Supervised vs. unsupervised

- **Supervised**
  - Learn $\hat{y} = f(x)$, generally to minimize some loss function $L(y, \hat{y})$
- **Unsupervised**
  - Only uses $x$ – clustering, dimensionality reduction

What examples have we seen of each?

---

Generative vs. Discriminative

- **Generative**
  - Learn $p(x)$
  - Or, equivalently, $p(y|x)$
- **Discriminative**
  - Learn $p(y|x)$
  - Or minimize $L(y, f(x, \delta))$

Given $p(y|x)$, can one estimate $p(y|x)$?

Given $p(y|x)$, can one estimate $p(y,x)$?

---

Parametric vs. non-parametric

**Parametric:** $y = f(x, \theta)$, where the form of $f$ is specified.

- K-nn
- Decision trees
- Linear regression
- Linear regression with complexity penalty
- Naïve Bayes

Which of these is parametric?