

Big Data Management and Analytics
WS 2016/17

Tutorial 4: MapReduce

Assignment 4-1 *Matrix multiplication*

Given two matrices $A \in \mathbb{R}^{i \times j}$ and $B \in \mathbb{R}^{j \times k}$

(a) Describe the steps which are required to perform a matrix multiplication using MapReduce.

(b) Let now $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ and $B := \begin{pmatrix} 7 & 8 \\ 9 & 10 \\ 11 & 12 \end{pmatrix}$

Multiply A and B using MapReduce.

Assignment 4-2 *Word count*

The goal of the word count task is to count the number of occurrences of each word in a set of documents.

(a) Extend the word count task by computing the average occurrences for every word. Describe the steps which are required for the task using MapReduce.

(b) Now compute the standard deviation given the number of occurrences of every word. Describe the steps which are necessary for the task using MapReduce.

(c) Given the following sentence:

How much ground would a groundhog hog, if a groundhog could hog ground? A groundhog would hog all the ground he could hog, if a groundhog could hog ground.

Compute the average occurrences of the words using MapReduce.