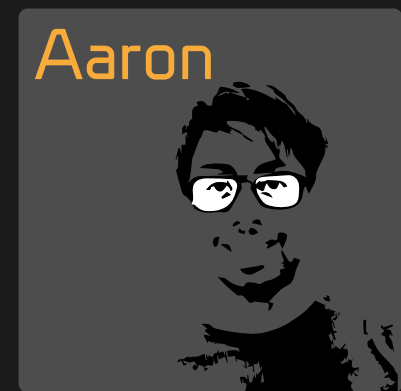
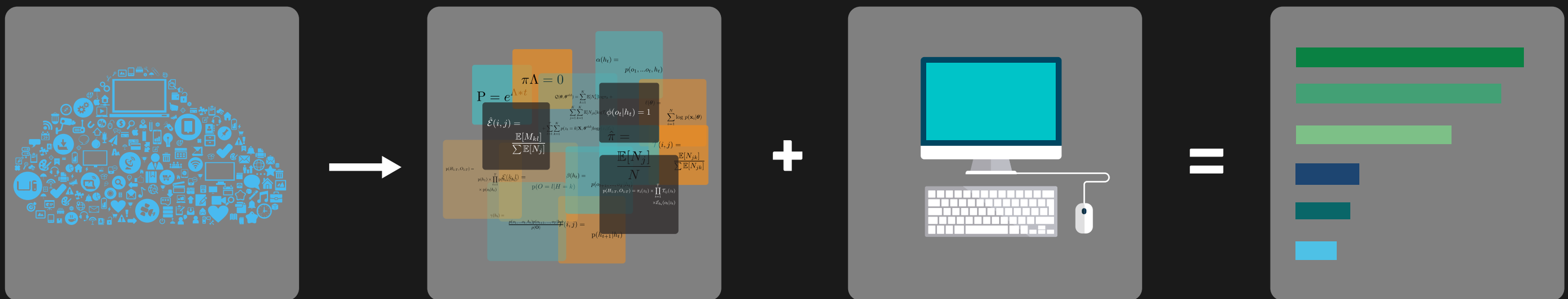


# Machine Learning 101



**STATISTICS**

**ALGORITHMS**

**REGRESSION**

**OPTIMIZATION**

**PROGRAMMING**

**INFORMATION THEORY**

**ARTIFICIAL INTELLIGENCE**

**DENSITY ESTIMATION**

**COMPUTERS**

**CLASSIFICATION**

**PATTERN RECOGNITION**

**DATA SCIENCE**

**CLUSTERING**

**PREDICTION PROBABILITY**

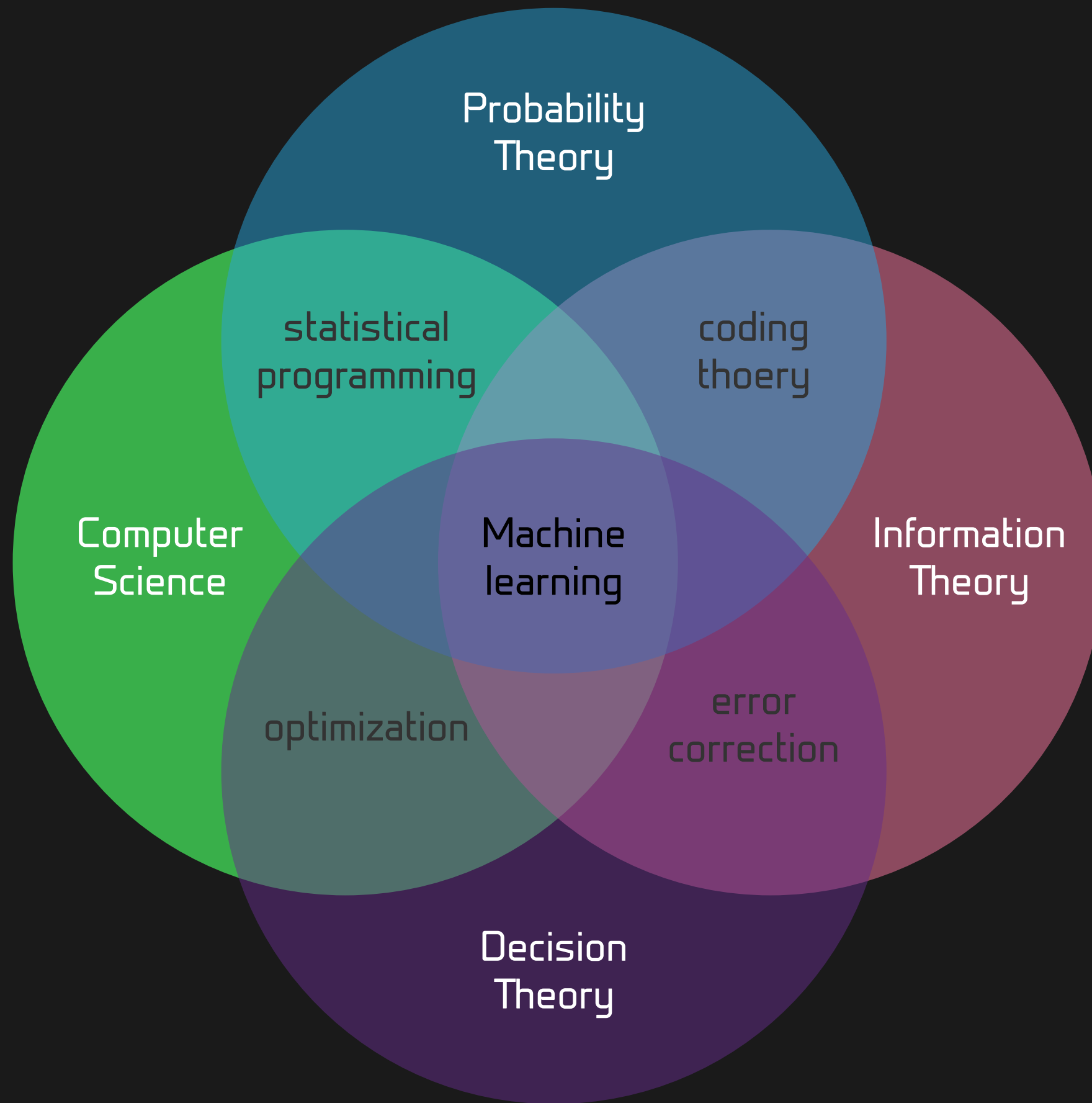
**MATH MODELLING**

**ANALYTICS**

# Goal

set of methods

detect patterns in data



learn mapping  
input to output

## Supervised

$X \longrightarrow Y$

$p(y|x)$

interesting  
patterns in data

## Unsupervised

Knowledge Discovery

$p(x)$

# PARAMETRIC

fixed no. parameters

faster. much faster.

strong assumptions

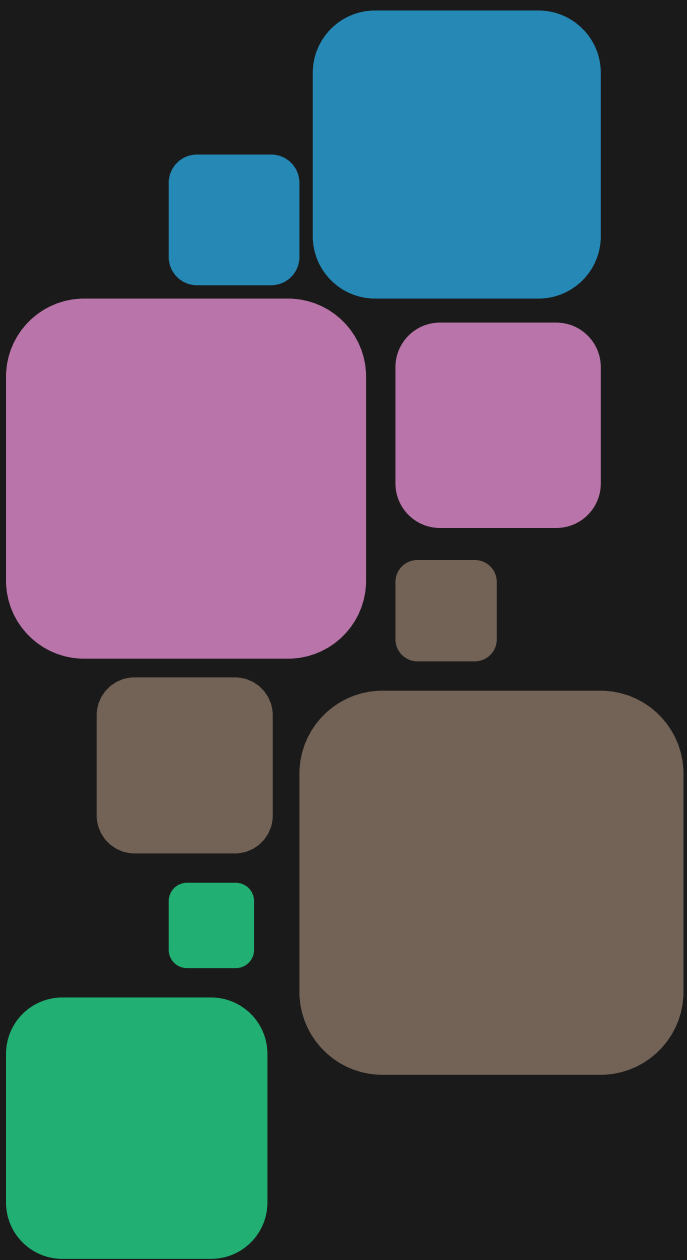
# NON-PARAMETRIC

parms. grow w/ data

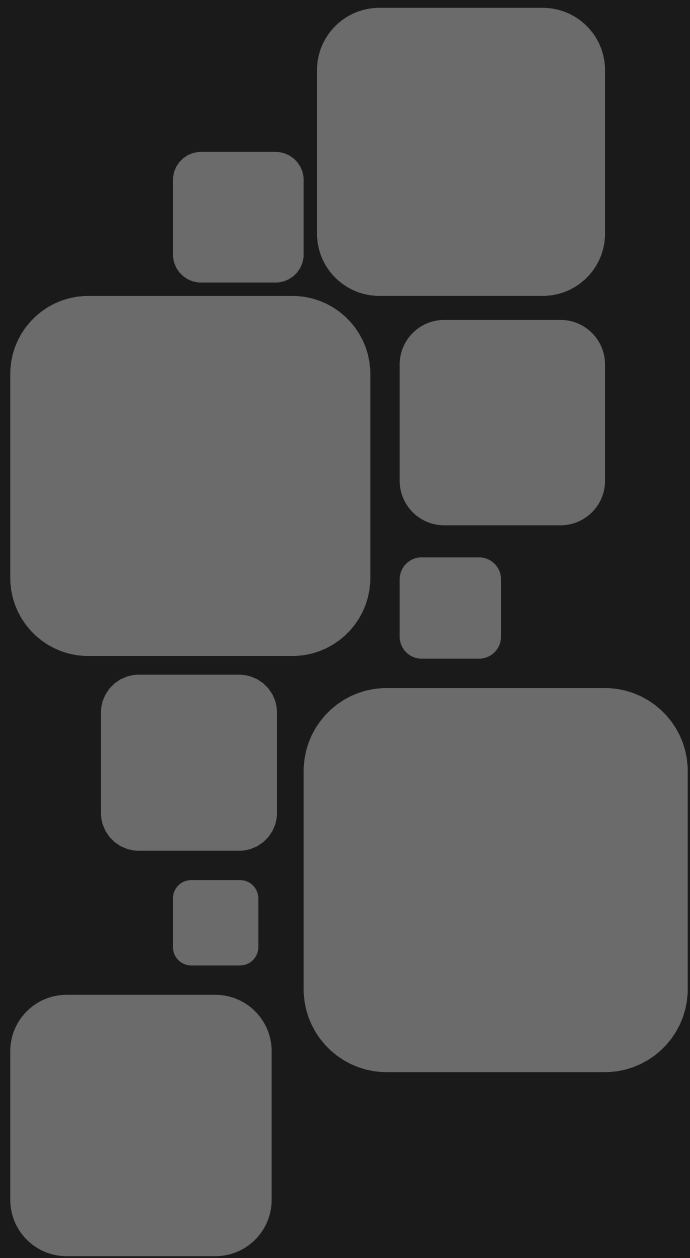
slow. very slow.

no assumptions

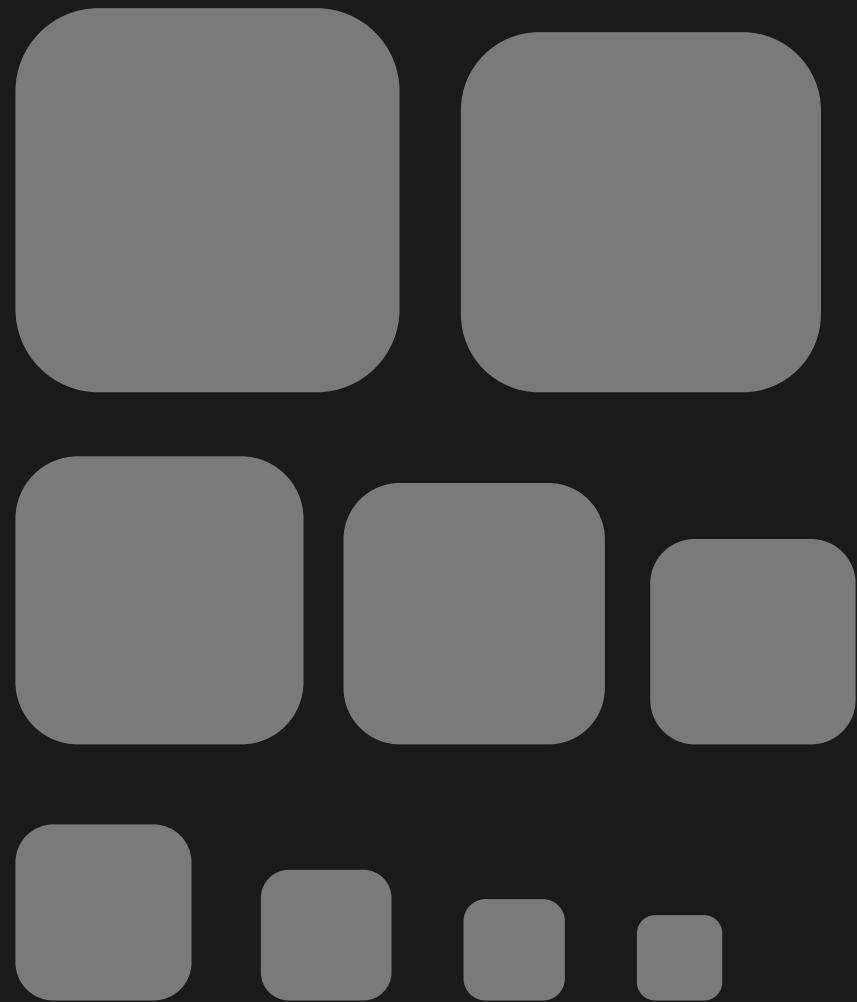
CLASSIFICATION



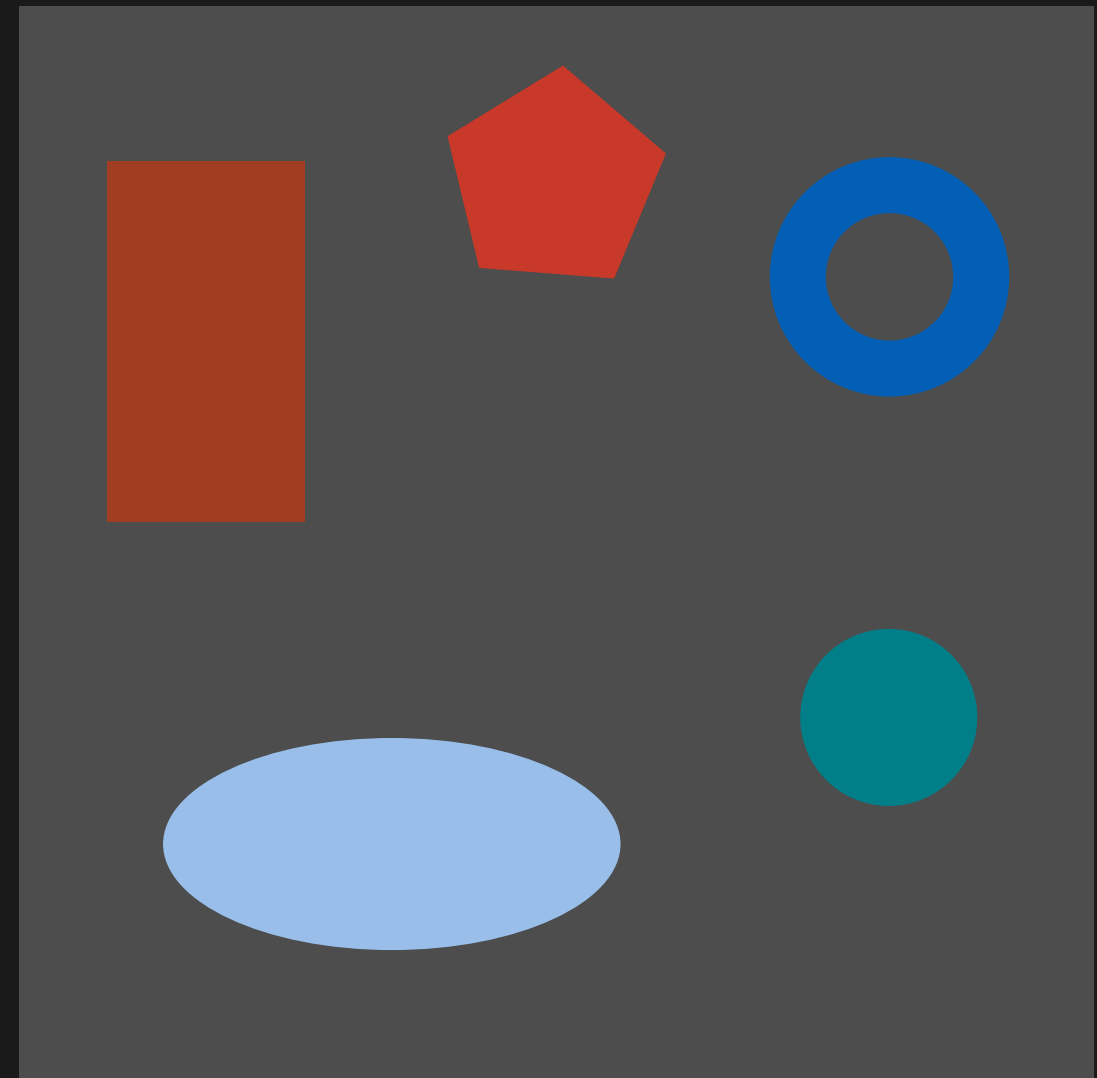
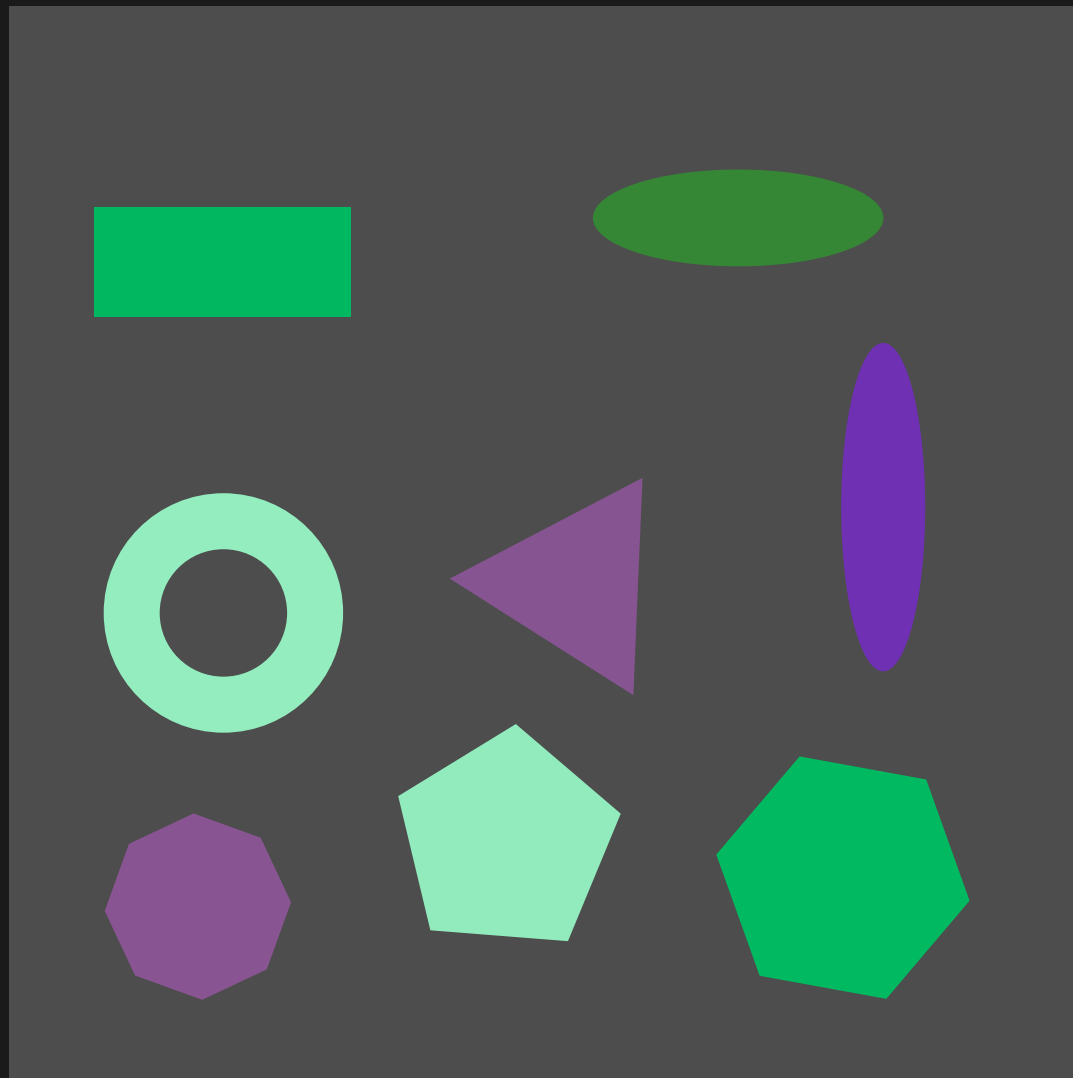
DATA



REGRESSION

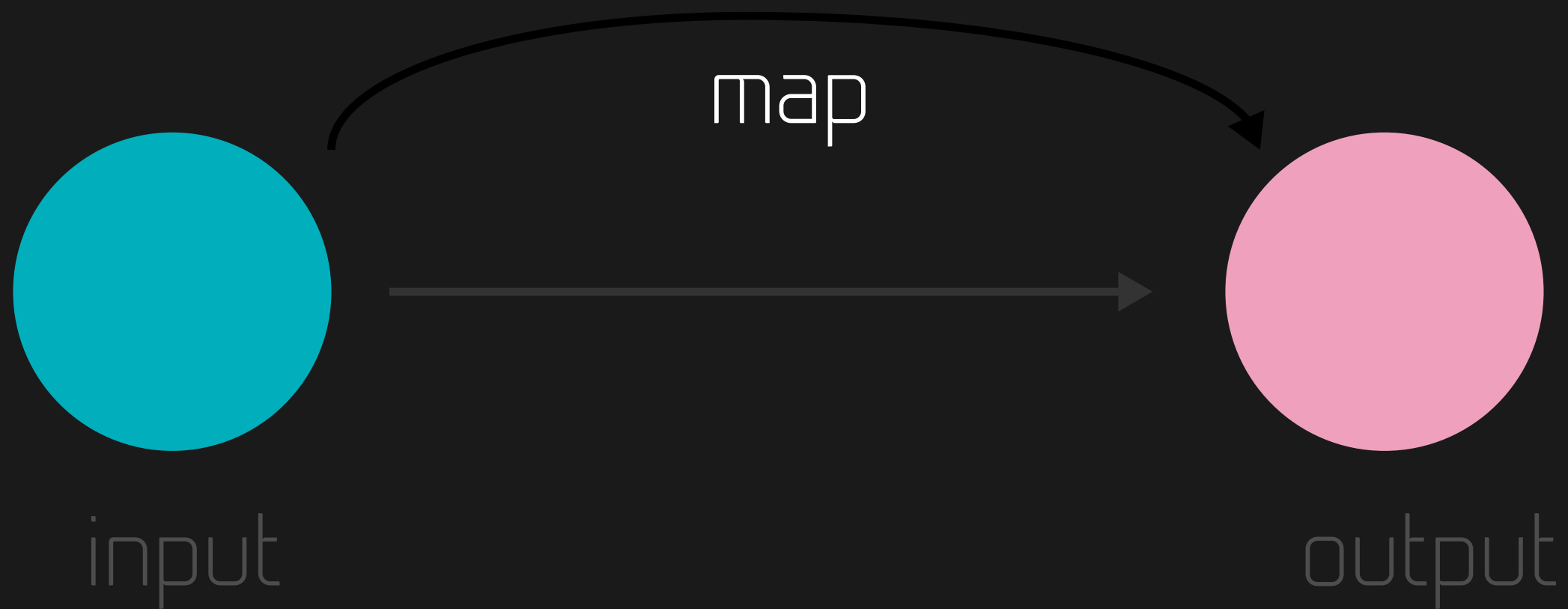


# CLUSTERING

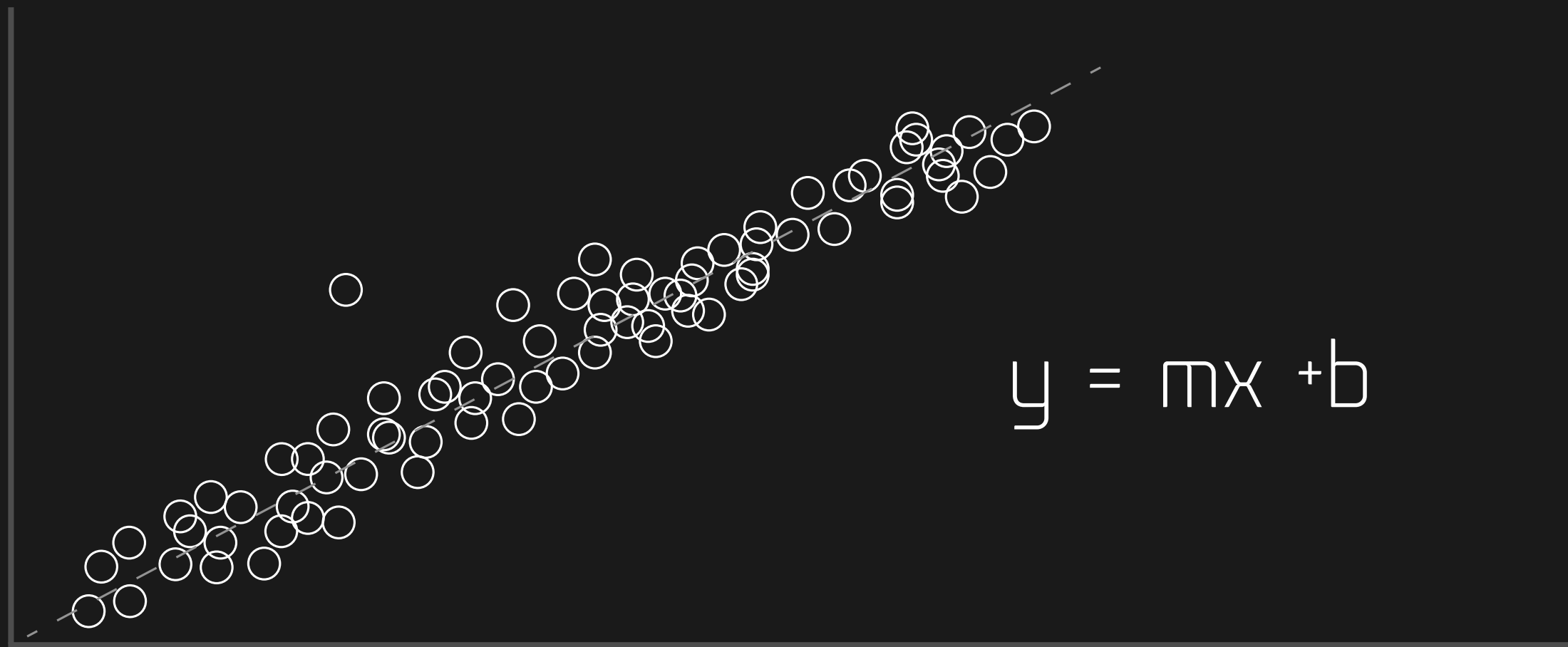




# LINEAR REGRESSION



# LINEAR REGRESSION



$$y = mx + b$$

# LINEAR REGRESSION

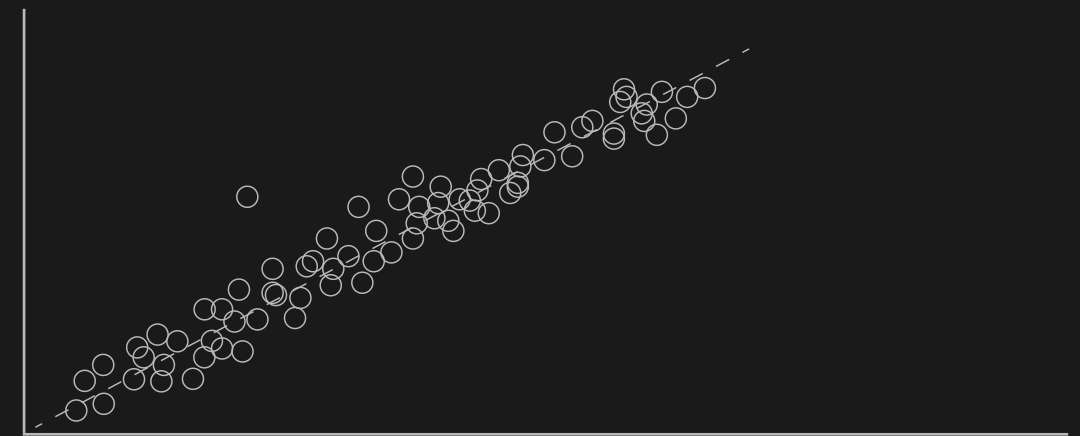
$$y = mx + b$$

$$y = b + mx$$

$$y = w_0 + w_1x$$

$$y = w_0 + w_1x_1 + w_2x_2 + \dots + w_Nx_N$$

$$y(x, \mathbf{w}) = w_0 + \sum_{n=1}^N (w_n x_n)$$



# LINEAR REGRESSION

How do you find  $w$ ?

minimize Error Function!

# LINEAR REGRESSION

$$E(\mathbf{w}) = \frac{1}{2} \sum_{n=1}^N \{y(x_n, \mathbf{w}) - t_n\}^2$$

$dE/d\mathbf{w}$  >>> closed form soln!



