

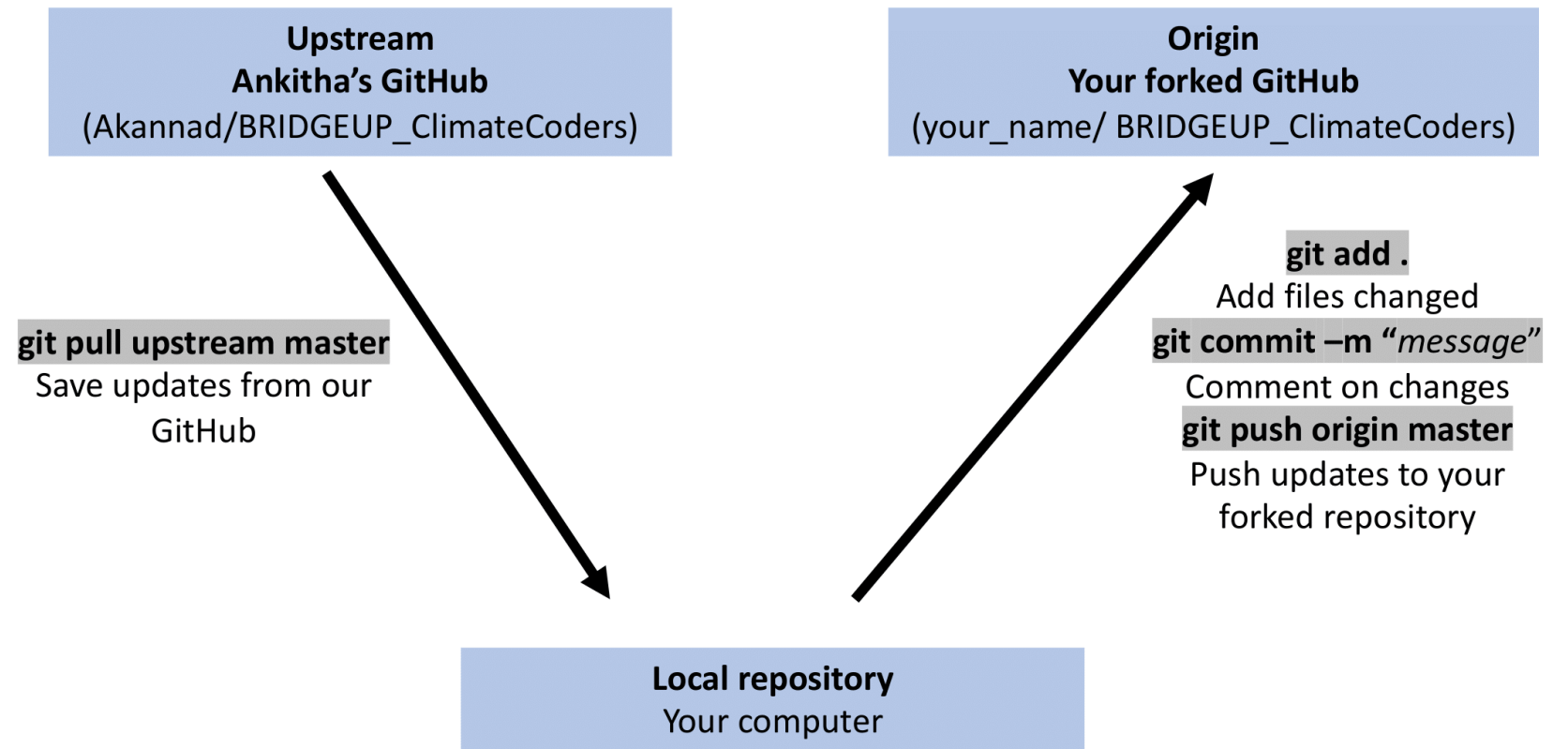
# LINEAR REGRESSION ON PYTHON

UNIT 3: RECONSTRUCTING  
CORAL CORE DATA

APRIL 2<sup>ND</sup> 2020

# HOUSEKEEPING

- Personal check-in
- Final presentations are cancelled
- Restructuring class



## PLAN FOR TODAY

- ❑ Complete time conversion function
- ❑ Linear regression in Python
- ❑ Update lab notes
- ❑ Exit survey

## CONCEPT CHECK-IN

$$A = 7.5891$$

Which of the rounding functions would give you the following results:

❖ 8.0

❖ 7.0

❖ 7.69

❖ 7

# FUNCTION

`def convert_dates ( any dataframe ):`

**your script**

`return` new dataframe with an additional year and month column

**Challenge:** Add in a “Day” column where all the entries are 1

```
>>> def add(a, b):  
    c = a+b  
    return c
```

```
>>> print("The sum of a and b is")  
The sum of a and b is  
>>> print(add(4, 5))  
9
```



SHARE OUR  
TEACHABLE  
MACHINE  
PROJECTS

# LINEAR REGRESSION = MACHINE LEARNING?

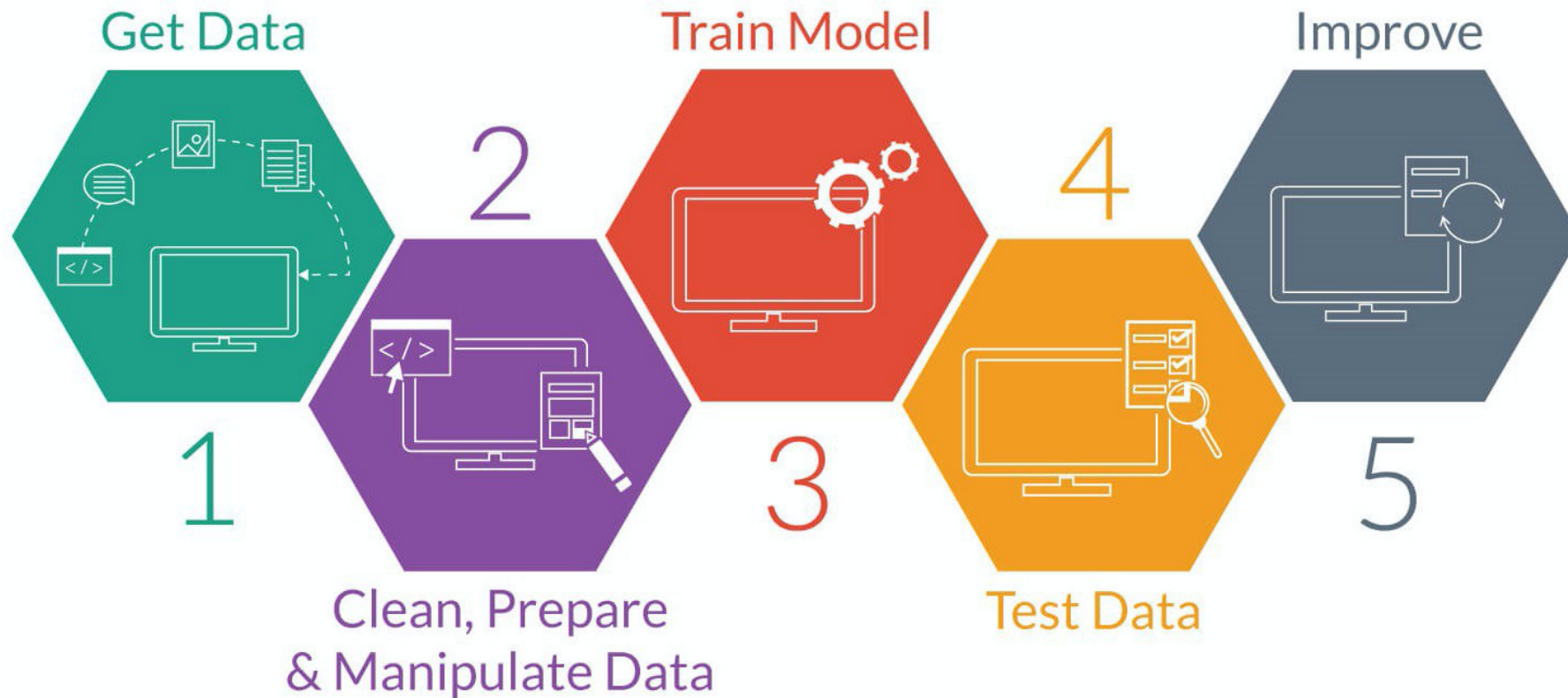
In break-out rooms,

- Use your machine learning diagram from last class, to brainstorm the steps that go into a machine learning algorithm based on linear regression
- Draw/ present your ideas on the whiteboard

Working offline:

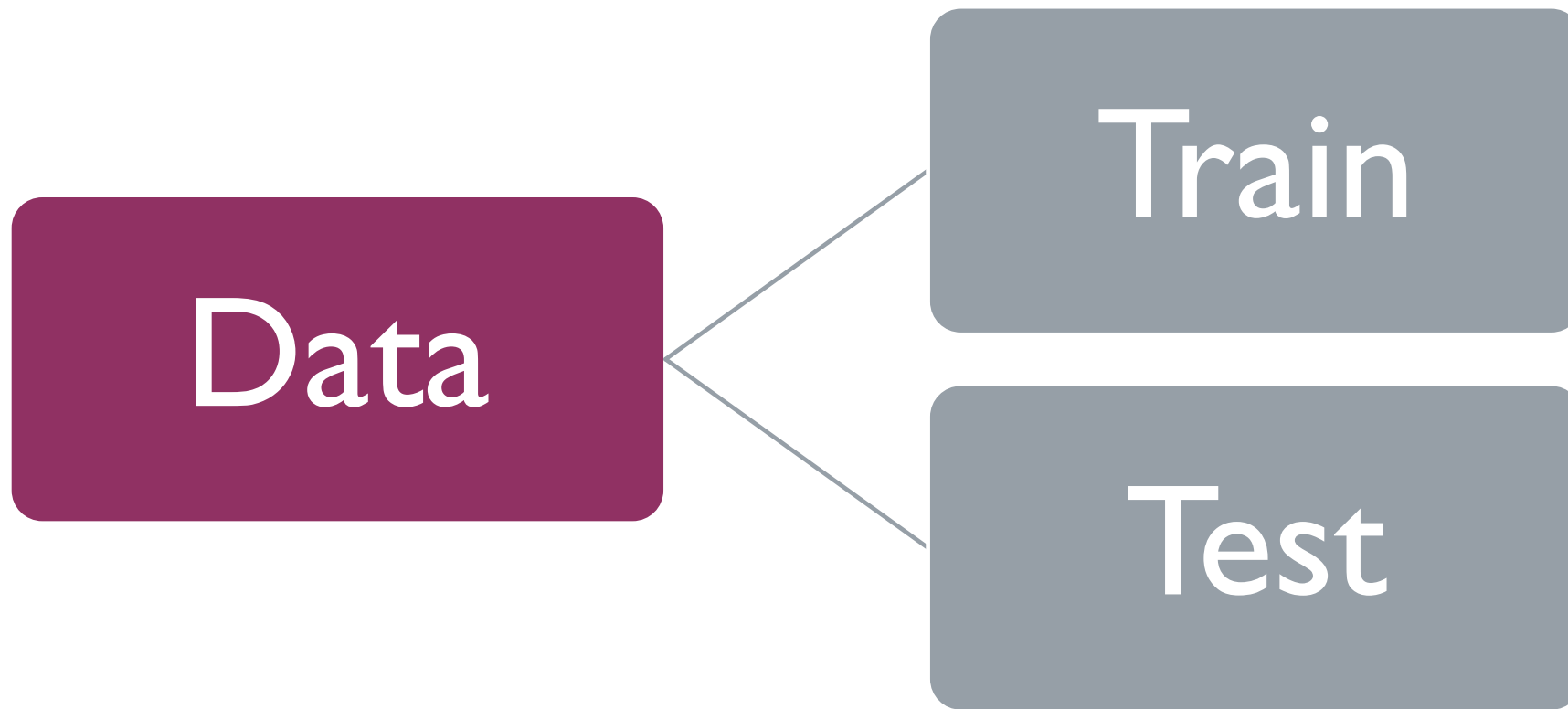
- [https://www.youtube.com/watch?v=Gv9\\_4yMHFhI&list=PLblh5JKOoLUICTaGLRoH\\_QDuF\\_7q2GfuJF](https://www.youtube.com/watch?v=Gv9_4yMHFhI&list=PLblh5JKOoLUICTaGLRoH_QDuF_7q2GfuJF) [2:30 – 8:03]

# MACHINE LEARNING STEPS



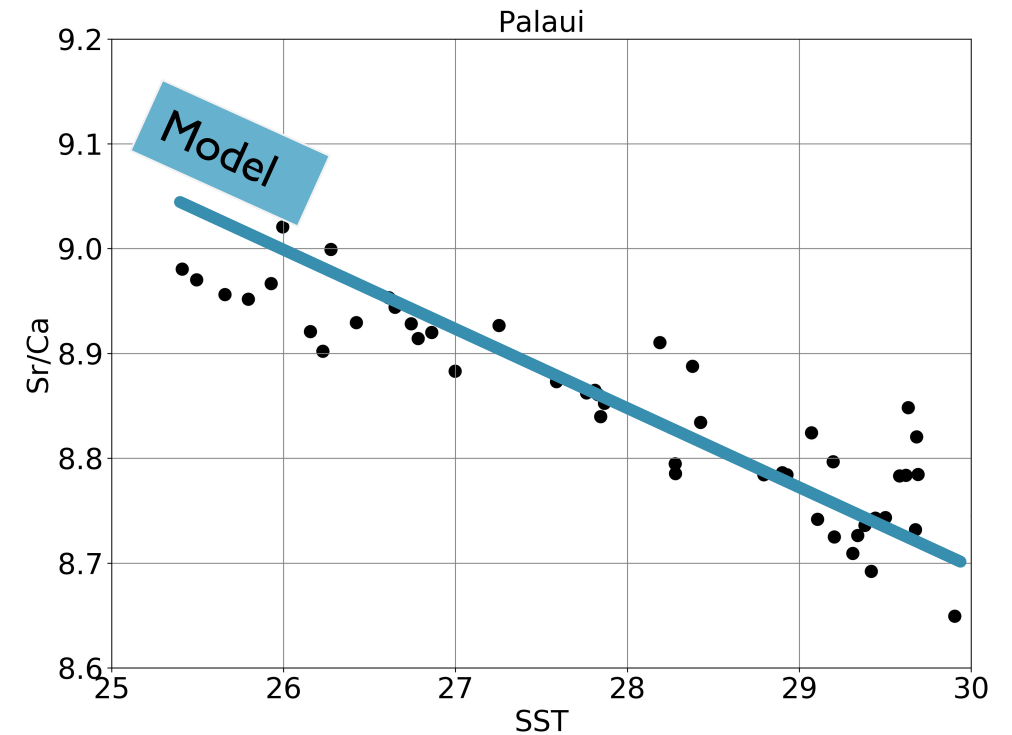
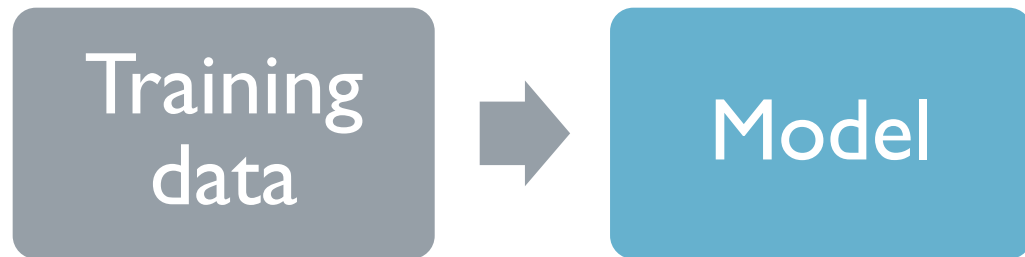


## PREPARE DATA

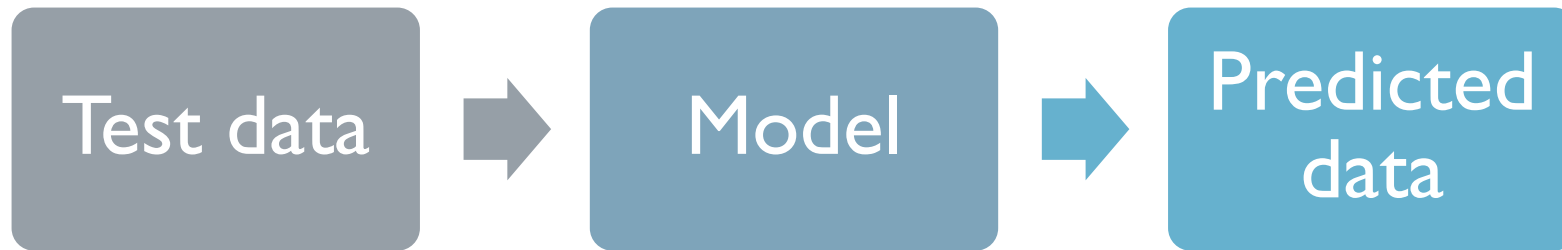


# TRAIN MODEL

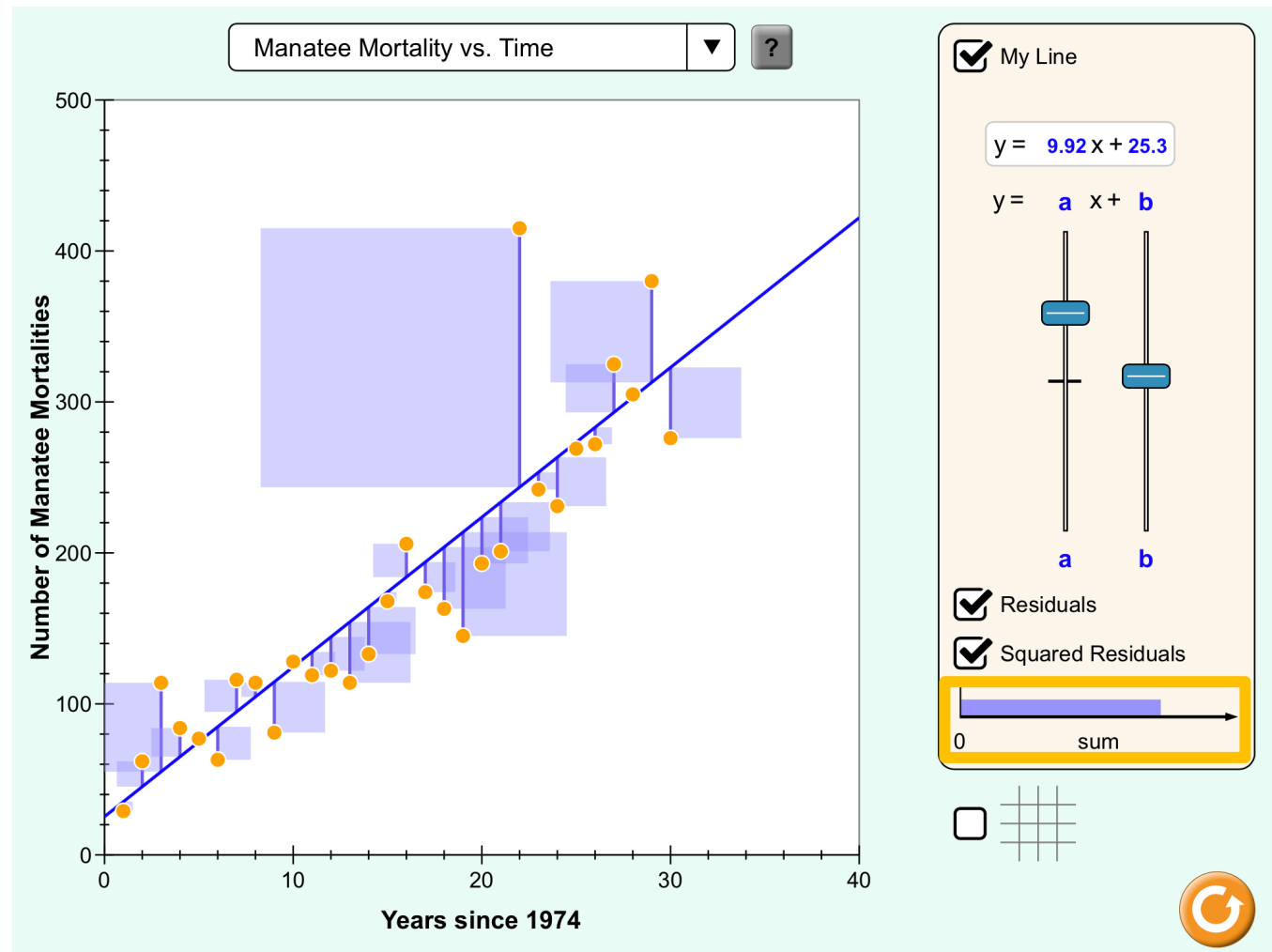
We know our data has some linear relationship but we want to find the equation of our best-fit line. This is where **linear regression** comes in to create a model that best represents our data



## TEST MODEL



# HOW DID THE MODEL DO?



# HOW DID THE MODEL DO?

$N$  = number of test data points

Average squared residual =  $\frac{\text{sum of squared residuals}}{N}$  → mean square error

Average residual =  $\sqrt{\text{mean square error}}$  → root mean square error (RMSE)



UPDATE LAB NOTES

The image features a dark gray background with a light blue horizontal bar at the top and bottom. A vertical pink line is positioned to the left of the text. The text "EXIT SURVEY" is centered in a white, sans-serif font.

# EXIT SURVEY