
Karina, Jules, Pollyana, Salim

Shark Attacks QUEST

Team: White Sharks



Summary

1. Introduction: business context and goals, hypothesis
2. Original dataset presentation
3. Cleaning steps
4. Cleaned dataset presentation
5. Findings and conclusions

Business opportunities (*KPSJ Inc.*)

As an insurance company, we want to analyze shark attack statistics in order to evaluate and assess the incidence rates based on various features and factors.

Business goal is to develop a fair method to determine insurance rates for shark attack coverage using data from shark attack incidents.

Hypothesis:

The easiest way to get attacked by a shark is by a certain type of activity

Shark attacks occur more frequently in certain countries

Examining the [Shark Attack dataset](#), understanding its structure and formulating a hypothesis or several hypotheses about the data

Presenting the original data

The dataset consists in :

- **6947** Observations (**rows**)
- **23** Features (**columns**)

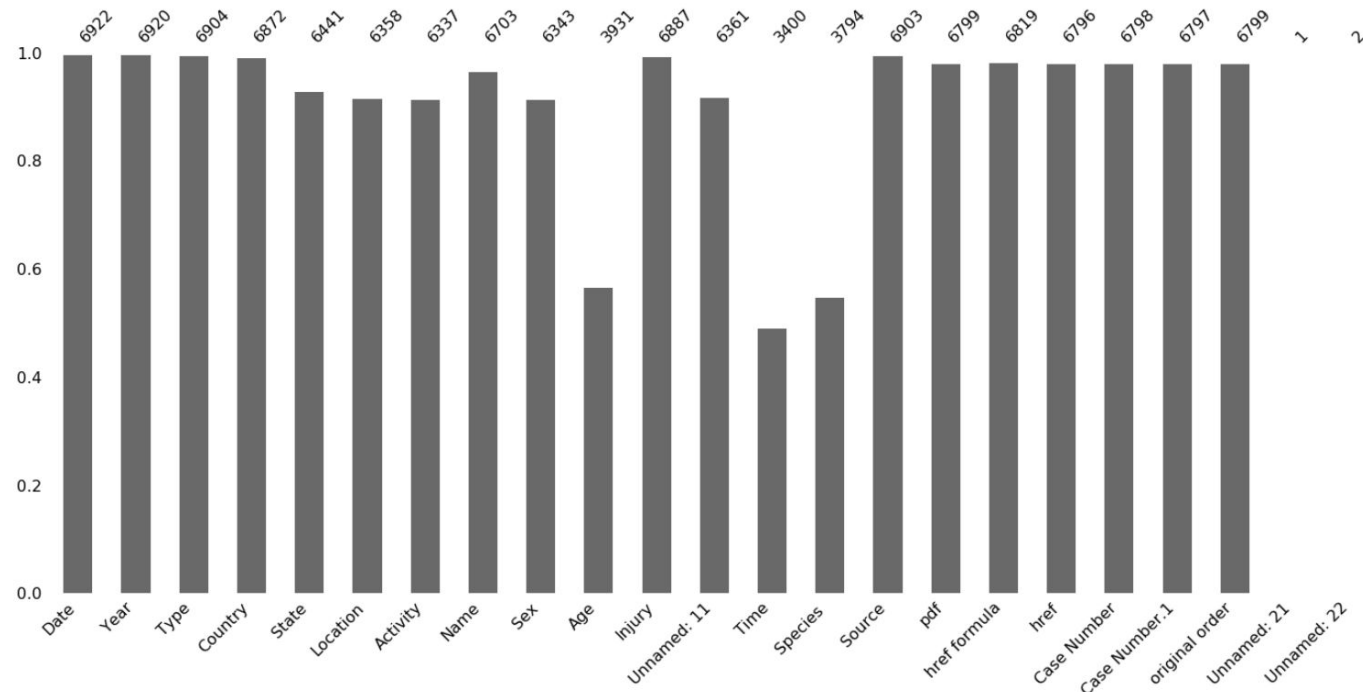
Date	5983
Year	258
Type	11
Country	224
State	896
Location	4495
Activity	1585
Name	5669
Sex	9
Age	243
Injury	4071
Unnamed: 11	12
Time	409
Species	1671
Source	5284
pdf	6789
href formula	6785
href	6776
Case Number	6777
Case Number.1	6775
original order	6797
Unnamed: 21	1
Unnamed: 22	2

Illustrative examples of the original dataset

	Date	Year	Type	Country	State	Location	Activity	Name	Sex	Age	...	Species	Source	pdf	nner formula	href	Case Number.1	original order	Unnamed: 21	Unnamed: 22
0	08 Dec-2023	2023.0	Unprovoked	AUSTRALIA	Queensland	1770	Swimming	malle	20s	NaN	...	NaN	B. Myatt, GSAF	NaN	NaN	NaN	NaN	NaN	NaN	NaN
1	04 Dec-2023	2023.0	Unprovoked	BAHAMAS	New Providence Isoad	Sandals Resort	Paddle boarding	Lauren Erickson Van Wart	F	44	...	NaN	NBC News, 12/4/2023	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	02 Dec-2023	2023.0	Unprovoked	MEXICO	Jalisco	San Patricio Melaque	Swimming	Maria Fernandez Martinez Jimenez	F	26	...	NaN	News Channel 21, 12/3.2023	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	30 Nov-2023	2023.0	Unprovoked	AUSTRALIA	Queensland	Clack Island	Swimming	Matthew Davitt	M	21	...	1.8m bull shark	ABC Net, 11/30/2023	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	21 Nov-2023	2023.0	Unprovoked	BAHAMAS	Grand Bahama Island	Tiger Beach	Scuba diving	female	F	47	...	NaN	Eye Witness News, 11/22/2023	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Missing values

Date	0.359868
Year	0.388657
Type	0.618972
Country	1.079603
State	7.283720
Location	8.478480
Activity	8.780769
Name	3.512307
Sex	8.694400
Age	43.414423
Injury	0.863682
Unnamed: 11	8.435296
Time	51.058011
Species	45.386498
Source	0.633367
pdf	2.130416
href formula	1.842522
href	2.173600
Case Number	2.144811
Case Number.1	2.159205
original order	2.130416
Unnamed: 21	99.985605
Unnamed: 22	99.971211



Cleaning steps

Removing: unnecessary/empty/ unusable columns and rows; duplicates

Formatting: using lowercase letters and grouping similar words or phrases with different spelling in the same group. Example:

Column Sex:

Before cleaning: 8 unique values

After cleaning: 3 unique values

Column Activities:

Before: 1585 values

After: 776 values

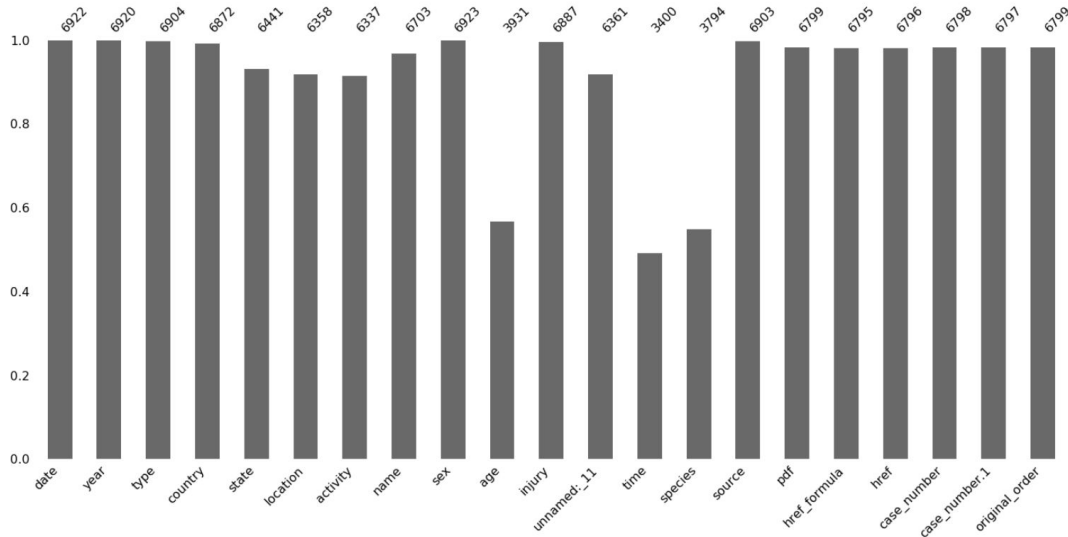
Cleaned Dataset

The dataset now consists in :

- **6923** Observations (**rows**)
- **21** Features (**columns**)

Sex:

M	5565
F	773
unknown	609

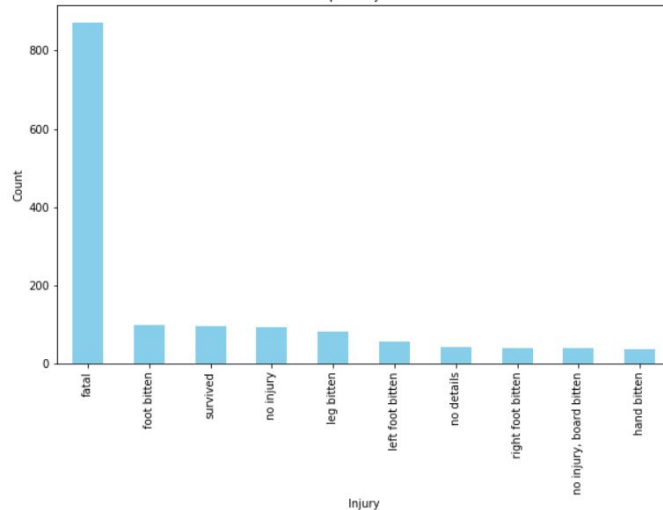


Date	5983	date	5983
Year	258	year	258
Type	11	type	6
Country	224	country	224
State	896	state	896
Location	4495	location	4495
Activity	1585	activity	799
Name	5669	name	5669
Sex	9	sex	3
Age	243	age	243
Injury	4071	injury	3981
Unnamed: 11	12	unnamed:_11	12
Time	409	time	409
Species	1671	species	1671
Source	5284	source	5284
pdf	6789	pdf	6789
href formula	6785	href_formula	6785
href	6776	href	6776
Case Number	6777	case_number	6777
Case Number.1	6775	case_number.1	6775
original_order	6797	original_order	6797
Unnamed: 21	1		
Unnamed: 22	2		

Findings and Conclusions

Top 10 Shark Attack Injuries

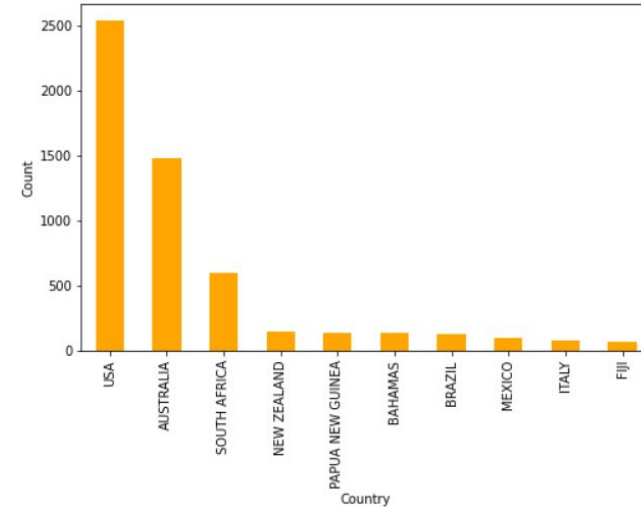
Injuries: 6887



Top 10 Countries with Shark Incident

Volusia County: 352

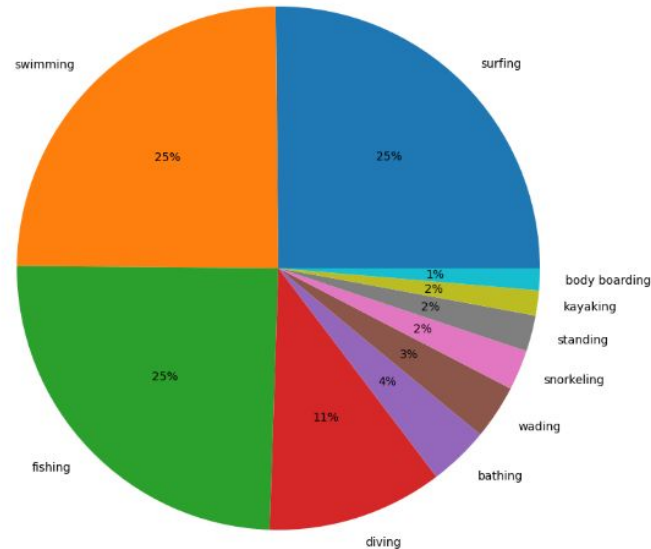
Brevard County: 100



Findings and Conclusions

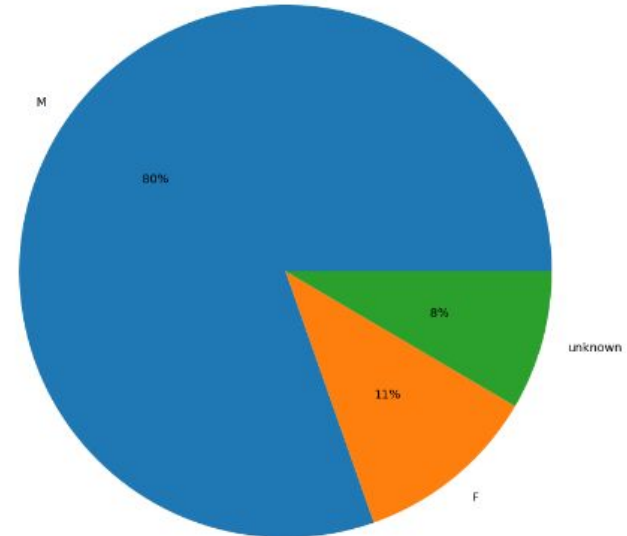
Top 10 Activities related to Shark Attacks

75% of reported shark attacks are related to one of the activities: swimming, surfing or fishing.



Gender x Reported Shark Attacks

80% of attacks involves men.



Karina, Jules, Pollyana, Salim

THANK YOU!

QUESTIONS?

