**Lecture 2:**

* Data Mining: Process of **discovering patterns in large data sets** involving methods at the intersection of machine learning, statistics, and database systems.

Steps of Data Analysis:

1. Make sense of the data.
2. Visualize the data to draw from our powerful visual system.
   1. Find visual encodings that:
      1. Guide viewer’s attention.
      2. Communicate data to the viewer.
      3. Let the viewer calculate with data.

Main Approaches:

* Visualization using t-SNE:
  + Pass a high-dimensional dataset to get a lower-dimensional representation.
  + Cluster sizes in a t-SNE plot mean nothing.
* Visualization using UMAP:
  + Faster than t-SNE.