



Aufgabe 1 (20 Punkte)

Beantworten Sie die folgenden Fragen:

- a) Was sind Gründe für die historische Trennung von OLTP und OLAP in verschiedene Systeme? (6 Punkte)
- analytische Anfragen verlangsamen das transaktionale System
 - verschiedene Datenbankschemata notwendig
 - inkompatible SQL-Dialekte
 - kein Bedarf an Vereinigung
 - performante und flexible analytische Anfragen möglich
 - unterschiedliche Workload-Eigenschaften
- b) Warum und wie ist es heutzutage möglich OLTP- und OLAP-Anfragen in einem System effizient zu bearbeiten? (6 Punkte)
- Schnellere Hardware (Moore's Law)
 - Entwicklung spaltenbasierter Hauptspeicher-DBSs
 - Entwicklung von Quantencomputer
 - größerer und günstigerer Hauptspeicher
 - Entwicklung der Blockchain
 - Multicore-CPU's
- c) Warum werden Inserts in Unternehmensanwendungen durch den Umstieg auf spaltenbasierte Hauptspeicherdatenbanken nicht unbedingt langsamer als in einer Zeilendatenbank? (3 Punkte)
- Inserts werden unter keinen Umständen schneller
 - Weil redundante materialisierte Aggregate weggelassen und nicht aktualisiert werden müssen
 - Weil ein spaltenbasiertes Speicherlayout Inserts besser unterstützt
- d) Erklären Sie warum die Vereinigung von OLTP und OLAP die Entwicklung neuer Unternehmensanwendungen vereinfacht. (2 Punkte)
- Einfacheres Schema mit konsistenten Daten.
 - Alle transaktionalen Daten können für Analysen genutzt werden (keine Auswahl von materialisierten Aggregaten).
- e) Welche Vorteile haben spaltenorientierte gegenüber zeilenorientierten Datenbanken? (3 Punkte)
- Bessere Komprimierung und Verarbeitung auf komprimierten Daten gut möglich.
 - Ein spaltenbasiertes Speicherlayout ist für alle Anfragetypen besser geeignet
 - Wenn SQL-Anfragen viele Werte weniger Attribute lesen, erlaubt eine Spaltenorientierung besseres Caching.



Aufgabe 2 (5 Punkte)

- a) Nennen Sie je eine spezifische Datenbankabfrage (SQL oder textuelle Beschreibung), für welche ein spalten- bzw. zeilenorientiertes Datenlayout vorteilhaft ist. Erklären Sie, warum diese Vorteile existieren. (3 Punkte)

- b) Erläutern Sie je einen Vor- und Nachteil eines hybriden Tabellenlayouts. (2 Punkte)

Vorteil: _____

Nachteil: _____

Aufgabe 3 (3 Punkte)

Berechnen Sie die Größe des folgenden unkomprimierten Vektors.
Berechnen Sie anschließend die Größe der einzelnen Datenstrukturen unter Verwendung von Dictionary-Kodierung. Bei der Komprimierung wird die minimal notwendige Anzahl Bits zur Repräsentation der Wert-IDs verwendet.

Anzahl der Werte des Vektors:	3.000
Anzahl der distinkten Werte des Vektors:	142
Fixe Größe der zu kodierenden Werte:	20 Bytes

(3 Punkte)



Abgabeanweisung

Die Aufgaben **müssen** in Zweierteams bearbeitet werden.

Geben Sie die Gruppenmitglieder in der Abgabedatei an.

Die Abgabe erfolgt über **Moodle**:

<https://moodle.hpi.de/course/view.php?id=182>

Reichen Sie eine **pdf-Datei** mit ihren Lösungen ein.

Abgabefrist: 8. Juli 2021 8:00Uhr.

Die pünktliche Abgabe und das Bestehen der Übungsblätter sind Voraussetzung für die Zulassung zur Klausur.

14 von 28 Punkte müssen zum Bestehen des Übungsblattes erreicht werden.