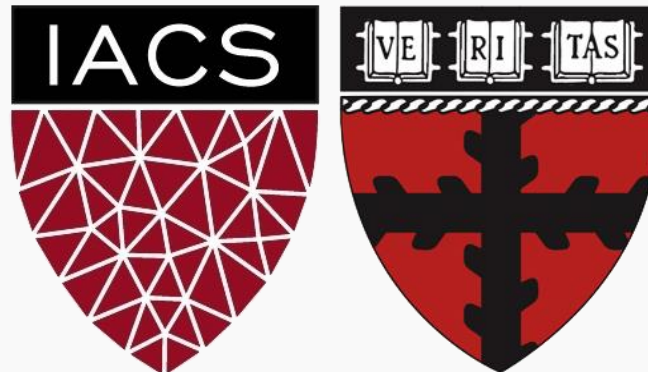# Lecture #4: Dask

## AC295 Advance Practical Data Science

Pavlos Protopapas

# Lecture Outline

- Scalable computing (and why it matter)

- Introduction to Dask

  - Work with structured data

- ML with Dask

- Scaling and deploying (Exercise)

# Scalable computing

- 1

- 2

- 3

- 4

- Introduction to Dask

  - 1

  - 2

  - 3

  - 4

- Work with structured data

  - Sample code + more into jupyter notebook

- # ML with Dask

  - Summarize literature

    - Application 1

    - Application 2

    - Application 3

    - Application 4

- Scaling and deploying

    - Explanation: build a Dask cluster on Amazon AWS with Docker

        - Direction 1

        - Direction 2

        - Direction 3