M348W — 2/1/21

"Models" are approximations/abstractions to reality/absolute truth/systems/phenomena.

| Model | Phenomenon |
|---|---|
| model airplane | real airplane |
| Street map | actual roads |
| "early to bed, early to rise makes a man healthy, wealthy, wise" | human health<br>human wealth<br>human wisdom |

"All models are wrong, but some are useful" - George Box, 1984

Wrong → by definition approximations which are not reality.

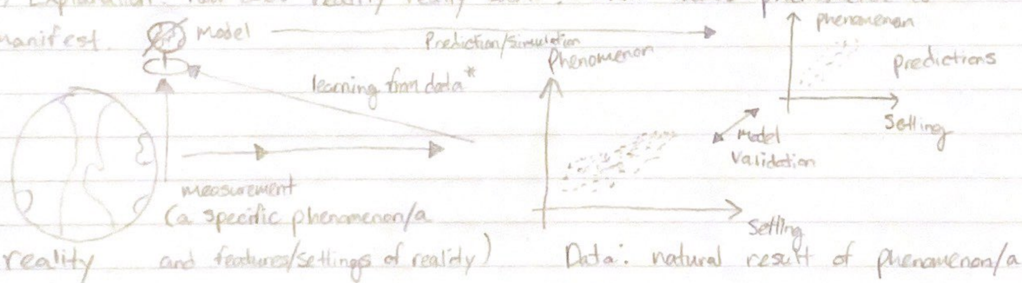
useful → are good enough to be used for a practical purpose.

$C \approx 3.14d$ wrong


Models are generally used for two goals.

1) Prediction: Can the model tell us what will happen in a certain phenomenon in a certain setting. *

2) Explanation: how does reality really work? What causes phenomena to manifest.




Data: natural result of phenomenon/a being measured.

Presteps to modeling:

1) identify a phenomenon/a you wish to predict/explain. This is your target of the modeling procedure.

2) Figure out a way to measure it.

3) Measure features/settings of the system/reality.

"Early to bed, early to rise makes a man heathy, wealthy and wise"

Phenomena: human health, wealth, and wisdom

Features/settings: bedtime, waketime

This model is ambiguous! We don't know how to measure the settings and phenomena. In order to make this model unambiguous, we need to establish "metrics". Metrics are well-defined ways to numerically gauge phenomena/settings.

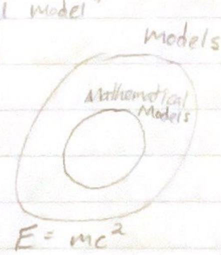
| Feature/phenoma | Metric | Symbol |
|---|---|---|
| bedtime | avg bedtime ages 18-60 measured in hours past 5pm | $b$ |
| waketime | avg waketime measured in hours past 4am | $w$ |
| health | longevity/lifespan, QOL metric | $\ell$ |
| wealth | net worth at time of death | $n$ |
| wisdom | take a test about what you would do in situations and have a panel of old people provide answers | $s$ |

$$f\left(\begin{bmatrix} b \\ w \end{bmatrix}\right) = \begin{bmatrix} \ell \\ n \\ s \end{bmatrix}$$

model — two settings (inputs) — three phenomena (outputs)

prediction of ← in this class we only build one output.

Since the inputs/outputs are numerical, $f$ is called a "mathematical model"




Mathematical models are not physical.

They are themselves ideas and abstraction.

But they are extremely useful! $a = F/m$, $E = mc^2$

For the purposes of this class, we will assume the universe is mathematical:

Assume: a phenomenon denoted y can be expressed as: $y = f(\underbrace{z_1, z_2, \ldots, z_t}_{\substack{\text{Causal inputs:} \\ \text{the true drivers} \\ \text{of the phenomenon.} \\ \text{In reality we don't} \\ \text{know them.}}})$

$y$: Phenomenon, response, outcome, endpoint, dependent variable

Let's examine the phenomenon y = pays back loan on time.

$y \in \{0, 1\} = \mathcal{Y}$ output space

0: did not pay on time

1: paid back on time (Convention: 1 is the "positive" event or thing you want to happen)

Models with output spaces of cardinality 2 are called "binary classification models."

The causal inputs are features or characteristics of the individual person. We don't know the causal model why people pay / don't pay back loans. We are going to make one up just as an illustration.