



More

# SUMMER VACATION IDEAS

Beach and Sea Adventures

Presented by: Ng Juat Teng  
8 April 2025



# And how to not get eaten by a shark

Presented by: Ng Juat Teng  
8 April 2025

# Problem

---

**Australia** is the only continent fully surrounded by water and has a booming tourist industry relating to beach and water activities.

Australia is also the second-highest number of shark attacks in the world, with the first being USA.

There is an increasing trend in shark bite incidents over the years. **How can we make it safer for water activity lovers and the sharks as well?**





June is an avid scuba diver who is planning a shark diving adventure in Australia. However, her family and friends are concerned about her safety due to recent reports of shark attacks globally.

June has read that it is safer to scuba dive as compared to swimming or surfing at a beach where sharks are known to be found, and would like find out if that is true.

Derek is an avid surfer from Germany and would like to conquer the waves in Australia one day. He noticed a lot of online articles regarding shark bites on surfers in Australia and is wondering how to achieve his dream while being safe.

He would like to know which beach is safer to go to, and if there is a particular time of the year to avoid when planning for his trip.



# Objective

---

- To reduce shark attack rates in Australia by raising awareness and educating the general public on when to avoid visiting beaches where sharks are known to frequent
- Target Audience:  
Domestic and international beach-goers, surfers and scuba divers

# Methods & Techniques

## Data Cleaning on Excel

- Remove unused columns
- Regroup / reclassification
- Fill blanks
- Correct spelling mistakes

## Data Wrangling and EDA on Python

- Drop outlier rows
- Univariate EDA analysis
- Further grouping

## Visualization on Tableau

- Heat map
- Key trends
- Key insights

## Classification Predictive Model

- Logistic Regression
- Random Forest
- XGBoost



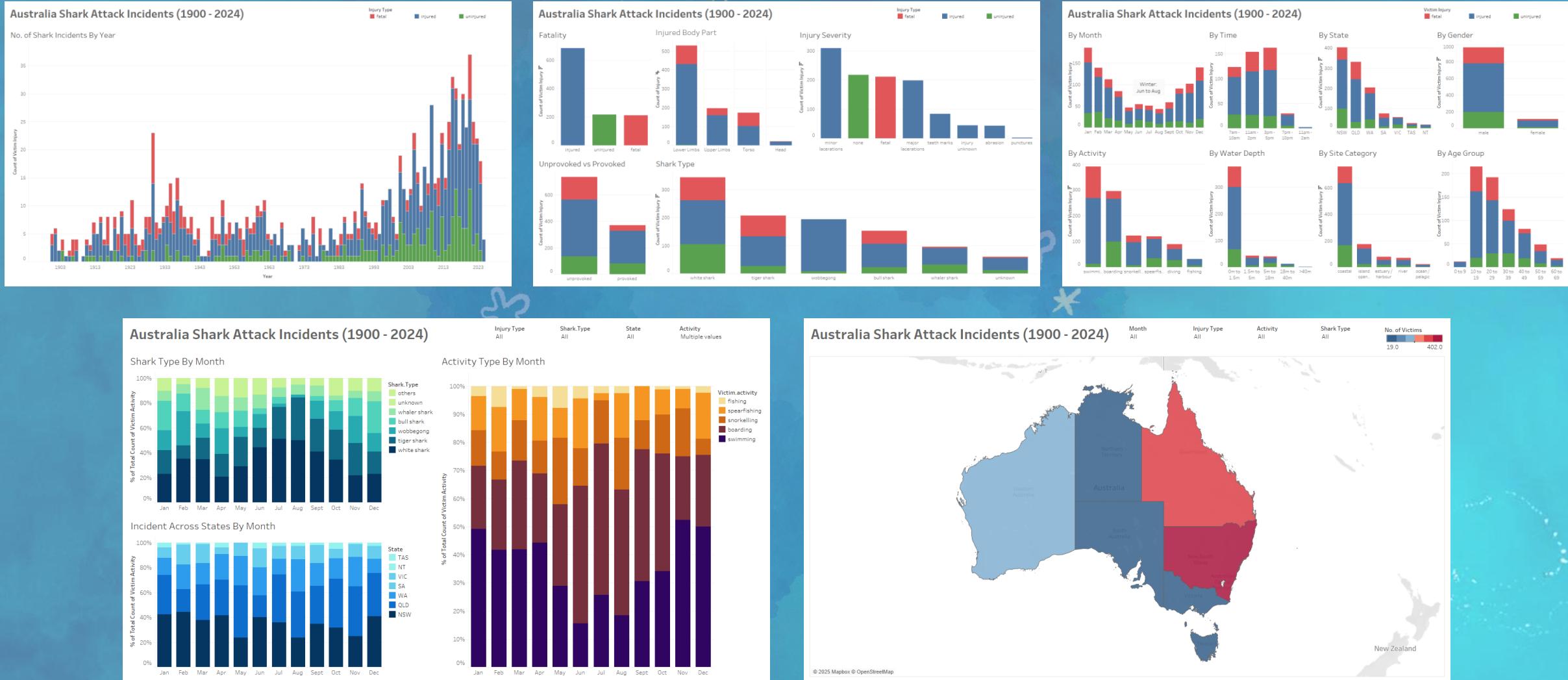
# Dataset

UIN	Incident.month	Incident.year	Victim.inj	Sta	Latitude	Longitude	Site.category	Shark.common	Provoked/	Victim.activity	Injury.location	Injury.severity	Victim.gend	Victim.ag	Hour.of.incide	Depth.of.incident	Depth.category
2	3	1803	injured	WA	-25.83333333	113.88333333	coastal	tiger shark	unprovoked	swimming		injury unknown	male	999	unknown	0	0m to 1.5m
3	1	1807	injured	NSW	-33.86666667	151.2	estuary/harbour	bull shark	unprovoked	swimming	arm, hand	minor lacerations	male	999	unknown	999	Unknown
4	1	1820	fatal	TAS	-42.8	147.53333333	coastal	unknown	provoked	swimming	leg	fatal	male	999	unknown	1	0m to 1.5m
5	1	1825	injured	NSW	-33.85	151.21666667	estuary/harbour	bull shark	unprovoked	swimming	leg	minor lacerations	male	15	unknown	999	Unknown
6	1	1830	injured	TAS	-43.28333333	147.33333333	island open ocean	white shark	unprovoked	swimming	arm	major lacerations	female	999	unknown	999	Unknown
7	1	1832	injured	NSW	-33.86666667	151.18333333	estuary/harbour	bull shark	provoked	swimming	leg	minor lacerations	male	999	unknown	999	Unknown
8	6	1832	injured	NSW	-33.83333333	151.26666667	coastal	white shark	provoked	fishing	leg	major lacerations	female	999	unknown	999	Unknown
9	1	1836	injured	SA	-34.95	138.5	coastal	white shark	unprovoked	swimming		injury unknown	male	999	unknown	999	Unknown
10	8	1836	fatal	QLD	-25.48333333	152.98333333	coastal	whaler shark	unprovoked	swimming	other: body not recovered	fatal	male	999	unknown	999	Unknown
11	8	1836	fatal	QLD	-25.48333333	152.98333333	coastal	whaler shark	unprovoked	swimming	other: body not recovered	fatal	male	999	unknown	999	Unknown
12	1	1837	fatal	NSW	-30.93333333	153.01666667	river	bull shark	unprovoked	swimming	leg	fatal	male	12	unknown	1	0m to 1.5m
13	8	1839	fatal	VIC	-38.98333333	146.45	coastal	white shark	provoked	swimming	leg	fatal	male	999	unknown	0	0m to 1.5m
14	12	1840	fatal	NSW	-33.85	151.21666667	estuary/harbour	bull shark	unprovoked	swimming	leg	fatal	male	999	unknown	999	Unknown
15	4	1841	fatal	NSW	-33.83333333	151.16666667	estuary/harbour	bull shark	unprovoked	swimming	thigh	fatal	male	999	unknown	0	0m to 1.5m

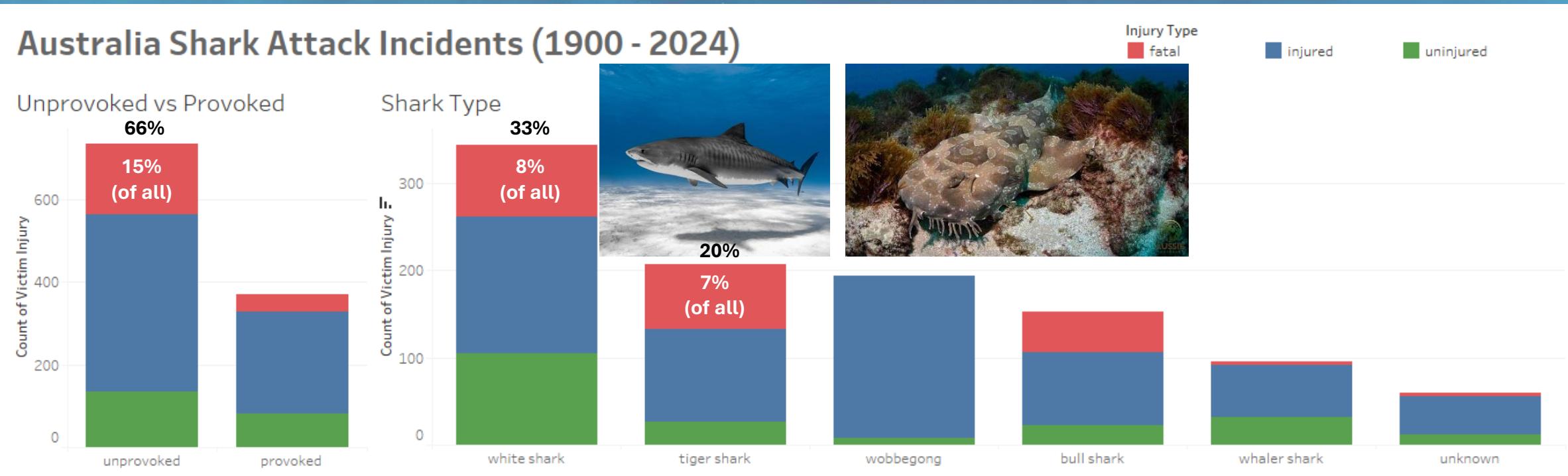
Data Source: Australian Shark-Incident Database

- 1,233 records x 59 features -> selected 18 features
- Primarily categorical features
- Many blanks – filled in with ‘unknown’
- Regrouped into categories: depth and time
- Reclassified: activities and shark type
- Dropped data from before 1820 for predictive modelling and before 1900 for visualization.

# Tableau Dashboard



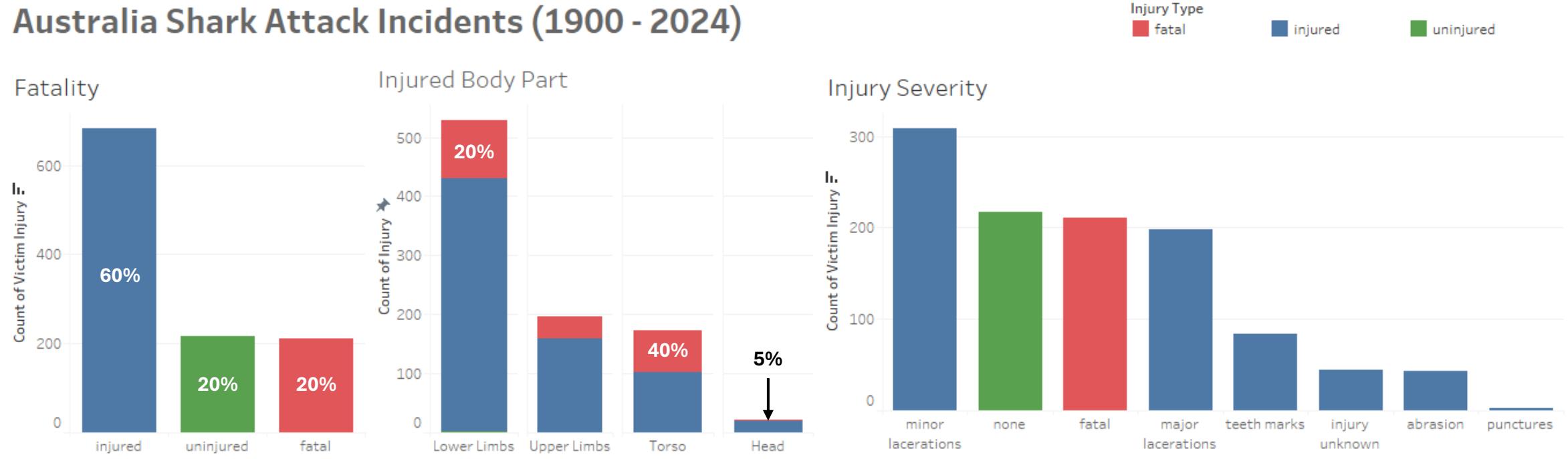
# Key Trends - Shark Type



- 66% of incidents were unprovoked
- Great whites and tiger sharks account for 53% of incidents, unsurprisingly.
- Wobbegong (not known for being aggressive) accounts for nearly 20% of all incidents albeit without fatalities.

# Key Trends - Injury

## Australia Shark Attack Incidents (1900 - 2024)

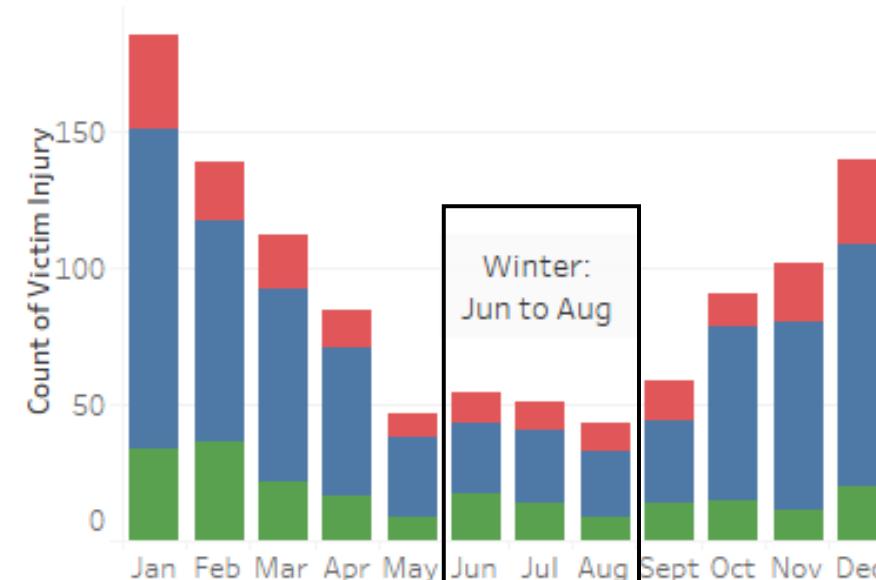


- 60% of incidents result in injury
- Same percentage for uninjured vs fatality
- Lowest limbs are attacked the most
- Highest rate of fatalities are when torso is attacked

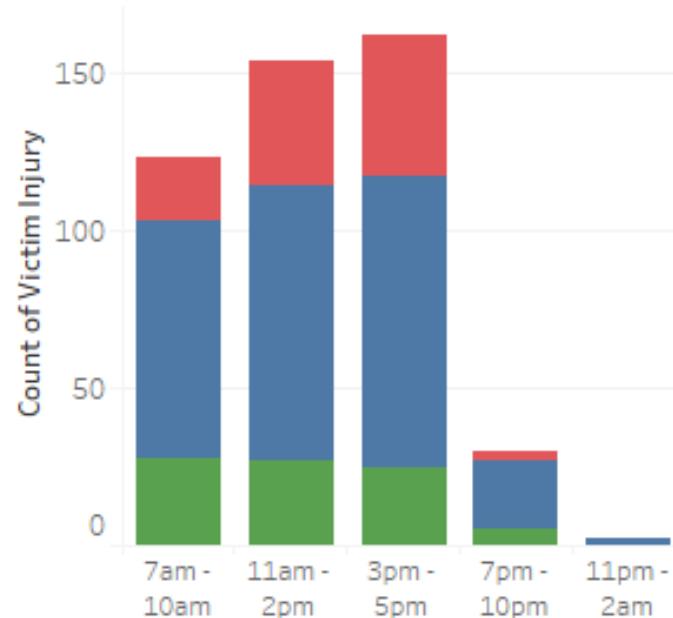
# Key Trends - Time & Location

## Australia Shark Attack Incidents (1900 - 2024)

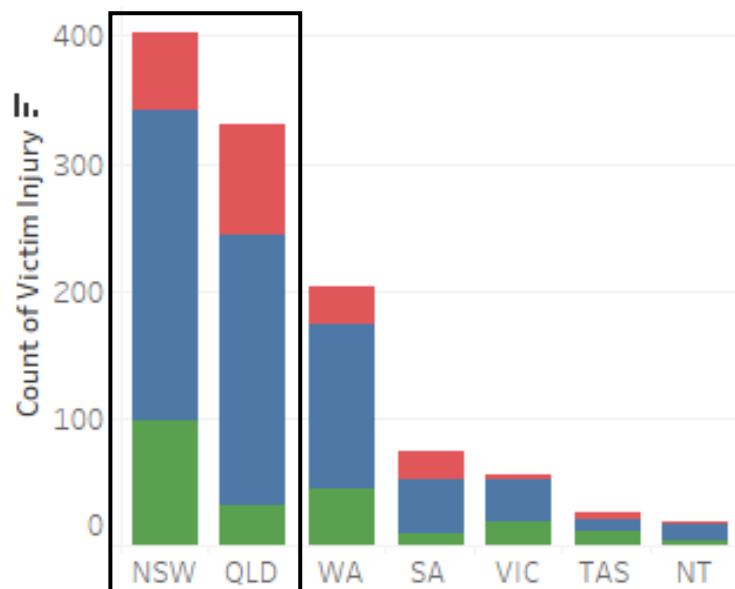
By Month



By Time



By State

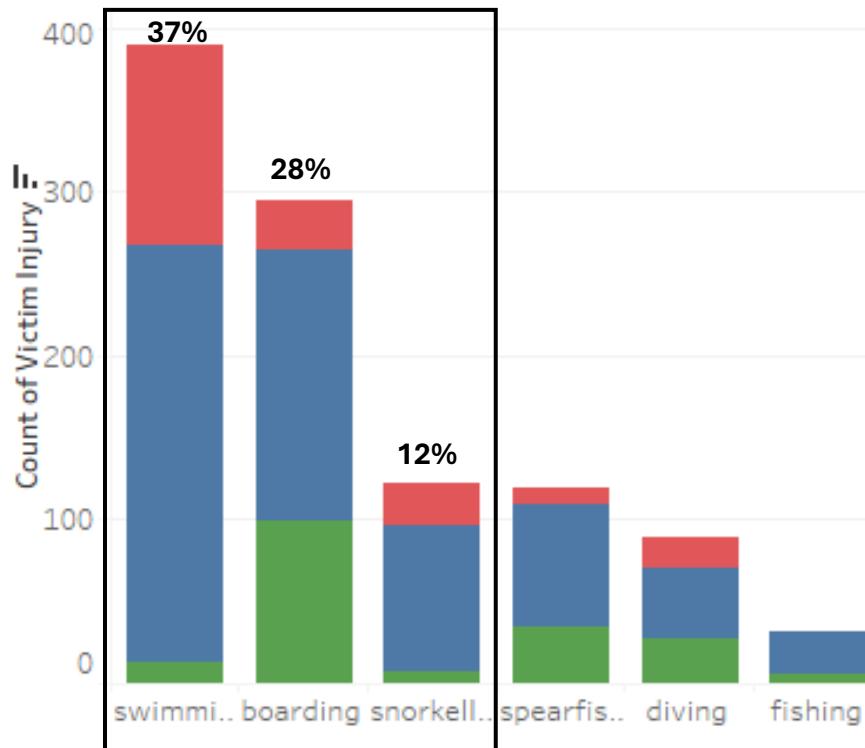


- Attack incidents generally follow Australia seasons, with lesser attacks in winter (less people partaking in in-water activities)
- Attacks generally happen during daylight hours.
- New South Wales and Queensland have significantly higher attacks than other states.

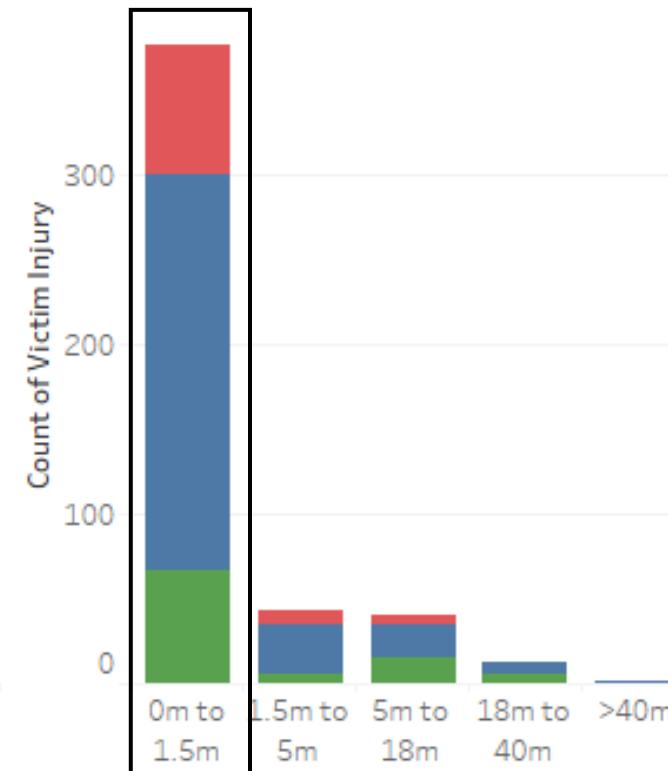
# Key Trends - Activity

## Australia Shark Attack Incidents (1900 - 2024)

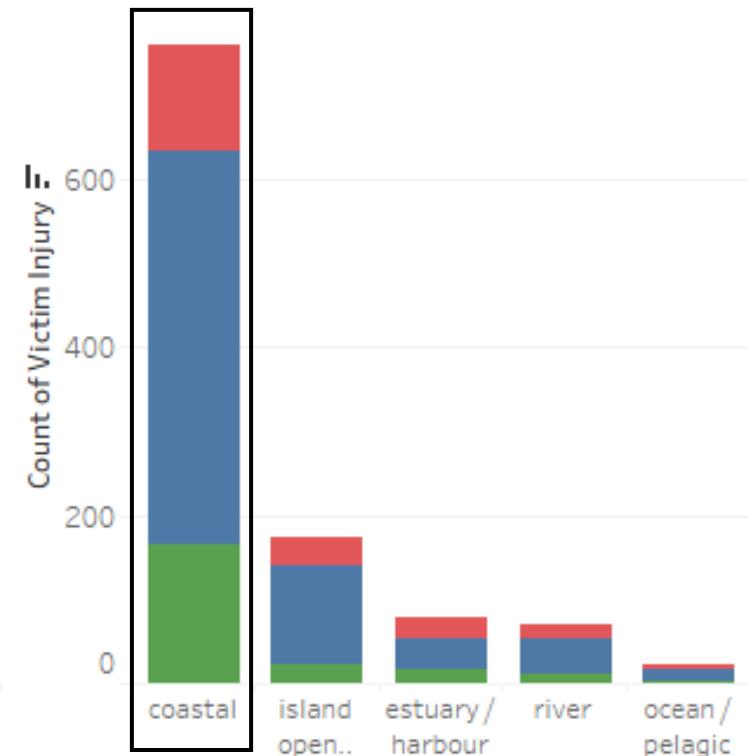
By Activity



By Water Depth



By Site Category



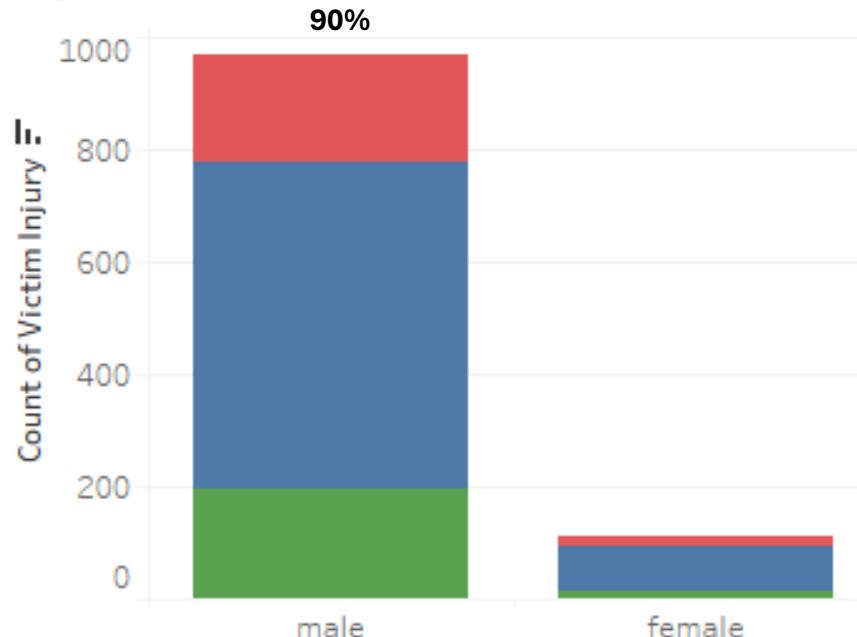
- Most attacks happen on the water surface or shallow waters in coastal regions

# Key Trends - Victim Profile

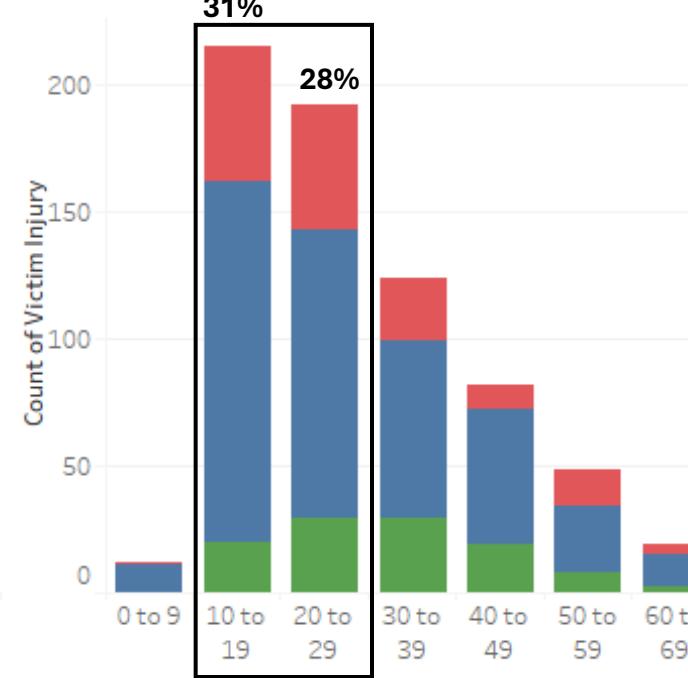
## Australia Shark Attack Incidents (1900 - 2024)

Injury Type  
fatal      injured      uninjured

### By Gender



### By Age Group

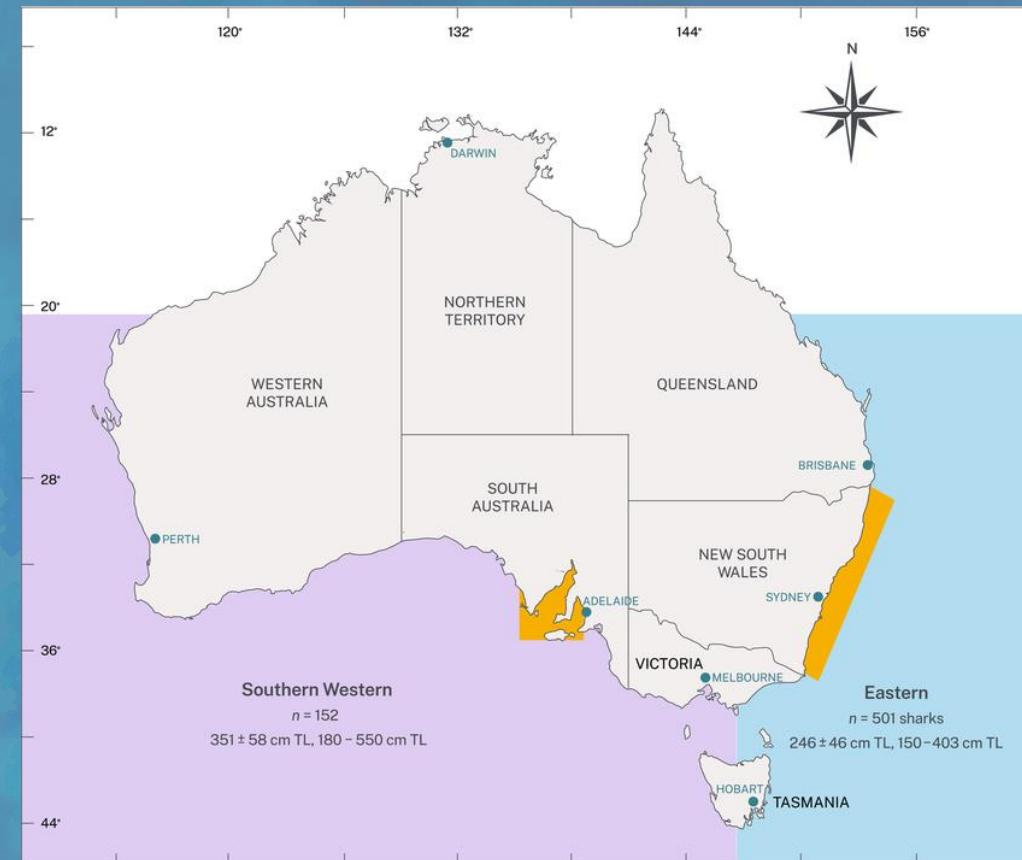


- Majority of victims are males – 70% of adult surfers in Australia are males.
- More than 50% of victims are between 10 – 30 years old.

# Insights - White Shark

There are two known populations of white sharks in Australia – the Southern Western group and the Eastern group.

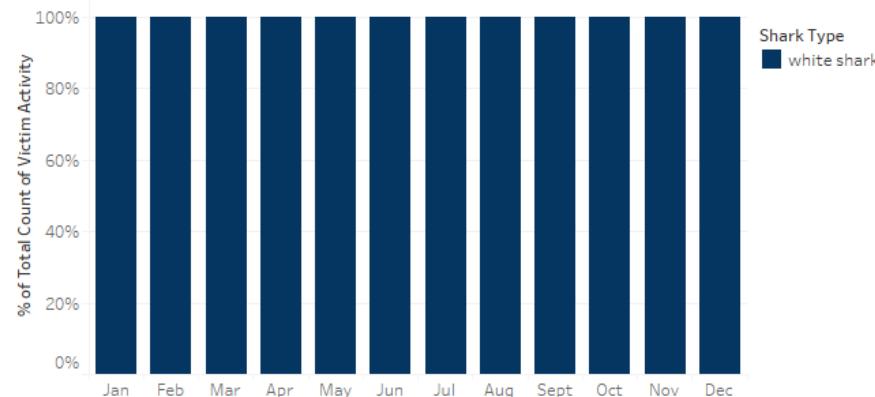
The Eastern group is also known to travel greater distances, all the way to Caledonia, etc.



# Insights - White Shark

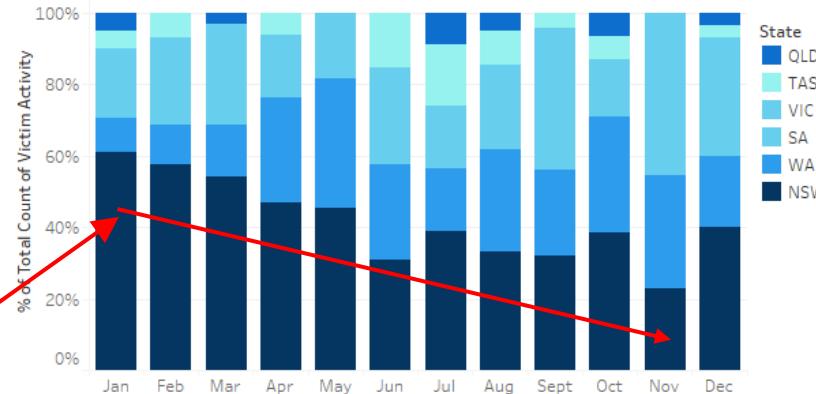
## Australia Shark Attack Incidents (1900 - 2024)

Shark Type By Month



Spread across  
3 – 5 states

Incident Across States By Month



Top States:  
1. NSW  
2. WA  
3. SA

Trending downwards

Injury Type  
All

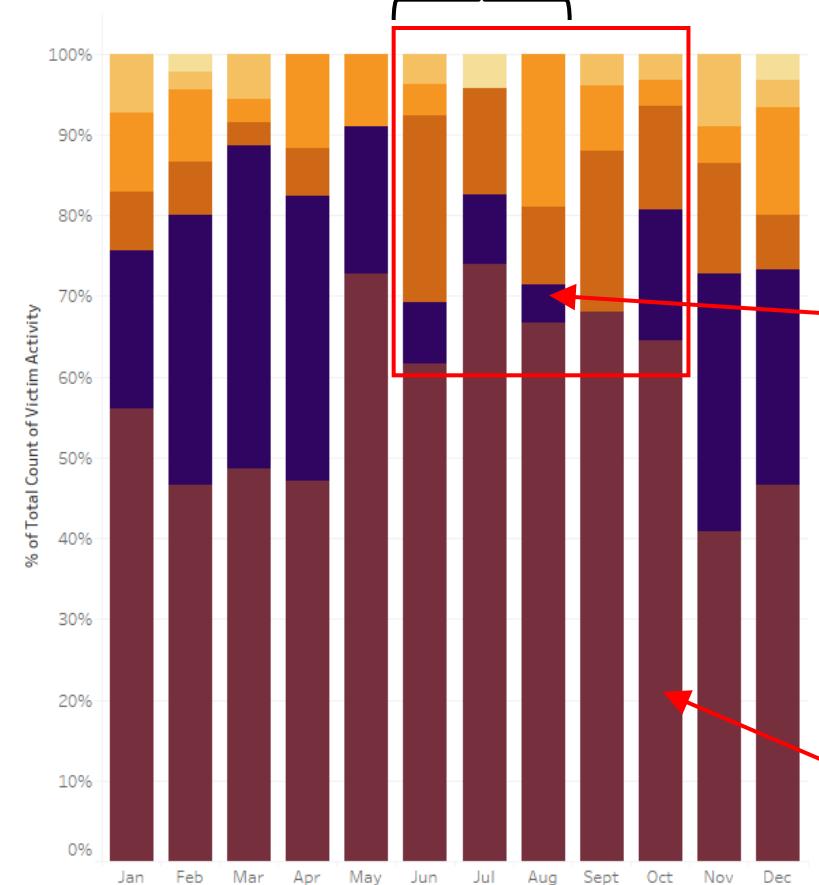
Shark Type  
white shark

State  
All

Activity  
Multiple values

Winter

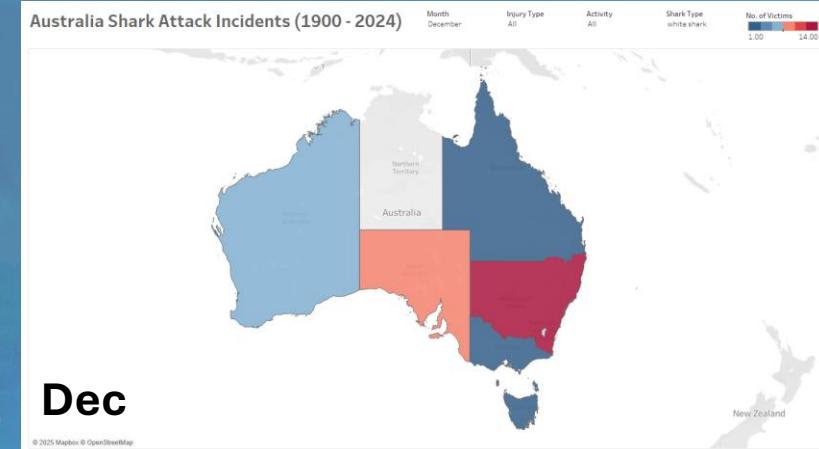
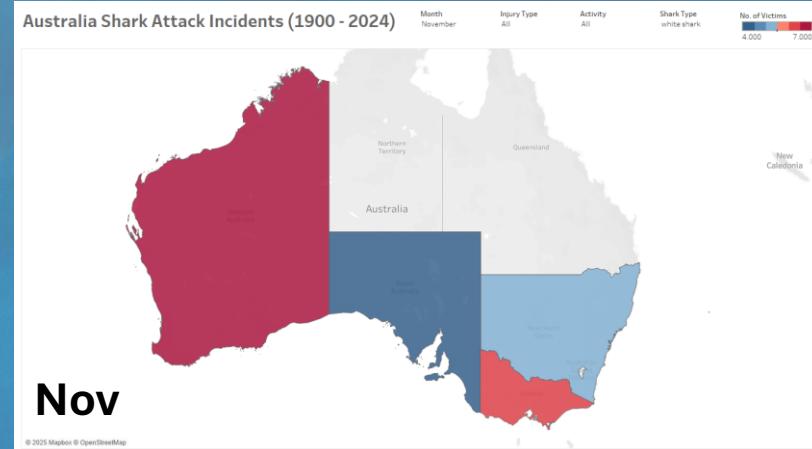
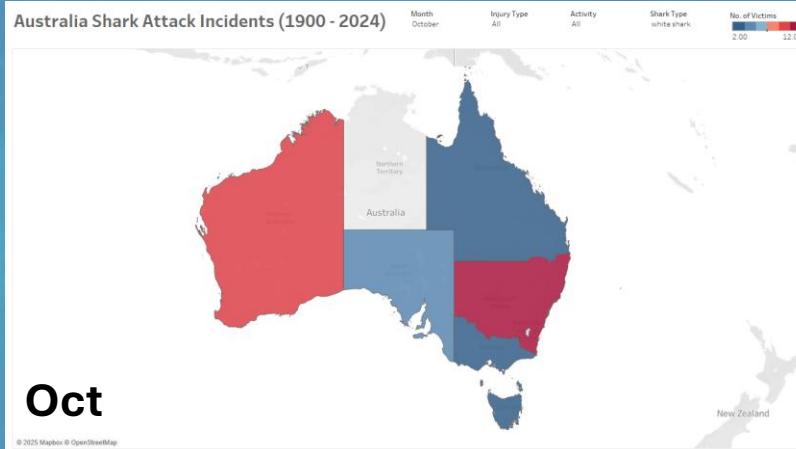
Activity Type By Month



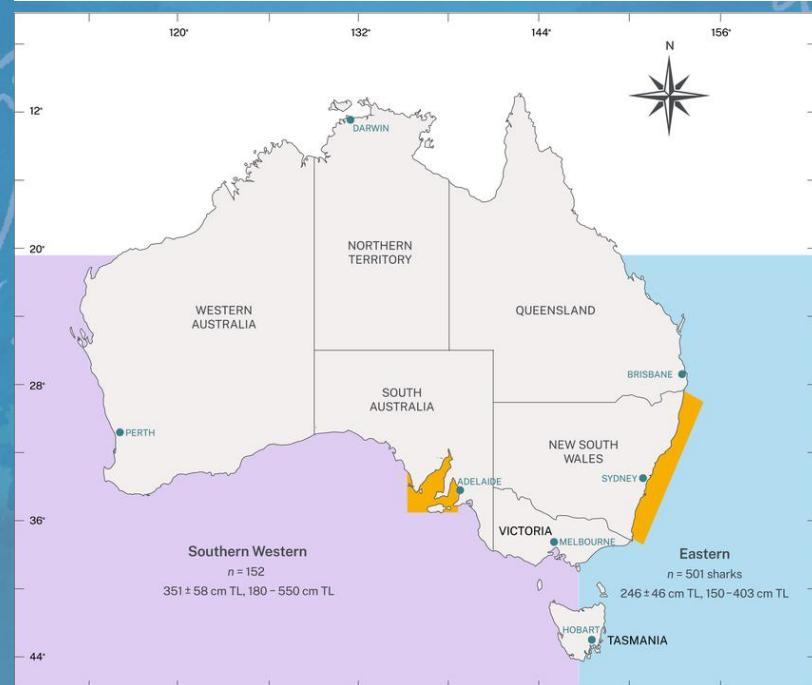
More divers,  
Less swimmers

Top Activity:  
1. Boarding  
2. Swimming

# Insights - White Shark



Looking at the number of shark attack incidents across different states by month, is there a link with the travel / migratory routes of the two white shark populations?



# Insights - Wobbegong



Spotted Wobbegong



Tasselled Wobbegong

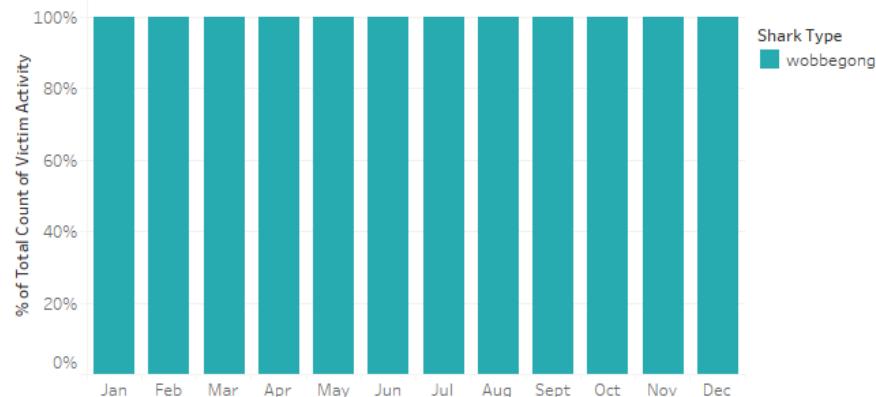
## Fun Facts

- Carpet sharks primarily found in Australia
- Not known to be aggressive
- Wobbegong = “shaggy beard” in Aboriginal language

# Insights - Wobbegong

## Australia Shark Attack Incidents (1900 - 2024)

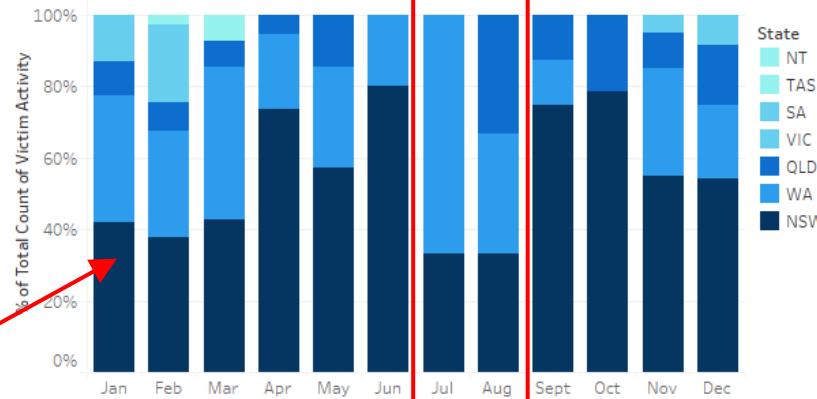
Shark Type By Month



Spread across  
2 – 4 states

**Top States:**  
1. NSW  
2. WA

Incident Across States By Month



No visible trend except  
for low points in Jul/Aug

Injury Type  
All

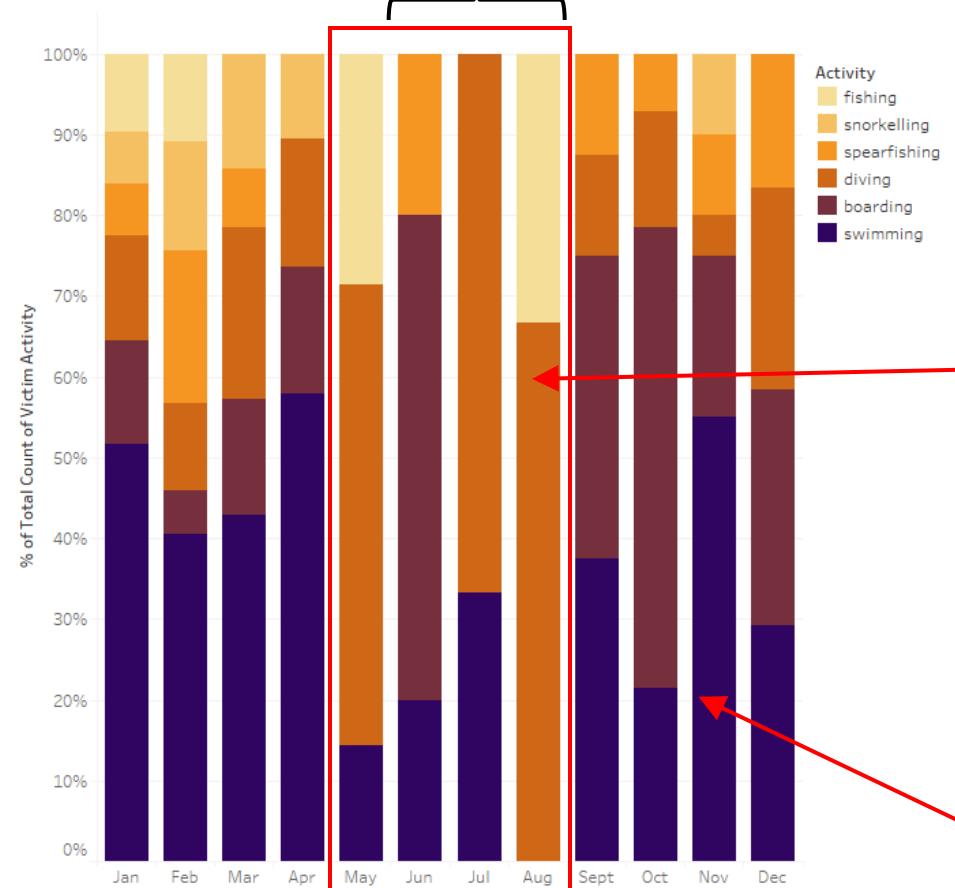
Shark Type  
wobbegong

State  
All

Activity  
Multiple values

Winter

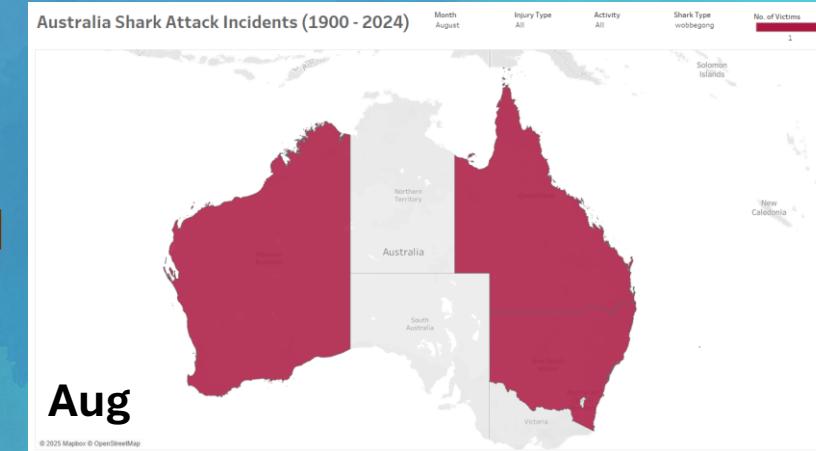
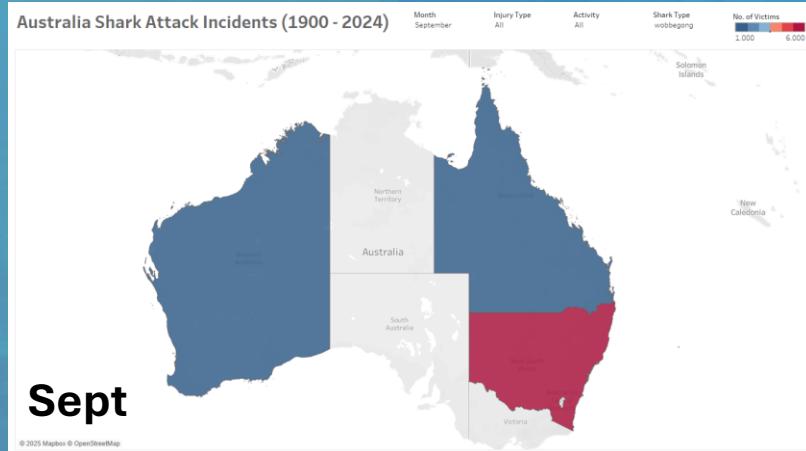
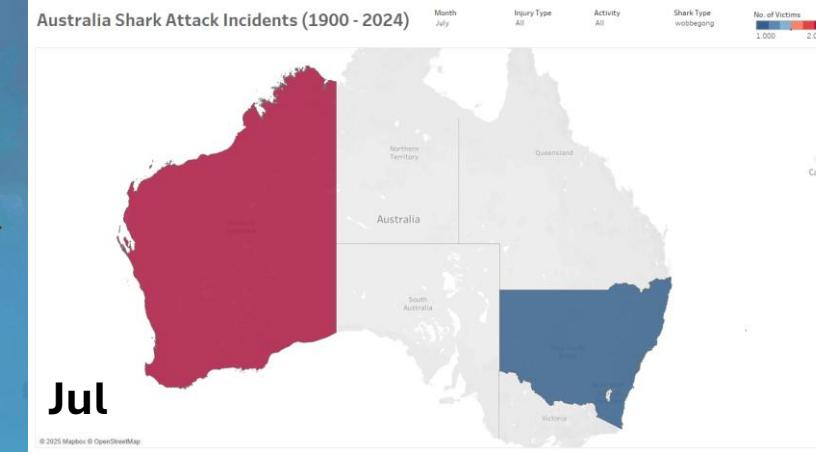
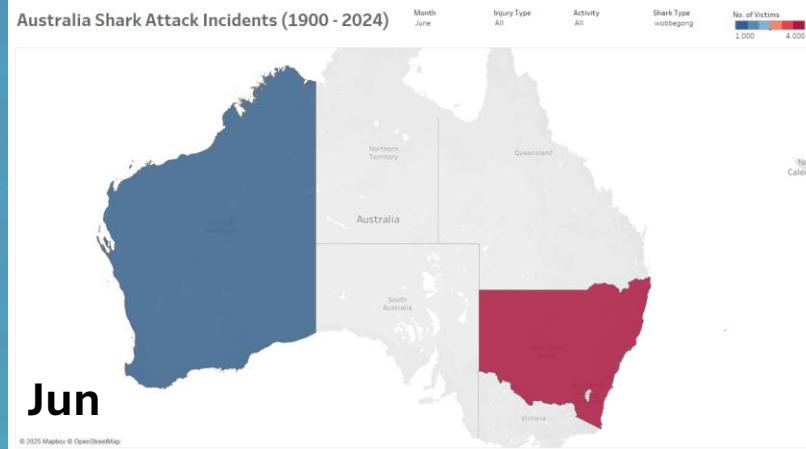
Activity Type By Month



More divers,  
little to no  
swimmers and  
boarders

**Top Activity:**  
1. Swimming  
2. Boarding

# Insights - Wobbegong



# Predictive Model

Model	Recall (Class 2)	Precision (Class 2)	F1-Score (Class 2)	F1-Score (Class 1)	Accuracy	Key Advantage
RF (Tuned)	0.64	<b>0.48</b>	<b>0.55</b>	<b>0.76</b>	<b>0.67</b>	Best overall accuracy
XGBoost (Tuned)	<b>0.74</b>	0.37	0.49	0.69	<b>0.60</b>	Best fatality detection
LR (Scaled)	0.32	0.39	0.35	0.74	0.63	Worst fatality detection
Baseline	-	-	-	-	<b>0.62</b>	

## Random Forest Classifier:

1. **Best Overall Accuracy (67%)**
2. **Balanced Fatality Detection**
  - **64% recall** - catches most fatalities.
  - **48% precision** - good precision (48%) considering the high recall
3. **Robust Performance**
  - Best F1-score for both fatalities (55%) and injuries (76%).

# Limitations

## ➤ Limited Dataset

- Small dataset of ~1,200 samples
- Imbalanced classes

## ➤ Missing Factors and Data

- Environmental factors (water temperature, weather)
- Human activity metrics (tourist density, activity participation)
- Shark behaviour data (migration patterns, breeding seasons)

# Recommendations

The image shows a screenshot of a web application titled "Shark Attack Risk Predictor (XGBoost 3.0)". The page has a header "Predict injury risk using XGBoost 3.0". Below the header are five input fields: "Month" (November), "State" (NSW), "Activity" (Swimming), "Time of Day" (23:00 to 02:00), and "Age Group" (20 to 29). At the bottom are two buttons: "Clear" and "Submit". To the right of the form, there are two overlapping modal windows. The top window is titled "Risk Prediction" and the bottom one is titled "Confidence Scores", both displaying a red "Error" message.

## ➤ Predictive Model

- Deploy model with confidence threshold
- Improve model accuracy by factoring environmental and human activity metrics
- Work with marine life experts

## ➤ Other than chatbot:

- Standardized reporting protocol for community or hospitals
- Community reporting apps for near-miss incidents
- Mobile app (Telegram) for lifeguards with location-based risk alerts

# Conclusion

## Key Findings

1. Attacks peak during **warmer months**.
2. High-risk zones: **NSW** and **Queensland**.
3. Most attacks are **unprovoked**.
4. **Swimming/surfing had higher risks** than diving

## Predictive Model

1. Best performance: **Random Forest Classifier**
2. **Highest accuracy** (67%)
3. **Best fatality detection**
  - 64% recall for fatalities (identifying most high-risk cases)
  - 48% precision (fewer false alarms)
4. **Most robust performance**
  - Best F1-score for both fatalities (55%) and injuries (76%)



“Humans are friends,  
not food!”

POSEIDON