

ONLINE MASTERS IN DATA SCIENCE

DSC 257R - UNSUPERVISED LEARNING

ONLINE UNSUPERVISED LEARNING

SANJOY DASGUPTA, PROFESSOR

UC San Diego

COMPUTER SCIENCE & ENGINEERING
HALICIOĞLU DATA SCIENCE INSTITUTE

Online Unsupervised Learning

- Endless stream of data: x_1, x_2, x_3, \dots
- Repeat forever:
 - Get data point x
 - Update model

Spaced allowed: $O(\text{size of model})$

Online Computation of the Mean

- $\mu = 0$ (mean so far)
- For $t = 1, 2, 3, \dots$:
 - Get x_t
 - Update μ

Why is this correct?

Online Computation of the Median

How to do this?