Big data: architectures and data analytics

Spark - Exercises
Exercise #47

- Full station identification in real-time
- Input:
  - A textual file containing the list of stations of a bike sharing system
    - Each line of the file contains the information about one station
      - id\longitude\latitude\name
  - A streaming of readings about the status of the stations
    - Each reading has the format
      - StationId, # free slots, #used slots, timestamp

Exercise #47

- Output:
  - For each reading with a number of free slots equal to 0
    - print on the standard output timestamp and name of the station