Capstone Project Data Analysis

Tou Lo

MIS480 – Capstone Business Analytics and Information Systems

Colorado State University – Global Campus

Doctor Justin Bateh

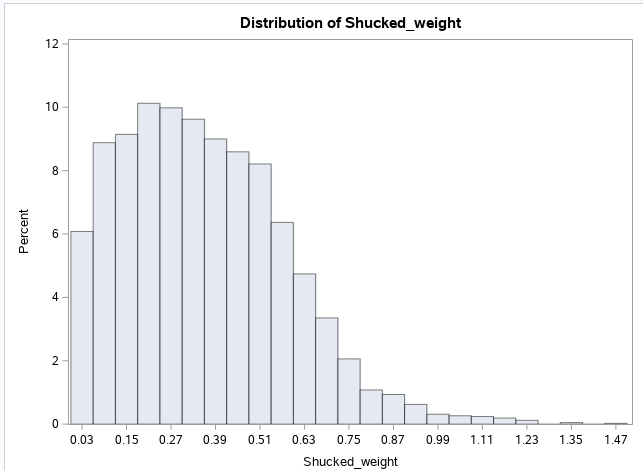
March 15, 2020

Capstone Project Data Analysis

Utilizing business intelligence tools can improve the efficiency of the operations of any organization. For abalone farms, utilizing historical harvest data can point to ways that operations can decrease waste, improve effectiveness, and research new ways to do things that keep an organization on the cutting edge of farming. The main tool that will be used in the analysis of this data will be SAS Studio. This tool has the capability of delving deeply into the data with many tools and also very customizable through coding the perfect data models. By utilizing SAS Studio correctly and harnessing its power, actionable plans will be able to be created from this data. With this power, the operations of the organization can be improved in every way.

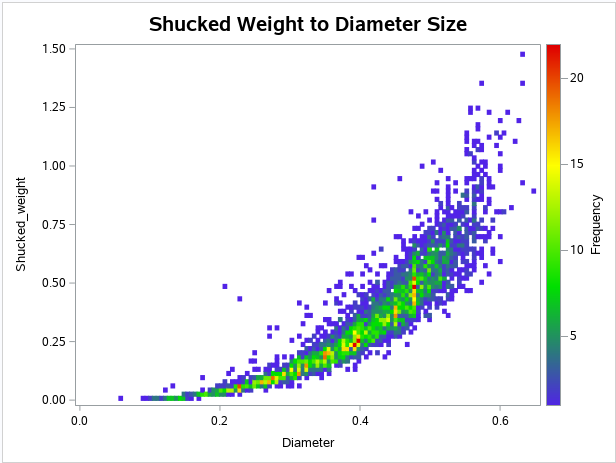
**Analysis of Data**

With the definition of the target made, which is to improve the operations of abalone farming, we can move onto the next step of data analysis. One of the first steps of using data analytics to explore the dataset to discover if there is relevant information within the data. Data exploration should be an iterative process to find answers to a query (Bagozi, Bianchini, De Antonellis, Garda, & Marini, 2019). To explore the data a few models were run on the data.



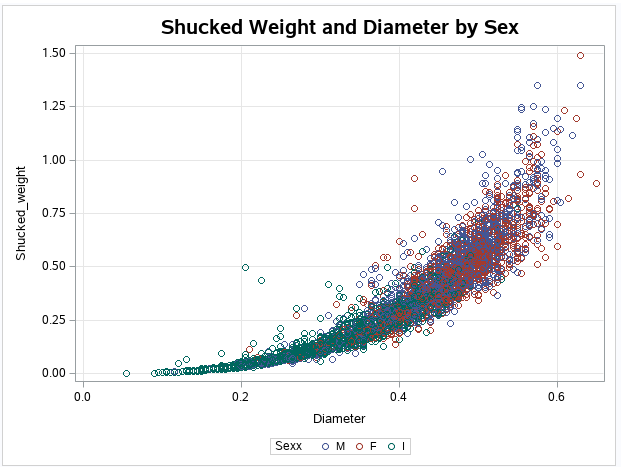
*Figure 1.* A distribution of the weight of shucked meat from abalone.

In Figure 1, we can see that the majority of the shucked meat from abalone comes in at .15 to .51 grams. This data shows us the range of the weight of meat that we should be aiming for when growing abalone. To find the variables that correlate to these weight numbers, we run some more models.



*Figure 2.* Weight of meat harvested by diameter of abalone.

Using a heat map, in Figure 2, we discover that there is a correlation of the diameter of an abalone to the amount of meat that comes from it. Analyzing this model compared to the distribution of meat harvested we can see that there is an ideal size for harvesting the most meat.



*Figure 3.* Adding sex to shucked weight and diameter model.

Adding the sex of abalone to discover if the sex of the animal correlates to any of the data reveals that there is opportunity here. Figure 3 shows that infant size abalone are a terrible group to harvest. This data can be attacked from many angles such as the relation of meat harvested to size and sex of abalone. Age of the abalone can be included as size and amounts of rings can be used to estimate the age of abalone (Naylor, 2015). Just within these three models, we can see that there is enough relevant and useful data that can be discovered within this dataset to accomplish the goals of this project.

**Using BI Tools in the Project**

SAS Studio is a very powerful tool that can extract information from data by creating graphs and models from that data. Looking at the examples in the figures above, it is proven that this BI tool has the capability of taking large amounts of data and modeling it in ways that will help extract actionable strategies. This will be a key part of this project since it will identify areas that operations at the abalone farm can be improved.

**Challenges of Analyzing the Data**

One of the challenges of this data is that there are so many ways it can be modeled to be useful. This will require that more data discovery be performed on this data. Bagozi, et al., (2019) suggests that this is normal and the exploration of data is a multistep process requiring many iterations to focus in on the correct information. By continually analyzing the data new and exciting information will jump out from the jumble.

**Conclusion**

Analysis of data will be an important part of this project and it is vital that actionable information can be extracted from this data. Ensuring that the analysis is thorough will be an important part of extracting good and complete information for this project. By completing these goals, the processes and operations of abalone farming can be improved. With a good tool such as SAS Studio, this process will be made much easier.

References

Bagozi, A., Bianchini, D., De Antonellis, V., Garda, M., & Marini, A. (2019). A Relevance-based approach for Big Data Exploration. *Future Generation Computer Systems*, *101*, 51–69. https://doi.org/10.1016/j.future.2019.05.056

Naylor, R. (2015). Literature review of abalone ageing techniques. *Ministry for Primary Industries.* Retrieved from https://www.mpi.govt.nz/dmsdocument/10583-far-201577-literature-review-of-abalone-ageing-techniques