**Algorithms / Solution technologies:**

* We plan on using Google Maps API to translate shark attack locations to Latitude & Longitude data for use in our algorithms/methods
* Density-based clustering algorithms
  + Research reveals DBSCAN as a potential algorithm for this
* Hot spot analysis methods
  + Hierarchical clustering?
  + location-based statistics can be used as examples (city crime data for example)

**Risks:**

* Not enough meaningful ‘provoked’ data – We have about 600 rows of this type of attack & around 10x more ‘unprovoked’ data.
* Poor data quality – some data is very old, some of the dates are not clear, some of the locations are unclear, shark species aren’t usually written, and many other data entries are NaN.

**Challenge:**

**Citations:**

Haven’t seen related projects but here are some resources that should help us code our project:

<https://scikit-learn.org/stable/modules/generated/sklearn.cluster.DBSCAN.html>

<https://www.mygreatlearning.com/blog/dbscan-algorithm/>

<https://glenbambrick.com/2016/01/21/what-is-hotspot-analysis/>

**Plan for Completion:**