

# Project Introduction: Analyzing U.S. Beaches

Calleigh Smith

Due 3/2/2021

## Overview

According to the National Ocean Service, “Our oceans and coasts touch every American every day - providing us with places to live, food to eat, jobs, commerce, recreation, energy, even medicines that heal” [1]. Unfortunately, waters globally are highly vulnerable to human actions and exposed to threats such as pollution, climate change, algae blooms, coastal development, and more. In order to ensure the health of these natural environments and those who depend on them, it is essential that we understand the current status of coastal waters and which beaches are at risk for the future.

This paper is important because it has the potential to identify risk factors for the environmental degradation of beaches, which can in turn guide legislators to develop regulations regarding these bodies of water. Additionally, environmental protection of the oceans and coasts is futile without the education and engagement of the public. Public knowledge of coastal conservation is limited. According to a survey by Lotze et al., 70% of respondents believe the marine environment is under threat from human activities, but only 15% thought the ocean’s health was poor or threatened. By statistically analyzing data relating to the oceans and the coasts and presenting it in a way that can be understood by a general audience, this project has the ability to illustrate a clear picture of the threats to aquatic bodies in the U.S. and inform marine managers, policy makers, conservation practitioners, and educators to improve marine management and conservation programs.

In our analysis, we will look at the following research questions:

- **Measuring Risk Factors:** Can we predict the number of beach actions (beach-specific advisories or closings issued by the reporting state or local governments) based on a beach’s characteristics? Similarly, are certain variables more valuable in predicting beach actions?
- **Measuring Pollution:** How are the different types of pollution (algal, animal, sewer line, etc.) distributed across U.S. waters? Are certain regions more prone to certain pollution types of pollution than others?
- **Measuring Government Communication with Locals:** How do local governments communicate with citizens when there is a beach advisory or closing? Are governments transparent in reporting the risks associated with local bodies of water to residents?

## Data

Our data was obtained through the Environmental Protection Agency (EPA) as part of their BEACON 2.0 online system [2]. The EPA created the BEach Advisory and Closing Online Notification (BEACON) to meet the Agency’s requirement to provide the public a database of pollution occurrences for coastal recreational waters. Under the Beaches Environmental Assessment and Coastal Health (BEACH) Act of 2000, EPA provides annual grants to coastal and Great Lakes states, territories, and eligible tribes to help local authorities monitor their coastal and Great Lakes beaches and notify the public of water quality conditions that may be unsafe for swimming. BEACON contains state-reported beach monitoring and notification data and is fully available online without the need for special permissions.

The BEACON 2.0 online tool provides many reports on all beaches in the United States. Available datasets

include those related to each action for a beach, beach attributes and profiles, beach days, beach monitoring frequency, possible pollution sources, water quality, tier 1 beach information state summaries, and more. Using BEACON 2.0's user guide found [here](#), one can view a full data dictionary, information on how to use the database, and basic summary statistics on some of the variables.

In this analysis, we will solely focus on Tier 1 beaches, which tend to have the highest risk and the most available data across all of the datasets. States and territories designate their significant public beaches as Tier 1 beaches, which is a requirement of the BEACH Act grant program. In order to deal with issues of spatial correlation, we will take a stratified random sample of 50% of the beaches from each water body where data was available from the year 2020, which results in about 1,070 beaches in this analysis.

For our first research question on measuring risk factors, we will look at either the number of beach actions or the number of days under action in a given year for each beach as the response variable. Features I would like to consider include action reason, beach ownership (private/public), beach length, water body type, beach monitoring frequency, water quality, and number of swim days. For our second research question, we will look at states, names of body waters, possible pollution sources, and water quality standards. For our third research question, we will look at local action decision procedures, which includes how governing bodies communicated beach actions with residents, and the reason/severity of the beach action.

## **EDA**