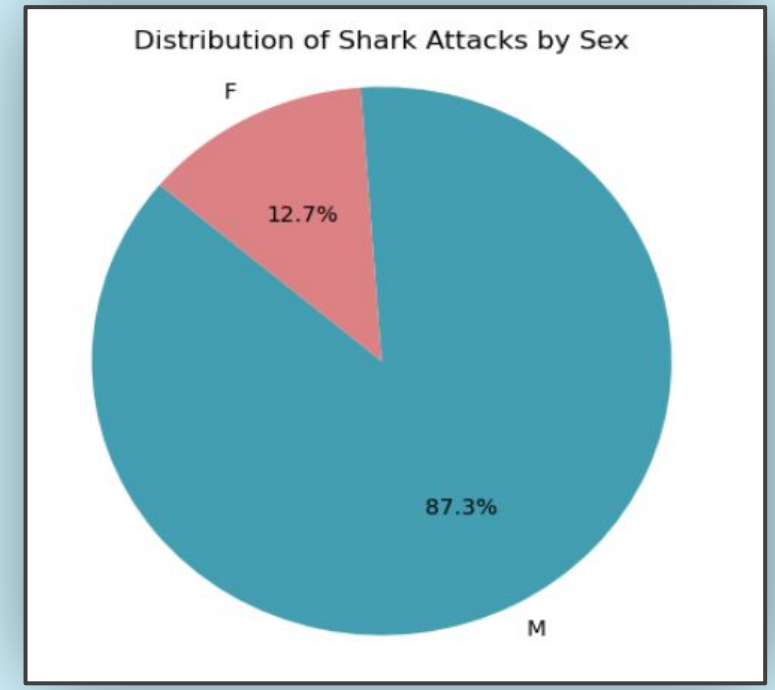
Shark Attacks: A Deeper Dive

Bradley Newell, Grant Hagen, Kelsey Mersinas, and Lindsey Krempa

Data Breakdown and Cleaning

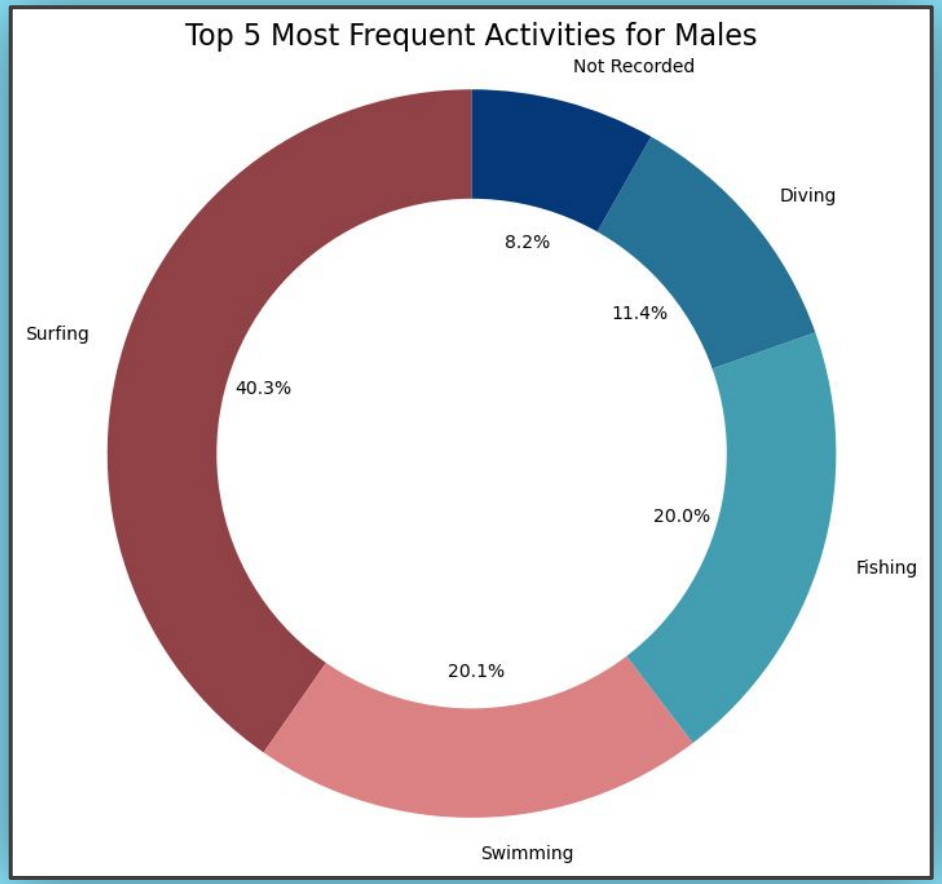
- Turning large diverse data into descriptive categories using dictionaries
- Refining columns down to the usable information
- Creating new columns utilizing already existing data: hemisphere, season, month
- Replacing null values with “Unknown” rather than dropping
- Are there any trends we can glean?

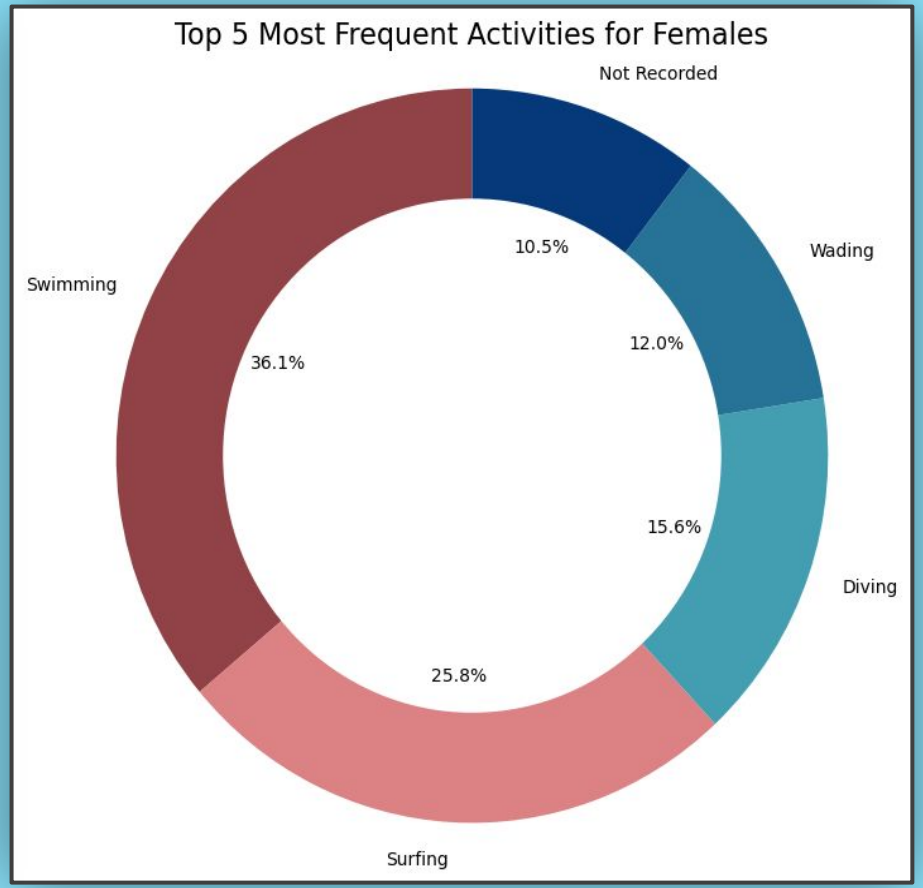
Is there a relationship between gender and activity?



- **Hypothesis:** The number of male victims in shark attacks will be higher than female victims.
- **Conclusion:** The number of male victims were higher, regardless of activity type.
- **Limitations:** some questionable data, i.e. “other”, “unknown”, and “accident” all listed as activity types. Number of males and females in each activity including non victims of attack.
- **Further Learning:** exploring other factors of the dataset while comparing gender, such as injury type and age.

What were the most frequent activities for each gender?

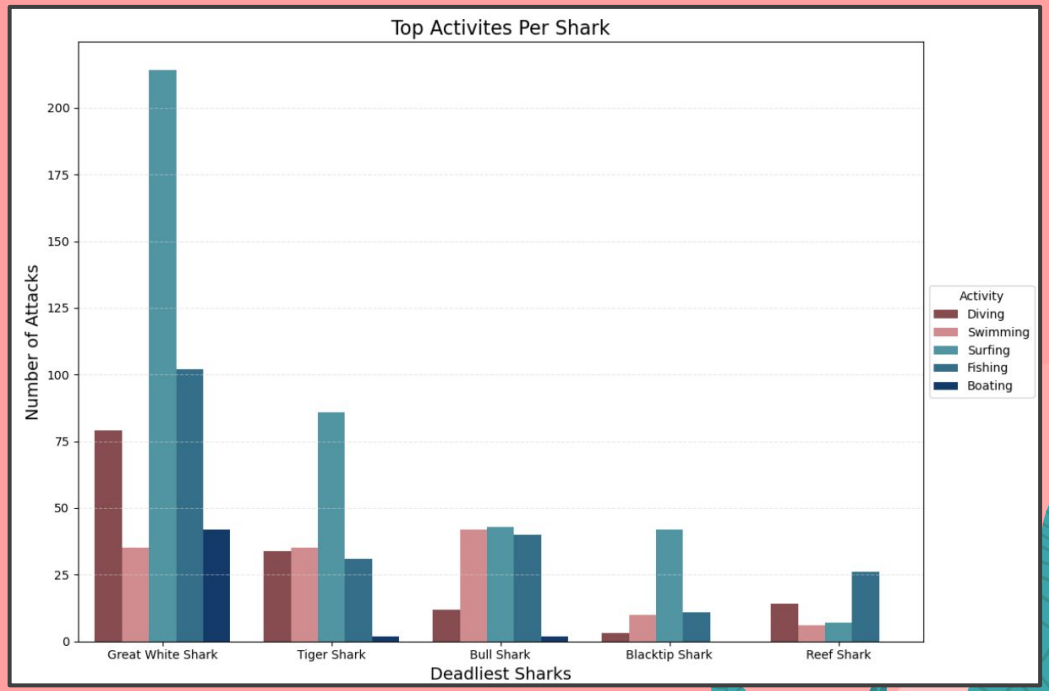




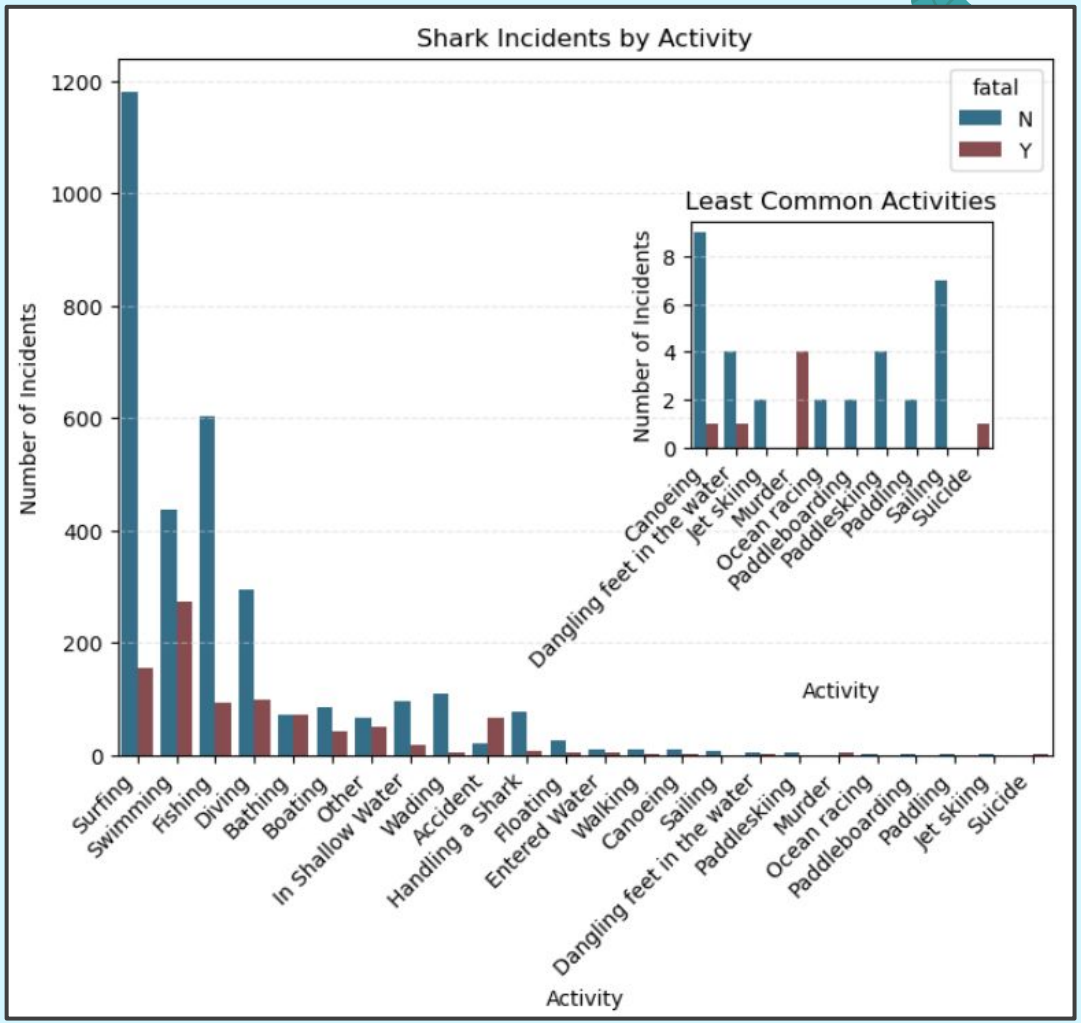
Is there an activity related to shark attacks?

Hypothesis - Surfing and other activities that are further from the shore should represent the most shark attacks of our data

If surfing increases the likelihood of shark attacks it should have the highest fatalities

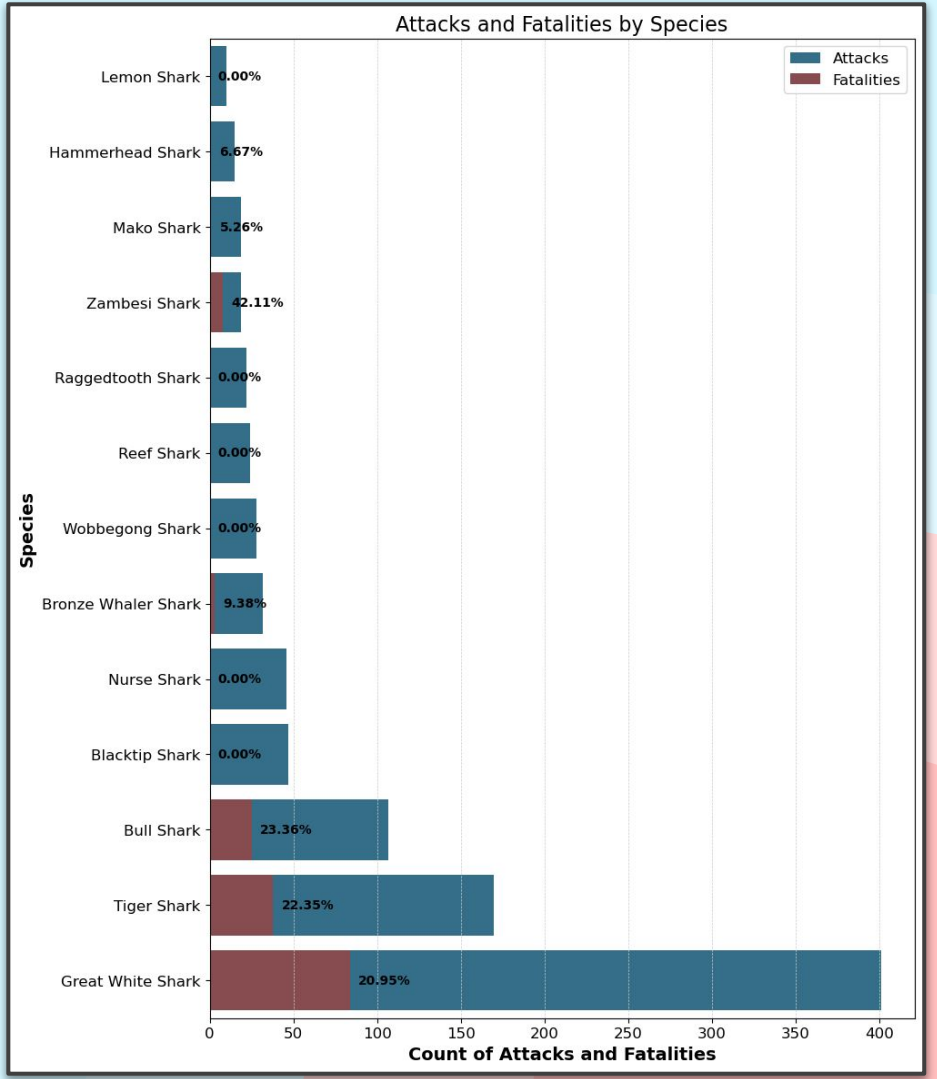


Conclusion: Surfing, swimming and fishing are the top activities related to shark attacks. Swimming had less encounters than surfing but had the most fatal attacks. Majority of activities were non fatal encounters.

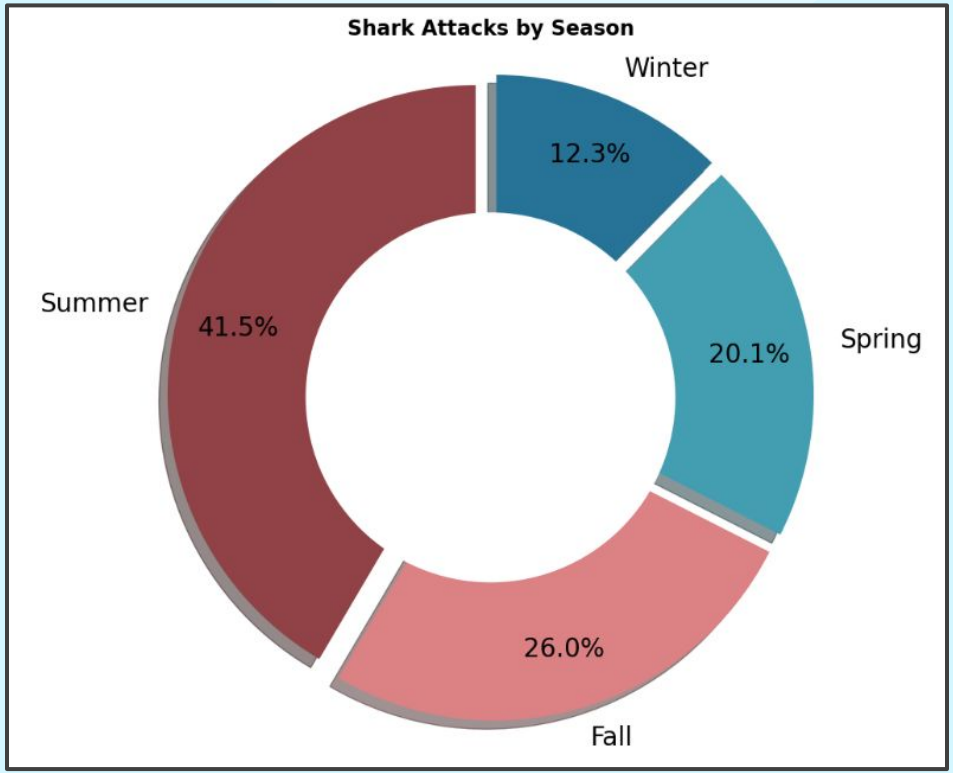


Attacks by Species

- **Hypothesis:** We will see a correlation between species and fatalities. Will the number of attacks relate to number of attacks?
- **Conclusion:** Great White Sharks had the highest frequency of attacks and fatalities. Zambesi Sharks, although less frequently involved in attacks, exhibited a higher fatality rate.
- **Limitations:** Additionally, inaccuracies were observed in certain data points, such as erroneous years extending up to 2060, and columns containing incorrect information.
- **Future Considerations:** Future investigations could include data on shark size, number of bites, and injury locations. This additional information could provide deeper insights into the factors influencing shark attacks and fatalities.
    - Great White Sharks emerged as the species with the highest frequency of attacks.
    - Zambesi Sharks displayed the highest fatality rate, followed by Bull Sharks, Tiger Sharks, and Great White Sharks.

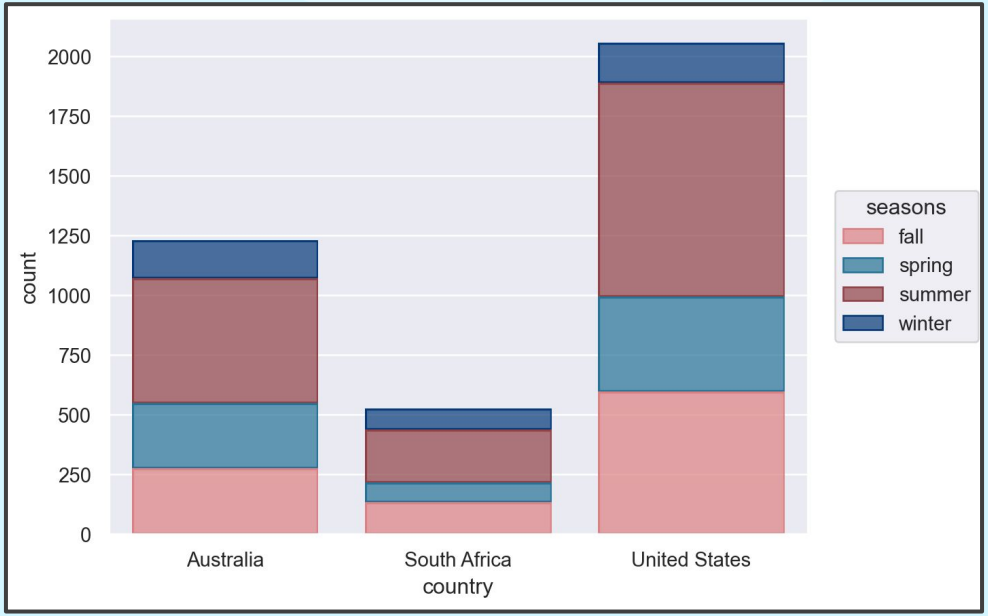


Is there seasonality to shark attacks?



A Global Look at Attacks

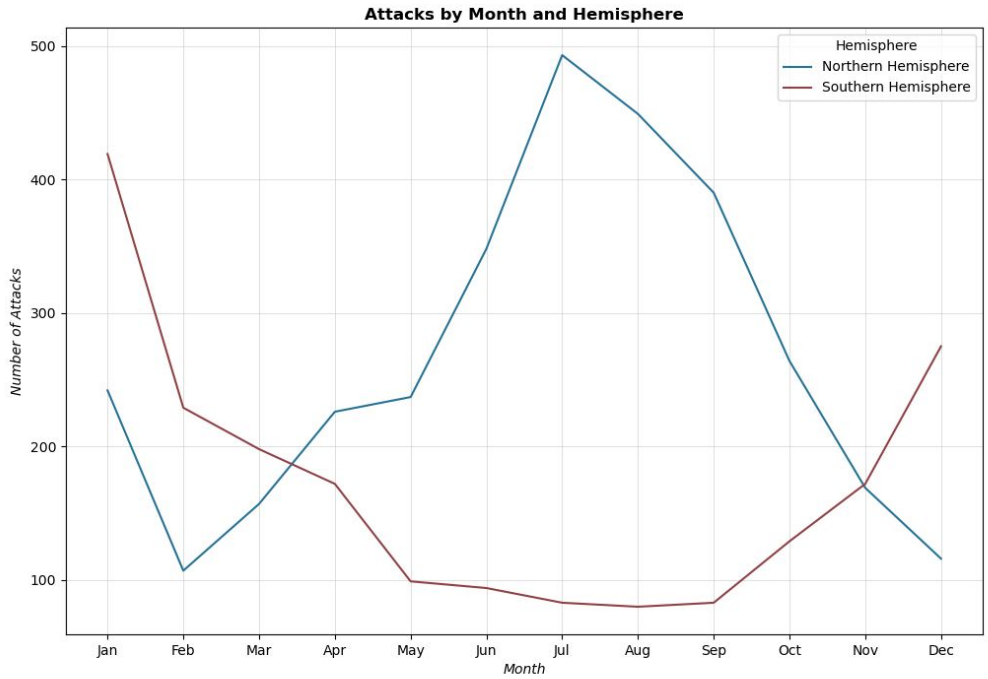
Top Countries Per Season



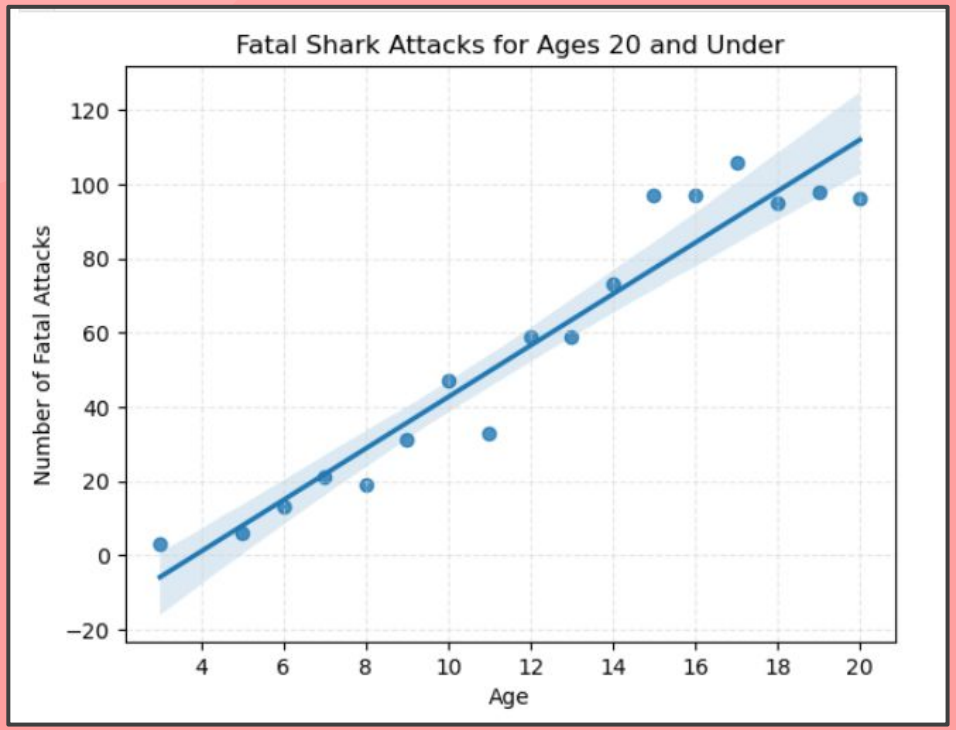
Hemisphere Breakdown

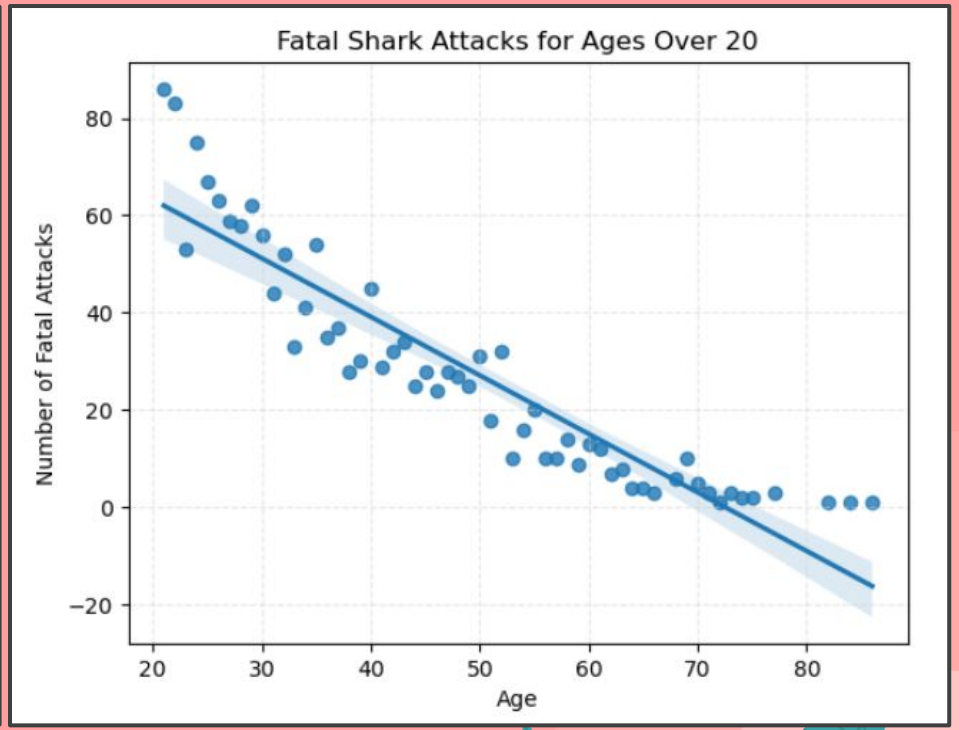
Northern Hemisphere: Climbs from February to its peak in July and declines.

Southern Hemisphere: Climbs September to its peak in January.



Is there a relationship between age and fatality?





Bias, Limitations, and Future Work

- **Bias**
  - **Omitted variables**
  - **Reliance on 3rd part knowledge**
- **Limitations**
  - **Inconsistent data structure**
  - **Many null values**
- **Future Work**
  - **Better structure within the dataset**
  - **Additional columns: lat, lon, hemisphere, weather**

Work cited

Data Set: https://www.kaggle.com/datasets/teajay/global-shark-attacks

Palette: https://www.color-hex.com/color-palette/112421

PowerPoint Theme: https://slidesgo.com/theme/stop-shark-massacre#search-Shark&position-2&results-8

Q & A