What is MapReduce?

- Terms are borrowed from Functional Language (e.g., Lisp)

Sum of squares:

- (map square '(1 2 3 4))
    - Output: (1 4 9 16)

    [processes each record sequentially and independently]

- (reduce + '(1 4 9 16))
    - (+ 16 (+ 9 (+ 4 1)))
    - Output: 30

    [processes set of all records in batches]

- Let's consider a sample application: Wordcount
    - You are given a <u>huge</u> dataset (e.g., Wikipedia dump or all of Shakespeare's works) and asked to list the count for each of the words in each of the documents therein