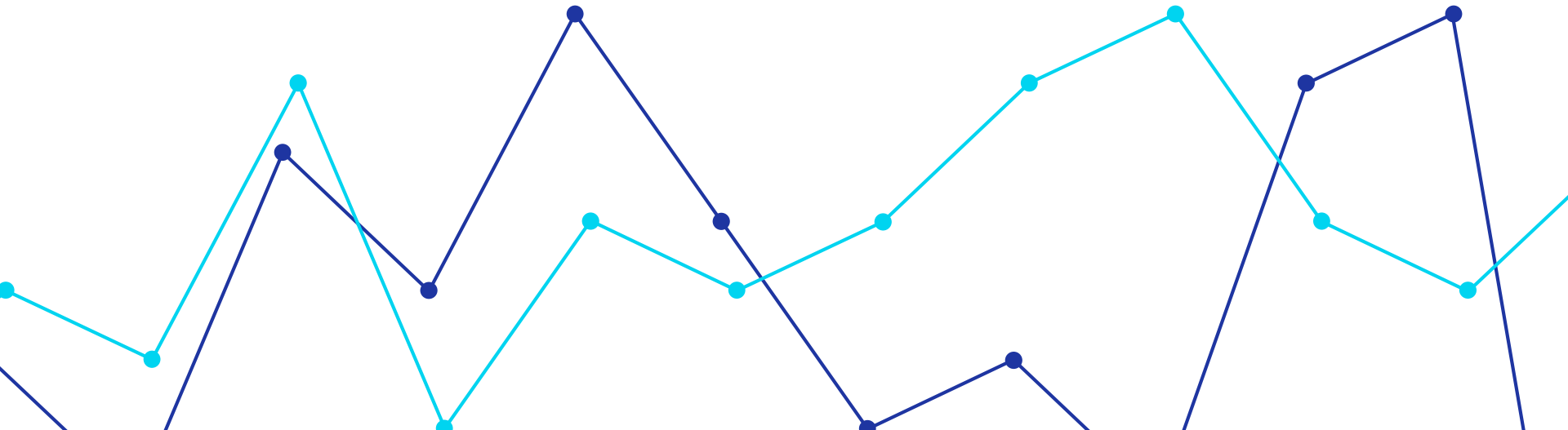


Abalone

Predicting the age of abalone from physical measurements

Rawan Balubaid | Project 1



Whats is abalone ?

Abalone is a popular choice of seafood—a shellfish to be precise—that lives in cold coastal waters around the world. Biologically, abalone is a mollusk belonging to the Gastropod class. In plain English, this means that abalone is technically a type of marine snail.



Project Background

Motivation

The age of abalone is determined by physical measurements. cutting the shell through the cone, staining it, and counting the number of rings through a microscope, a boring and time-consuming task.

Objective

predict the age of the abalone through its physical measurements

Data Sources

UCI



Machine Learning Repository

Center for Machine Learning and Intelligent Systems

This project originated from uci.edu
which contain 4177 of different
abalone information With 9 features
for each of them

Data comes from an original (non-
machine-learning) study : The
Population Biology of Abalone
(_Haliotis_ species) in Tasmania.

Overview

	Sex	Length	Diameter	Height	Whole weight	Shucked weight	Viscera weight	Shell weight	Age
0	0	0.455	0.365	0.308221	0.5140	0.2245	0.1010	0.150	16.5
1	0	0.350	0.265	0.300000	0.2255	0.0995	0.0485	0.070	8.5
2	1	0.530	0.420	0.367423	0.6770	0.2565	0.1415	0.210	10.5
3	0	0.440	0.365	0.353553	0.5160	0.2155	0.1140	0.155	11.5
4	2	0.330	0.255	0.282843	0.2050	0.0895	0.0395	0.055	8.5

`Sex`: M (male), F (female), I (infant)

`Length`: longest shell measurement (in mm)

`Diameter`: measurement perpendicular to length (in mm)

`Height`: with meat in shell (in mm)

`Whole weight`: the whole abalone (in grams)

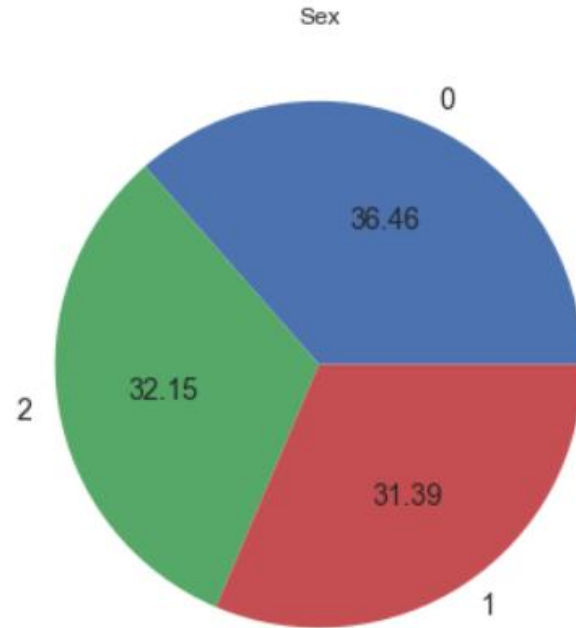
`Shucked weight`: weight of the meat (in grams)

`Viscera weight`: gut weight after bleeding (in grams)

`Shell weight`: after being dried (in grams)

`Rings`: + 1.5 gives the age in years (the value to predict)

Abalone gender dataset is male and female and infant

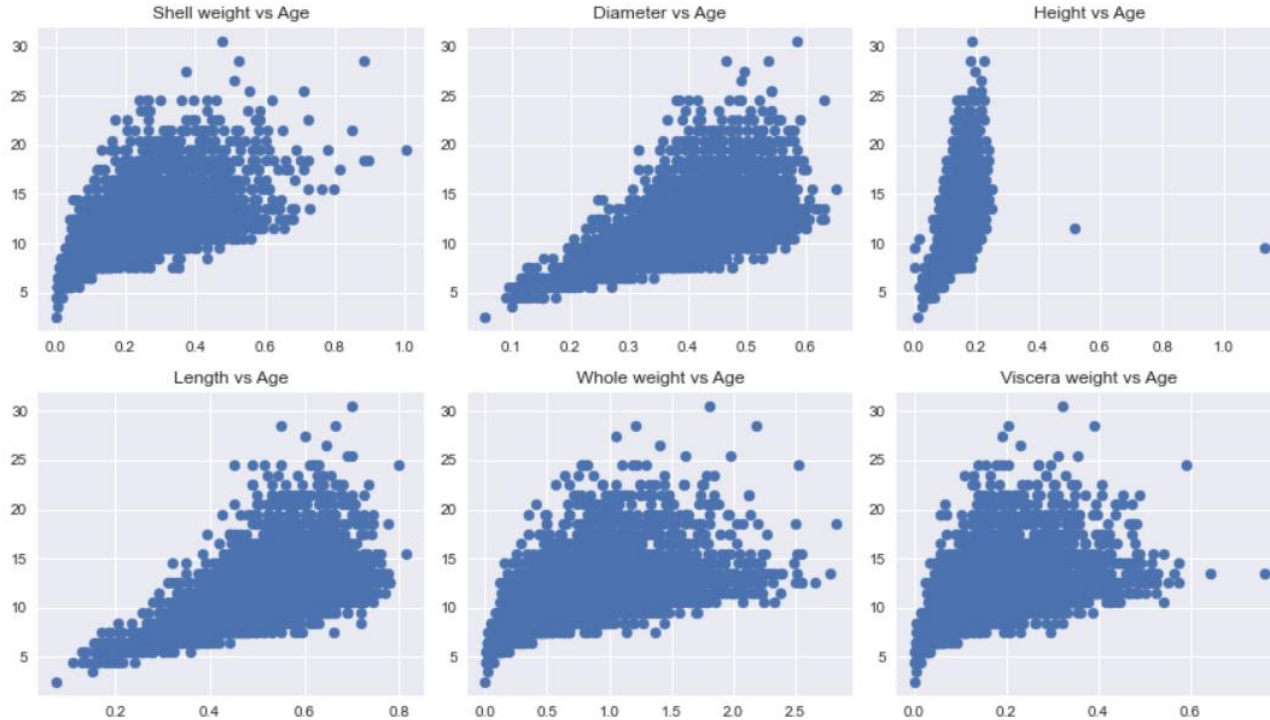


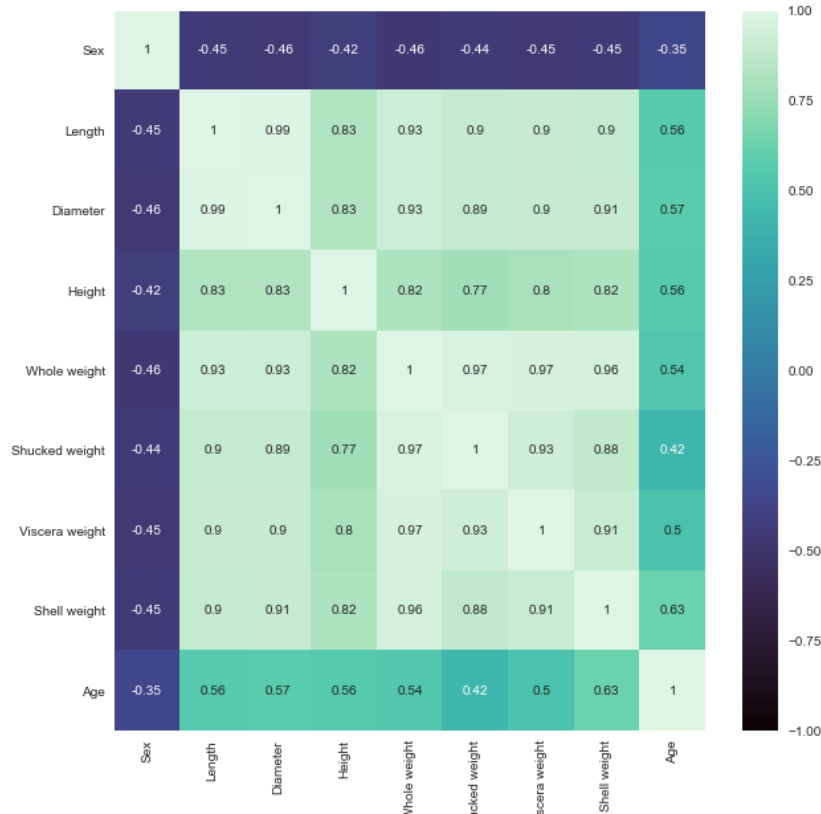
0-> Male

1-> Female

2-> infant

Characteristics analysis and their compatibility with age





In this section we analyzed how the dataset attributes are related and how the independent variables influences the target variable. Our first step in the multivariate analysis was to visualize the correlation matrix in a heatmap

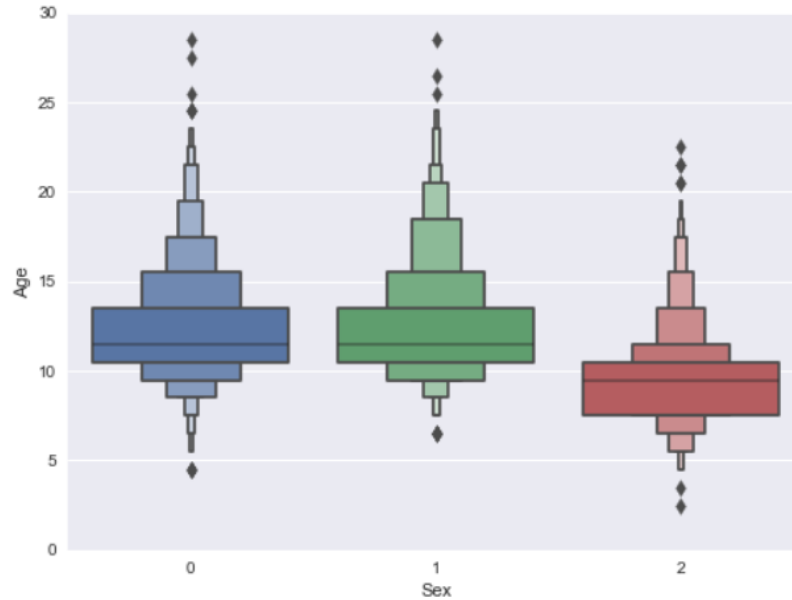
- no negative correlation

- High correlation between Length & Diameter //

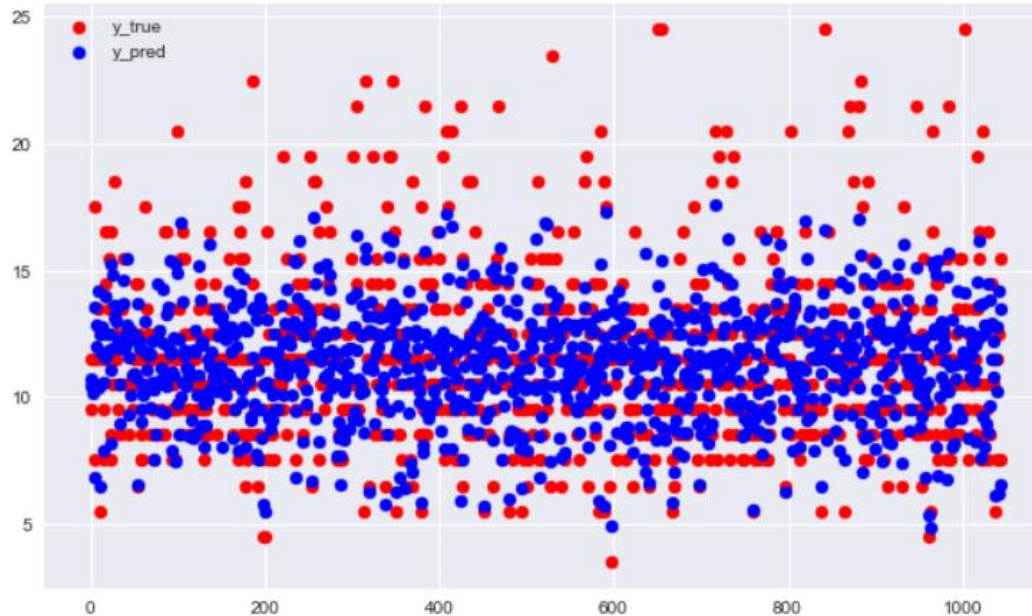
Whole weight & Shucked weight // Whole weight &

Viscera weight // Whole weight & Shell weight

same Distribution between male and female , and most of them age are between 8 to 19



Using Random Forest regression to predict an age of abalone



Train rmse : 1.5269239609800958

Test rmse : 2.1385971780892463

Conclusion

By observing the correlation between the target attribute Age and the independent variables, we conclude that it is possible to build a model to predict the target value in function of the independent attributes.

The weight and height of abalones varies accordingly to age until the adult age, after adult life size and weight stops varying, and after 16.5 years (15 rings) these measurements aren't correlated.

There's no significant differences in size, weight and age between male/female abalones