Data Science, Machine Learning, & Artificial Intelligence,

An Introduction Illustrated with Linear Regression

A guide by Blake Hament

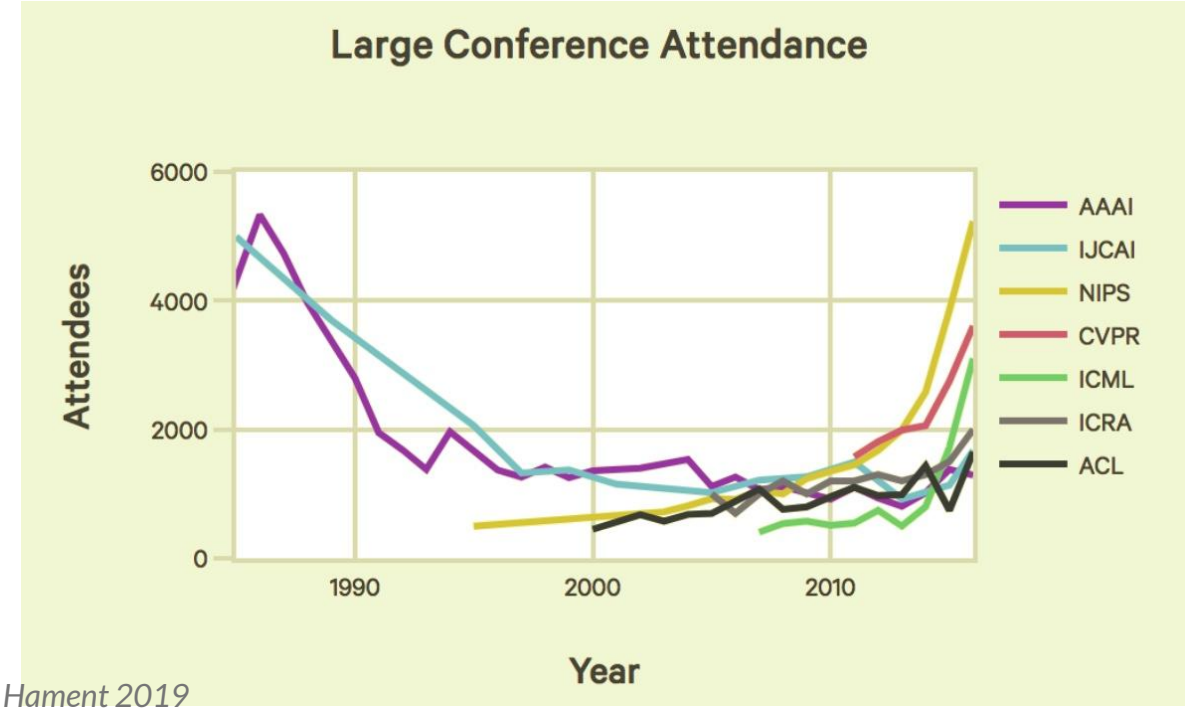
1. Download Data & Code from:

https://github.com/dsinnovated

2. Open an iPython Notebook

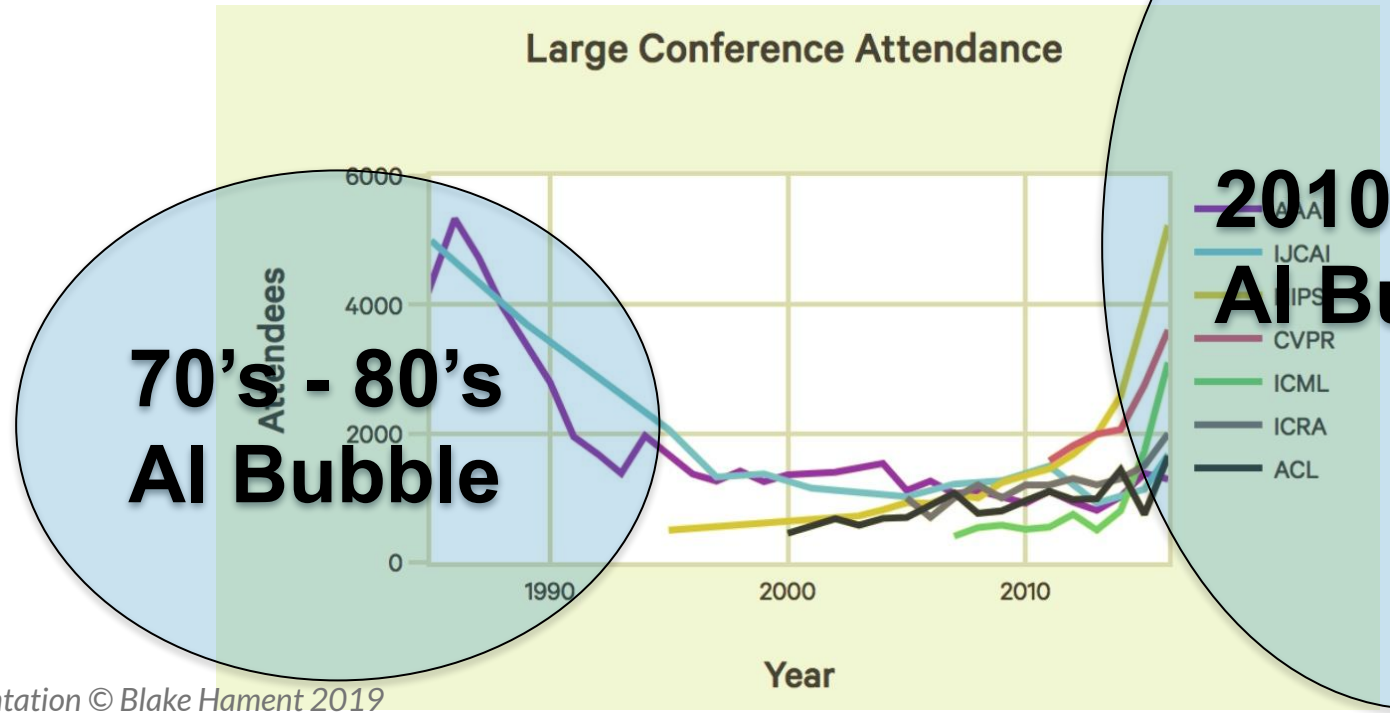
- **Option 1:** Anaconda and Jupyter Notebooks
- **Option 2:** Google Drive and Google Colab Notebooks

AI is “Bubbling” AGAIN



AI Index https://aiindex.org/2017-report.pdf

AI is "Bubbling" AGAIN



AI Index https://aiindex.org/2017-report.pdf

What is A.I.?

What is A.I.?

Code that
LEARNS
and
ADAPTS

What is A.I.?

GENERAL vs SPECIFIC Intelligence (human vs squirrel)

What is A.I.?

```
DRC Fails
```

What is A.I.?

GENERAL vs SPECIFIC Intelligence (human vs squirrel)

—

DS ==

ML ==

AI ==

DS == find and analyze patterns in data

ML ==

AI ==

DS == find and analyze patterns in data

ML == algorithmic learning with computers (to find patterns in data)

AI ==

DS == find and analyze patterns in data

ML == algorithmic learning with computers (to find patterns in data)

AI ==

Tip

Algo's are just "lists of instructions".

Ex-- Simple Algo for "Waking Up":

1. Check window
2. If light go to 3. Else, sleep until awake, then go back to 1.
3. Get out of bed

DS == find and analyze patterns in data

ML == algorithmic learning with computers (to find patterns in data)

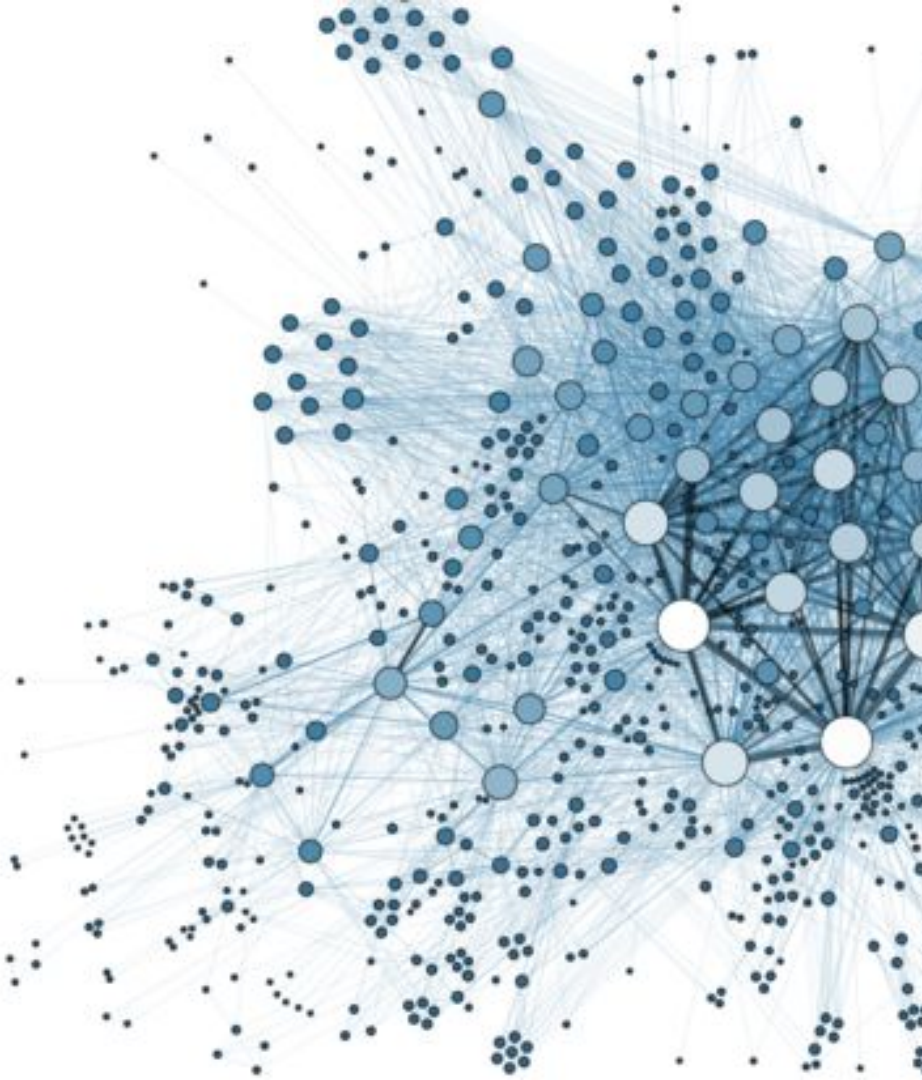
AI == create systems that learn (using ML techniques) and adapt (change their algorithms over time)

DS == find and analyze patterns in data

ML == automated DS

AI == automated ML

What types of
PATTERNS can we
look for in data?



Patterns:

- **Metrics** count, mean, mode, std, skewness, etc.
- **Regression** model **relationships** between features and target (linear, polynomial, logistic, much more!)
- **Classification**
  Group your data points together to make different classes or clusters

Patterns:

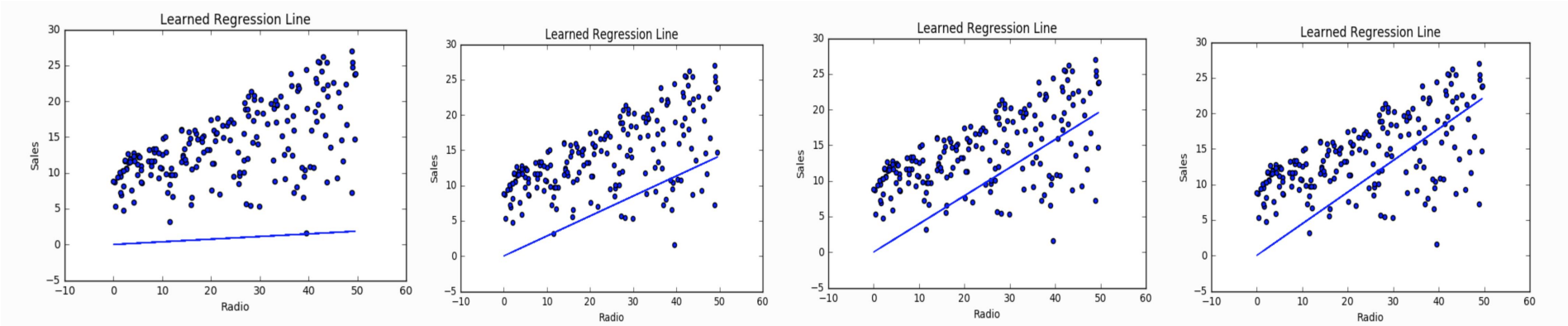
- ➔ **Metri** [illegible], mode, [illegible], skewness, etc.
- ➔ **Regression** model **relationships** between features and target (linear, polynomial, logistic, much more!)
- ➔ **[illegible]tion** **Group** your data points together to make different classes or clusters

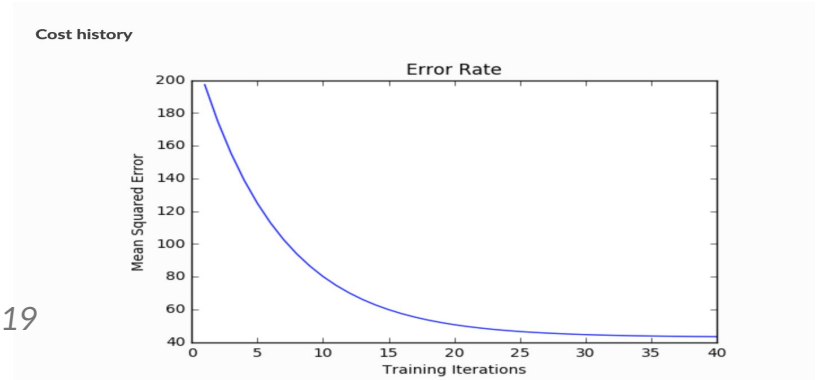
Let's run some code....

Update “a” and “b” to find best

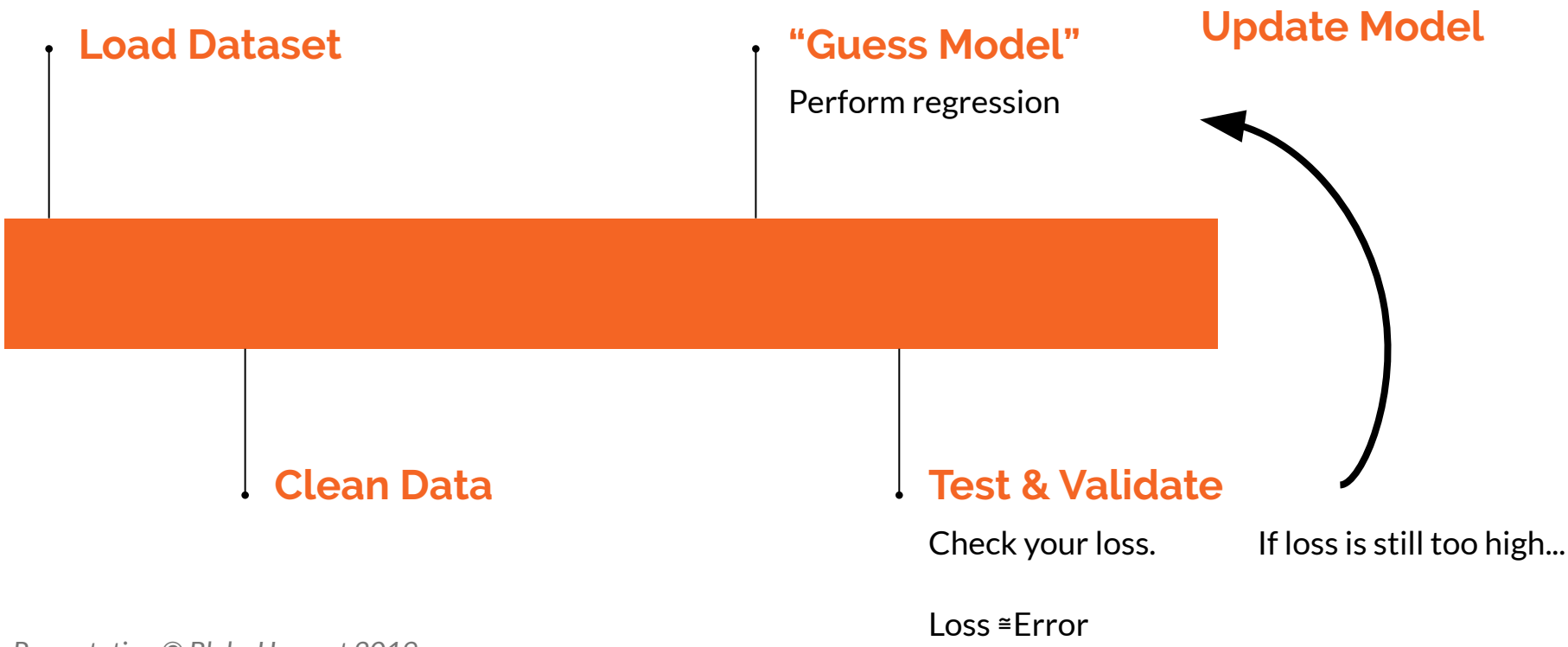
Y = aX + b

Regression with ML



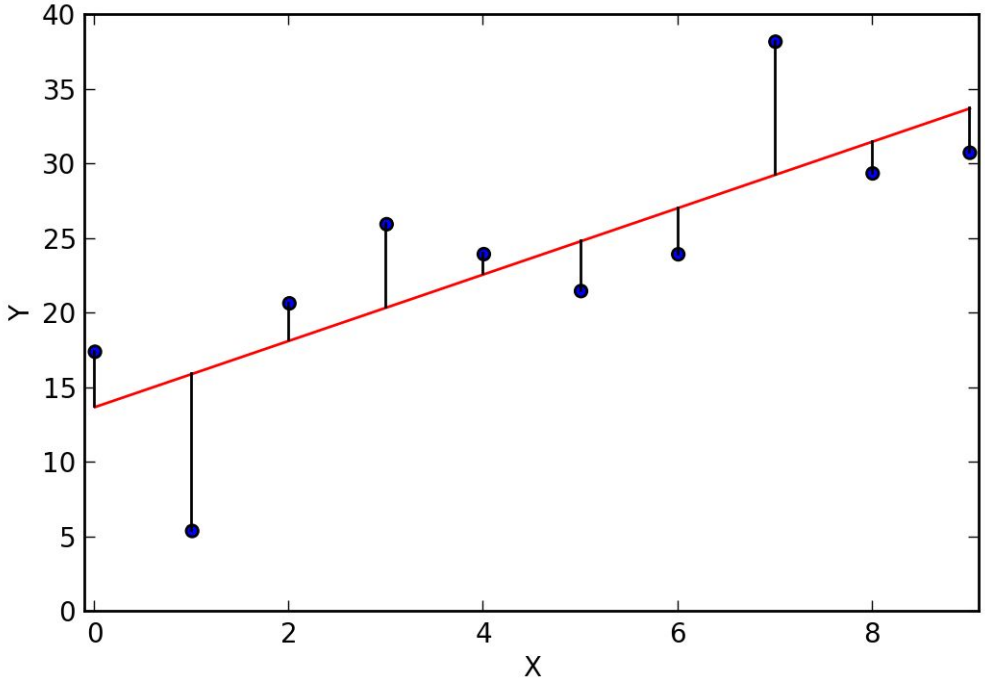


ML Pipeline



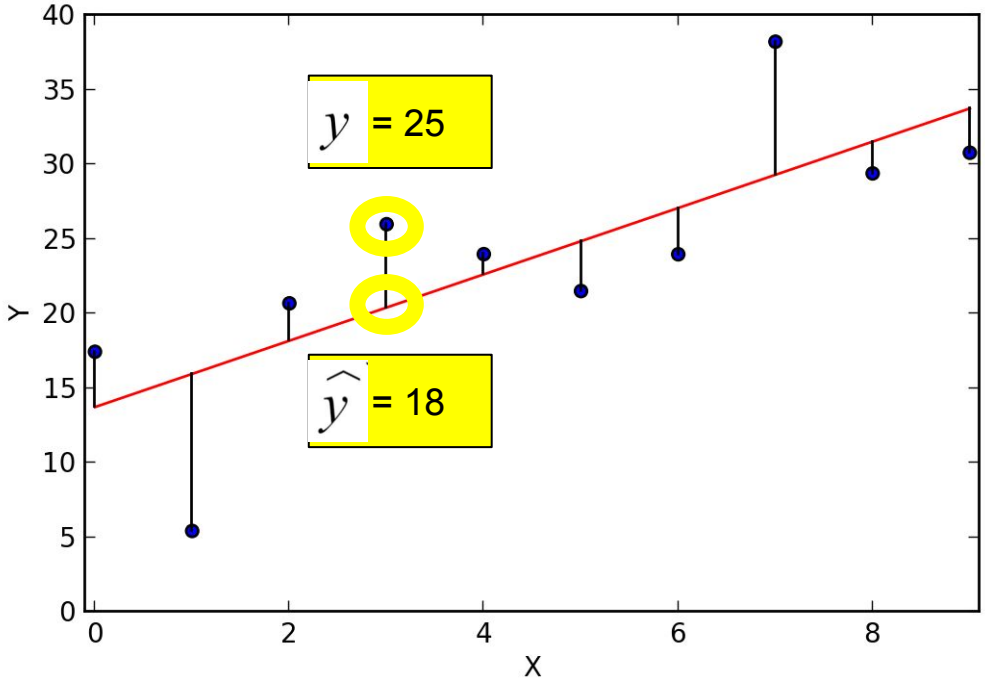
Loss == Error

$$MSE = \frac{1}{n}\Sigma\underbrace{\left(y - \hat{y}\right)^2}_{\text{The square of the difference between actual and predicted}}$$



Loss == Error

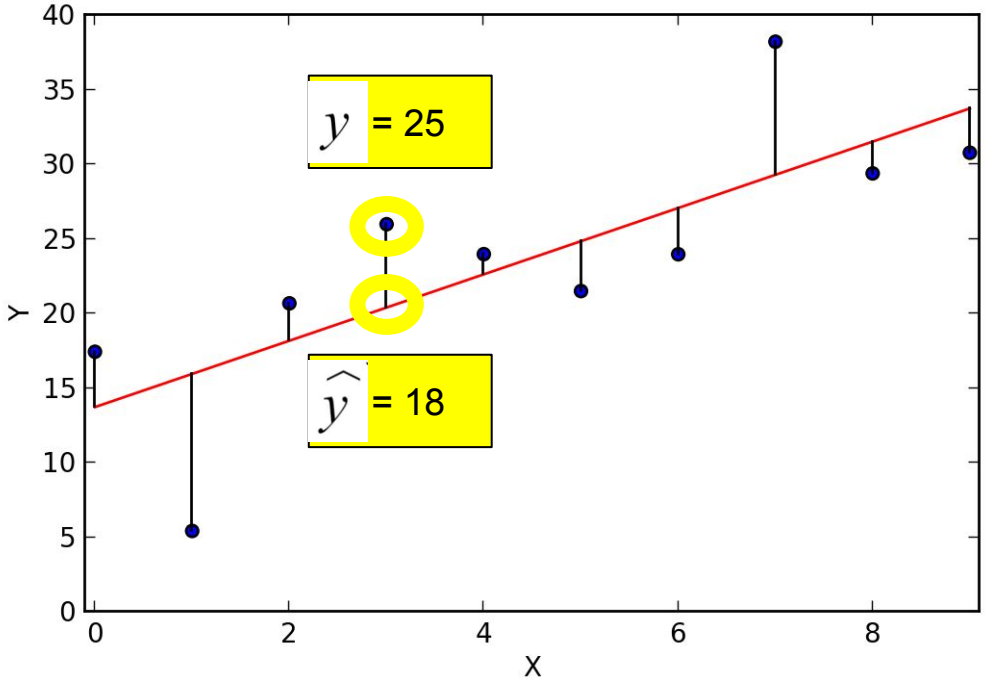
$$MSE = \frac{1}{n}\Sigma\underbrace{\left(y - \hat{y}\right)^2}_{\text{The square of the difference between actual and predicted}}$$



SE for point #4 = $(25\text{-}18)^2 = 49$

Loss == Error

$$MSE = \frac{1}{n}\Sigma\underbrace{\left(y - \hat{y}\right)^2}_{\text{The square of the difference between actual and predicted}}$$



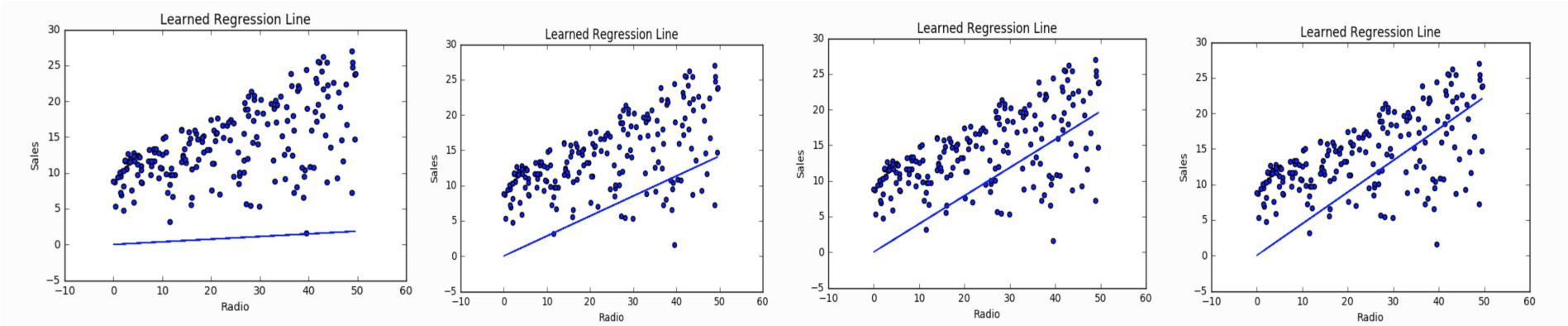
SE for point #4 = $(25-18)^2 = 49$

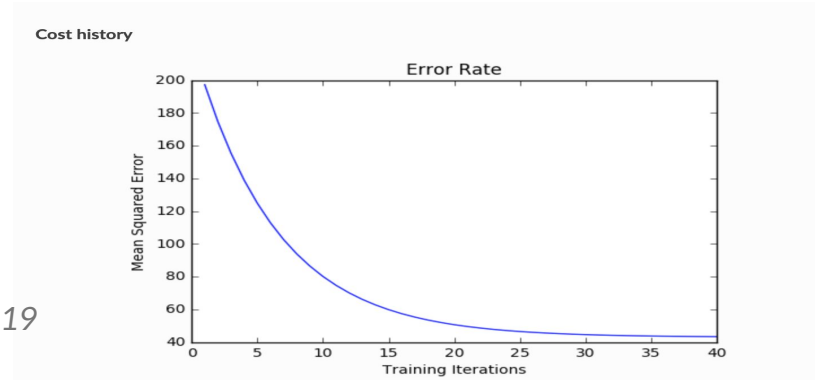
MSE = (SE for point #1 +
SE for point #2 +
SE for point #3 +
... +
SE for point #10)/10

Update “a” and “b” to find best

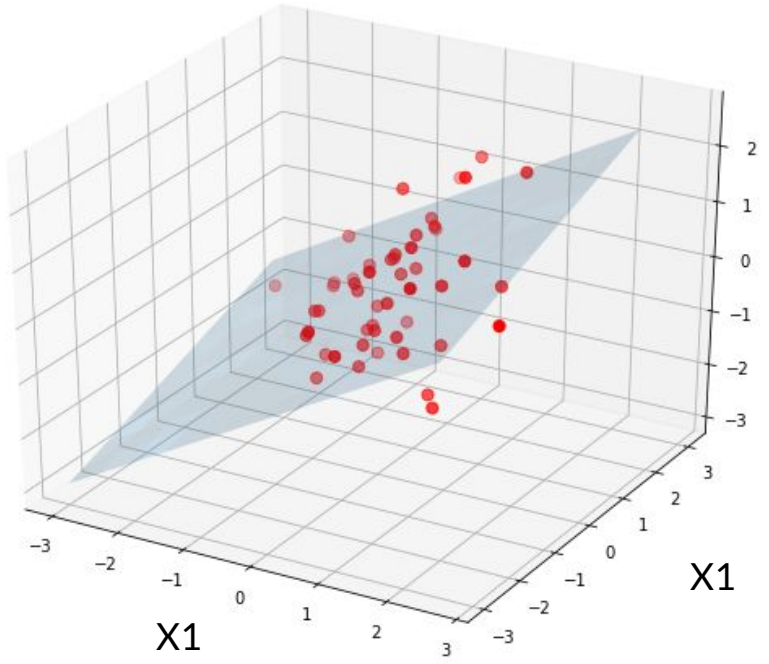
Y = aX + b

Regression with ML





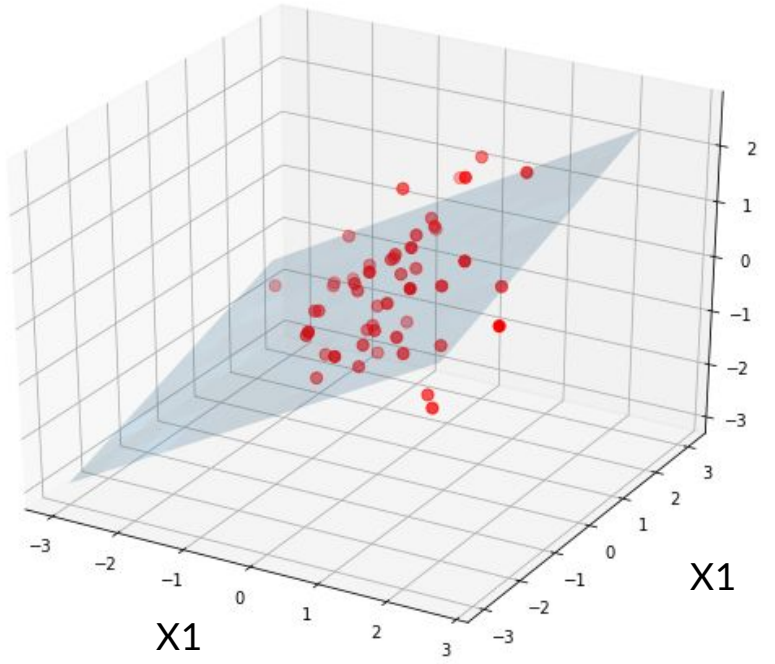
Multivariate Regression



2 features:

$$Y = a(X1) + b(X2)$$

Multivariate Regression



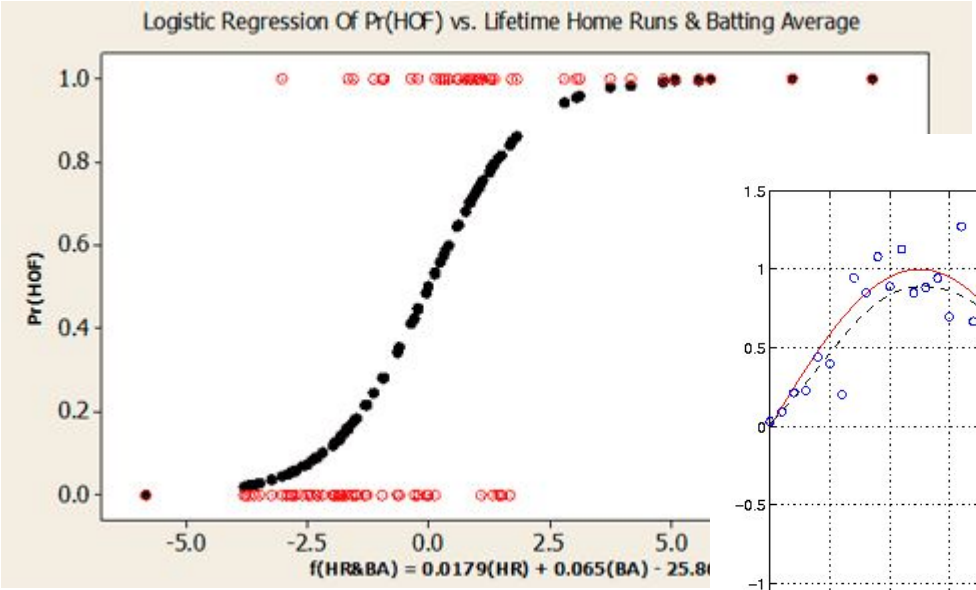
2 features:

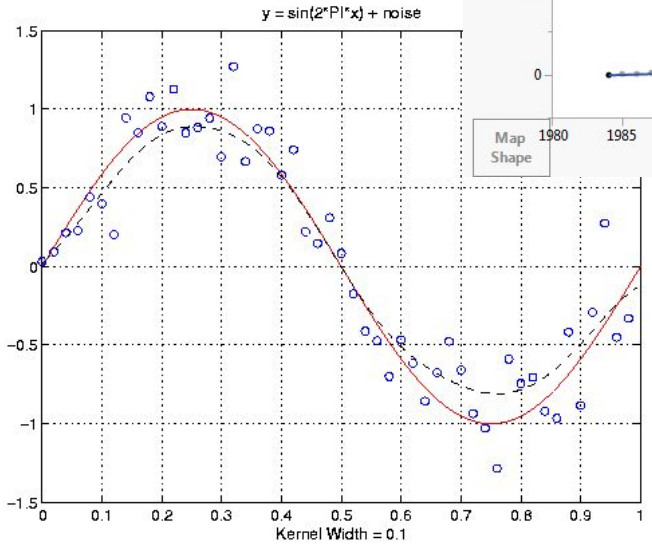
$$Y = a(X1) + b(X2)$$

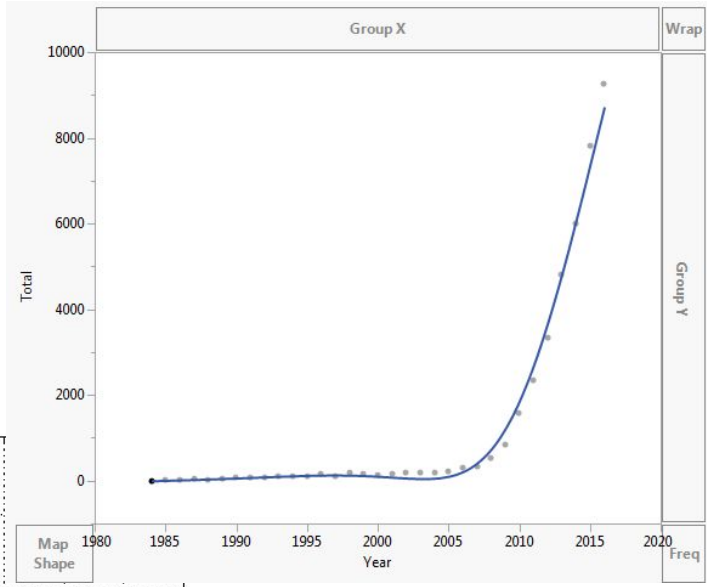
n features:

$$Y = a(X1) + b(X2) + c(X3) + \ldots + z(Xn)$$

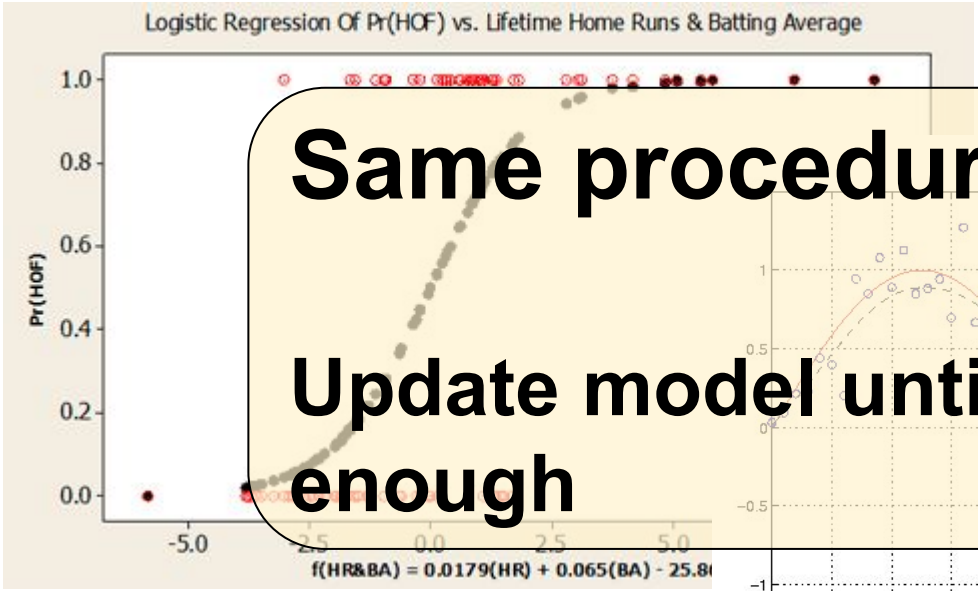
NonLinear Regression

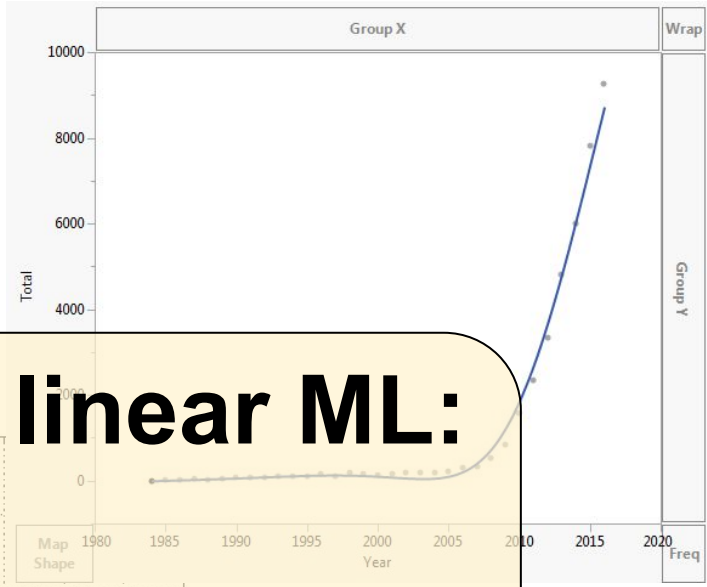






NonLinear Regression

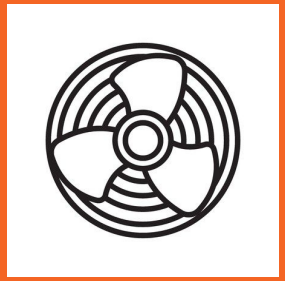


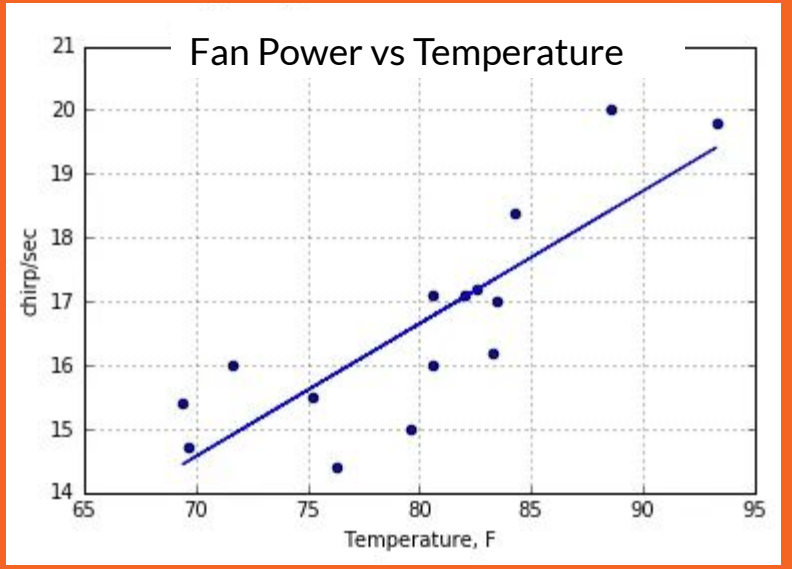


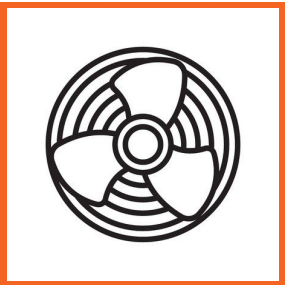
Same procedure as linear ML:

Update model until loss is low enough

Who Cares???

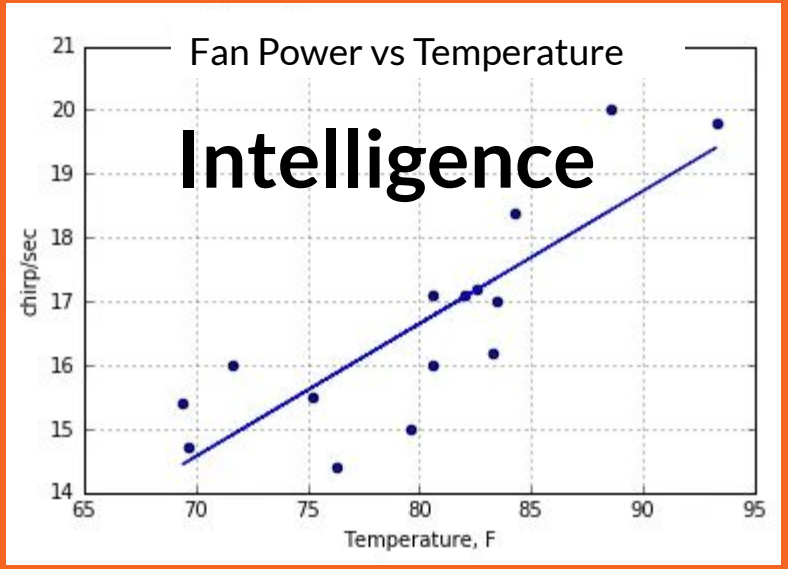


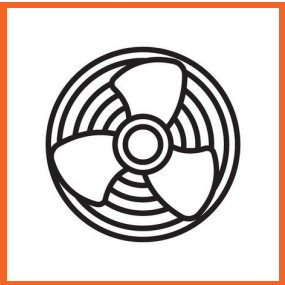




AI System

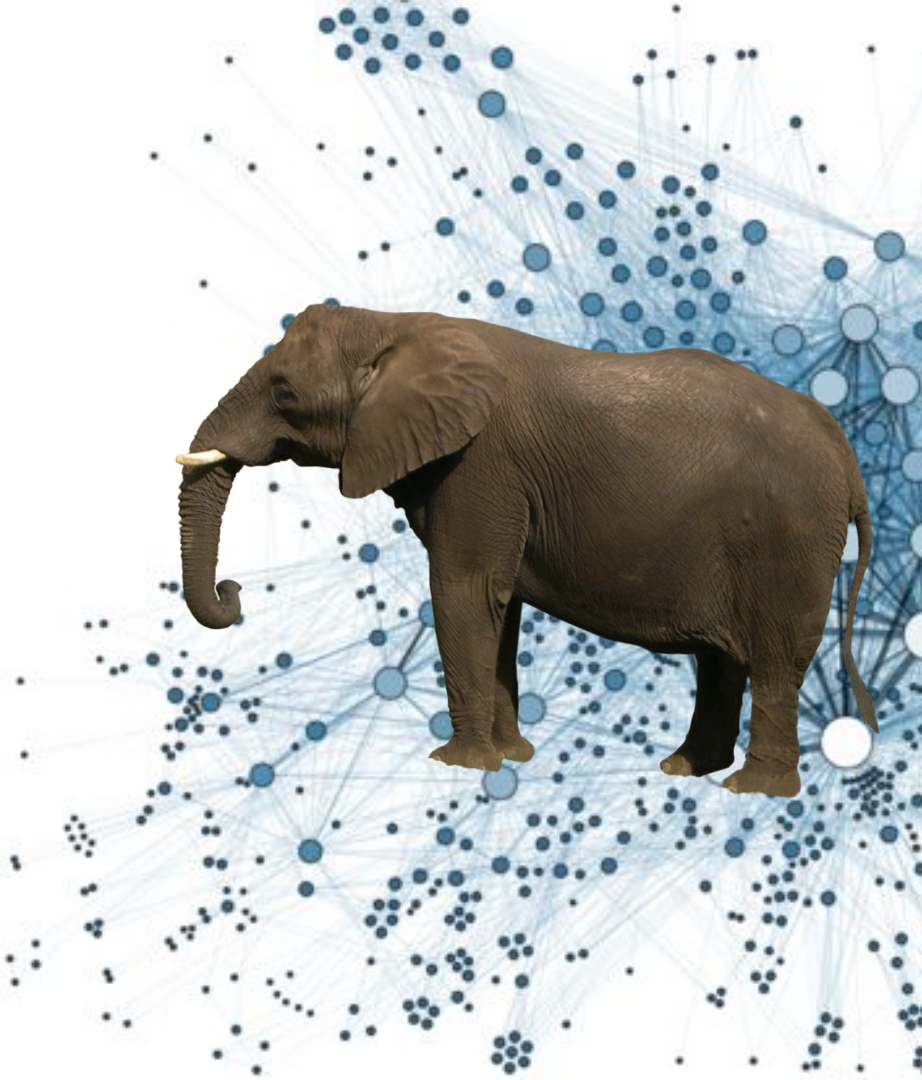
Sensing





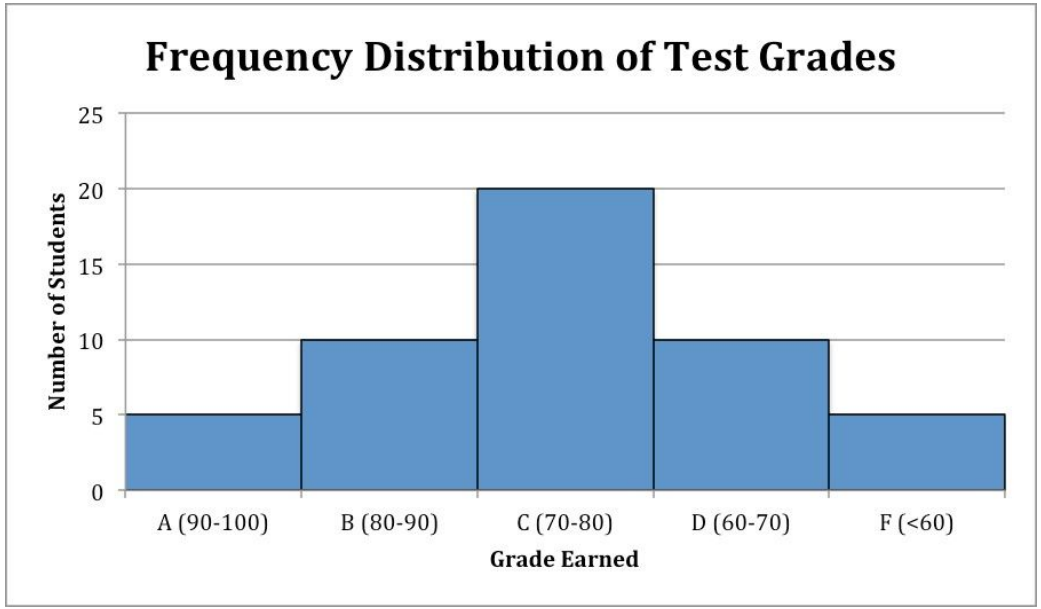
Actuation

How do you **EAT** an **ELEPHANT**?

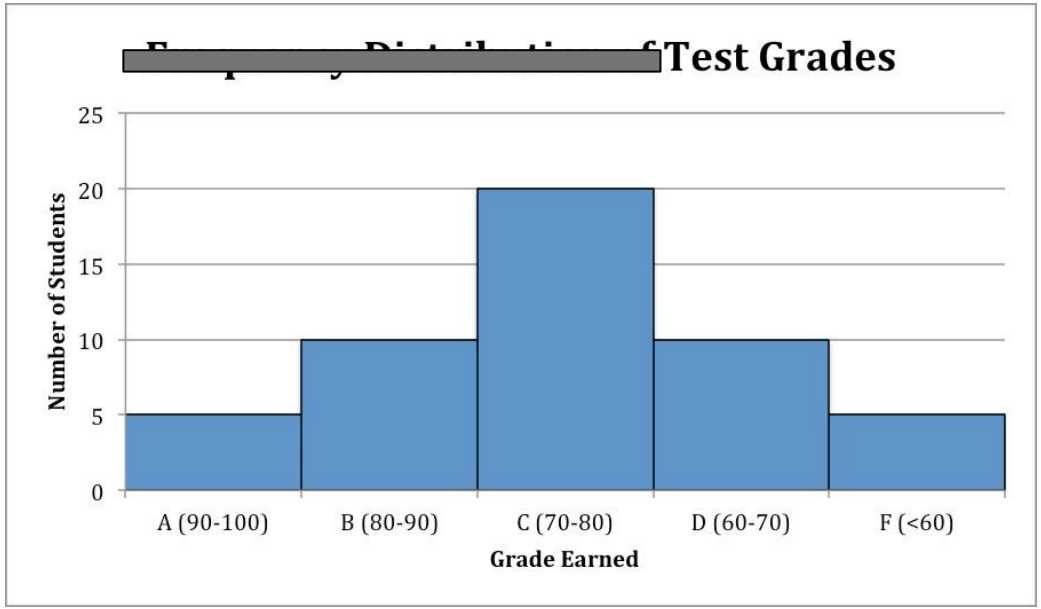


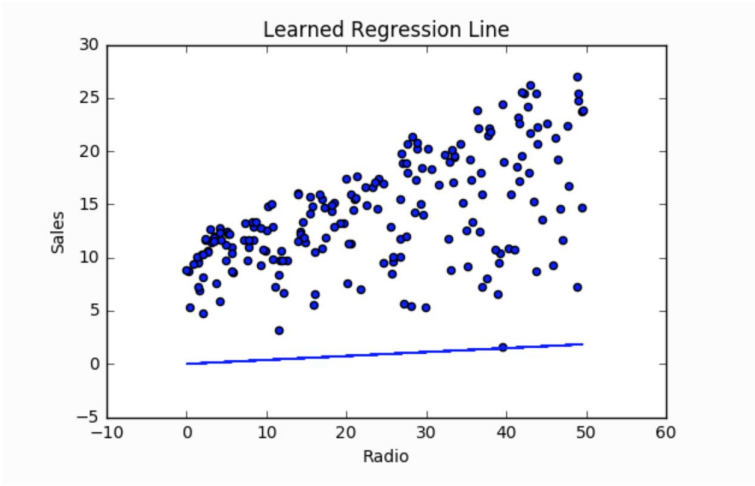
Appendix

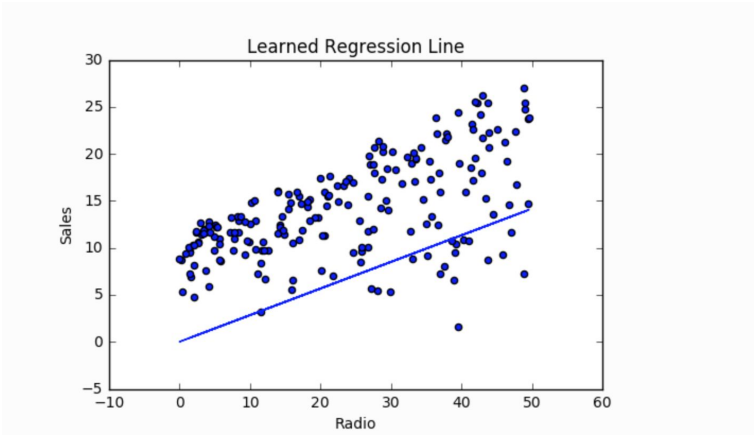
What's a "histogram"?

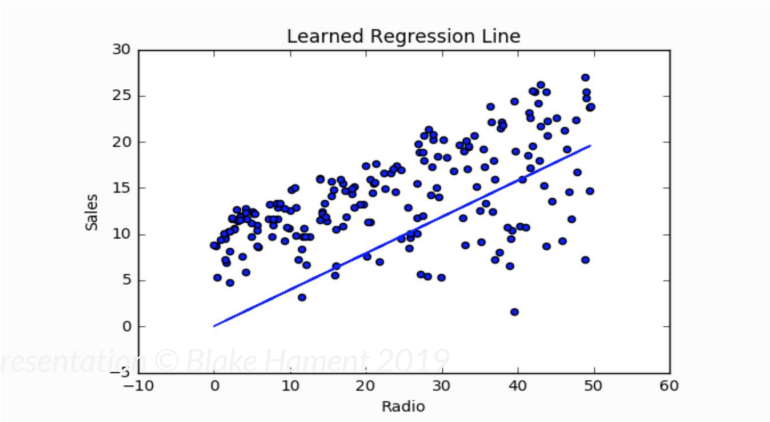


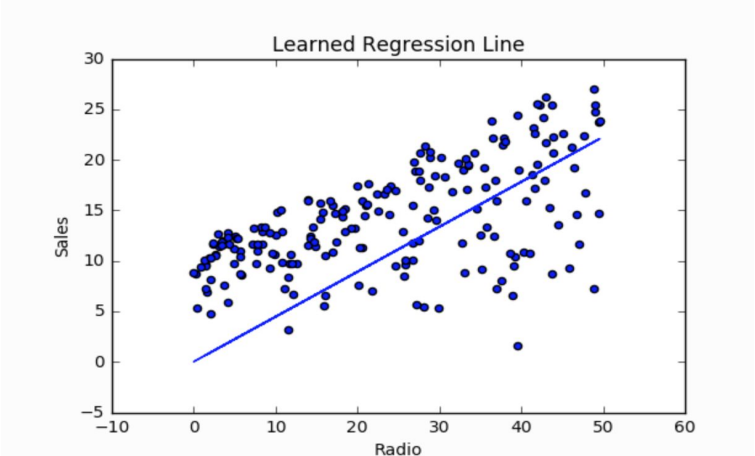
What's a "histogram"?



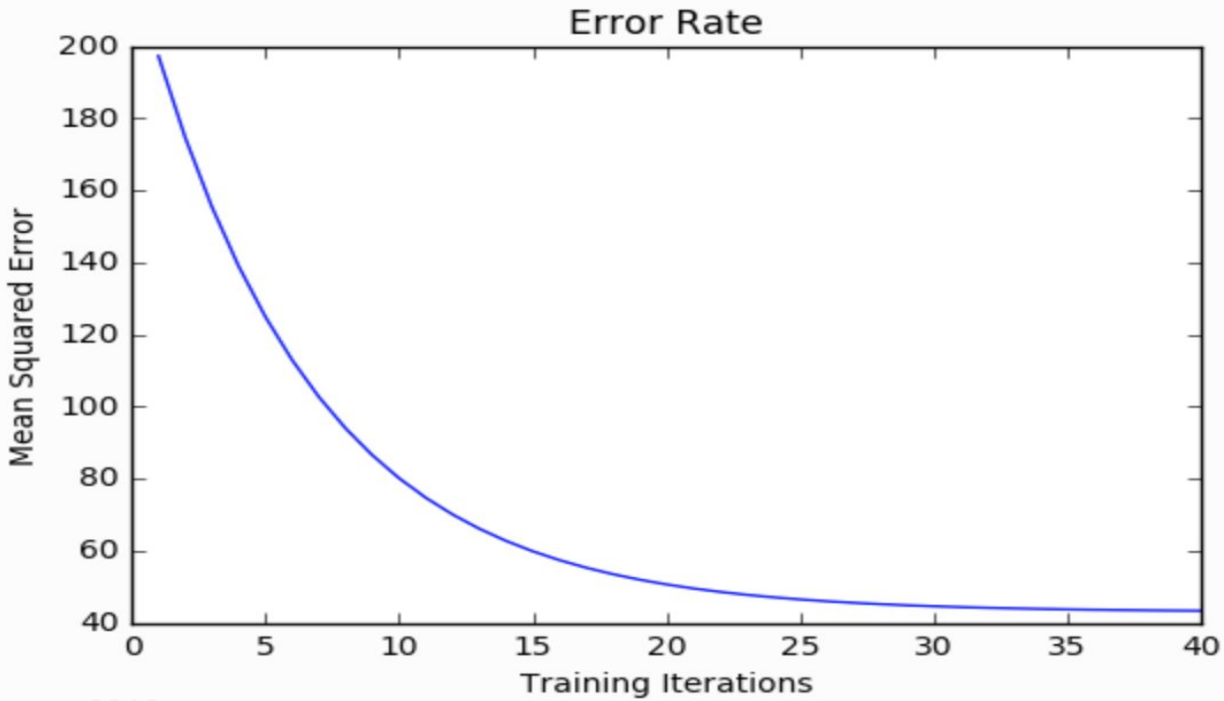






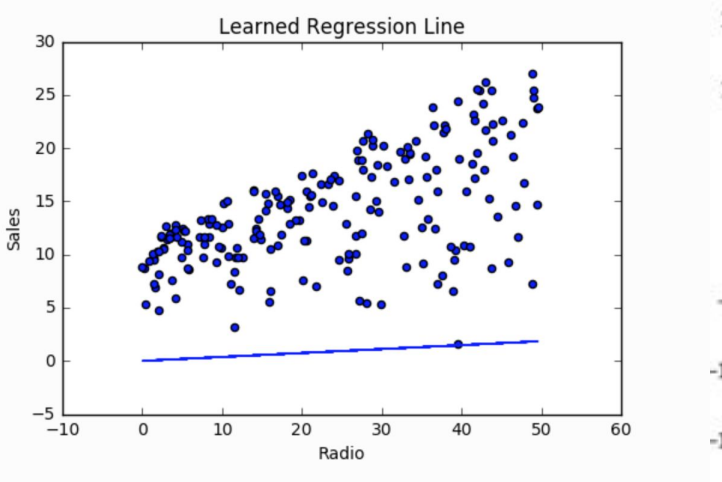


Cost history

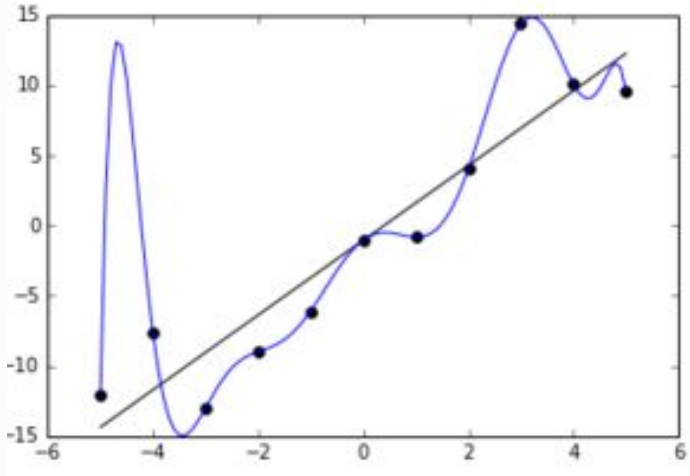


Training & Testing

“Under” and “Over” Fitting



Under



Over

Tip

If we “over fit”our model... we will have BAD performance when estimating new values.

Under fitting usually just means we’re not done yet.

Is your model “good”?


With great power
comes great responsibility.