

a) numeric data

\* discrete and continuous

discrete: eg whole numbers like player's score

Continuous: → all sort of numbers, any number within a range

Categorical Data:

\* represent characteristics

\* ordinal data → sorted or ordered

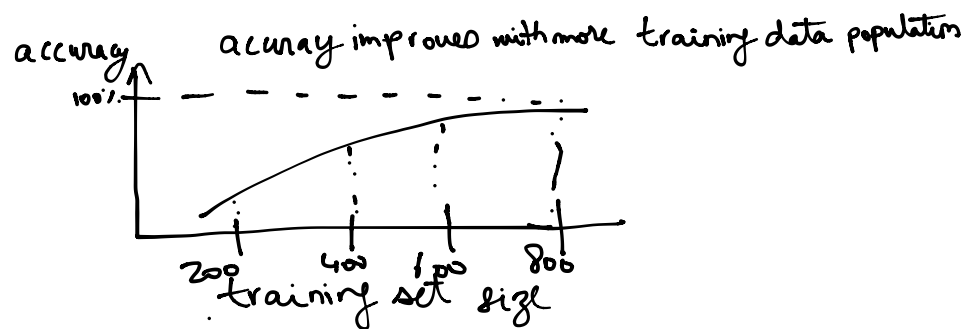
Time series: sequence of data taken at time intervals

---

Enron data set:

Person of Interest (POI)

size of Training set has an effect on the accuracy of the prediction



\* Truism of machine learning: More data gets better results than a fine-tuned algorithm.