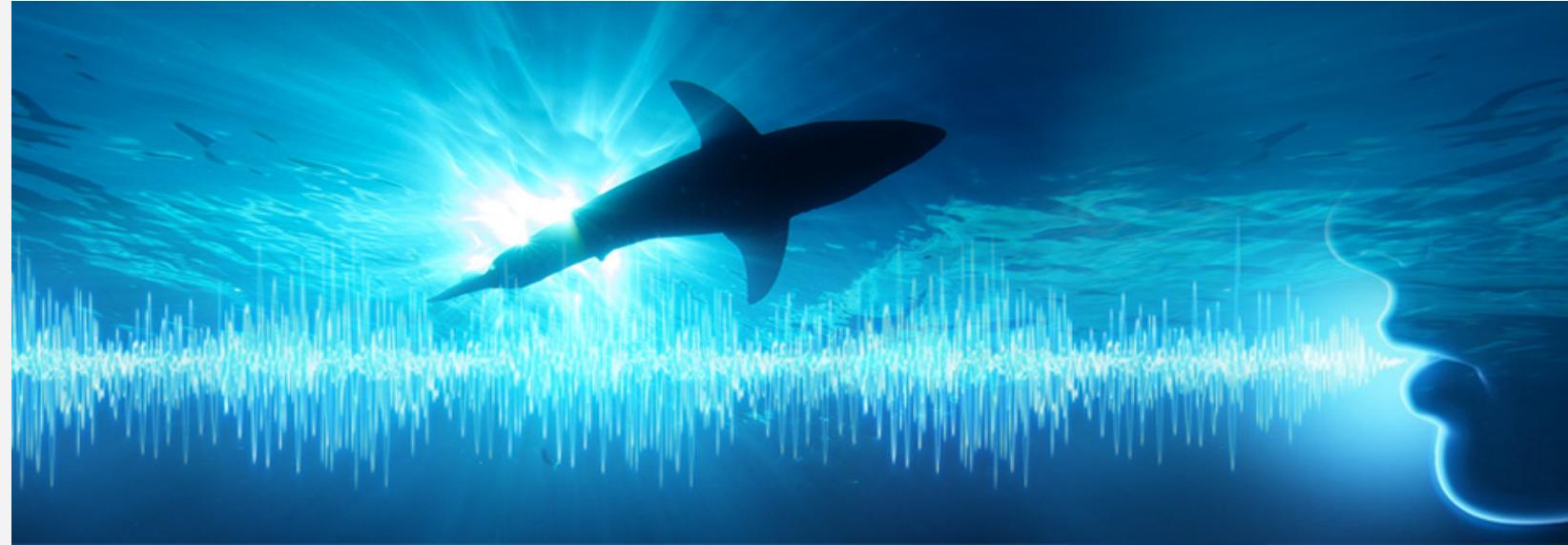




DALHOUSIE
UNIVERSITY

FACULTY OF
COMPUTER SCIENCE



Analysis of Global Wildlife Trade Data Using Named Entity Recognition to Extract Top Reason for Sharks Being Endangered Species

Koustav Pain (Tukai) (B00 795292)

Why Sharks Are Endangered?

Sharks have been swimming in Earth's oceans since before the time of the dinosaurs.

In recent years Sharks have faced rapid population declines and most species are considered in danger of extinction.

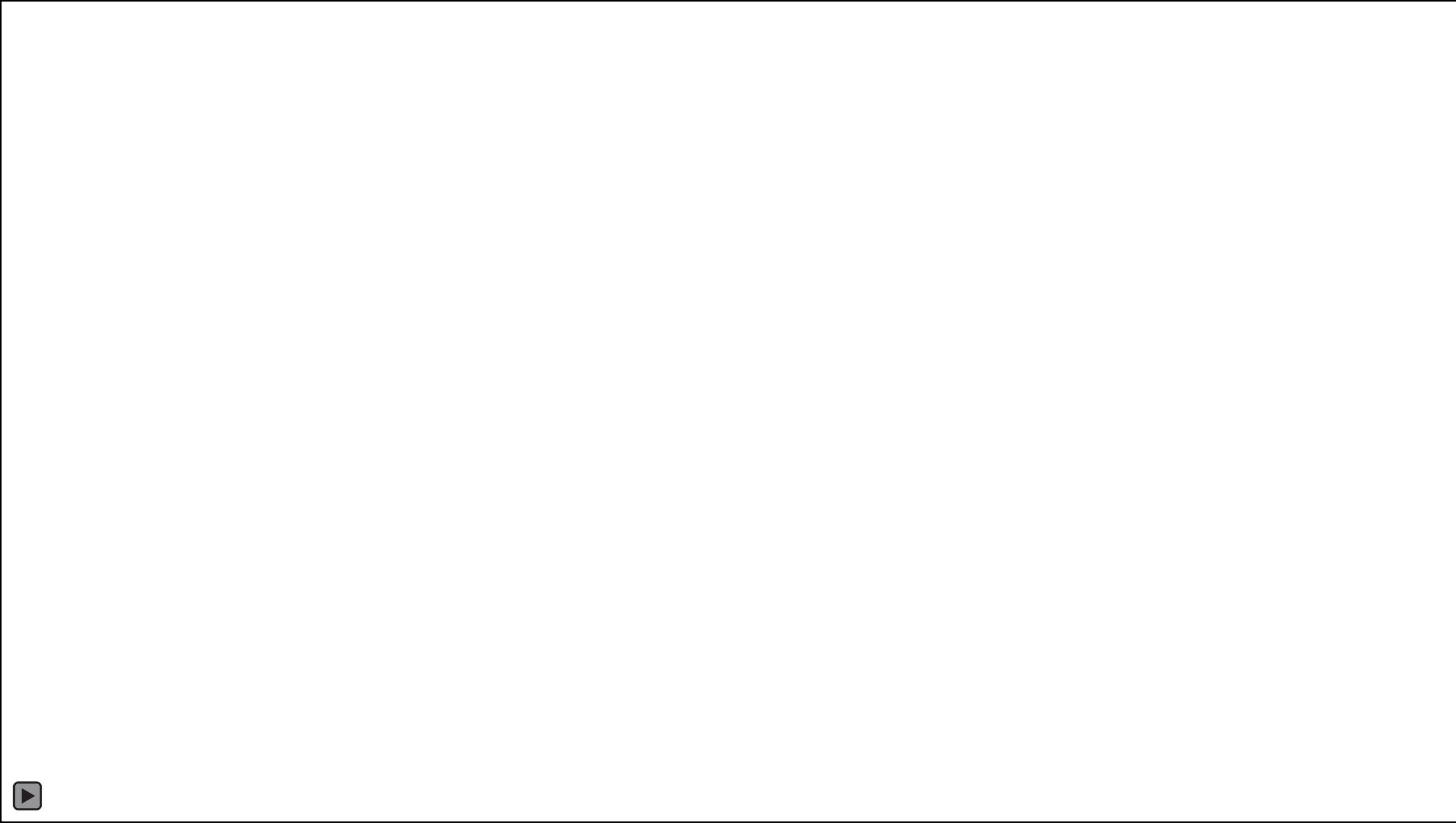
Sharks are being over-fished to satisfy demand of shark fin soup, Chinese tradition that dates back to the second century B.C.

This leads to the Shark Finning practice.



Shark Finning

Shark finning is a cruel practice where after fishing up a shark, it's killed, their fins are cut off and then whole body thrown back into the sea.



Research Questions

- How to obtain only shark products trade data from global wildlife trade data?
- How to obtain illegal shark products trade data?
- Which countries have the most shark fin trade and does the trade affect the population of the shark to the point of being endangered?

How to obtain only shark products trade data from global wildlife trade data?

Hypothesis: Named Entity Recognition (**NER**) trained with full shark taxonomy has higher **F-score** than string matching algorithm to find only shark trade datasets from **CITES** global wildlife trade dataset.



Software > Stanford Named Entity Recognizer (NER)

Stanford NER is a Java implementation of a Named Entity Recognizer. Named Entity Recognition (NER) labels sequences of words in a text which are the names of things, such as person and company names, or gene and protein names. It comes with well-engineered feature extractors for Named Entity Recognition, and many options for defining feature extractors.

Shark Taxonomy

🔗 <https://en.wikipedia.org/wiki/Shark>

Taxonomy

Sharks belong to the superorder Selachimorpha in the subclass Elasmobranchii in the class Chondrichthyes. The Elasmobranchii also include rays and skates; the Chondrichthyes also include Chimaeras. It was thought that the sharks form a polyphyletic group: some sharks are more closely related to rays than they are to some other sharks,^[21] but current molecular studies support monophyly of both groups of sharks and batoids.^{[22][23]}

The superorder Selachimorpha is divided into Galea (or Galeomorphii), and Squalea (or Squalomorphii). The Galeans are the Heterodontiformes, Orectolobiformes, Lamniformes, and Carcharhiniformes. Lamnoids and Carcharhinoids are usually placed in one clade, but recent studies show the Lamnoids and Orectoloboids are a clade. Some scientists now think that Heterodontoids may be Squalean. The Squalceans are divided into Hexanchiformes and Squalomorpha. The former includes cow shark and frilled shark, though some authors propose both families to be moved to separate orders. The Squalomorpha contains the Squaliformes and the Hypnosqualea. The Hypnosqualea may be invalid. It includes the Squatiniformes, and the Pistorajea, which may also be invalid, but includes the Pristiophoriformes and the Batoidea.^{[21][24]}

There are more than 470 species of sharks split across twelve orders, including four orders of sharks that have gone extinct:^[24]

- Carcharhiniformes: Commonly known as ground sharks, the order includes the blue, tiger, bull, grey reef, blacktip reef, Caribbean reef, blacktail reef, whitetip reef, and oceanic whitetip sharks (collectively called the requiem sharks) along with the houndsharks, catsharks, and hammerhead sharks. They are distinguished by an elongated snout and a nictitating membrane which protects the eyes during an attack.
- Heterodontiformes: They are generally referred to as the bullhead or horn sharks.
- Hexanchiformes: Examples from this group include the cow sharks and frilled sharks, which somewhat resembles a marine snake.
- Lamniformes: They are commonly known as the mackerel sharks. They include the goblin shark, basking shark, megamouth shark, the thresher sharks, shortfin and longfin mako sharks, and great white shark. They are distinguished by their large jaws and ovoviparous reproduction. The Lamniformes also include the extinct megalodon, *Carcharodon megalodon*.
- Orectolobiformes: They are commonly referred to as the carpet sharks, including zebra sharks, nurse sharks, wobbegongs, and the whale shark.
- Pristiophoriformes: These are the sawsharks, with an elongated, toothed snout that they use for slashing their prey.
- Squaliformes: This group includes the dogfish sharks and roughsharks.
- Squatiniformes: Also known as angel sharks, they are flattened sharks with a strong resemblance to stingrays and skates.
- † Cladoselachiformes
- † Hybodontiformes
- † Symmoriida
- † Xenacanthida (Xenacantiformes)





CITES (the Convention on International Trade in Endangered Species of Wild Fauna and Flora) is an international agreement between governments. Its aim is to ensure that international trade in specimens of wild animals and plants does not threaten their survival.

We are using data from 2005 to 2018

The screenshot shows the official website for the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES). The header features the CITES logo on the left, a red bar with a lock icon and the URL 'https://trade.cites.org' in the center, and language links 'Espanol' and 'Francais' on the right. Below the header, there's a large image of two rhinoceroses in a grassy field. The main content area includes the text 'Convention on International Trade in Endangered Species of Wild Fauna and Flora' and a call-to-action button 'Download here' with a file size of '237MB'.

Full CITES Trade Database download available

[Download here](#)
File size: 237MB

F1-score

- The reliability and validity are measured in NLP data mining by a method called F1 score. A good F1 score means that you have low false positives and low false negatives, so you're correctly identifying real threats and you are not disturbed by false alarms. An F1 score is considered perfect when it's 1, while the model is a total failure when it's 0.
- $\text{F1 score} = 2 \times \text{Precision} \times \text{Recall} / (\text{Precision} + \text{Recall})$
- Where, $\text{Precision} = \text{True positives} / (\text{True positives} + \text{False positives})$
- And $\text{Recall} = \text{True positives} / (\text{True positives} + \text{False negatives})$
- A true positive is an outcome where the model correctly predicts the positive class. Similarly, a true negative is an outcome where the model correctly predicts the negative class.
- A false positive is an outcome where the model incorrectly predicts the positive class. And a false negative is an outcome where the model incorrectly predicts the negative class.

How to obtain illegal shark fin trade data?

Hypothesis:

Spacy performs with a better F-score than Stanford NER and Polyglot to extract illegal shark fin trade information from news articles.

News articles: Nat geo, Animal traffic – NY Times.



- Non-destructive tokenization
- Named entity recognition
- Support for 51+ languages
- 16 statistical models for 9 languages
- Pre-trained word vectors
- State-of-the-art speed
- Easy deep learning integration
- Part-of-speech tagging
- Labelled dependency parsing
- Syntax-driven sentence segmentation
- Built in visualizers for syntax and NER
- Convenient string-to-hash mapping
- Export to numpy data arrays
- Efficient binary serialization
- Easy model packaging and deployment
- Robust, rigorously evaluated accuracy

Which countries have the most shark fin trade and does the trade affect the population of the shark to the point of being endangered?

Hypothesis correlation :

Increasing global shark fin trade affects global shark population

Rapid finning i.e. shark death ratio > shark reproduction

Lazy finning i.e. shark death ratio < shark reproduction rate

IUCN Redlist

https://www.iucnredlist.org



More than 28,000 species are threatened with extinction

That is 27% of all assessed species.

AMPHIBIANS

40%



MAMMALS

25%



CONIFERS

34%



BIRDS

14%



SHARKS &
RAYS

30%



REEF CORALS

33%



SELECTED CRUSTACEANS

27%



feedback

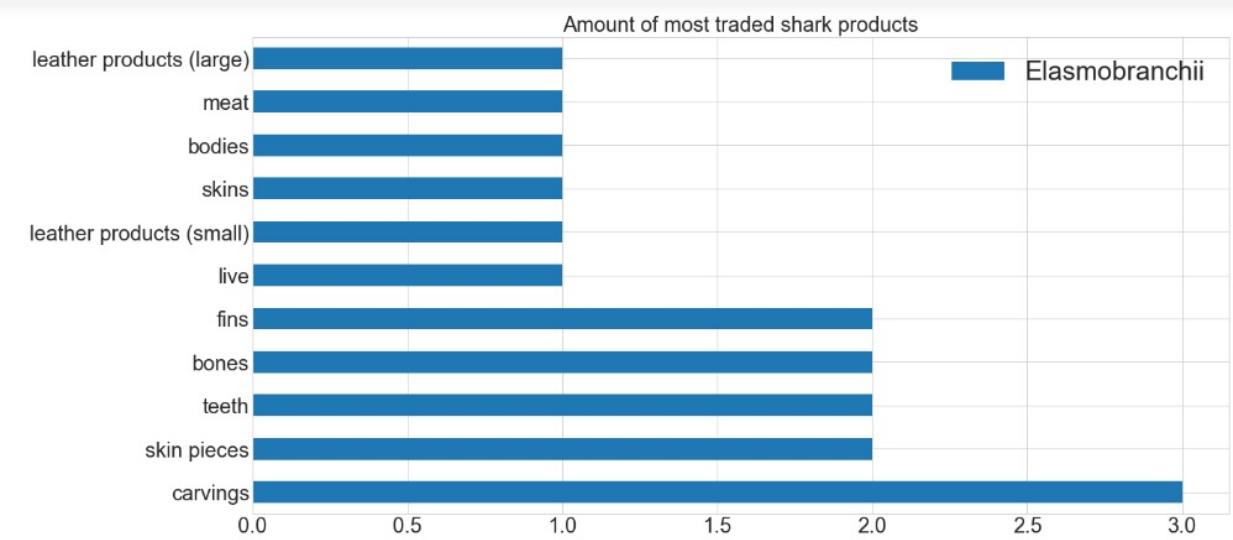
Take action

Help us make The IUCN Red List a more complete barometer of life.

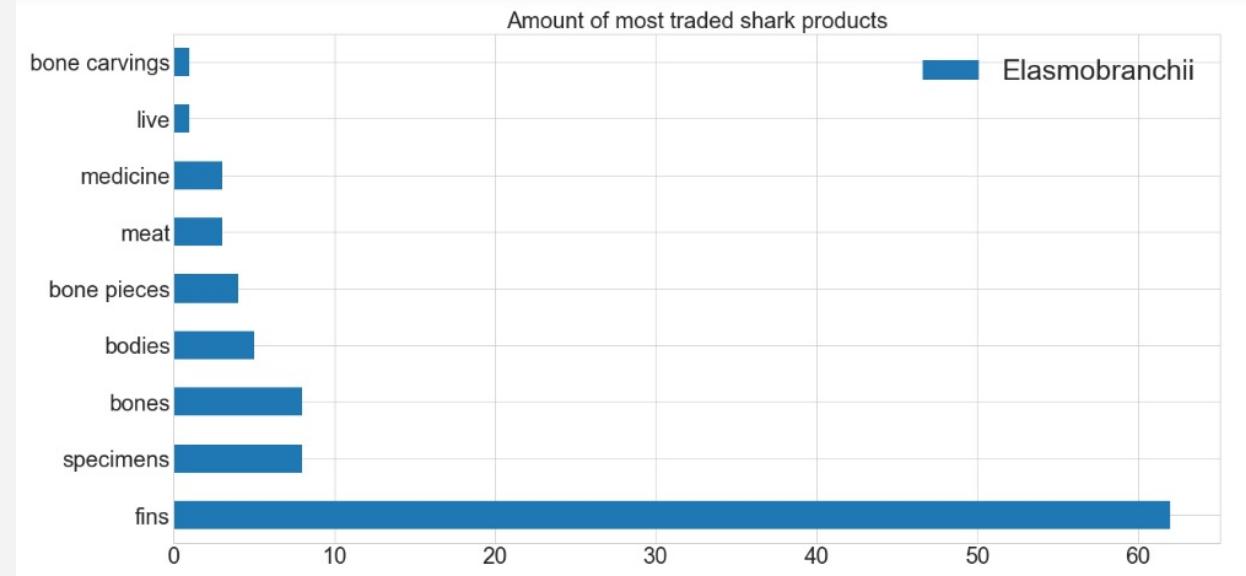
Visualization of Results



Shark Products Comparison

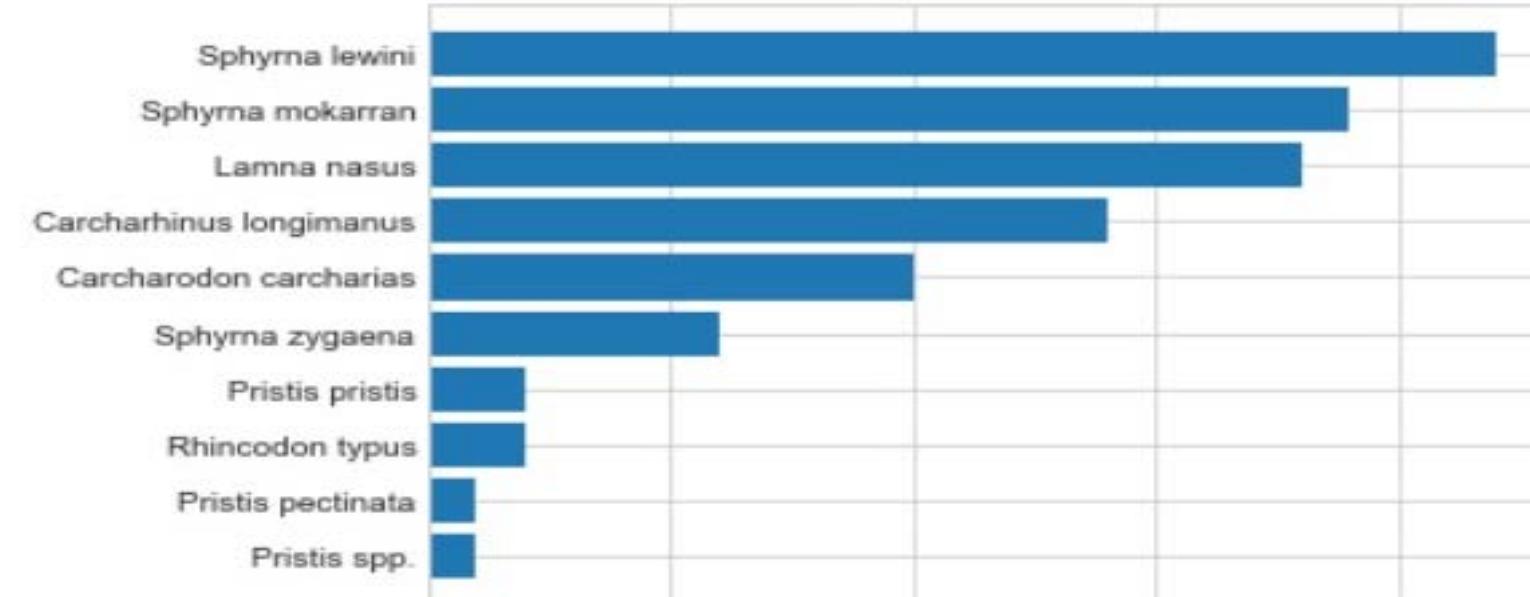


2005



2016 -2017

Top Traded Taxons



THE IUCN RED LIST
OF THREATENED SPECIES™



ANIMALIA - CHONDRICHTHYES

Scalloped Hammerhead

Sphyrna lewini

Unknown



ENDANGERED

EN



ANIMALIA - CHONDRICHTHYES

Great Hammerhead

Sphyrna mokarran

Decreasing



ANIMALIA - CHONDRICHTHYES

Porbeagle

Lamna nasus

Decreasing

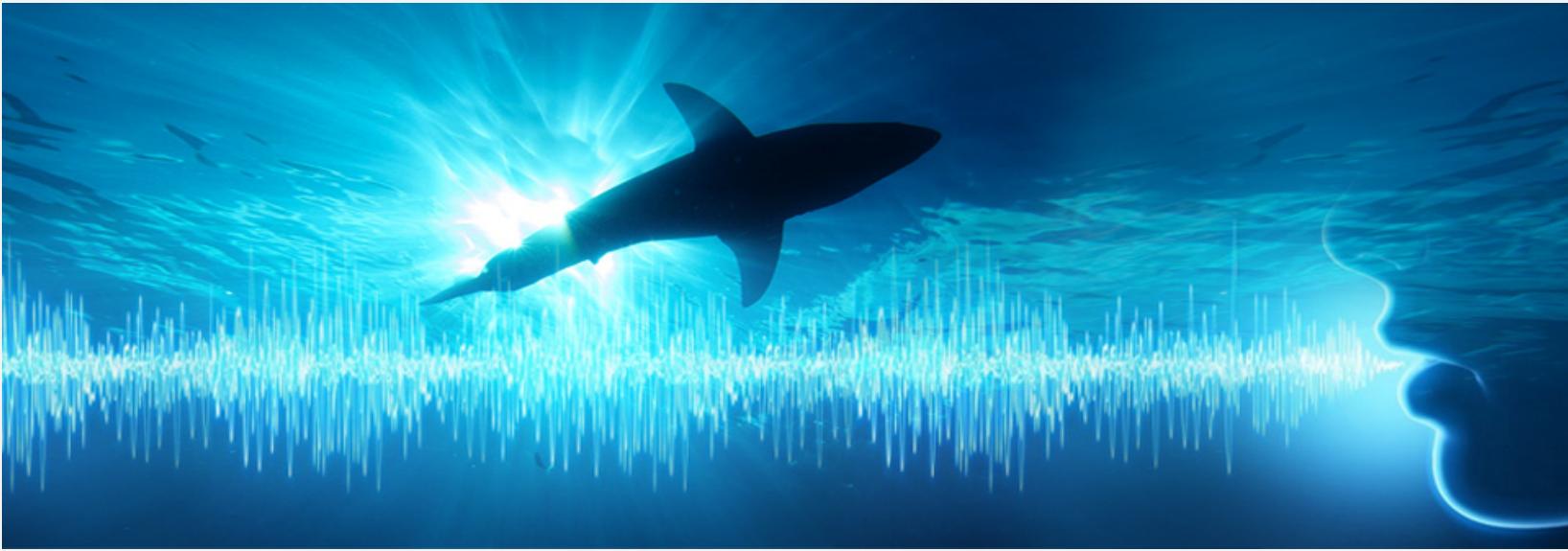


VULNERABLE

VU

Responsibilities

Very important aspect of this project is to consider that it is not to attack one's culture, or way of things such as social or religious activities that involves shark products. Shark finning is one of the most cruel things to do to a free animal of the Ocean and should be highly discouraged. The shark fin soup is just considered a good omen and is not beneficial to the human body any more than a mixed vegetable soup. Sharks are a crucial part of the ocean's ecosystem. As the apex predator, sharks keep the marine food chain in balance. Without sharks other predatory fish and marine animals will thrive, which will deplete the amount of food fishes. Without these food fishes, algae and other plants will thrive causing coral reefs to disappear and making the water toxic for all life.



Thank You!
