Lecture 1 Outline:

1. What is statistical learning?
   - Intersects: statistics, applied math, computer science.
   - How is statistical learning different from machine learning / artificial intelligence?
   - *The art of learning on current data to predict responses, labels, or properties of future data.*

2. Statistical Learning Tasks:
   (a) Prediction.
   (b) Exploratory Analysis (Data-Driven Discoveries).

3. “Big Data”.
   - Volume, Velocity, Variety (complexity).

4. Data Matrix.
   - Big $n$.
   - Big $p$.

5. Types of statistical learning:
   - Supervised Learning.
     - Data with labels or a response.
   - Unsupervised Learning.
     - Data with no labels.
   - Others:
     - Semi-supervised learning, recommender systems, online learning, network models, sequential learning, text mining, etc.

6. Who uses statistical learning?

7. What do we cover in this course?

8. Course overview and expectations.

9. A few basics:
   - Training & test sets.
   - Model complexity, overfitting, & training and test errors.
   - High-dimensional data & curse of dimensionality.

    - Predict the label based on the labels of the $K$ closest observations to the data point of interest.
    - What happens to the test prediction error when $K$ is small? Large?
    - How does KNN behave in high-dimensions? With possibly noise variables?