Numerical Analysis Exercise Pandas – Shark Attack Dataset:

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

Attribute Information:

The attributes recorded in the dataset are as follows:

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

Open this file using Pandas read_csv() function. The data file is stored in a different encoding format so you can use the following line to read the data into a dataframe.

df = pd.read_csv('attacks.csv', encoding = "ISO-8859-1")

(i)
What location globally has the highest number of shark attacks? (ii)

Read the shark attack dataset into a Pandas Dataframe.
Determine the six countries that have experienced the highest number of shark attacks.

(iii)

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

(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".

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

(vi)

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. Your code should only consider countries where the number of non-fatal and fatal attacks are greater than 0.