**Data Analytics Capstone Topic Approval Form**

**Student Name:** Jessica Schmidt

**Student ID:** 010774455

**Capstone Project Name:** Alaskan Eastern Bering Sea Snow Crab Geospatial Abundance Multiple Linear Regression Analysis

**Project Topic**: A Multiple Linear Regression Analysis to determine whether there are any statistically significant factors associated with the size of a haul of snow crab (catch per unit effort measure of abundance).

**This project does not involve human subjects research and is exempt from WGU IRB review.**

**Research Question:** To what extent does snow crab gender, year of haul, bottom depth, surface temperature, bottom temperature, latitude, and longitude affect the catch per unit effort of snow crab?

**Hypothesis**: **Null hypothesis**- Snow crab gender, year of haul, bottom depth, surface temperature, bottom temperature, latitude, and longitude do not statistically significantly affect the catch per unit effort of snow crab. **Alternate Hypothesis**- Snow crab gender, year of haul, bottom depth, surface temperature, bottom temperature, latitude, and longitude do statistically significantly affect the catch per unit effort of snow crab.

**Context:** Snow Crab Fishermen of the Bering Sea in Alaska need to maximize the size of their hauls of snow crab. The National Oceanic and Atmospheric Association (NOAA) has tracked the size of snow crab hauls (a count of snow crab per haul) as well as other relevant factors (crab gender, bottom depth of haul, surface temperature, bottom temperature, year specimen was collected, latitude, and longitude) from 1975-2018. Therefore, it would be useful for snow crab fishermen to know which factors are associated with the size of a haul of snow crab. This Multiple Linear Regression seeks to determine which of these factors are statistically significant and therefore relevant in assisting and informing snow crab fishermen when it comes time to fish.

**Data:** The data for this analysis comes from the National Oceanic and Atmospheric Association (NOAA) and is called "Snow Crab Geospatial Data (1975-2018)" with a subtitle of "Alaskan Snow Crab Eastern Bering Sea Geospatial Data (1975-2018)". The data is publicly available on Kaggle and comes from the U.S. Government. "The dataset contains catch per unit effort data of commercial snow crab landings in the Alaskan Eastern Bering Sea. The catch per unit effort is an indirect measure of the abundance of a target species. The data was collected from NOAA then cleaned for data analysis." (Source: NOAA Snow Crab Data)

**Data Gathering:** The dataset is publicly available via Kaggle and has been collected by the NOAA. The file itself is in csv format, and, therefore, it will be easy to create the data frame for analysis by using the read\_csv() function. This is very useful in that it is one of the most common data analytics tools and therefore highly accessible to anyone seeking to use this dataset. This dataset is also from a highly respectable source (the NOAA) and therefore should allow for high value analysis.

**Data Analytics Tools and Techniques**: The data analysis technique that I will use to analyze this data is a Multiple Linear Regression.

**Justification of Tools/Techniques:** A Multiple Linear Regression is appropriate for the analysis of this data because we are trying to see if there is a relationship between multiple variables and a quantitative numeric dependent variable (the catch per unit effort measure of abundance)

**Project Outcomes**: The key anticipated outcome of this analysis is that I will create a Multiple Linear Regression model to determine whether there are any factors/variables related to the catch per unit effort of snow crab in the Eastern Bering Sea snow crab dataset.

**Projected Project End Date**: 9/30/2024

**Sources**: NOAA Snow Crab Data: https://www.kaggle.com/datasets/mattop/snowcrab

**Course Instructor Signature/Date:**

The research is exempt from an IRB Review.

An IRB approval is in place (provide proof in appendix B).

Course Instructor’s Approval Status: Approved

Date: 9/24/2024

Reviewed by:

Comments: Click here to enter text.

