original dataset: https://www.sharkattackfile.net/incidentlog.htm

SHARK ATTACKS

ANALYSIS OF GLOBAL DATA SET ON SHARK ATTACKS USING PYTHON

BY CLEMENS, HAYDN & TIMO

Data Wrangling & Cleaning

Input Errors

Missing Data

Formatting

Name Variations

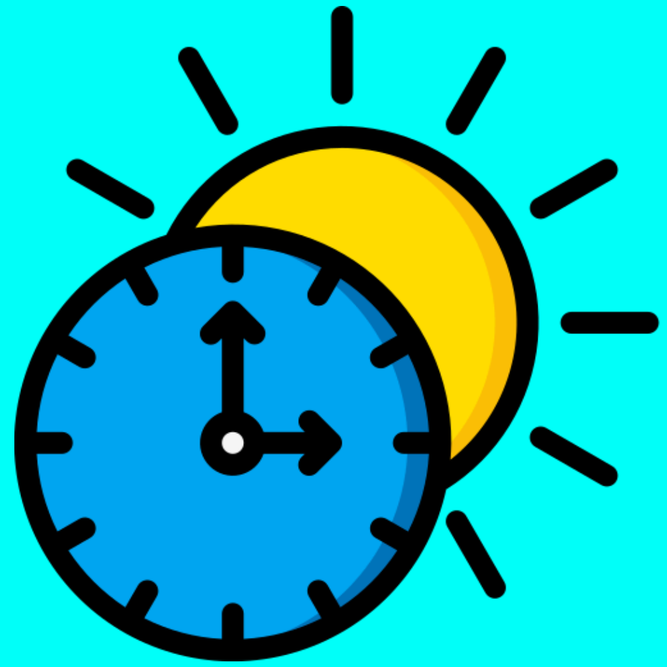
Major Headache

Time & Date Formatting

=

Exploratory Data Analysis

- Are Great White sharks deadlier?
- Are sharks deadlier during certain times of day?

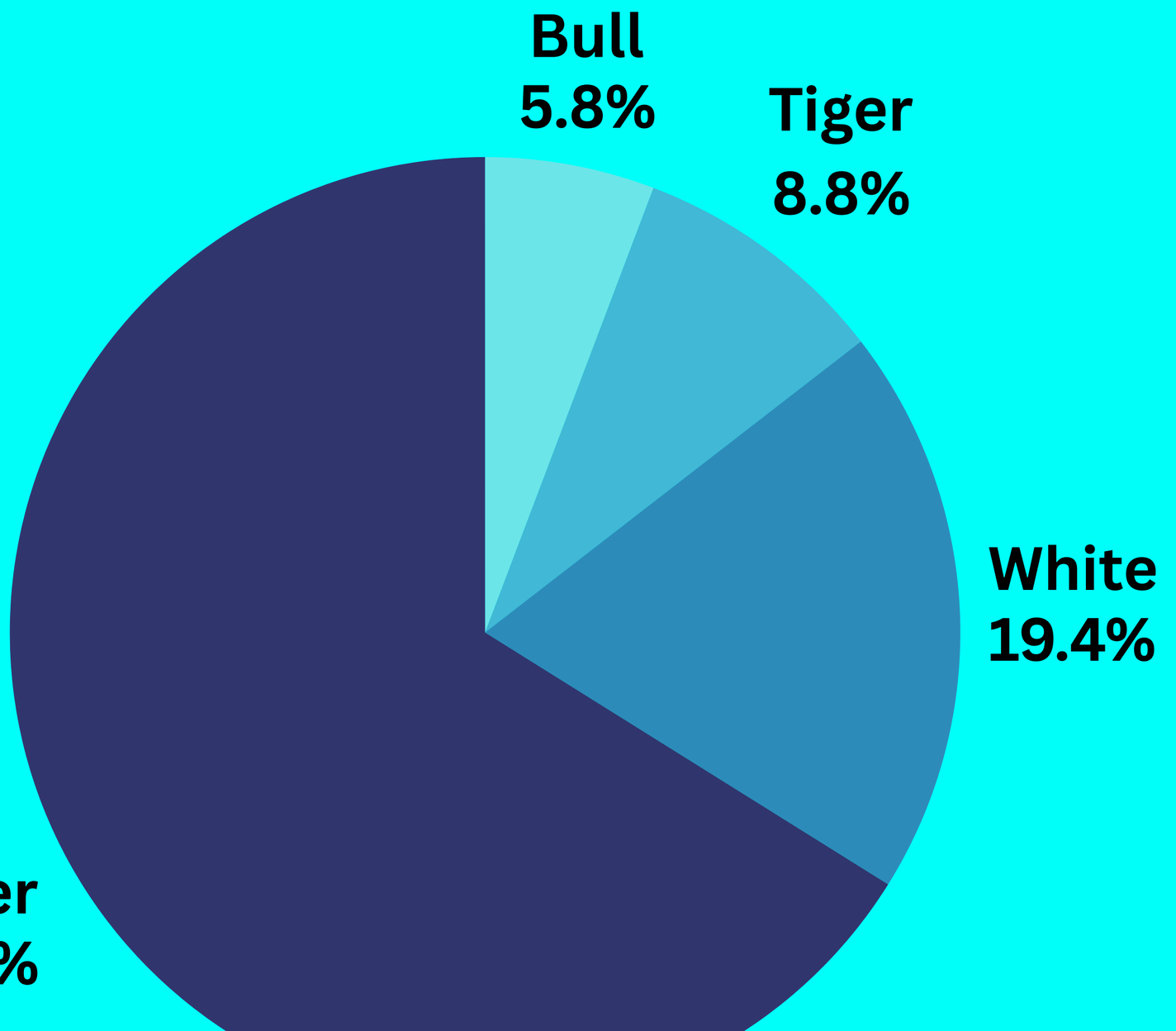


Hypothesis #1:

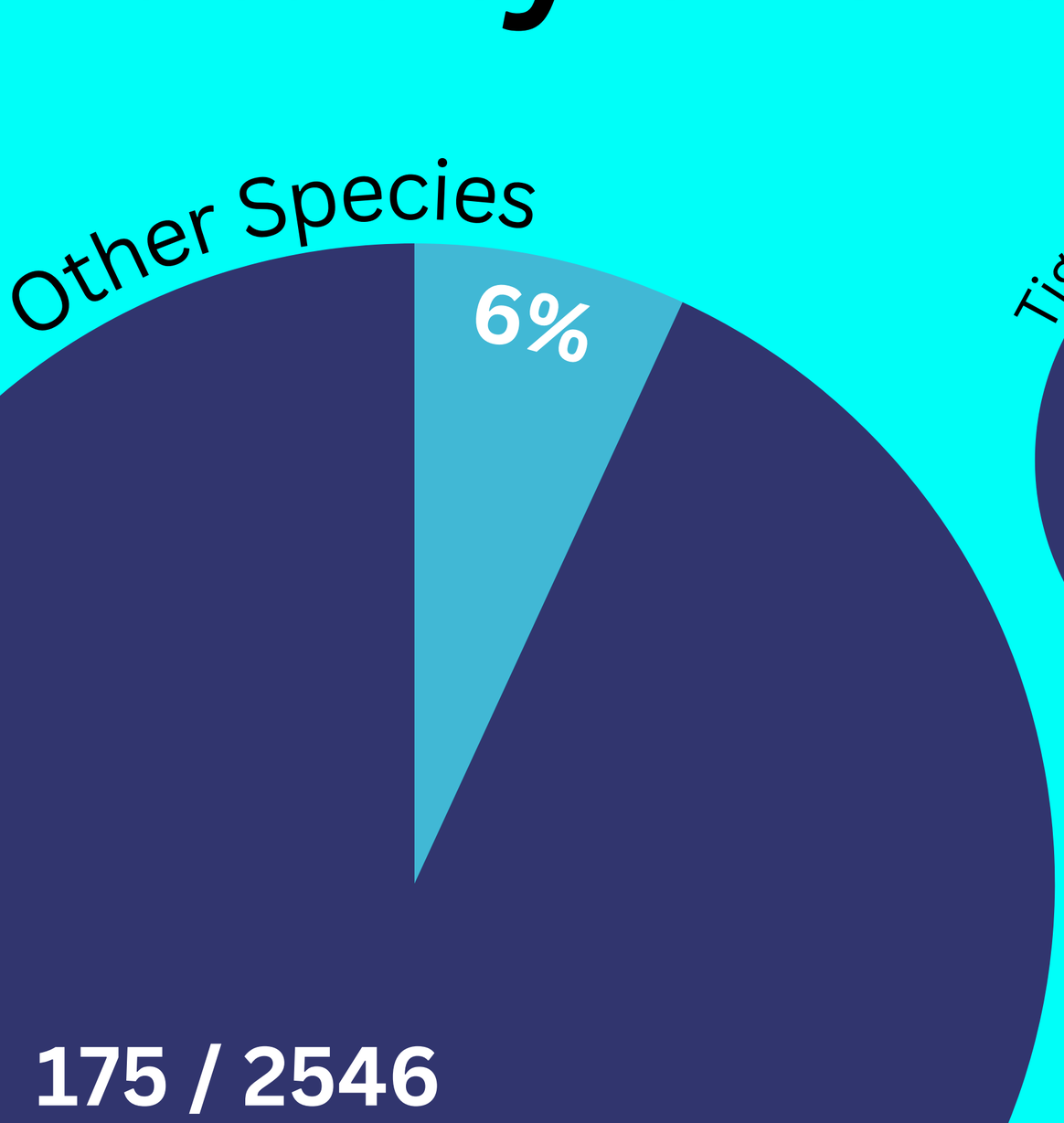
Great White Shark attacks result in highest fatalities.

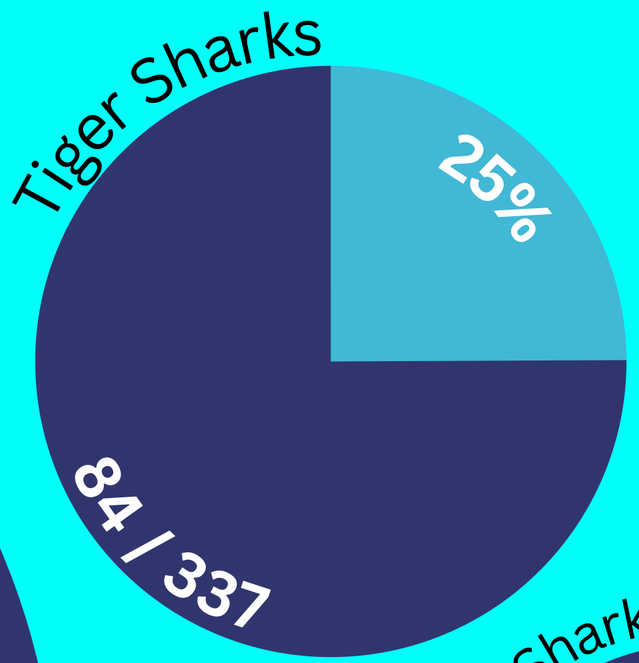
Total Shark Attacks:

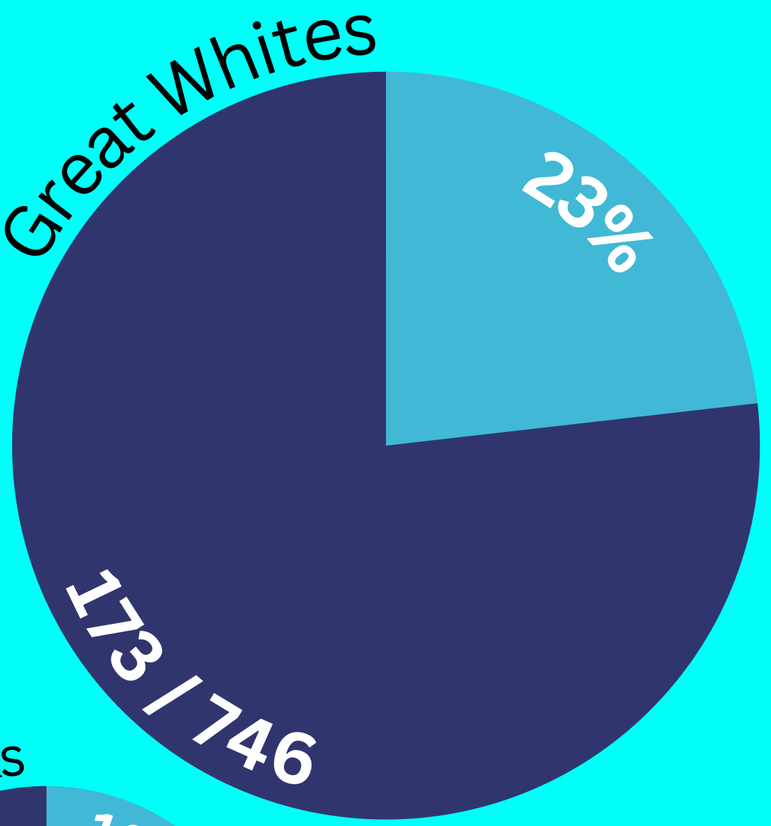
3851

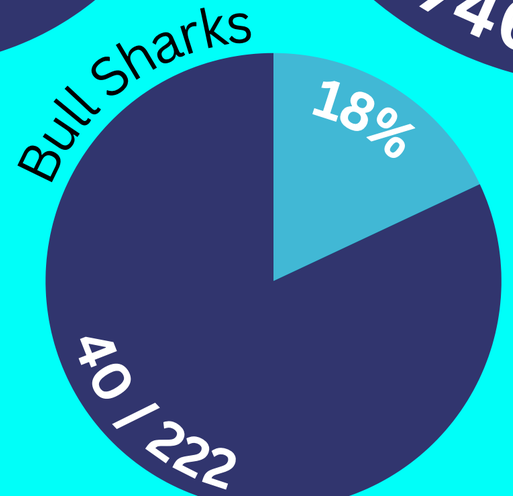


Fatality of Shark Attacks



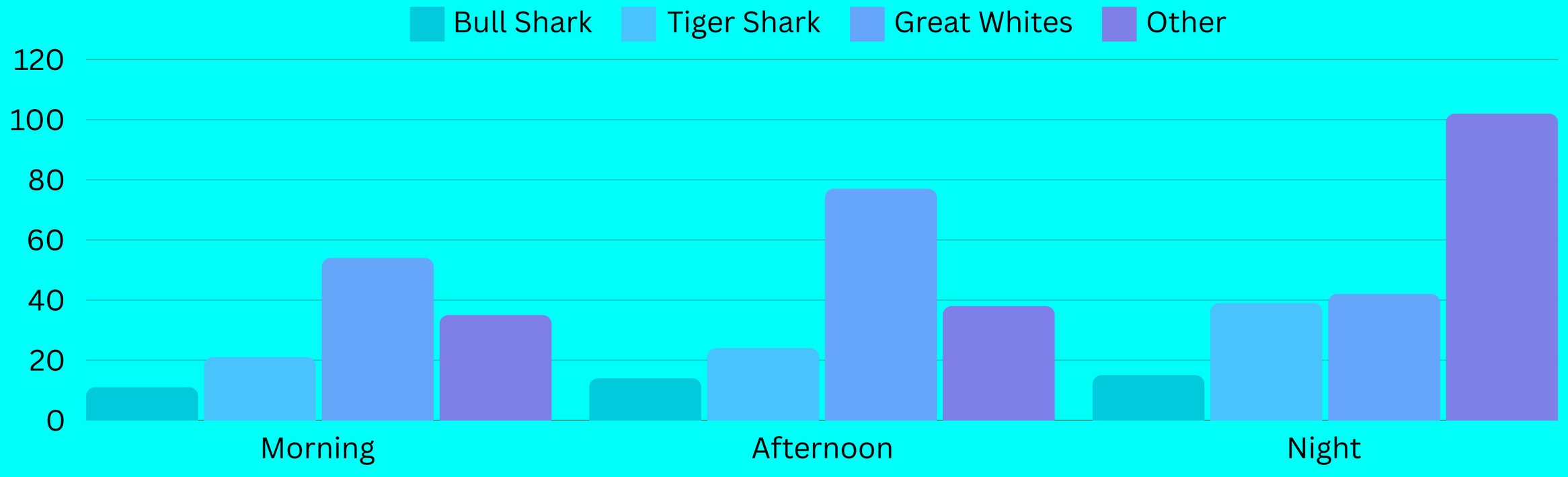






Major Obstacle: Lack of Data

Are there certain times of day that sharks are deadlier? (Hypothesis #2)



Conclusion & Insights

consistency is key in data tables
try to avoid manual entries

multiple data analysts need to sync very well to avoid errors

initial hypothesis #1 was supported
hypothesis #2 could not be analyzed

Thank You