1. **Formalizing prediction problems.**

In class, we defined a *supervised learning problem*: given a collection of data \((x_i, y_i) \in (X \times Y)\) for \(i = 1, \ldots, n\), find a mapping \(f : X \rightarrow Y\) so that

\[
y_i \approx f(x_i).
\]

What might \(X\), \(Y\), and \(f : X \rightarrow Y\) be in the following cases? Say what kinds of data (continuous, discrete, nominal, ordinal) make up \(X\) and \(Y\). Is any of the data likely to be missing? Why?

- **Medical treatment planning:** You receive a patient’s medical record containing a complete medical history, and want to identify which treatment will work best.
- **Electoral campaigning:** Given a voter’s voting history, you want to predict whether they’re likely to support your candidate.
- **Time series forecasting:** You’d like to predict how your favorite stock will perform tomorrow.
- **Handwriting recognition:** The post office would like an automated procedure to understand which zip code is written on an envelope.
- **Class placement:** You’re in charge of determining class sections in a middle school. You receive a file on each student with their previous course history and exam results, and want to place them in the appropriate math class.
- **Pick your own problem:** Define \(X\), \(Y\), and \(f\) for a big messy prediction problem you’d like to solve.

2. **Coding experience.**

- For every student in the class, let \(x \in \{0, 1, 2\}^6\) be a vector describing coding experience. Each entry of \(x\) corresponds to a programming language, and is 0 if you have no experience with that language, 1 if you’ve written at least one script or program in that language, and 2 if you’d be able to write a simple program
• Your TAs have written the following code to process your coding experience vector. What are they trying to do?

```plaintext
n = 0
for j = 1:6
    if x[j] > 0
        n += 1
    end
end
return n
```

• We would like to identify which students have taken a computer science class. To formalize our problem, let’s say that the feature space $X = \mathbb{R}^6$, and the space of outcomes is $Y = \{\text{has taken a CS class}, \text{has not taken a CS class}\}$. Suppose we have found a vector $w \in \mathbb{R}^6$ so that $w^T x > 0$ whenever a student with coding experience vector $x$ has taken a computer science class, and $w^T x \leq 0$ otherwise. Write down a piecewise definition of the function $f : X \rightarrow Y$ mapping coding experience vectors to the labels in $Y$. (You can also write pseudocode if you prefer.)

• Do you think $w_5$ is positive, negative, or 0? Why?

*Note: this problem does not have a unique right answer...!