

Principles of E-Commerce I - Big Data Platforms

Assignment 3 (4)

Todor Ivanov

Dr. Karsten Tolle

<http://www.bigdata.uni-frankfurt.de/>



Assignment 3

- Assignments are done in groups of 2 persons (one exception of 3 persons possible)
- ... same group for the 3rd and 4th assignments
- Results (code & description) need to be submitted latest by **July 6, 2015 (23:55h)** – one submission per group!
- Submission of Assignment 3+4 together as one ZIP file!
- Results need to be presented on: **July 8th, 14th and 15th, 2015** (about 15 minutes presentation time + 5 minutes discussion).

Submissions via OLAT

- Registering: <https://olat.server.uni-frankfurt.de/olat/url/RepositoryEntry/3031171083/CourseNode/91531840833533>
- ... only one person per group!



Assignment 3 – Pig & Hive

- Part A - Complete the Pig & Hive Labs:
 - Pig Lab 5: **Analyzing Ad Campaign Data with Pig**
 - Pig Lab 6: **Analyzing Disparate Data Sets with Pig**
 - Pig Lab 6: **Extending Pig with Streaming and UDFs**
 - Hive Lab 7: **Data Management with Hive**
 - Hive Lab 8: **Data Transformation with Hive**
- Part B - Using Pig & Hive:
 - rewrite your MapReduce programs from Assignment 1**XOR**
 - implement a simple but meaningful new use case

Assignment 3 – Submission

- Data of Group members including: „**Name, Matrikelnummer und Studiengang**” in a file called “group_data.txt”.
- Folder: „Assignment3_A“ including your solutions for **part A** in text files following the structure of the lab exercises. Include appropriate comments where necessary.
- Folder: „Assignment3_B“ your slides in order to present part B. (Presentation will be on: July 8th, 14th and 15th.)

Assignment 4

... details follow ...

area will be In-Memory Databases (SAP-Hana)

