* Import the data (first downloaded from website and uploaded in GitHub)
* Drop columns and rows that are not needed for answering the questions
* Rename columns ‘Fatal (Y/N)’ and ‘Species ‘
* Look what the non-default missing values are
  + Type
    - Invalid
    - Questionable
    - Boat
    - Boating
    - Boatomg
  + Activity
    - ' '
    - ' '
    - '.'
  + Age
    - ''
    - ' '
    - ' '
    - 'X'
    - 'MAKE LINE GREEN'
    - \xa0
  + Injury
    - No non-default missing values
  + Fatal
    - 'M'
    - 'UNKNOWN'
    - '2017'
    - ' N'
    - 'N '
    - 'y'
  + Species
    - ''
    - ' '
    - '\xa0 '
* I assumed that ‘Sea disaster’ in the ‘Type’ column was not provoked or unprovoked, because a sea disaster is not a shark attack (More like feeding the sharks that are already there). This could create a bias, because maybe the disaster itself was a provoking of the sharks?
* I assumed that ´Questionable´, ´confirmed´, ´involvement´, ´Invalid´ and ´No shark´ in the ‘Species’ column didn´t involve a shark attack and left them out. This could create a bias, because maybe there was a shark attack in there.

1. What are the most dangerous types of sharks to humans?

* I want to know how many of each species are responsible for the attacks and split them in fatal and non-fatal (to determine “most” dangerous).
* I don’t think I made assumptions on this one, except the one mentioned above (about questionable etc).
* Conclusion:

The most dangerous type of shark is ‘Unknown shark’ which was 918 times fatal and 2599 non-fatal. Total of 3517 attacks.

The next most dangerous type of shark is ‘White shark’ which was 126 times fatal and 427 non-fatal. Total of 553 attacks.

1. Are children more likely to be attacked by sharks?

* What age are considered children?

“In Dutch youth policy, youth is generally defined as the age group 0-25 years. Within this age group a distinction is made (in policy, not in laws) between Children (0-12 years), Young people (12-18 years) and Young Adults (18-25 years).” (Source: <https://www.protection-of-minors.eu/en/country/NL#answer1>)

* I want to know how many attacks per age have taken place.
* I assumed that the column for ‘Age’ only contained ages in numbers. This could have created a bias, because there could have been more attacks at a certain age.
* Conclusion:

Children are not more likely to be attacked by sharks. Young people and Young adults are more likely to be attacked by sharks than children or older adults.

1. Are shark attacks where sharks were provoked more or less dangerous?

* I assumed boat, boating and boatingomg were unprovoked, because they were just there boating and were not there for the sharks. This could create a bias, because maybe they were provoking from the boat.
* I assumed that the remaining values in the column Injury after ‘Fatal’ and ‘No injury’ were ‘injured’. This could create a bias, because they could have been minor injuries that actually were ‘No injury’.
* Conclusion:

Shark-attacks that were unprovoked were more dangerous than provoked.

Sharks that were provoked caused 2.61% fatalities, 85.71% injuries and 11.67% non-injuries. Sharks that were unprovoked caused 23.88% fatalities, 62.72% injuries and 12.99% non-injuries.

Fatalities+injuries and non-injuries were almost the same for both provoked and unprovoked. But there were many more fatalities to unprovoked than to provoked.

1. Are certain activities more likely to result in a shark attack?

* I want to know how many attacks per activity have taken place.
* I don’t think I made assumptions on this one.
* Conclusion:

Certain activities are more likely to result in a shark attack.

Top 6 of activities that resulted in a shark attack:

* + 1. Surfing, Surf-skiing, Body-surfing, Kite-surfing, Windsurfing
    2. Fishing
    3. Swimming
    4. Diving
    5. Shark-fishing, Tagging sharks, Feeding sharks
    6. Unknown activity