



Hi Crab, How Old Are You?

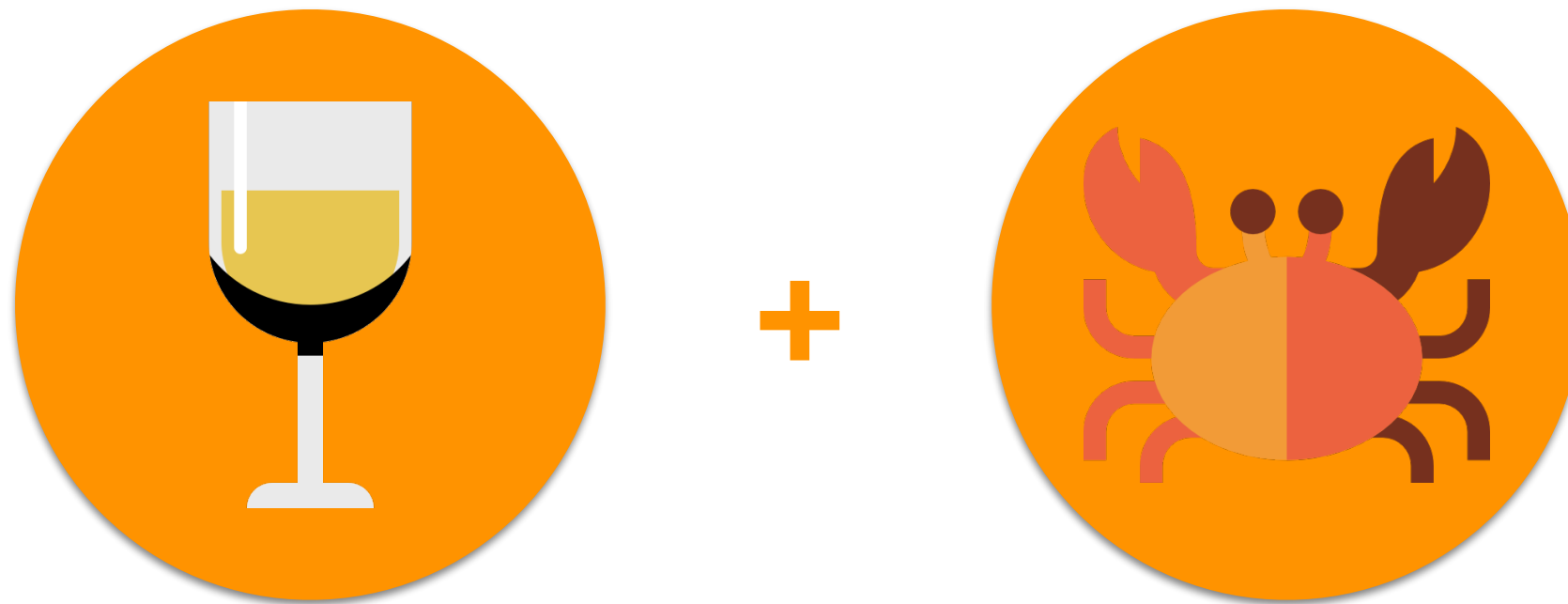
Robbin

A Restaurant in Boston...



Alice: Oh, what a good 2010 Chablis!

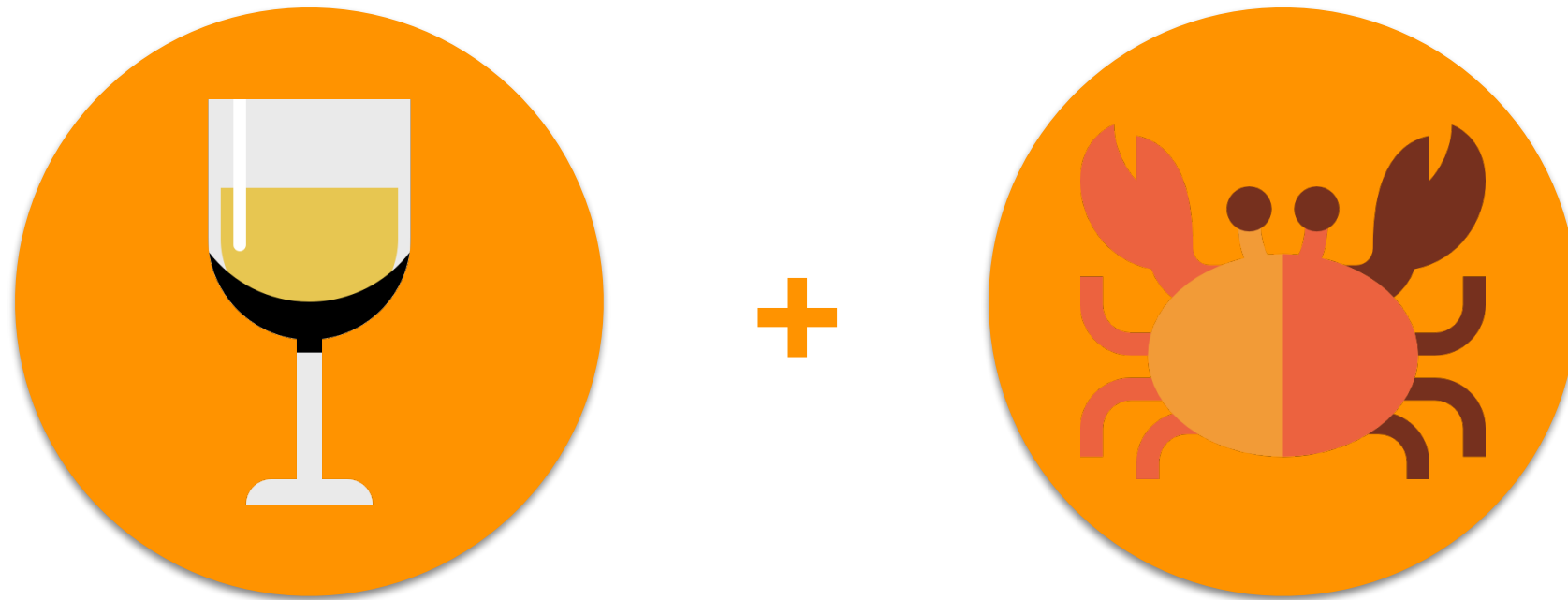
A Restaurant in Boston...



Alice: Oh, what a good 2010 Chablis!

Bob: Perfect match with this 2015 crab.

A Restaurant in Boston...



Alice: Oh, what a good 2010 Chablis!

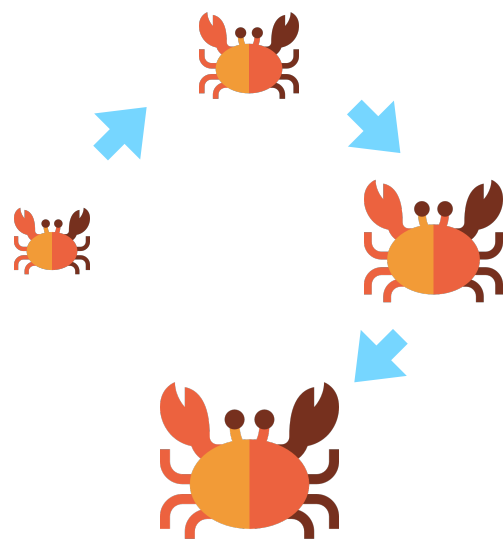
Bob: Perfect match with this 2015 crab.

Alice: Oh, you are genius! How do you know the age of a crab?

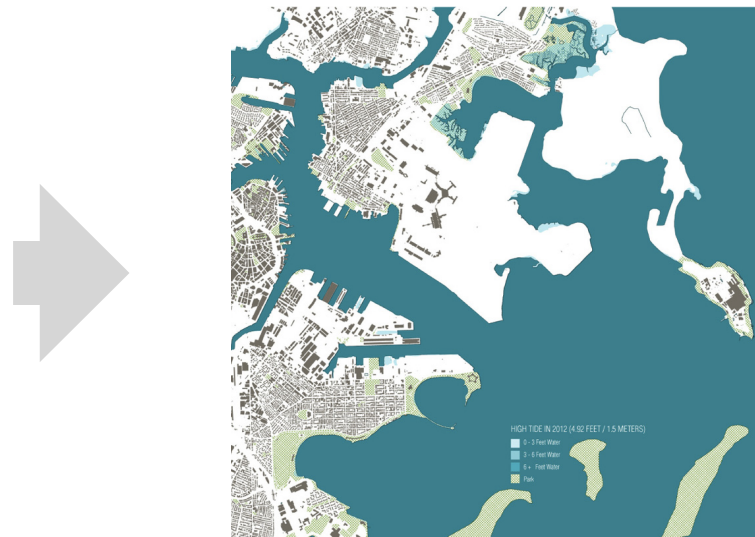
How to predict the age of a crab?



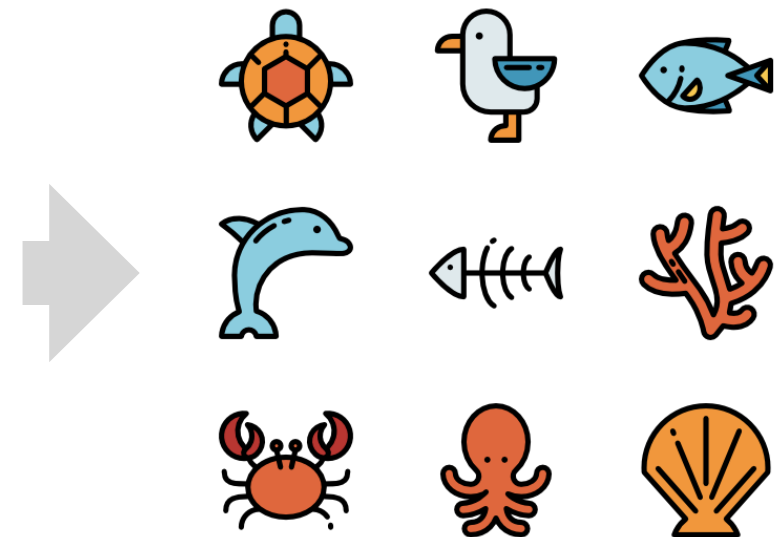
Why Do We Predict the Age of Crabs?



Crab's Lifecycle



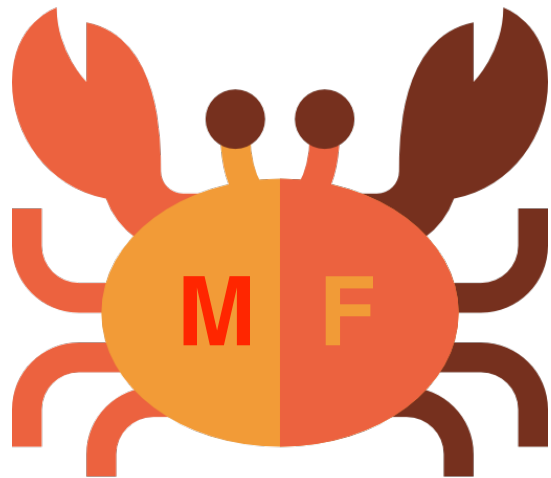
Understand Ecosystem



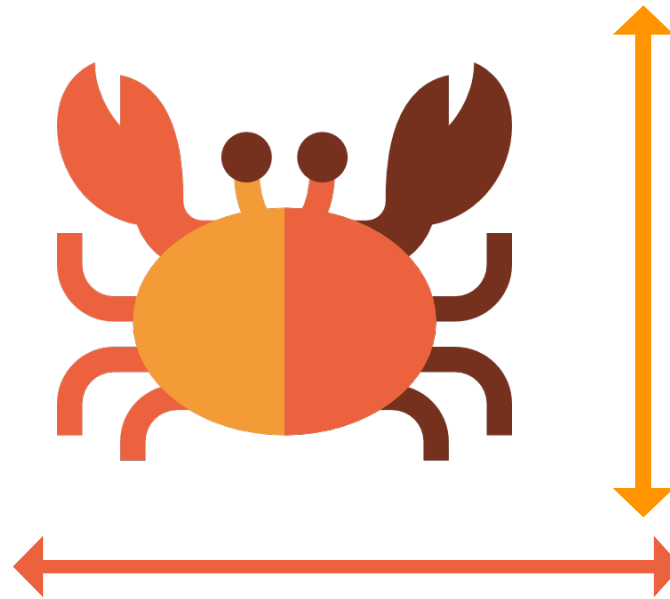
Environmental Protection

What is Used to Predict?

Gender



Size

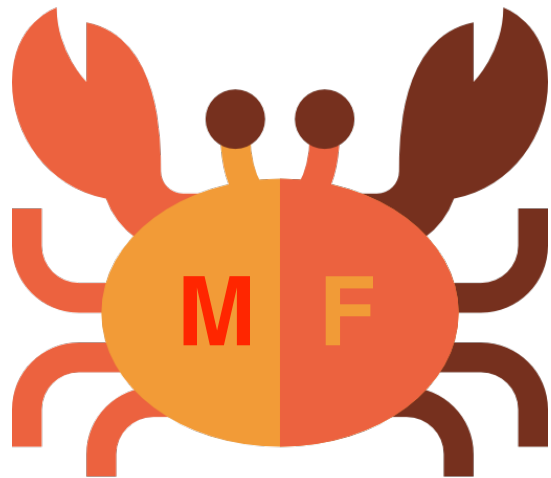


Weight



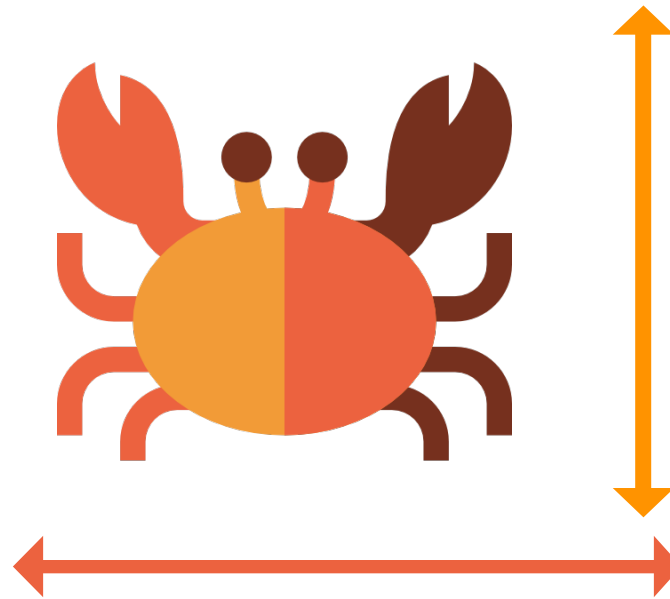
Let's Try!

Gender



Female can live longer?

Size



The larger the older?

Weight



The heavier the older?

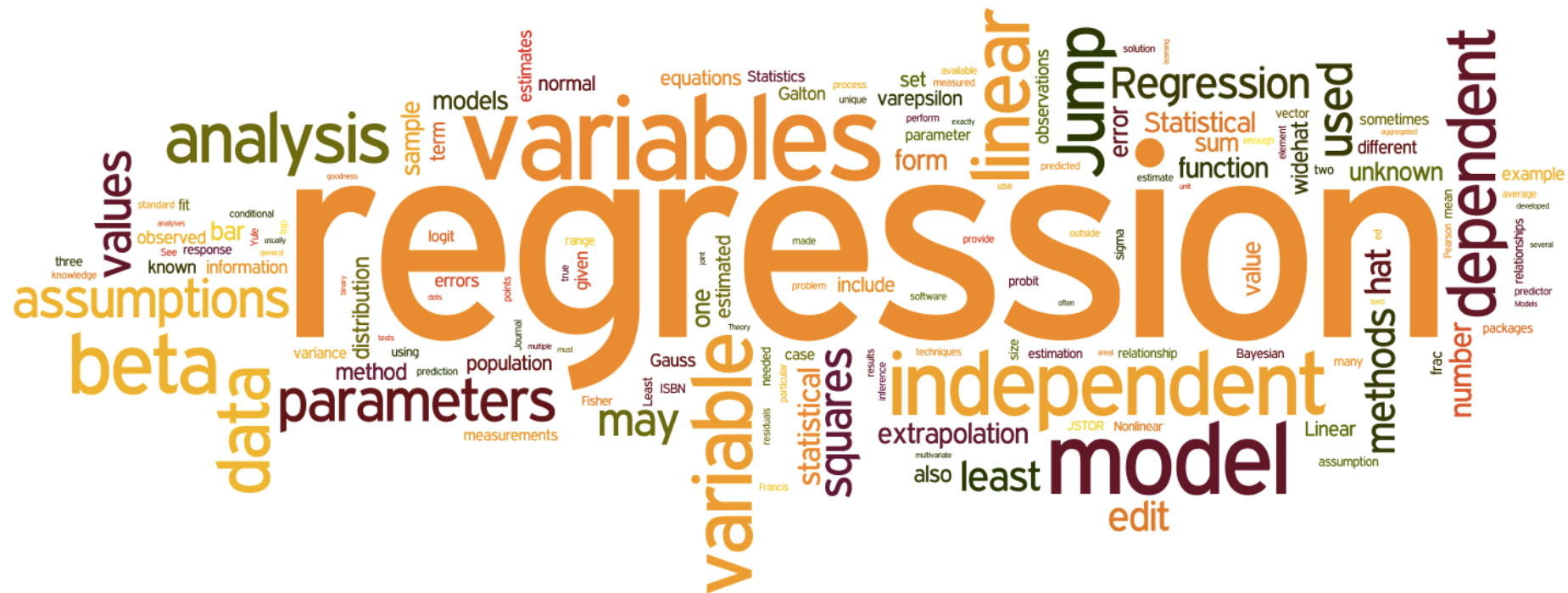
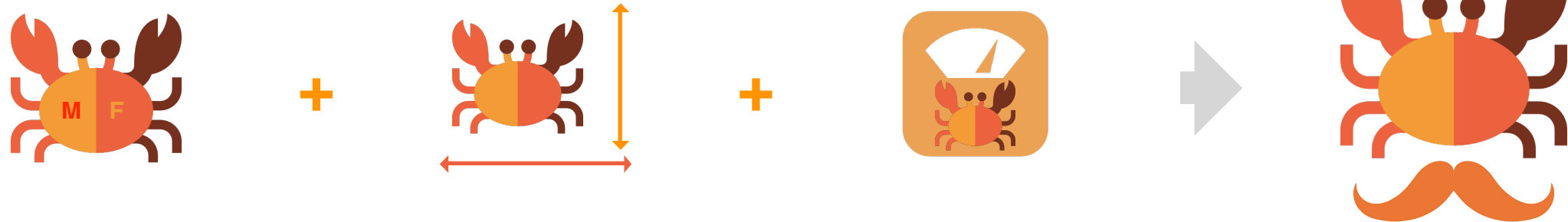
Why Not Use Machine Learning!

Gender

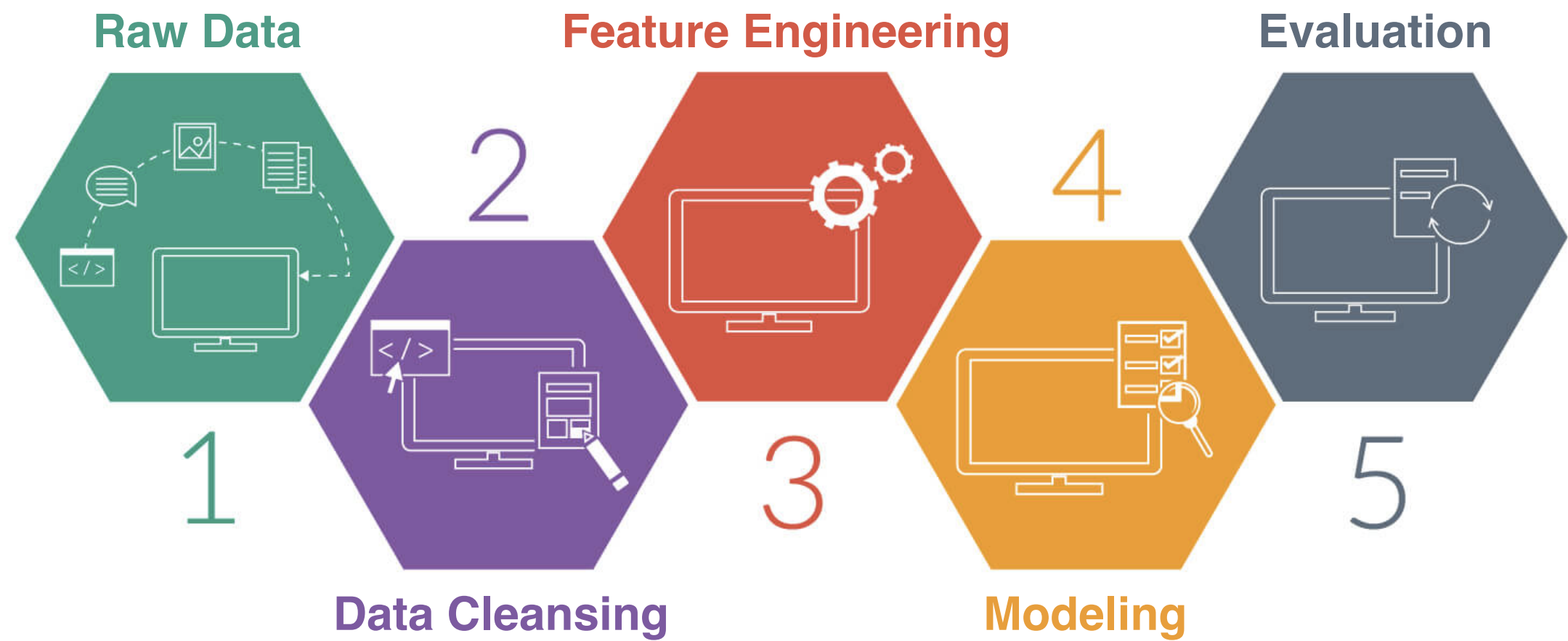
Size

Weight

Age



Workflow



Raw Data

PDF



data.pdf

OCR Processing



OCR tools:

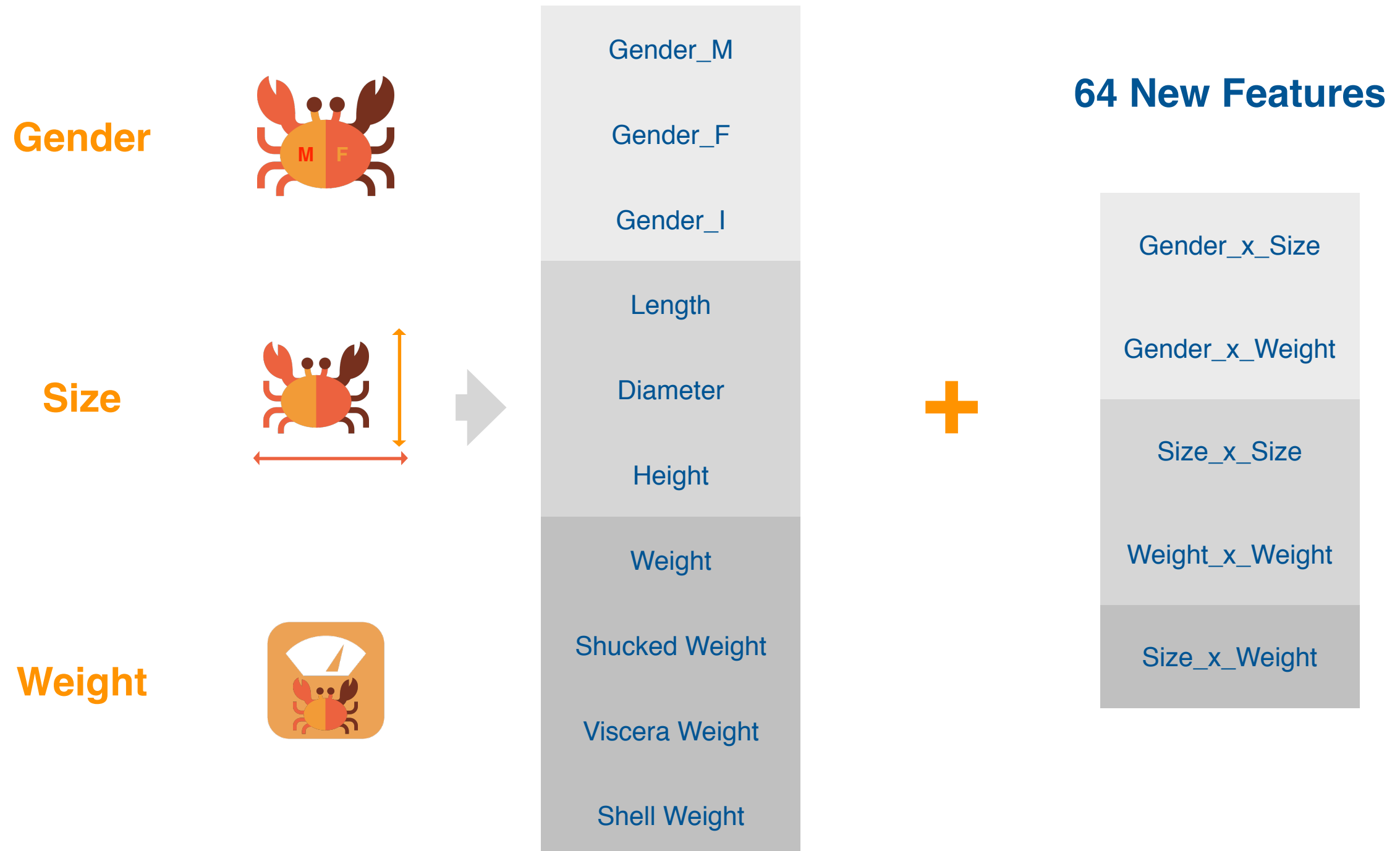
- tabula-py
- pdfplumber

CSV



data.csv

Data Cleansing and Feature Engineering

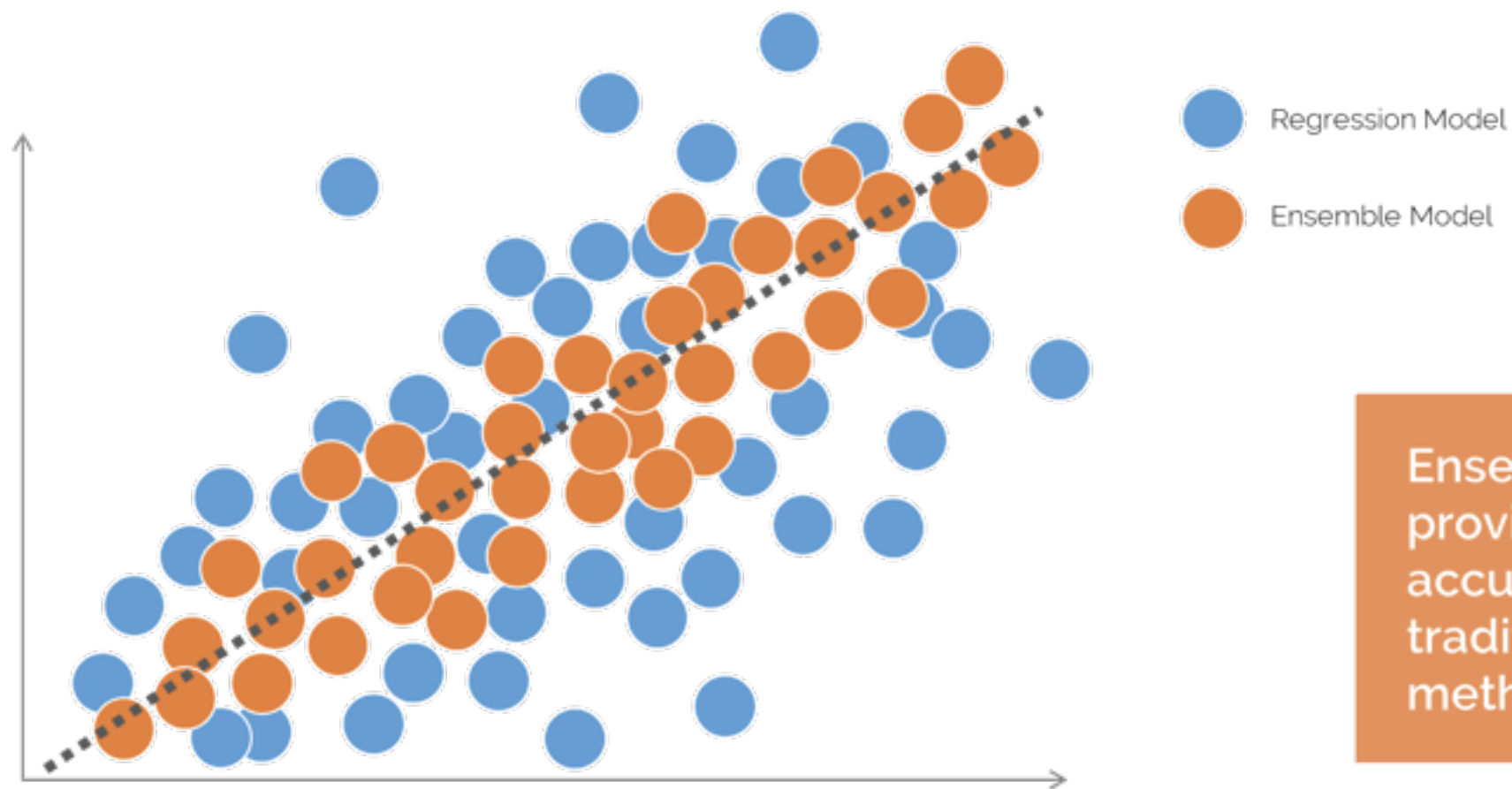


Model Methodology

Ridge + Lasso + XGBoost + BayesianRidge

||

Ensemble Model



Ensemble models provide more accuracy than traditional regression methods

Performance

R^2

58% of
explained
variance

58%

Mean Absolute Error

1.44 difference
between real age
and prediction
on average

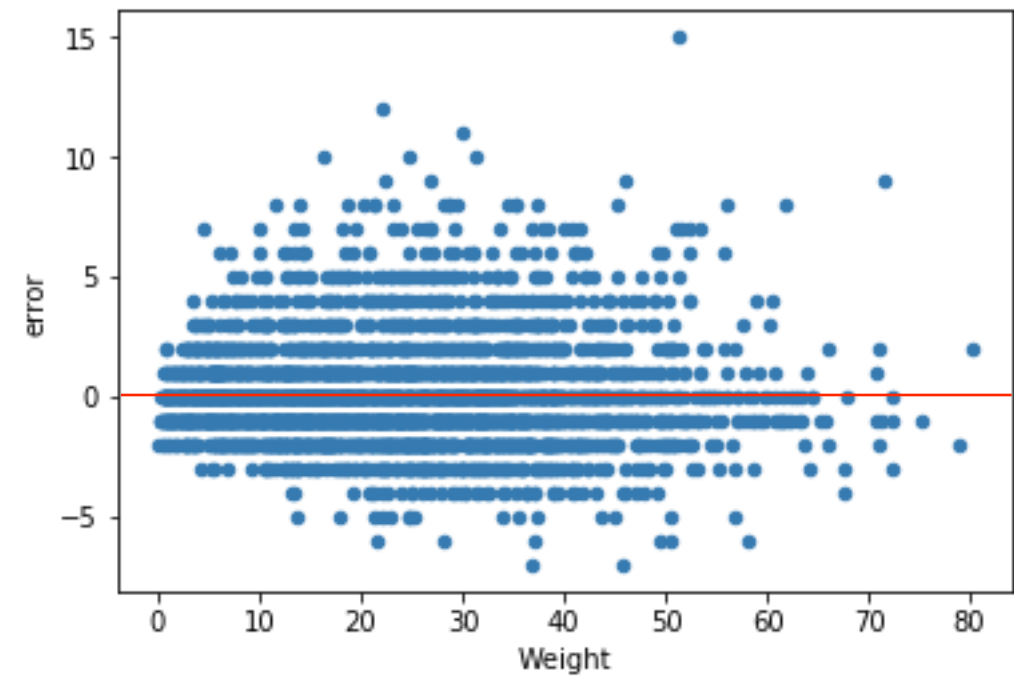
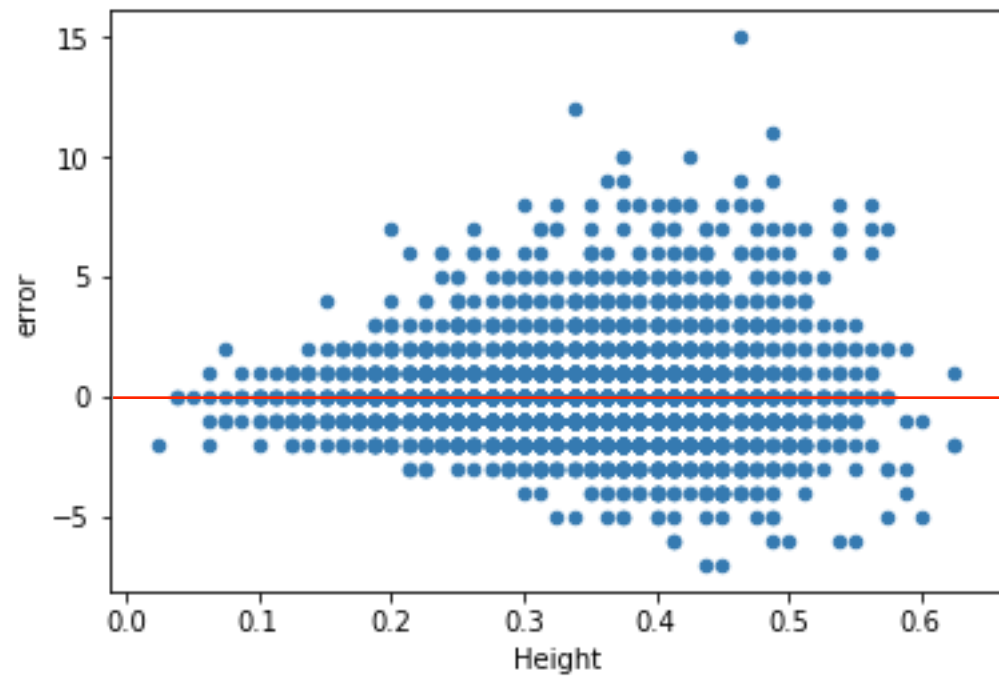
1.44

Root Mean Square Error

2.08 for the
average of root
mean square
error

2.08

Appendix - Residual Plot



Thank you!

Q & A