZIL            103
```

(iii)

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

```
AUSTRALIA          342
USA                250
SOUTH AFRICA       137
PAPUA NEW GUINEA    56
MEXICO              44
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  931
Numbers of attack when Scuba Diving  74
```

Question 2.

(i)

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

```
Percentage of attacks that are:
6.11
```

(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.

```
The percentage of fatal attacks:  AUSTRALIA 27.01421800947867
The percentage of fatal attacks:  USA 11.742602160638798
The percentage of fatal attacks:  UNITED KINGDOM 18.181818181818183
The percentage of fatal attacks:  BAHAMAS 11.881188118811881
The percentage of fatal attacks:  MEXICO 53.01204819277108
The percentage of fatal attacks:  SOUTH AFRICA 24.29078014184397
The percentage of fatal attacks:  REUNION 49.152542372881356
The percentage of fatal attacks:  NEW ZEALAND 22.764227642276424
The percentage of fatal attacks:  BRAZIL 39.603960396039604
……
```

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 1931.0  is  9
------
Number of attacks in  AUSTRALIA  during 1930.0  is  10
------
Number of attacks in  AUSTRALIA  during 1929.0  is  24
------
Number of attacks in  AUSTRALIA  during 1928.0  is  9
------
Number of attacks in  AUSTRALIA  during 1927.0  is  11
------
Number of attacks in  AUSTRALIA  during 1926.0  is  5
------
Number of attacks in  AUSTRALIA  during 1925.0  is  4
------
Number of attacks in  AUSTRALIA  during 1924.0  is  6


……
```