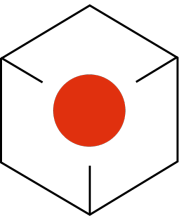


Wie ging das schon wieder mit diesen Normalformen?

Dr. Andrea Kennel
InfoPunkt Kennel GmbH
Dübendorf-Schweiz
November 2023



Dr. Andrea Kennel

SYM^{L2}



Consultant

Dozentin für Datenbanken

Coach für Project Management

Fachhochschule Nordwestschweiz

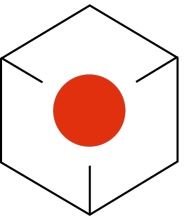
Brugg/Windisch, Schweiz



andrea.kennel@fhnw.ch

andrea@infokennel.ch

www.infokennel.ch



Ein Dialog in der Nanobrauerei

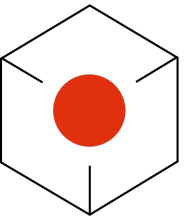
M: Da ist eine Rechnung zurück gekommen.

B: Die sind doch umgezogen, habe das bei der letzten Rechnung so angepasst.

M: Uups. Da habe ich eine zu alte Rechnung als Vorlage genommen.

B: Vielleicht brauchen wir doch mal eine Datenbanklösung, in der die Rechnungen dann gleich mit der neusten Adresse generiert werden.

Redundante Daten können problematisch sein.



Begriff Redundanz

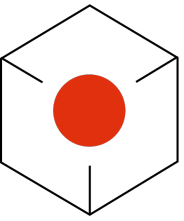
Bedeutung gemäss Duden:

Das Vorhandensein von eigentlich überflüssigen, für die Information nicht notwendigen Elementen

Bedeutung in der Informationstheorie:

Der Begriff der Redundanz beschreibt in der Informationstheorie diejenigen Informationen oder Daten, die in einer Informationsquelle mehrfach vorhanden sind. Eine Informationseinheit ist dann redundant, wenn sie ohne Informationsverlust weggelassen werden kann.

Quelle: [https://de.wikipedia.org/wiki/Redundanz_\(Informationstheorie\)](https://de.wikipedia.org/wiki/Redundanz_(Informationstheorie))

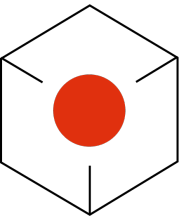


Begriff Normalisierung

Bedeutung:

Normalisierung ist ein Entwurfsansatz für relationale Datenbanken mit dem Zweck, redundante Speicherung von Informationen und damit Inkonsistenz und Anomalien zu vermeiden. Die Methode strukturiert die Daten anhand einer Folge von Regeln – „Normalformen“ genannt – die aufeinander aufbauen und formale Anforderungen an das Schema bestimmen.

Quelle: [https://de.wikipedia.org/wiki/Normalisierung_\(Datenbank\)](https://de.wikipedia.org/wiki/Normalisierung_(Datenbank))



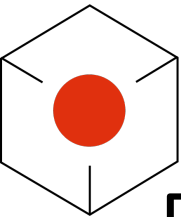
Folgerung

Wenn Informationen nur einmal in der Datenbank vorkommen, können sich diese nicht widersprechen.

Beispiel Nanobrauerei:

Wenn ich Kundenadressen nur einmal speichere, kann ich nicht eine alte Adresse erwischen.

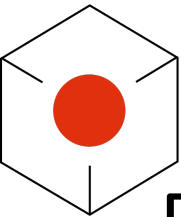
Die Normalformen beginnen aber schon dabei, die Daten so zu strukturieren, dass wir diese auch sinnvoll nutzen können.



Ein Datenbeispiel

Wie viele Lazariter
haben wir verkauft?

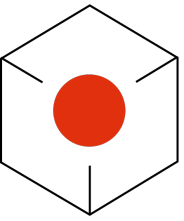
<u>RechnungNr</u>	<u>Datum</u>	<u>Kunde</u>	<u>Adresse</u>	<u>Bestellung</u>	<u>Preis total</u>
101	02.10.22	KinoBar	Filmstrasse 77, 8600 Dübendorf	5 Schwarze Kobra 22.5 CHF; 5 Lazariter 37.5 CHF; 10 Depotflaschen 0.5 5 CHF;	65 CHF
102	02.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	10 Blauer Pirat 75 CHF;	75 CHF
103	02.10.22	NachBar	Hopfenweg 12, 8057 Zürich	2 Lazariter 15 CHF; 3 Schwarze Kobra 13.5 CHF; 2 TrippTrap 14 CHF; 5 Depotflaschen 0.5 2.5 CHF; 2 Depotflaschen 1.0 2 CHF;	47 CHF
104	09.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	20 Blauer Pirat 150 CHF;	150 CHF



Ein Datenbeispiel

Wie viele Lazariter
haben wir verkauft?

RechnungNr	Datum	Kunde	Adresse	Bestellung	Preis total
101	02.10.22	KinoBar	Filmstrasse 77, 8600 Dübendorf	5 Schwarze Kobra 22.5 CHF; 5 Lazariter 37.5 CHF; 10 Depotflaschen 0.5 5 CHF;	65 CHF
102	02.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	10 Blauer Pirat 75 CHF;	75 CHF
103	02.10.22	NachBar	Hopfenweg 12, 8057 Zürich	2 Lazariter 15 CHF; 3 Schwarze Kobra 13.5 CHF; 2 TrippTrap 14 CHF; 5 Depotflaschen 0.5 2.5 CHF; 2 Depotflaschen 1.0 2 CHF;	47 CHF
104	09.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	20 Blauer Pirat 150 CHF;	150 CHF

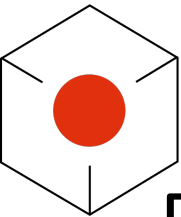


Frage an ChatGPT: Was ist in Datenbanken die erste Normalform?

Die erste Normalform (1NF) ist ein Konzept in der Datenbanktheorie, das bestimmte Anforderungen an die Struktur von Datenbanktabellen stellt, um Redundanzen und Anomalien zu vermeiden. Die erste Normalform besteht aus den folgenden Regeln:

1. **Atomare Werte**: Jedes Attribut (Spalte) einer Tabelle enthält nur atomare (unzerlegbare) Werte. Das bedeutet, dass jedes Feld einer Spalte nur einen einzelnen Wert enthält und keine mehrwertigen Attribute oder wiederholten Gruppen von Werten enthält.
2. **Eindeutige Bezeichner**: Jeder Datensatz (Zeile) in der Tabelle muss eindeutig identifiziert werden können. Dazu gehört die Verwendung eines Primärschlüssels, der eindeutig ist und keine Duplikate enthält.
3. **Keine Reihenfolge** oder Sortierung: Die Reihenfolge der Datensätze in der Tabelle ist nicht relevant und sollte keine Auswirkungen auf die Bedeutung der Daten haben.

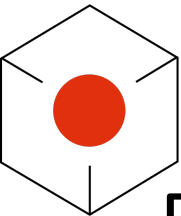
Wenn diese Regeln erfüllt sind, befindet sich die Tabelle in der ersten Normalform. Die Einhaltung der ersten Normalform ist die grundlegende Anforderung für die Datenmodellierung in relationalen Datenbanken.



Ein Datenbeispiel

Bestellpositionen aufteilen
und für Kunde und Artikel
eindeutige Nummern einführen

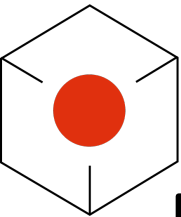
RechnungNr	Datum	Kunde	Adresse	Bestellung	Preis total
101	02.10.22	KinoBar	Filmstrasse 77, 8600 Dübendorf	5 Schwarze Kobra 22.5 CHF; 5 Lazariter 37.5 CHF; 10 Depotflaschen 0.5 5 CHF;	65 CHF
102	02.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	10 Blauer Pirat 75 CHF;	75 CHF
103	02.10.22	NachBar	Hopfenweg 12, 8057 Zürich	2 Lazariter 15 CHF; 3 Schwarze Kobra 13.5 CHF; 2 TrippTrap 14 CHF; 5 Depotflaschen 0.5 2.5 CHF; 2 Depotflaschen 1.0 2 CHF;	47 CHF
104	09.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	20 Blauer Pirat 150 CHF;	150 CHF



Ein Datenbeispiel

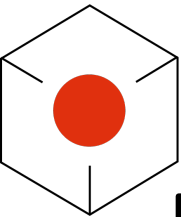
Bestellpositionen aufteilen
und für Kunde und Artikel
eindeutige Nummern einführen

RechnungNr	Datum	Kunde	Adresse	Bestellung	Preis total
101	02.10.22	KinoBar	Filmstrasse 77, 8600 Dübendorf	5 Schwarze Kobra 22.5 CHF; 5 Lazariter 37.5 CHF; 10 Depotflaschen 0.5 5 CHF;	65 CHF
102	02.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	10 Blauer Pirat 75 CHF;	75 CHF
103	02.10.22	NachBar	Hopfenweg 12, 8057 Zürich	2 Lazariter 15 CHF; 3 Schwarze Kobra 13.5 CHF; 2 TrippTrap 14 CHF; 5 Depotflaschen 0.5 2.5 CHF; 2 Depotflaschen 1.0 2 CHF;	47 CHF
104	09.10.22	MyBar	Dorfstrasse 5, 8600 Dübendorf	20 Blauer Pirat 150 CHF;	150 CHF



Ein Datenbeispiel

<u>Rechnung</u> Nr	Datum	KundeNr	Kunde	Strasse	PLZ	Ort	ArtikelNr	Artikel	Artikelgruppe	Artikelkate gorie	Anzahl	Währung	Preis Position	Preis total
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	2	Schwarze Kobra	Stout	Bier	5	CHF	22.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	5	Lazariter	Belgisches Klosterbier	Bier	5	CHF	37.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	8	Depotflaschen 0.5	Depotflasche	Gebinde	10	CHF	5	65
102	02.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	10	CHF	75	75
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	5	Lazariter	Belgisches Klosterbier	Bier	2	CHF	15	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	2	Schwarze Kobra	Stout	Bier	3	CHF	13.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	3	TrippTrap	Belgisches Klosterbier	Bier	2	CHF	14	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	8	Depotflaschen 0.5	Depotflasche	Gebinde	5	CHF	2.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	9	Depotflaschen 1.0	Depotflasche	Gebinde	2	CHF	2	47
104	09.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	20	CHF	150	150

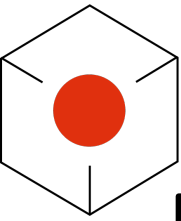


Ein Datenbeispiel

Nun haben wir eindeutig Redundanz

Rechnung Nr	Datum	KundeNr	Kunde	Strasse	PLZ	Ort	ArtikelNr	Artikel	Artikelgruppe	Artikelkate gorie	Anzahl	Währung	Preis Position	Preis total
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	2	Schwarze Kobra	Stout	Bier	5	CHF	22.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	5	Lazariter	Belgisches Klosterbier	Bier	5	CHF	37.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	8	Depotflaschen 0.5	Depotflasche	Gebinde	10	CHF	5	65
102	02.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	10	CHF	75	75
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	5	Lazariter	Belgisches Klosterbier	Bier	2	CHF	15	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	2	Schwarze Kobra	Stout	Bier	3	CHF	13.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	3	TrippTrap	Belgisches Klosterbier	Bier	2	CHF	14	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	8	Depotflaschen 0.5	Depotflasche	Gebinde	5	CHF	2.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	9	Depotflaschen 1.0	Depotflasche	Gebinde	2	CHF	2	47
104	09.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	20	CHF	150	150

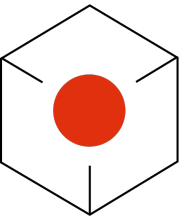
Wieso das?



Ein Datenbeispiel

Rechnung Nr	Datum	KundeNr	Kunde	Strasse	PLZ	Ort	ArtikelNr	Artikel	Artikelgruppe	Artikelkate gorie	Anzahl	Währung	Preis Position	Preis total
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	2	Schwarze Kobra	Stout	Bier	5	CHF	22.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	5	Lazariter	Belgisches Klosterbier	Bier	5	CHF	37.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	8	Depotflaschen 0.5	Depotflasche	Gebinde	10	CHF	5	65
102	02.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	10	CHF	75	75
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	5	Lazariter	Belgisches Klosterbier	Bier	2	CHF	15	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	2	Schwarze Kobra	Stout	Bier	3	CHF	13.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	3	TrippTrap	Belgisches Klosterbier	Bier	2	CHF	14	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	8	Depotflaschen 0.5	Depotflasche	Gebinde	5	CHF	2.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	9	Depotflaschen 1.0	Depotflasche	Gebinde	2	CHF	2	47
104	09.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	20	CHF	150	150





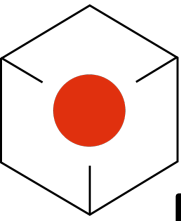
Die zweite Normalform (2NF)

1NF ist erfüllt und

Jedes Attribut ist voll funktional abhängig vom Primärschlüssel

Jedes Attribut ist durch den Schlüssel eindeutig bestimmt

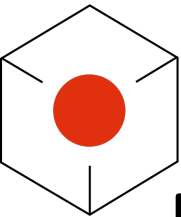
Die zweite Normalform stellt sicher, dass alle Attribute vom ganzen Schlüssel und **nicht** nur **von einem Teil** abhängig sind.



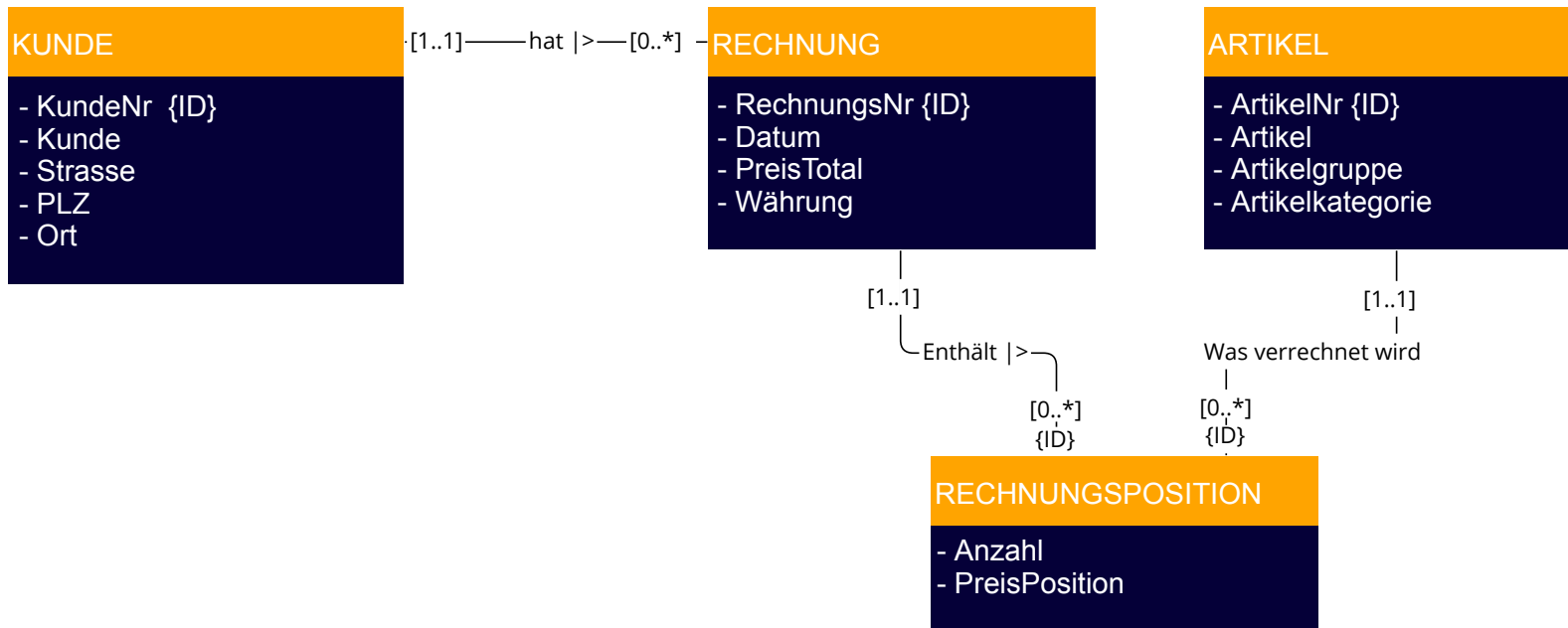
Ein Datenbeispiel

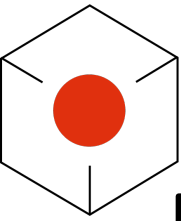
Rechnung Nr	Datum	KundeNr	Kunde	Strasse	PLZ	Ort	ArtikelNr	Artikel	Artikelgruppe	Artikelkate gorie	Anzahl	Währung	Preis Position	Preis total
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	2	Schwarze Kobra	Stout	Bier	5	CHF	22.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	5	Lazariter	Belgisches Klosterbier	Bier	5	CHF	37.5	65
101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	8	Depotflaschen 0.5	Depotflasche	Gebinde	10	CHF	5	65
102	02.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	10	CHF	75	75
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	5	Lazariter	Belgisches Klosterbier	Bier	2	CHF	15	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	2	Schwarze Kobra	Stout	Bier	3	CHF	13.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	3	TrippTrap	Belgisches Klosterbier	Bier	2	CHF	14	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	8	Depotflaschen 0.5	Depotflasche	Gebinde	5	CHF	2.5	47
103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	9	Depotflaschen 1.0	Depotflasche	Gebinde	2	CHF	2	47
104	09.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	20	CHF	150	150





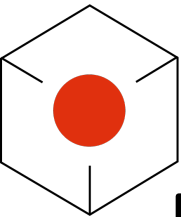
Das ERD dazu (CHEN und UML)





Ein Datenbeispiel

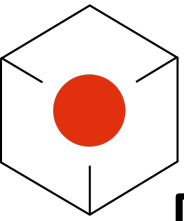
Kunde					Rechnung					Artikel			
KundeNr	Kunde	Strasse	PLZ	Ort	RechnungNr	KundeNr	Datum	Preis total	Währung	ArtikelNr	Artikel	Artikelgruppe	Artikelkategorie
201	KinoBar	Filmstrasse 77	8600	Dübendorf	101	201	02.10.22	65	CHF	2	Schwarze Kobra	Stout	Bier
202	MyBar	Dorfstrasse 5	8600	Dübendorf	102	202	02.10.22	75	CHF	3	TrippTrap	Belgisches Klosterbier	Bier
203	NachBar	Hopfenweg 12	8057	Zürich	103	203	02.10.22	47	CHF	4	Blauer Pirat	IPA	Bier
					104	202	09.10.22	150	CHF	5	Lazariter	Belgisches Klosterbier	Bier
										8	Depotflaschen 0.5	Depotflasche	Gebinde
										9	Depotflaschen 1.0	Depotflasche	Gebinde
					Rechnungsposition								
					RechnungNr	ArtikelNr	Anzahl	Preis	Position				
					101	2	5	22.5					
					101	5	5	37.5					
					101	8	10	5					
					102	4	10	75					
					103	5	2	15					
					103	2	3	13.5					
					103	3	2	14					
					103	8	5	2.5					
					103	9	2	2					
					104	4	20	150					



Ein Datenbeispiel

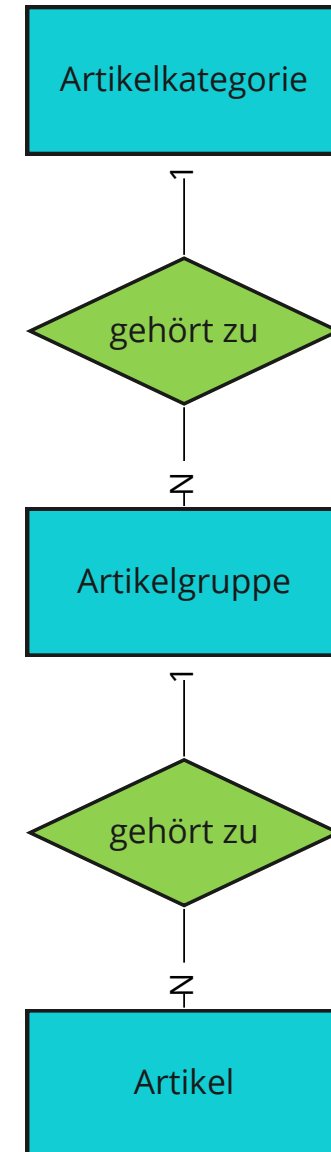
Artikel			
<u>ArtikelNr</u>	Artikel	Artikelgruppe	Artikelkategorie
2	Schwarze Kobra	Stout	Bier
3	TrippTrap	Belgisches Klosterbier	Bier
4	Blauer Pirat	IPA	Bier
5	Lazariter	Belgisches Klosterbier	Bier
8	Depotflaschen 0.5	Depotflasche	Gebinde
9	Depotflaschen 1.0	Depotflasche	Gebinde

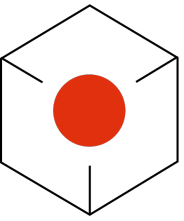
Gibt es noch Redundanz?



Ein Datenbeispiel

Artikel			
<u>ArtikelNr</u>	Artikel	Artikelgruppe	Artikelkategorie
2	Schwarze Kobra	Stout	Bier
3	TrippTrap	Belgisches Klosterbier	Bier
4	Blauer Pirat	IPA	Bier
5	Lazariter	Belgisches Klosterbier	Bier
8	Depotflaschen 0.5	Depotflasche	Gebinde
9	Depotflaschen 1.0	Depotflasche	Gebinde





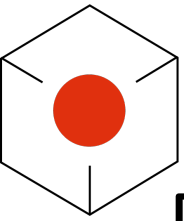
Die dritte Normalform (3NF)

2NF erfüllt und

Kein Attribut ist transitiv abhängig

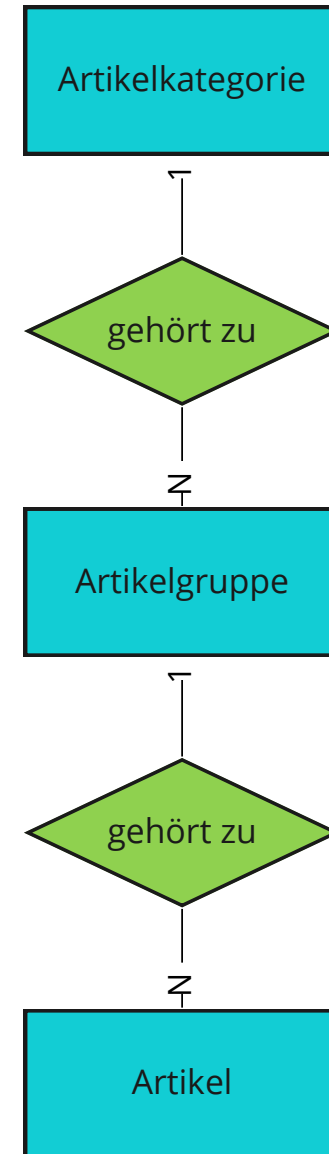
C ist transitiv abhängig von A, wenn B von A und C von B funktional abhängig ist

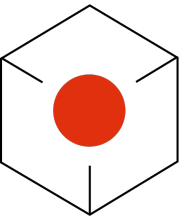
Die dritte Normalform stellt sicher, dass alle Attribute direkt vom Schlüssel abhängig sind.



Ein Datenbeispiel

Artikel			
<u>ArtikelNr</u>	Artikel	Artikelgruppe	Artikelkategorie
2	Schwarze Kobra	Stout	Bier
3	TrippTrap	Belgisches Klosterbier	Bier
4	Blauer Pirat	IPA	Bier
5	Lazariter	Belgisches Klosterbier	Bier
8	Depotflaschen 0.5	Depotflasche	Gebinde
9	Depotflaschen 1.0	Depotflasche	Gebinde
A	B	C	D



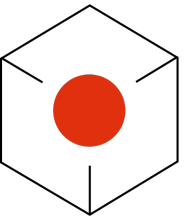


Weitere Normalformen

Weitere Normalformen sind

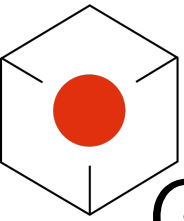
- **Boyce-Codd-Normalform**
- **Vierte Normalform (4NF)**
- **Fünfte Normalform (5NF)**

Diese Normalformen sind in der Praxis nicht relevant



Zusammenhang der Normalformen





Grenzen der Normalisierung

Kunde				
<u>KundeNr</u>	Kunde	Strasse	PLZ	Ort
201	KinoBar	Filmstrasse 77	8600	Dübendorf
202	MyBar	Dorfstrasse 5	8600	Dübendorf
203	NachBar	Hopfenweg 12	8057	Zürich

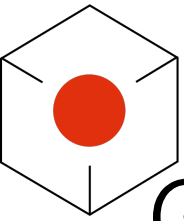
Kann Ort normalisiert werden?

Wäre das sinnvoll?

Wie oft ändern PLZ?

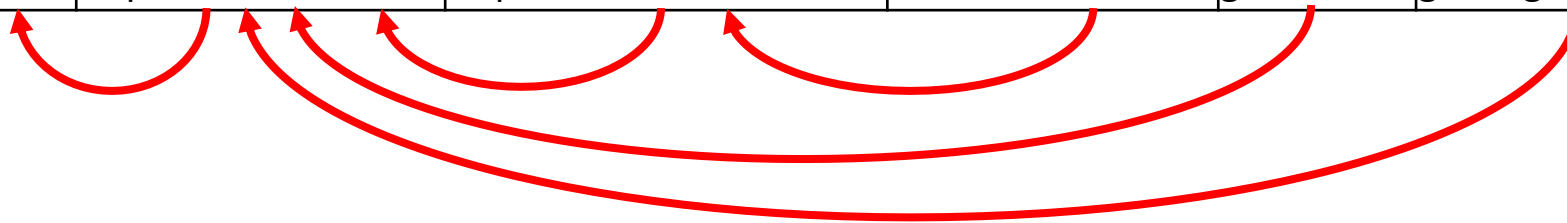
PLZ	Ort
8044	Zürich
8044	Gockhausen
8006	Zürich

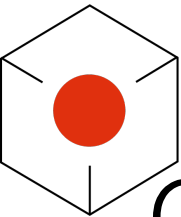
Kann man überhaupt noch normalisieren?



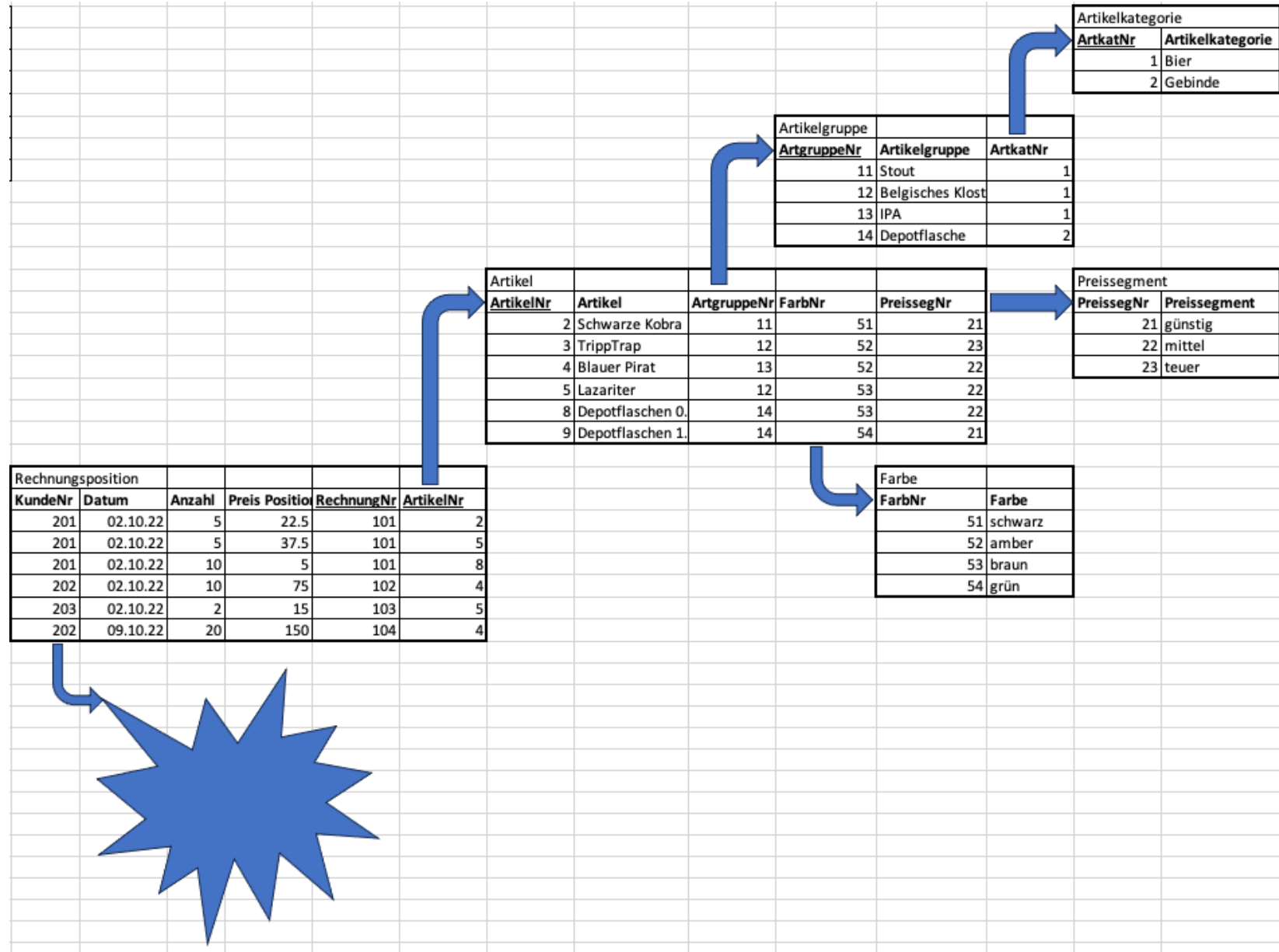
Grenzen der Normalisierung

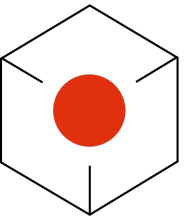
Artikel					
<u>ArtikelNr</u>	Artikel	Artikelgruppe	Artikelkategorie	Farbe	Preissegment
2	Schwarze Kobra	Stout	Bier	schwarz	günstig
3	TrippTrap	Belgisches Klosterbier	Bier	amber	teuer
4	Blauer Pirat	IPA	Bier	amber	mittel
5	Lazariter	Belgisches Klosterbier	Bier	braun	mittel
8	Depotflaschen 0.5	Depotflasche	Gebinde	braun	mittel
9	Depotflaschen 1.0	Depotflasche	Gebinde	grün	günstig





Grenzen



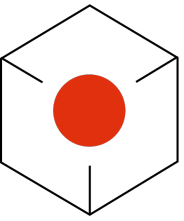


Komplexität durch Views reduzieren

Die Tabellen normalisiert anlegen.

Wo sinnvoll mit Lookup Tabellen arbeiten.

Gegen aussen Views, die Komplexität reduzieren.



Begriff Redundanz

Eine Informationseinheit ist dann redundant, wenn sie ohne Informationsverlust weggelassen werden kann.

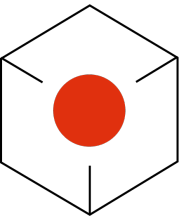
Warum ist das wichtig?

Beim Schreiben, Ändern können Daten inkonsistent werden.

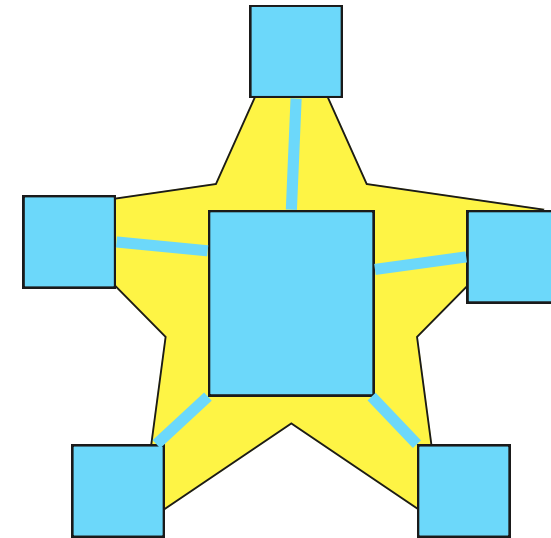
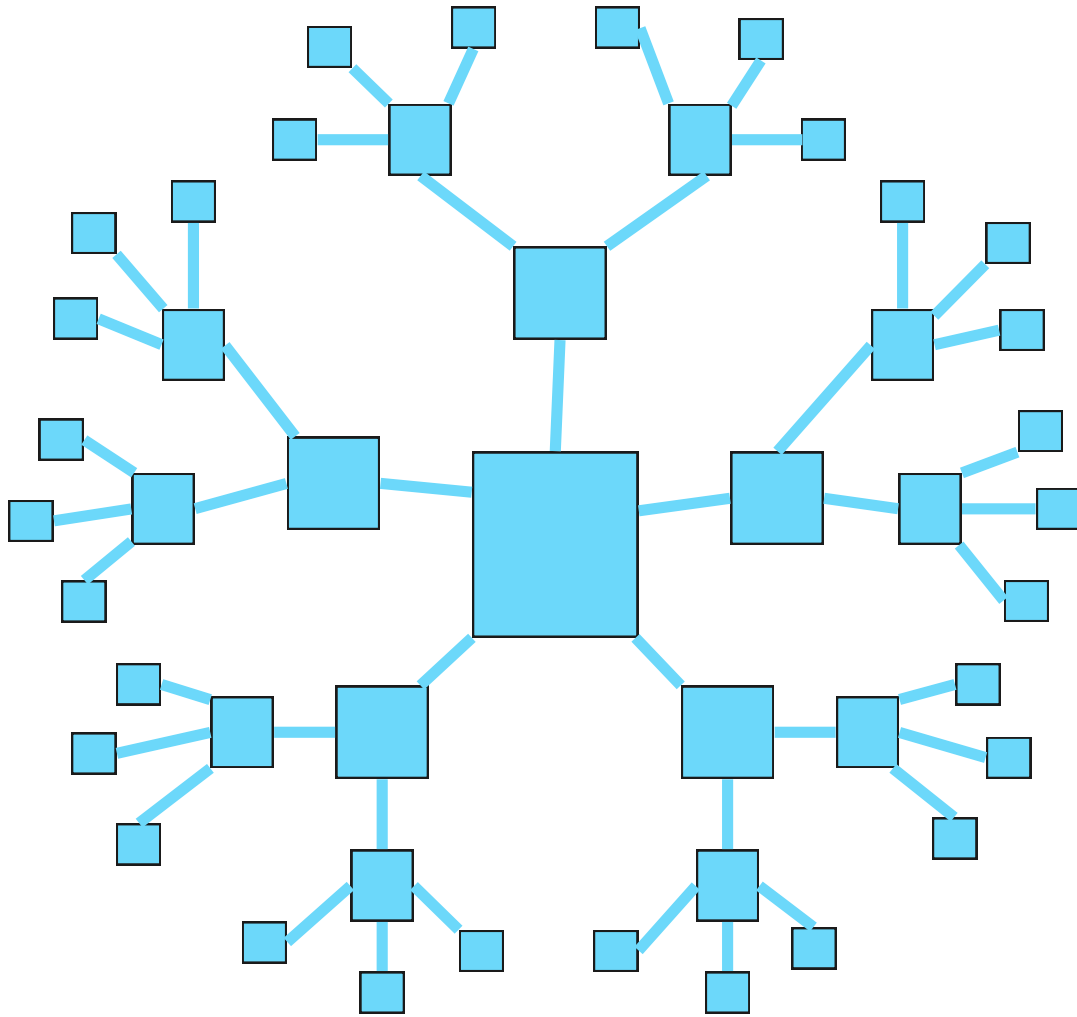
➔ OLTP

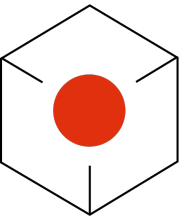
Wie sieht das bei OLAP aus?

Redundanz hilft Abfragen beschleunigen.

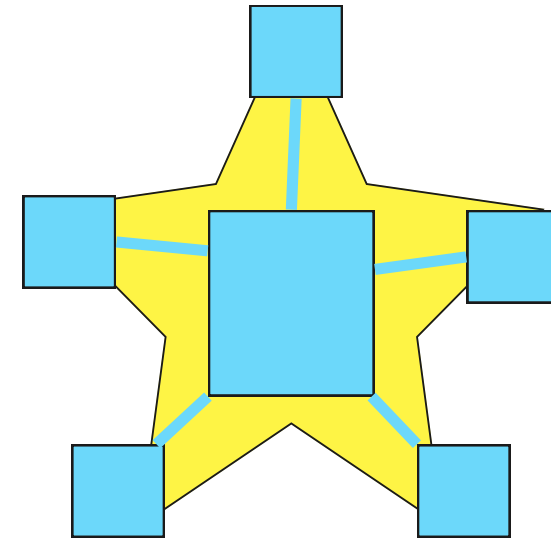
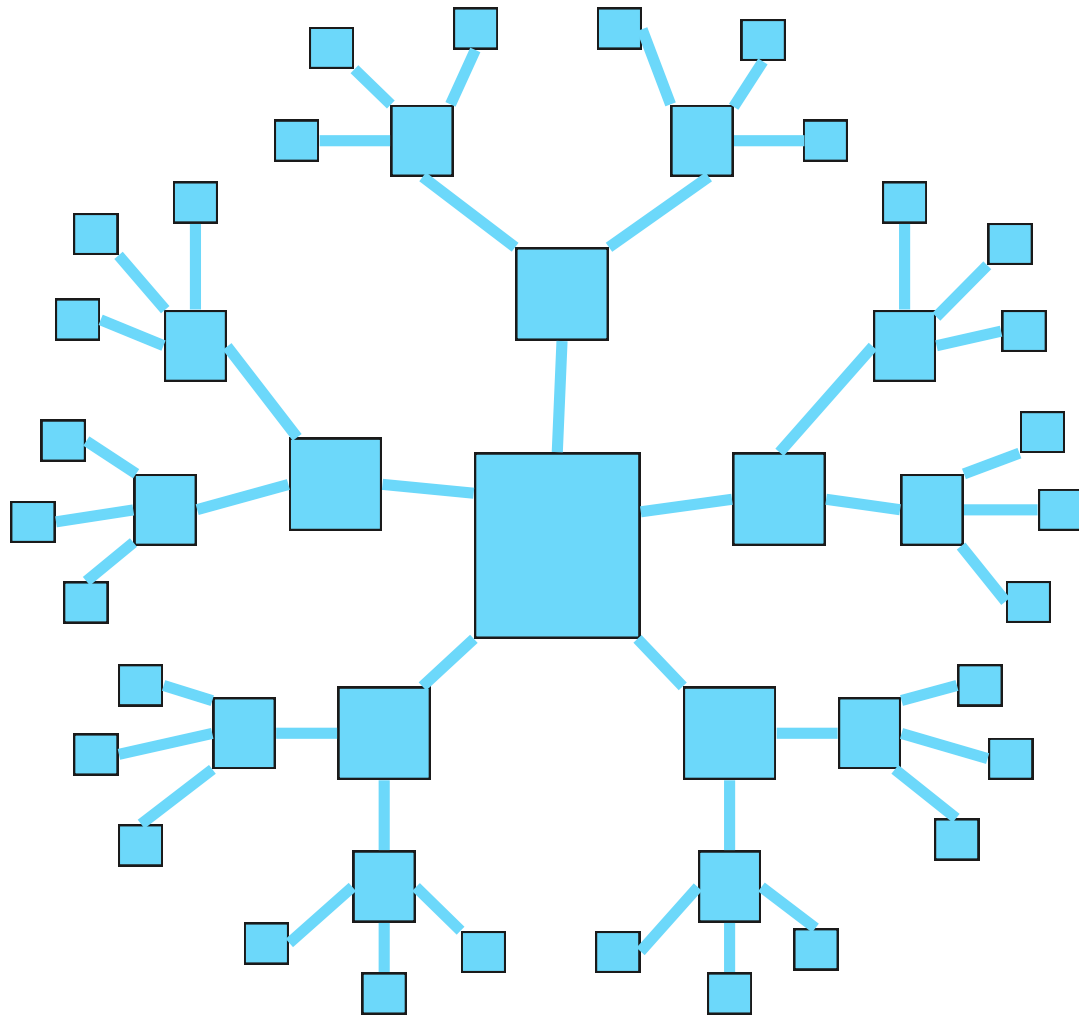


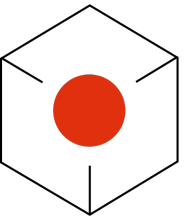
DWH ohne Redundanz und mit Redundanz





Snowflake (Schneeflocke) und Star (Stern)





Folgerung

Normalformen sind in relationalen Datenbanken wichtig.

OLTP

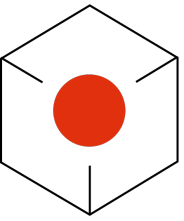
2NF wichtig

3NF mehrheitlich beachten

OLAP

2NF und 3NF weniger wichtig

Wichtiger: bewusste Redundanz für schnellere Abfragen



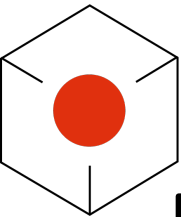
War's das schon?

Wer hat schon mal was gehört von

- Superschlüssel
- Schlüsselkandidat
- Funktionale Abhängigkeit

Wie hängen diese mit den Normalformen zusammen?

Braucht man das nicht für Normalformen?



Ein Datenbeispiel

ID	Rechnung Nr	Datum	KundeNr	Kunde	Strasse	PLZ	Ort	ArtikelNr	Artikel	Artikelgruppe	Artikelkate gorie	Anzahl	Währung	Preis Position	Preis total
1	101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	2	Schwarze Kobra	Stout	Bier	5	CHF	22.5	65
2	101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	5	Lazariter	Belgisches Klosterbier	Bier	5	CHF	37.5	65
3	101	02.10.22	201	KinoBar	Filmstrasse 77	8600	Dübendorf	8	Depotflaschen 0.5	Depotflasche	Gebinde	10	CHF	5	65
4	102	02.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	10	CHF	75	75
5	103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	5	Lazariter	Belgisches Klosterbier	Bier	2	CHF	15	47
6	103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	2	Schwarze Kobra	Stout	Bier	3	CHF	13.5	47
7	103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	3	TrippTrap	Belgisches Klosterbier	Bier	2	CHF	14	47
8	103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	8	Depotflaschen 0.5	Depotflasche	Gebinde	5	CHF	2.5	47
9	103	02.10.22	203	NachBar	Hopfenweg 12	8057	Zürich	9	Depotflaschen 1.0	Depotflasche	Gebinde	2	CHF	2	47
10	104	09.10.22	202	MyBar	Dorfstrasse 5	8600	Dübendorf	4	Blauer Pirat	IPA	Bier	20	CHF	150	150

Welches ist der Schlüssel?

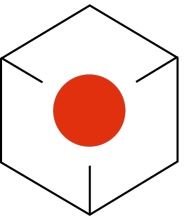
Sind alle Attribute zusammen ein Schlüssel?

Gibt es Schlüssel, die nicht reduzierbar sind?

Das sind **Schlüsselkandidaten**

Ja, ein **Superschlüssel**

Ja: ID oder RechnungsNr und ArtikelNr zusammen

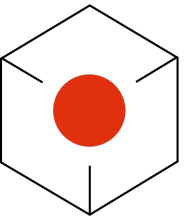


Die zweite Normalform (2NF)

Eine Relation ist genau dann in zweiter Normalform, wenn sie

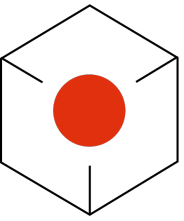
1. in der ersten Normalform ist und
2. für jedes Attribut α der Relation gilt:
 - α ist Teil eines Schlüsselkandidaten oder
 - α ist von einem Schlüsselkandidaten abhängig und
 - α ist nicht von einer echten Teilmenge eines Schlüsselkandidaten abhängig.

Quelle: [https://de.wikipedia.org/wiki/Normalisierung_\(Datenbank\)#Zweite_Normalform_\(2NF\)](https://de.wikipedia.org/wiki/Normalisierung_(Datenbank)#Zweite_Normalform_(2NF))



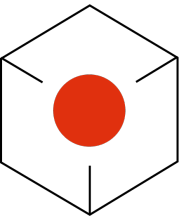
Funktionale Abhängigkeit (FD): Notation

- FD = functional dependency = funktionale Abhängigkeit
- ArtikelNr \rightarrow Artikelname, Farbe =
 - Artikelname und Farbe sind funktional abhängig von ArtikelNr
 - ArtikelNr determiniert funktional Artikelname und Farbe
- $X \rightarrow Y$
 - Sei $R(A_1, A_2, \dots, A_n)$ ein Relationenschema
 - $\text{Attr}(R) = \{A_1, A_2, \dots, A_n\}$ die Menge aller Attribute von R
 - Seien weiter X und Y nichtleere Teilmengen von $\text{Attr}(R)$



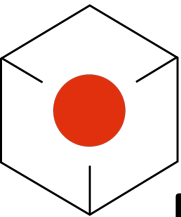
Funktionale Abhängigkeit (FD): Gültigkeit

- Die FD $X \rightarrow Y$ ist gültig in der Relation $r(R)$ wenn gilt:
 - stimmen beliebige Tupel von r im Wert von X überein, so stimmen sie auch im Wert von Y überein
 - Also: $\forall t1, t2 \in r : t1[X] = t2[X] \Rightarrow t1[Y] = t2[Y]$
 - Deutsch: die Attributswerte Y sind von den Attributswerte X abhängig
- Die FD $X \rightarrow Y$ ist gültig in der Relation $r(R)$ wenn gilt:
 - $\text{card}(\pi_X(r)) = \text{card}(\pi_{X,Y}(r))$
 - Deutsch: Die Mengen der Attributswerte von X ist gleich gross, wie die Mengen der Attributswerte von XY



Funktionale Abhängigkeit (FD): Regeln

- Seien $R(A_1, A_2, \dots, A_n)$ ein Relationenschema und X , Y und Z nichtleere Teilmengen von $\text{Attr}(R)$:
 - a) Reflexivität: aus $\{Y\} \subseteq \{X\}$ folgt $X \rightarrow Y$
 - b) Erweiterung: aus $X \rightarrow Y$ folgt $XZ \rightarrow YZ$
 - c) Transitivität: aus $X \rightarrow Y$, $Y \rightarrow Z$ folgt $X \rightarrow Z$
- Das bedeutet: Z ist transitive abhängig von X



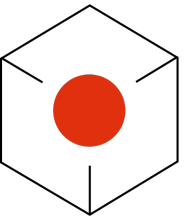
Ein Datenbeispiel

Artikel			
<u>ArtikelNr</u>	Artikel	Artikelgruppe	Artikelkategorie
2	Schwarze Kobra	Stout	Bier
3	TrippTrap	Belgisches Klosterbier	Bier
4	Blauer Pirat	IPA	Bier
5	Lazariter	Belgisches Klosterbier	Bier
8	Depotflaschen 0.5	Depotflasche	Gebinde
9	Depotflaschen 1.0	Depotflasche	Gebinde



ArtikelNr → Artikel und Artikel → Artikelgruppe

Das bedeutet Artikelgruppe ist transitiv von ArtikelNr abhängig

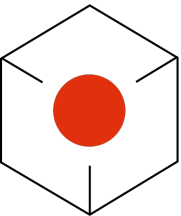


Folgerung

Die relationale Algebra hilft die Normalformen präzise zu definieren und überprüfen.

Dies ist aber nur möglich, wenn alle funktionalen Abhängigkeiten definiert sind und fachlich auch korrekt sind.

In der Praxis ist die relationale Algebra nicht zwingend, um Datenbanken zu normalisieren.



Dr. Andrea Kennel

SYM^{L2}



Consultant

Dozentin für Datenbanken

Coach für Project Management

Fachhochschule Nordwestschweiz

Brugg/Windisch, Schweiz



andrea.kennel@fhnw.ch

andrea@infokennel.ch

www.infokennel.ch