STAT 4355 Project Proposal

Names: Yoshita Narang, Maximus Buckelew, Kenan Stredic

Dataset: Shark Tank Companies

Team Name: The Sharks

INTRODUCTION

The Shark Tank Companies dataset is a valuable resource for investors, entrepreneurs, and business analysts. It offers a comprehensive analysis of companies that have appeared on the popular reality TV show Shark Tank, providing information on their funding, products or services, and overall market success. The dataset also includes details on the sharks themselves, such as their investment preferences and net worth.

By examining the success rates and trends of Shark Tank companies, the dataset can be used to identify successful patterns in pitches and investments, as well as to analyze the factors that contribute to a company's success after appearing on the show. Additionally, the dataset may have practical applications in marketing and advertising, as it provides insights into the types of products and services that are popular among consumers.

OBJECTIVE

We will analyze whether a proposal will get a 'deal' or not based on the category of the product/service, the location of the company, the amount of stake asked for, the amount of money asked for, the company valuation, and the episode number.

DATA VARIABLES

1. **Category (Discrete):** We use this to organize the products and services into factors of overall categories.
2. **Location (Discrete):** The location of the company creating the product or service.
3. **Percent of Stake Asked For (Ranges from 3% to 100%):** The percentage of the company that the entrepreneur wants to retain ownership over when negotiating with the investors.
4. **Amount of Money Asked For (Ranges from $10,000 to $5,000,000):** The amount of money the entrepreneurs request from the investors for funding.
5. **Company Valuation (Ranges from $40,000 to $30,000,000):** The current value of the entrepreneur's company before appearing on Shark Tank.
6. **Episode Number (Ranges from episode 1 to episode 29):** The current episode of Shark Tank.

COURSE OF ACTION

- ➔ **Data Cleaning -** First we will remove the description, entrepreneurs, and website since it should not impact our data given our goal and the lack of repeated values in these columns. Then, we will remove N/A variables, since we have an abundance of data and a low amount of N/A data.
- ➔ **Exploratory Data Analysis -** We plot each variable against the response variable to analyze correlations between each variable and the response individually.
- ➔ **Finding Strength of Correlation between Variables -** Check the variables' correlations with each other to see if it is great enough to affect our coefficient and p-value estimates in our fitted model. We'll do this via the Variance Inflation Factor test.
- ➔ **Build Multiple Linear Regression Model -** Now we use the information we obtained during the last step to fit our multiple linear regression model.
- ➔ **Anonva Fitting-** Using each model we will perform ANOVA with a significance value of .05 to find our best-performing model.
- ➔ **Draft a Conclusion Analyzing the Data -** Quantify the impact of each variable on whether or not our sharks grant contestants a 'deal' and draft a final conclusion.

RESPONSIBILITIES

Our team has devised a plan to execute our project efficiently. Our chosen dataset, Shark Tank Companies, comprises five key variables. We aim to determine the most appropriate variables to include in our regression model through a concerted effort. Each team member will be assigned one to two variables and conduct an individual analysis to establish their correlation with the response variable, which will determine whether or not an entrepreneur will receive a 'deal'.

After completing the analyses, we will evaluate the outcomes to determine the most suitable variables to integrate into our regression model. Subsequently, we will develop the most effective conclusions based on our analysis. Every team member is responsible for creating graphs and charts that relate to their variables and analyses. We will hold weekly meetings, every Sunday evening, throughout the project's duration to ensure successful project completion.