



# DATASET ANALYSIS

App identification based on data volume

**Huber Flores**

huber.flores@helsinki.fi

# Preliminaries

- Based on previous observations (part 1), answer the following issues.

# Questions

- Patterns in data volume samples
  - How significant is the data transferred rate of an app for distinguish it among the rest?
  - Is it possible to categorize application usage based on data transferred rate?
    - Provide a few examples
  - What is the minimal amount of samples to characterize data transferred rate?

# Questions

- Patterns in base station access
  - Are peak times of a base station influenced by a particular app?
    - Describe most likely behavior of a base station
  - Categorize users based on app usage
  - Is it possible to identify users 'mobility (trajectory) based on base station access
    - For instance, a user connects to base station  $y$ , after leaving base station  $x$ 
      - Overall dataset percentages, if none, then try a specific users

# Deliverables

- A oral presentation using slides, e.g., Powerpoint, LibreOffice, among others.