

Beleg 2

Ziel des zweiten Belegs ist es, die Basisfunktionen von Spark zu üben. Dies erfolgt im Rahmen einer Analyse eines Apache Webserverlog-File. Hierzu wird der frei verfügbare Apache Log der NASA aus dem Monat Juni 1995 verwendet.

Das Beleg-Projekt befindet sich im File Beleg2_Aufgab2.zip. Entpacken Sie das Projekt und füllen Sie die Funktionen der Klasse LogAnalyseFuns. Beschreibungen der Funktionen sind im Programmcode. Tests für die einzelnen Funktionen finden Sie im Paket test in der Klasse LogAnalyseFunsTest

Es befinden sich im Test-Ressourcen-Ordner zwei Logfiles, die einen Ausschnitt des Gesamtlogfiles beinhalten. Weiter befindet sich im Hauptressourcen-Ordner der gesamte Logfile. In der Klasse main.App wird dieser geladen und auf Basis Ihrer Funktionen einzelne Statistiken erstellt und innerhalb von Graphen visualisiert. Wird die App gestartet, so sollte der Virtuellen Maschine von Java mehr Speicher zugewiesen werden. Dies kann über das Kommando *env JAVA_OPTS="-Xmx4g" sbt run* erfolgen.

Hilfsfunktionen zum Parsen und bearbeiten des Logfiles befinden sich in der Klasse Utility. Dort wird auch der Aufbau der Log-Einträge beschrieben. Alles Weitere wird in der Übung erklärt.

Die Übungen einzeln gelöst werden, d.h. jeder sollte in der Lage sein, alle Aufgaben zu erklären. Die Abnahme erfolgt in der Übung am 13.07.2018.