

Praktikum 2: Datenbankabfragen

Ziel des Praktikums

Ziel des zweiten Praktikumstermins ist es, erste praktische Erfahrungen mit Abfragen in dokumentenorientierten NoSQL-DBMS – auch im Vergleich zu relationalen DBMS – zu sammeln.

Aufgabe 1 (Anlegen und Befüllen der Datenbanken)

Löschen Sie Ihre in Praktikum 1 angelegten Daten und fügen Sie die 1M-Datensätze in die jeweilige Datenbank ein. Diese finden Sie aufbereitet in den Verzeichnissen

- /pgpool/movielens/adjusted für PostgreSQL
- /mnt/datasets/Movielens/JSON für MongoDB und
- /mnt/datasets/Movielens/couchbase für Couchbase

Hinweis: Bei den aufbereiteten Datensätzen für MongoDB und Couchbase handelt es sich um eine *embedded* Modellierung (statt einer *referentiellen* Variante – auf die referentielle Variante werden wir in Praktikum 3 eingehen). Verwenden Sie die Daten so, wie sie bereitgestellt wurden, d.h. passen Sie ggf. Ihr Datenmodell aus Praktikum 1 entsprechend an die Daten an.

Aufgabe 2 (Datenbankabfragen)

Führen Sie auf den Daten aus Aufgabe 1 die folgende Datenbankabfragen aus und dokumentieren Sie diese im Praktikumsbericht:

1. Ausgabe aller(!) Filme mit zugehörigen Genres und Bewertungen.
2. Ausgabe der Anzahl aller Bewertungen, die abgegeben wurden.
3. Ausgabe des Filmtitels, für den die meisten Bewertungen abgegeben wurden.
4. Ausgabe aller Filmtitel und der durchschnittlichen Bewertung für alle Filme mit einer durchschnittlichen Bewertung ≥ 4 .
5. Ausgabe des besten Films (Title) anhand der Bewertungen (es müssen mindestens 100 Bewertungen für den Film vorhanden sein).
6. Ausgabe aller Action-Filme.
7. Ausgabe aller Filmtitel, die der Nutzer mit der ID = 10 bewertet hat.
8. Ausgabe des Benutzers, der die meisten Bewertungen abgegeben hat.

Hinweis: Die Daten sind so aufbereitet, dass für alle drei Datenbanksysteme die gleichen Ergebnisse erhalten. Überprüfen Sie das insbesondere für die Aufgaben 2, 3, 5 und 8. Außerdem wird Ihnen eine Datei mit den erwarteten Ergebnissen zur Kontrolle bereitgestellt.

Der Praktikumsbericht ist **eine Woche** nach dem jeweiligen Praktikumstermin per Mail als pdf-Datei an die jeweiligen Praktikumsbetreuer abzugeben.