# Group 13 Homework 1 Analysis

### Marco Annunziata

Sveva Turola

We denote as:

* N: size of the input. ;
* L: the number of partitions;
* , meaning: the maximum size across all partitions;

Design Goals for the MapReduce algorithm implemented:

* R (Number of rounds): 2
* ML (Local Space):
  + Round 1:
    - Map Phase: so that each element is transformed into a pair where is a point and is the length of the closest center;
    - Reduce Phase: , because by random partitioning the amount of space needed for the intermediate pairs is cut;
  + Round 2:
    - Map Phase:
    - Reduce Phase: , because of the operation of having to sum up the counts for each partition and then for all partitions, the needed allocated space increases linearly with the number of partitions

Overall Local Space:

If partitions, then by using random partitioning there is (high) probability

:

And

* MA (Aggregate Space): If and , then: