Data-Warehouse-Technologien
Aufgabenblatt 6

Aufgabe 1 Welche Datenqualitätsfehler sind in den Relationen Kunde und Bestellungzu finden?

KNr	Nachname	Vorname	Addresse	Stadt	Geburtstag	Vertragsfähig
555666	Maier	Thomas	First Avenue 12	New York	1983-10-10	true
123456	Muster	Max	Rue du Tour 1	Frankreich	1972-01-01	true
112233	Schulz	Maik	M.-Gorki-Str. 5	Machteburch	2000-12-03	true
445566	Thomas	Maier	Rue du Gare 11	Paris	NULL	true
123456	Schulz	Mike	Maxim-Gorki-Strasse 5		1985-08-08	true

Table 1: Relation Kunde

BNr	KNr	Artikel	Menge	Zugestellt
125	555666	4123649700201	1	T
512	123456	4222451689005	Zwei	1
699	112233	40815487990	3	0
730	555566	4900043174599	6	Nein
938	123456	3900004433901	Eins	Ja

Table 2: Relation Bestellung

Klassifizieren Sie die gefundenen Fehler und beschreiben Sie, wie diese Fehler verhindert werden können?

Aufgabe 2 Beschreiben sie die Phasen der Datenaufbereitung im Data Warehouse!

Aufgabe 3 Erläutern Sie das Problem der Duplikaterkennung? Welche Methoden können zur Duplikaterkennung verwendet werden? Welcher Unterschied existiert zum Differential Snapshot Problem?

Aufgabe 4 Erläutern Sie den Unterschied zwischen dem ETL und ELT Prozess.

 Viel Erfolg!