

## 2. Versuch: SQL-Anfragen, Sichten, Updates

### Hinweise

Sofern nicht anders angegeben, sollen die Aufgaben mit nur einer Anfrage ohne weitere Unterstützung (wie Sichten) gelöst werden. Vergessen Sie nicht, Ihre Lösungen sinnvoll zu *formatieren* und zu *kommentieren*.

#### Aufgabe 2.1 (Update; 15 P.)

- Vertauschen Sie die Länder der Kontinente Europa und Asien. Führen Sie die Vertauschung mit *nur einem* Update-Befehl aber ohne die Verwendung von Hilfskonstrukten, PL/SQL, DECODE, CASE etc. durch. **Hinweis:** Der neue Wert des Attributs *continent* muss gewissermaßen als Funktionswert des alten dargestellt werden. Machen Sie die Veränderung danach wieder rückgängig.
- Verwenden Sie nun CASE, um europäische Länder nach Asien, asiatische Länder nach Amerika und amerikanische Länder nach Europa zu verlegen, ebenfalls mit nur einem Update. Machen Sie auch diese Veränderung wieder rückgängig.

#### Aufgabe 2.2 (sym\_borders; 10 P.)

- Schreiben Sie eine Anfrage um zu überprüfen, ob die Relation *borders* asymmetrisch ist. Die Ausgabe soll entweder „ja“ oder „nein“ sein.
- Definieren Sie eine Sicht *sym\_borders*, die die symmetrische Hülle von *borders* enthält. Diese Sicht kann in allen folgenden Aufgaben verwendet werden.

#### Aufgabe 2.3 (Sichten für Zwischenergebnisse; 15 P.)

Bestimmen Sie die Bevölkerungsdichte der Region, die die Länder Algerien, Libyen und sämtliche Nachbarn dieser Länder umfasst. Vergleichen Sie das Ergebnis mit der Bevölkerungsdichte, die man erhält, wenn man die Wüsten als unbewohnbar berücksichtigt. Untersuchen Sie zuvor, ob es Wüsten gibt, die sowohl in als auch außerhalb der oben definierten Region liegen. Verwenden Sie in dieser Aufgabe Sichten zur sinnvollen Strukturierung des Lösungsweges.

#### Aufgabe 2.4 (Redundanz; 10 P.)

Die Informationen über Einwohner werden redundant gespeichert: Sowohl in *Country.population* als auch in *Province.population*. Untersuchen Sie die Daten auf Widersprüche:

- Ermitteln Sie alle Länder, bei denen die Einwohnersumme ihrer Provinzen um wenigstens fünf Prozent von der Einwohnerzahl abweicht. Zeigen Sie den Namen, die absoluten Zahlen und die Abweichung an.
- Berechnen Sie jetzt alle Länder, die Inkonsistenzen zwischen ihrer Einwohnerzahl und den Einwohnern ihrer Städte aufweisen. Hier soll jede Abweichung als fehlerhaft gelten.
- Finden Sie eine weitere mögliche Inkonsistenz bezüglich der Einwohnerzahlen, die bisher noch nicht abgedeckt ist und überprüfen Sie die Daten daraufhin.

#### Aufgabe 2.5 (Urbanisierung; 20 P.)

- Finden Sie heraus, ob der Urbanisierungsgrad und das Pro-Kopf-Einkommen eines Landes korrelieren. Verwenden Sie hierzu die Produkt-Moment-Korrelation  $r$ . Zur Berechnung sind nur die Standard-Aggregationsfunktionen (MIN, MAX, SUM, COUNT, AVG) erlaubt. Die Formel lautet:

$$r = \frac{\sum_{i=1}^n (x_i - m_x) \cdot (y_i - m_y)}{n \cdot s_x \cdot s_y}$$

Dabei sind  $x_i$  und  $y_i$  die einzelnen Meßwerte der beiden Merkmale (hier: Urbanisierungsgrad und Pro-Kopf-Einkommen),  $s_x$  und  $s_y$  deren Standardabweichungen, sowie  $m_x$  und  $m_y$  ihre arithmetischen Mittel. Es gibt positive (je mehr, desto mehr) und negative (je mehr, desto weniger) Korrelationen. Betrachten Sie einen Betrag von  $r$  über 0,1 als geringe, über 0,3 als mittelmäßige und über 0,5 als starke Korrelation.

**Hinweise:** Die Formel soll wie gegeben in eine SQL-Anfrage umgesetzt werden, die  $r$  ermittelt. Schließen Sie Länder mit widersprüchlichen Einwohnerzahlen bezüglich ihrer Städte von der Berechnung aus. Nehmen Sie der Einfachheit halber an, daß Orte ab 10000 Einwohnern als urbanisiert gelten.

- Versuchen Sie, Ihr Ergebnis in einigen kurzen Sätzen zu begründen.

**Aufgabe 2.6 (Durchschnittskordinaten; 10 P.)**

Ermitteln Sie die Durchschnittskordinaten der geographischen Verteilung der Weltbevölkerung.

- a) Bestimmen Sie dazu zunächst den Durchschnittswert der Koordinaten aller Städte.
- b) Versuchen Sie nun, die ungleichmäßige Verteilung der Städte in der Datenbank auszugleichen: Berechnen Sie den Durchschnitt der nach Landesbevölkerung gewichteten durchschnittlichen Koordinaten aller Städte pro Land. Sie berechnen also die Durchschnittskordinaten pro Land, gewichten sie nach Landesbevölkerung und berechnen dann den Gesamtdurchschnitt.

**Aufgabe 2.7 (Mengenvergleich; 10 P.)**

Geben Sie alle Paare von Ländern in Europa aus, die an die gleiche Menge von Meeren angrenzen. **Hinweis:** Mengen sind gleich, wenn sie genau die gleichen Elemente enthalten. Da ein Mengenvergleich in SQL nicht direkt ausgedrückt werden kann, muss die Gleichheit mit den Mengenoperatoren UNION, INTERSECT und MINUS (in Oracle für *except*) bestimmt werden.

Abgabe: 19.5.2010, 11h