

# PSTAT131-HW2

2022-10-15

```
library(tidyverse)
library(tidymodels)
```

```
abalone <- read.csv("abalone.csv")
```

1

```
abalone["age"] <- abalone$rings + 1.5

# Assess and describe the distribution of age.
```

2

```
set.seed(1234)

abalone_split <- initial_split(abalone, prop = 0.80, strata = age)
abalone_train <- training(abalone_split)
abalone_test  <- testing (abalone_split)
```

3

We should not use rings to predict age because the two variables would be very highly correlated. This is because age is equal to rings + 1.5.

4

5

6

7