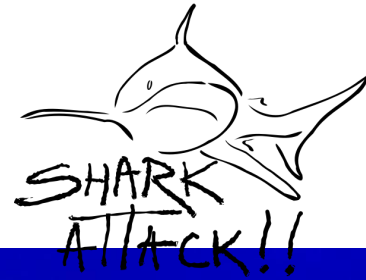# GROUP 2

Project: Shark Attacks

Members: Dimitra, Negar, Elio, Roy

HYPOTHESIS

1. **Hypothesis Statement:** The frequency of shark attacks is influenced by **geographical factors** such as country and state (or month).

2. **Hypothesis Statement**: Certain **activities**, such as surfing or spearfishing, are more prone to shark attacks compared to other water sports or recreational activities.

Data Cleaning

Relevant cleaning methods:

1. .drop
2. .dropna
3. .replace
4. .fillna
5. .to_numeric
6. .str.extract (r'([A-Za-z]{3})')

Justification of our imputation strategy (Why we did what)

- Filling missing values at "Year" with data from the Date column, by using pattern search - 4 consecutive digids.
- Creating "Month" column, using the "Date" column, by using pattern search - 3 consecutive letters. Cleaning the data prior to that to exclude words that are not months.
- Data relevance: keeping in the analysis data from the year 1947 and forward. These data consist the 75% of the total data.
- Dropping columns that we know will offer no insight to our analysis, either due to content or excessively missing data.
- Deleting rows with all empty values.
- Filling empty values in State, Species and activity with unknown or undefined, so as not to reduce our data too much. In age we use the average to fill in the non values, and in name we create anonymous1, 2, 3.
- Formatting data, capitalize everything, pass age as integer, remove extra spaces and sort by descending year.

Data analysis after cleaning

	Year	Type	Country	State	Activity	Name	Sex	Age	Injury	Species	Month
0	2024	Unprovoked	India	Maharashtra	Fishing	Vicky suresh govari	M	32	Calf of lower left leg injured	Bull shark, 7'	Feb
1	2024	Unprovoked	Australia	New__south_wales	Swimming	Lauren o'neill	F	29	Right leg bitten	Bull shark	Jan
2	2024	Unprovoked	Bahamas	Paradise_island	Swimming	Male	M	10	Right leg injured	Unknown	Jan
3	2024	Unprovoked	Australia	South_australia	Surfing	Murray adams	M	64	Leg bitten	White shark	Jan
4	2024	Provoked	South_africa	Eastern_cape_province	Fishing	Male	M	62	Provoked incident. lacerations and puncture wo...	Raggedtooth shark	Jan
5	2024	provoked	Trinidad	Undefined	Spearfishing	Male	M	28	Shoulder bitten	Blacktip reef shark	Feb
6	2023	Unprovoked	Australia	South_australia	Surfing	Simon baccanello	M	46	Fatal	White shark	May
7	2023	Unprovoked	Usa	New_jersey	Surfing	Maggie drozdowski	F	15	Lacerations to left foot and calf	Unknown	May
8	2023	Provoked	Usa	Florida	Fishing	Male	M	35	Provoked incident hooked shark bit his foot	Unknown	May
9	2023	Unprovoked	Usa	Florida	Spearfishing	Kevin blanco	M	20		Unknown	May
10	2023	Unprovoked	Usa	South_carolina	Swimming	Claire gugerty	F	24	Minor inuries to right knee	Unknown	May
11	2023	Unprovoked	Usa	Florida	Spearfishing	Ethan wilder	M	22	Injury to lower leg	Unknown	May
12	2023	Provoked	Usa	Hawaii	Kayak fishing	Scott haraguchi	M	28	Shark bit kayak, no inury to occupant	Tiger shark	May
13	2023	provoked	Cuba	Mayabeque	Spearfishing	Maro alejandro díaz jiménez	M	28	Fatal	Bull shark	May
14	2023	Unprovoked	Turks_and_caicos	Providenciales_island	Snorkeling	Alexandra truwit	F	22	Left foot severed	Caribbean reef shark	May
15	2023	Unprovoked	South_africa	Easten_cape_province	Surfing	Male	M	28	Leg injured	Unknown	May
16	2023	Unprovoked	South_africa	Easten_cape_province	Diving for crayfish	Nkululeko mphehlule	M	39	Shoulder injured	Unknown	May
17	2023	Questionable	Bahamas	Lucayan_lucayan_archipelago	Jumped overboard	Cameron robbins	M	18	Presumed fatal, body not recovered	Shark involvement not confirmes	May
18	2023	Unprovoked	Bahamas	Freeport	Scuba diving	Heidi ernst	F	73	Calf severely bitten	Caribbean rreef shark	Jun
19	2023	Unprovoked	New_caledonia	Poum	Spearfishing	Male	M	42	Fatal, bite to leg, shoulder and head	Unknown	May

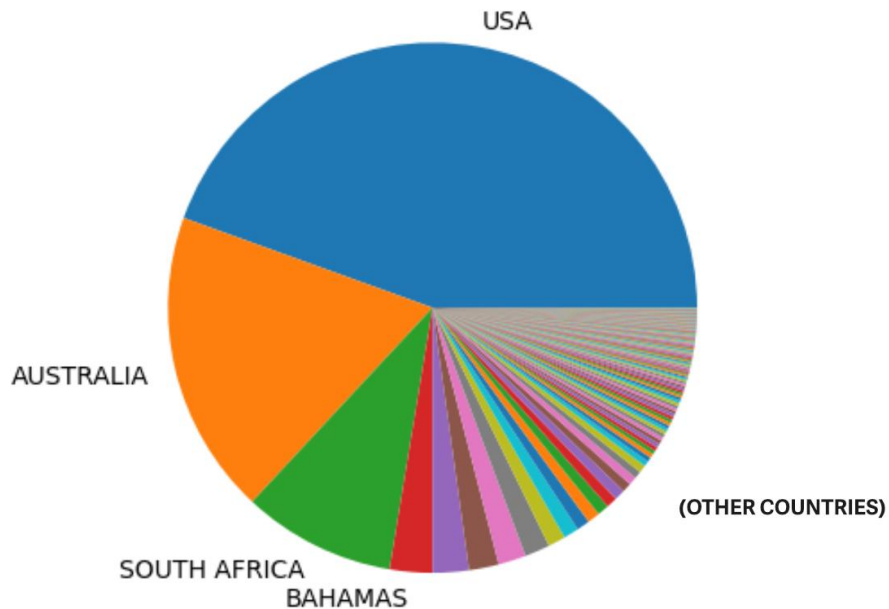# Major Obstacle during the project AND LEARNINGS

- Data Cleaning was a challenge – a lot of deleted data which we fear may impact the results of our analysis.

- Time constraints

- Challenge for division of tasks. Sometimes recycling tasks slowed us down because we had to get up to speed with what our teammates worked on earlier
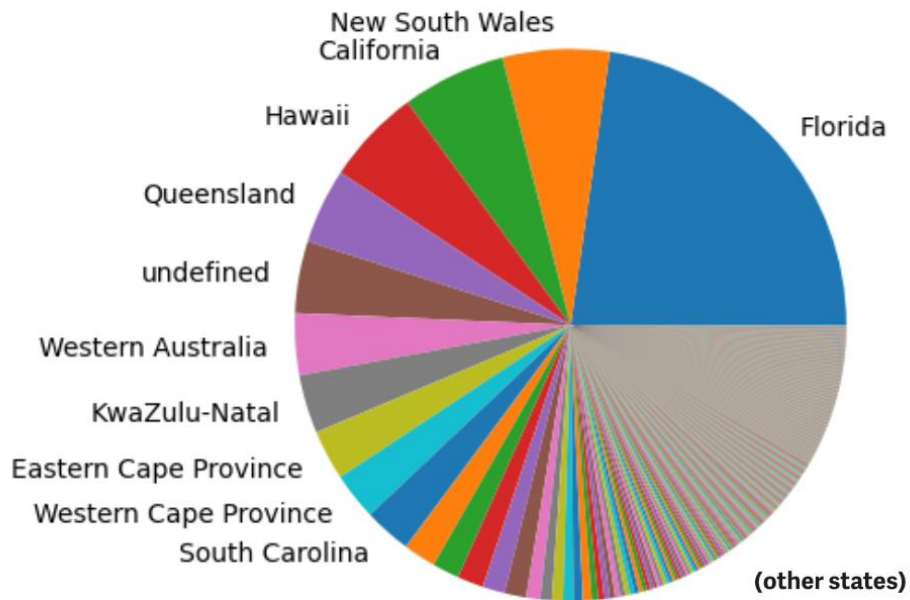
LEARNINGS

- ***WORKING ITERATIVELY AND WITH AN AGILE APPROACH!!***

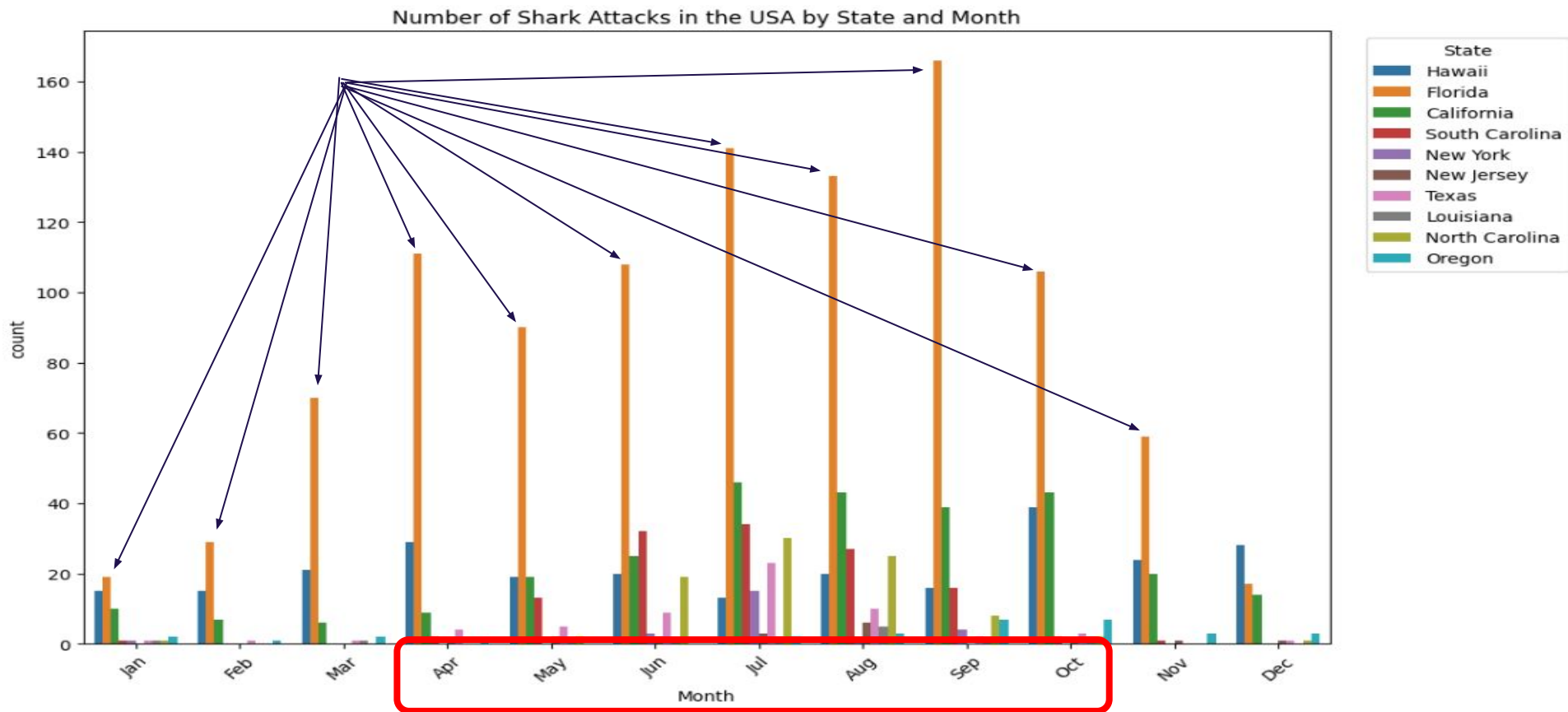# HYPOTHESIS 1– Frequency of Shark attacks correlates with geography(country/state)

Shark Attacks Distribution by Country

USA
AUSTRALIA
SOUTH AFRICA
BAHAMAS
(OTHER COUNTRIES)

Shark Attacks Distribution by State

New South Wales
California
Hawaii
Queensland
undefined
Western Australia
KwaZulu-Natal
Eastern Cape Province
Western Cape Province
South Carolina
Florida
(other states)

HYPOTHESIS 1- Frequency of Shark attacks is relative to geography(USA)

Number of Shark Attacks in the USA by State and Month

HYPOTHESIS 1- Frequency of Shark attacks is relative to geography(Australia)

Number of Shark Attacks in Australia by State and Month
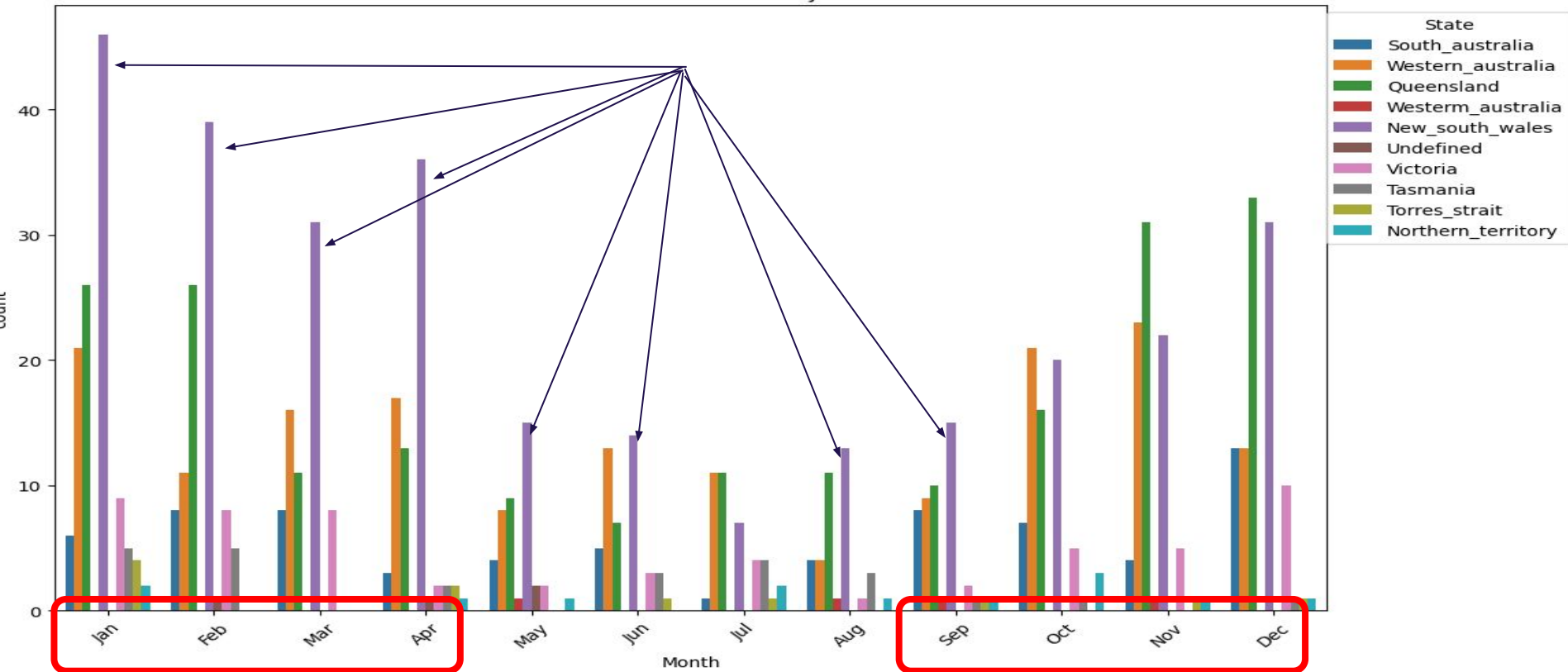
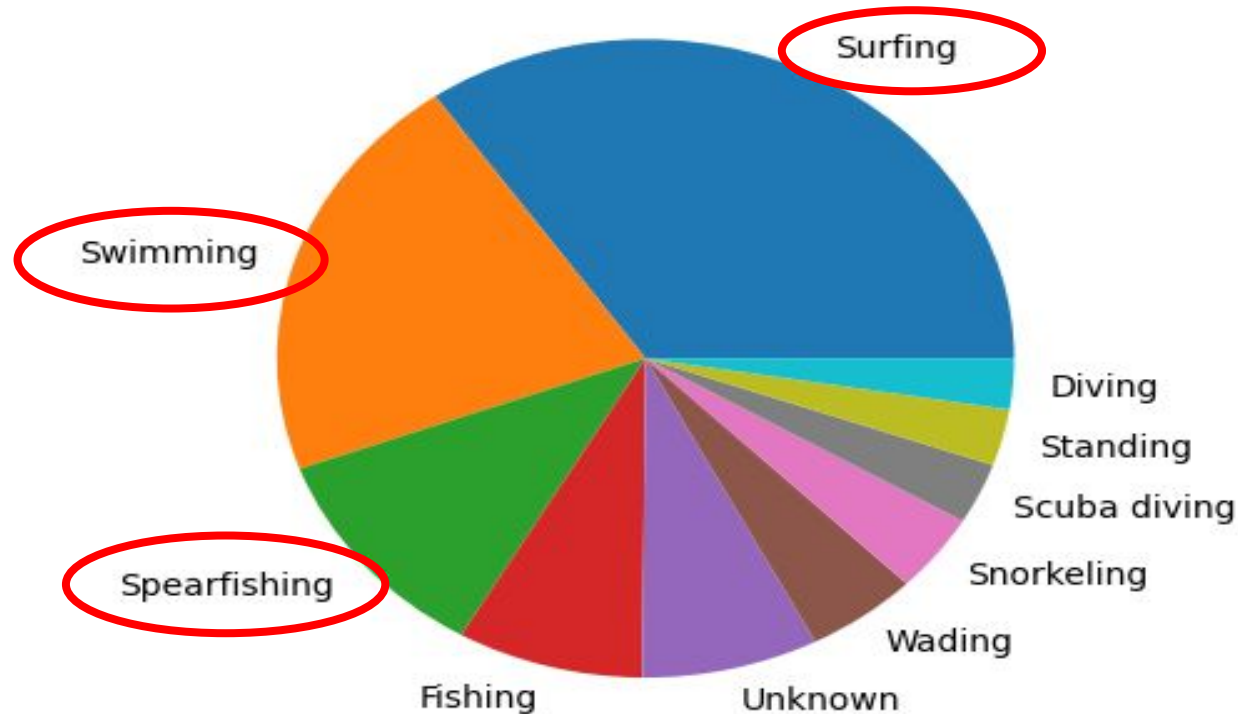# Hypotheses 1 validation & data-driven conclusions

- The **Hypothesis Statement 1:** The frequency of shark attacks is influenced by **geographical factors** such as country and state is **justified**.

- Business Opportunity: The travel agency, DREN co. specializing in beach vacations could use this information to offer **safer destinations for their customers** or offer **tours for shark cage dives**.

INSIGHTS

- Additionally, our plots indicate the seasonality of these attacks with respect to months of the year. This helps in more informed decision making for investments, and harm prevention.

- Beach resorts and local authorities can implement **targeted safety protocols** and lifeguard schedules **during high-risk months** to mitigate shark attack risks.

HYPOTHESIS 2- higher Shark atTacks occur during certain activities

Shark Attacks Distribution by ACTIVITY-TOP 10

Hypotheses 2 validation & data-driven conclusions

- The **Hypothesis Statement 2:** The Hypothesis Statement: Certain activities, such as surfing or spearfishing, are more prone to shark attacks compared to other water sports or recreational activities, is **justified**

- Business Opportunity: Companies selling water sports equipment (eg: safety gear) or offering adventure tourism packages can tailor their offerings and safety measures based on this analysis.

INSIGHTS

- The data suggests that surfing and spearfishing are more susceptible to shark attacks. This highlights the significance for companies in these sectors to **educate participants about shark attack prevention measures**, presenting an opportunity to promote safer experiences and enhance the feeling of security while enjoying these activities.

THANK YOU !!