# Big Data Analytics

**Exercise 5**
**Lucas Drumond, Rasoul Karimi**
**Information System and Machine Learning Lab (ISMLL)**
**University of Hildesheim**

## Deadline: 25.06.2014

1. Update the MapReduce application in exercise 4 to calculate the statistics for any given column.

   **Hint**: Give the column number as argument, save it in job configuration and read its value from job configuration in init() function.

2. Update the MapReduce in exercise 4 to calculate statistics for every column at the same time.

   **Hint**: the output key of map function can be like  key = quality + "column_name"

## Deliverables

1. x.java : the source code of task 1.
2. x. pdf:  it contains the results of all 11 columns. This file consists of 11 sections, each section reports the results of one column as the input argument. The sections must be in order from 1 to 11.
3. xx.java : the source code of task 2.
4. xx. pdf:  it consists of 6 sections. Each section has 11 lines that show the results of the corresponding quality for each column.