

Relationale Datenbanken

Cumulus-DB mit Excel

Name	Klasse	Sparte	Artikel	Preis	Laden
Tick Duck	12s	Lebensmittel	Brot	2.50	Coop Bahnhof Biel
Bart Simpson	12s	Lebensmittel	Milch	1.90	Migros Innenstadt
Lisa Simpson	14e	Kleider	Jeans	49.50	Migros Bahnhof Bern
Trick Duck	13d	Lebensmittel	Chips	3.70	Bahnhof-Coop in Biel
Homer Simpson	12s	Freizeit	Kinoticket	17.90	Alhambra Bern
Daisy Duck	14e	Kosmetik	Lippenstift	14.95	Migros Innenstadt
Donald Duck	14e	Kosmetik	Haargel	12.80	Coop Bahnhof Biel
Dagobert Duck	13d	lebensmittel	Mars	1.20	Migros Bahnhof Bern
Track Duck	12s	Freizeit	CD	19.90	City-Disk Biel
Gustav Gans	12s	Lebensmittel	M&Ms	4.20	Coop Bern

Alle Einkäufe im Coop?

Alle Einkäufe der Simpsons?

Name	Klasse	Sparte	Artikel	Preis	Laden
Tick Duck	12s	Lebensmittel	Brot	2.50	Coop Bahnhof Biel
Bart Simpson	12s	Lebensmittel	Milch	1.90	Migros Innenstadt
Lisa Simpson	14e	Kleider	Jeans	49.50	Migros Bahnhof Bern
Trick Duck	13d	Lebensmittel	Chips	3.70	Bahnhof-Coop in Biel
Homer Simpson	12s	Freizeit	Kinoticket	17.90	Alhambra Bern
Daisy Duck	14e	Kosmetik	Lippenstift	14.95	Migros Innenstadt
Donald Duck	14e	Kosmetik	Haargel	12.80	Coop Bahnhof Biel
Dagobert Duck	13d	lebensmittel	Mars	1.20	Migros Bahnhof Bern
Track Duck	12s	Freizeit	CD	19.90	City-Disk Biel
Gustav Gans	12s	Lebensmittel	M&Ms	4.20	Coop Bern

„Atomare“ Felder wählen!

Vorname	Nachname	Klasse	Sparte	Artikel	Preis	Laden	Ort
Tick	Duck	12s	Lebensmittel	Brot	2.50	Coop	Bahnhof Biel
Bart	Simpson	12s	Lebensmittel	Milch	1.90	Migros	Innenstadt Biel
Lisa	Simpson	14e	Kleider	Jeans	49.50	Migros	Bahnhof Bern
Trick	Duck	13d	Lebensmittel	Chips	3.70	Coop	Biel
Homer	Simpson	12s	Freizeit	Kinoticket	17.90	Alhambra	Bern
Daisy	Duck	14e	Kosmetik	Lippenstift	14.95	Migros	Innenstadt Biel
Donald	Duck	14e	Kosmetik	Haargel	12.80	Coop	Bahnhof Biel
Dagobert	Duck	13d	lebensmittel	Mars	1.20	Migros	Bahnhof Bern
Track	Duck	12s	Freizeit	CD	19.90	City-Disk	Biel
Gustav	Gans	12s	Lebensmittel	M&Ms	4.20	Coop	Bern

Schlampige Dateneingabe

Name	Klasse	Sparte	Artikel	Preis	Laden
Tick Duck	12s	Lebensmittel	Brot	2.50	Coop Bahnhof Biel
Bart Simpson	12s	Lebensmittel	Milch	1.90	Migros Innenstadt
Lisa Simpson	14e	Kleider	Jeans	49.50	Migros Bahnhof Bern
Trick Duck	13d	Lebensmittel	Chips	3.70	Bahnhof-Coop in Biel
Homer Simpson	12s	Freizeit	Kinoticket	17.90	Alhambra Bern
Daisy Duck	14e	Kosmetik	Lippenstift	14.95	Migros Innenstadt
Donald Duck	14e	Kosmetik	Haargel	12.80	Coop Bahnhof Biel
Dagobert Duck	13d	lebensmittel	Mars	1.20	Migros Bahnhof Bern
Track Duck	12s	Freizeit	CD	19.90	City-Disk Biel
Gustav Gans	12s	lebensmittel	M&Ms	4.20	Coop Bern

Zählt: Wie oft seht ihr das Wort „Lebensmittel“ ?

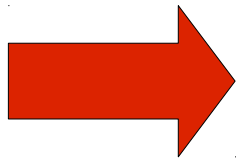
Name	Klasse	Sparte	Artikel	Preis	Laden
Tick Duck	12s	Lebensmittel	Brot	2.50	Coop Bahnhof Biel
Bart Simpson	12s	Lebensmittel	Milch	1.90	Migros Innenstadt
Lisa Simpson	14e	Kleider	Jeans	49.50	Migros Bahnhof Bern
Trick Duck	13d	Lebensmittel	Chips	3.70	Bahnhof-Coop in Biel
Homer Simpson	12s	Freizeit	Kinoticket	17.90	Alhambra Bern
Daisy Duck	14e	Kosmetik	Lippenstift	14.95	Migros Innenstadt
Donald Duck	14e	Kosmetik	Haargel	12.80	Coop Bahnhof Biel
Dagobert Duck	13d	lebensmittel	Mars	1.20	Migros Bahnhof Bern
Track Duck	12s	Freizeit	CD	19.90	City-Disk Biel
Gustav Gans	12s	lebensmittel	M&Ms	4.20	Coop Bern

Redundanz bedeutet höhere Fehleranfälligkeit!

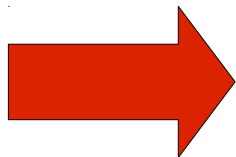
Name	Klasse	Sparte	Artikel	Preis	Laden
Tick Duck	12s	Lebensmittel	Brot	2.50	Coop Bahnhof Biel
Bart Simpson	12s	Lebensmittel	Milch	1.90	Migros Innenstadt
Lisa Simpson	14e	Kleider	Jeans	49.50	Migros Bahnhof Bern
Trick Duck	13d	Lebensmittel	Chips	3.70	Bahnhof-Coop in Biel
Homer Simpson	12s	Freizeit	Kinoticket	17.90	Alhambra Bern
Daisy Duck	14e	Kosmetik	Lippenstift	14.95	Migros Innenstadt
Donald Duck	14e	Kosmetik	Haargel	12.80	Coop Bahnhof Biel
Dagobert Duck	13d	lebensmittel	Mars	1.20	Migros Bahnhof Bern
Track Duck	12s	Freizeit	CD	19.90	City-Disk Biel
Gustav Gans	12s	lebensmittel	M&Ms	4.20	Coop Bern

Lösung: Eigene Tabelle für die Spartenbezeichnungen

Sparten-ID	Bezeichnung
1	Lebensmittel
2	Bekleidung
3	Freizeit
4	Kosmetik



Keine
Redundanz



Eindeutige Zuordnung

Verknüpfung über Feld „Sparte“

Vorname	Nachname	Klasse	Sparte	Artikel	Preis	Laden	Ort
Tick	Duck	12s	1	Brot	2.50	Coop	Bahnhof Biel
Bart	Simpson	12s	1	Milch	1.90	Migros	Innenstadt Biel
Lisa	Simpson	14e	2	Jeans	49.50	Migros	Bahnhof Bern
Trick	Duck	13d	1	Chips	3.70	Coop	Biel
Homer	Simpson	12s	3	Kinoticket	17.90	Alhambra	Bern
Daisy	Duck	14e	4	Lippenstift	14.95	Migros	Innenstadt Biel
Donald	Duck	14e	4	Haargel			
Dagobert	Duck	13d	1	Mars			
Track	Duck	12s	3	CD			
Gustav	Gans	12s	1	M&Ms			

Sparten-ID	Bezeichnung
1	Lebensmittel
2	Bekleidung
3	Freizeit
4	Kosmetik

Ist der Kunde eindeutig bestimmt?

Vorname	Nachname	Klasse	Sparte
Tick	Duck	12s	1
Bart	Simpson	12s	1
Lisa	Simpson	14e	2
Trick	Duck	13d	1
Homer	Simpson	12s	3
Daisy	Duck	14e	4
Donald	Duck	14e	4
Dagobert	Duck	13d	1
Track	Duck	12s	3
Gustav	Gans	12s	1

In der Praxis:

Ja, mit hoher Wahrscheinlichkeit

Garantiert?

Nein. Beispiel: In der 14e sitzen zwei Lisa Simpsons. Welche ist gemeint?

Und ausserdem: Weitere Reduktion von Redundanz

Vorname	Nachname	Klasse	Sparte	Artikel	Preis	Laden	Ort
Tick	Duck	12s	1	Brot	2.50	Coop	Bahnhof Biel
Bart	Simpson	12s	1	Milch	1.90	Migros	Innenstadt Biel
Lisa	Simpson	14e	2	Jeans	49.50	Migros	Bahnhof Bern
Trick	Duck	13d	1	Chips	3.70	Coop	Biel
Lisa	Simpson	14e	3	Kinoticket	17.90	Alhambra	Bern
Lisa	Simpson	14e	4	Lippenstift	14.95	Migros	Innenstadt Biel
Donald	Duck	14e	4	Haargel	12.80	Coop	Bahnhof Biel
Dagobert	Duck	13d	1	Mars	1.20	Migros	Bahnhof Bern
Lisa	Simpson	14e	3	CD	19.90	City-Disk	Biel
Lisa	Simpson	14e	1	M&Ms	4.20	Coop	Bern

Und ausserdem: Weitere Reduktion von Redundanz

Vorname	Nachname	Klasse
Tick	Duck	12s
Bart	Simpson	12s
Lisa	Simpson	14e
Trick	Duck	13d
Lisa	Simpson	14e
Lisa	Simpson	14e
Donald	Duck	14e
Dagobert	Duck	13d
Lisa	Simpson	14e
Lisa	Simpson	14e

Gleiche Kunden führen
in 3 Spalten zu
identischen Einträgen!

Wo ist das Feld „Klasse“ am besten aufgehoben?

Kunde	Sparte	Artikel
1	1	Brot
2	1	Milch
3	2	Jeans
4	1	Chips
3	3	Kinoticket
3	4	Lippenstift
7	4	Haargel

8	Vorname	Nachname	Klasse	Nr
3	Tick	Duck	12s	1
3	Bart	Simpson	12s	2
	Lisa	Simpson	14e	3

Kunde	Klasse	Sparte	Artikel
1	12s	1	Brot
2	12s	1	Milch
3	14e	2	Jeans
4	13d	1	Chips
3	14e	3	Kinoticket
3	14e	4	Lippenstift
7	14e	4	Haargel

8	13d	Vorname	Nachname	Nr
3	14e	Tick	Duck	1
3	14e	Bart	Simpson	2
		Lisa	Simpson	3

Schlecht (da redundant): Funktionale Abhängigkeit

Definition:

Wenn wir in einem Datensatz aufgrund eines Wertes in Feld A den Wert in Feld B *vorhersagen* können, ist B von A abhängig!

Kunde	Klasse	Sparte	Artikel		
1	12s	1	Brot		
2	12s	1	Milch		
3	14e	2	Jeans		
4	13d	1	Chips		
3	14e	3	Kinoticket		
3	14e	4	Lippenstift		
7	14e	4	Haargel		
8	13d	Vorname	Nachname	Nr	
3	14e	Tick	Duck	1	
3	14e	Bart	Simpson	2	
		Lisa	Simpson	3	

Finde Abhängigkeiten!

Vorname	Nachname	Adresse	PLZ	Ort	Telefon	Geburtsdatum
Pascal	Schuppli	Oberdorfstrasse 18	3054	Schüpfen	031 862 10 07	2. Feb. 1977

Autor	Titel	Seiten	Verlag	Publikation sort	Publikation sjahr	ISBN
King, Stephen	It	1104	Signet	New York City	1987	978-0451169518

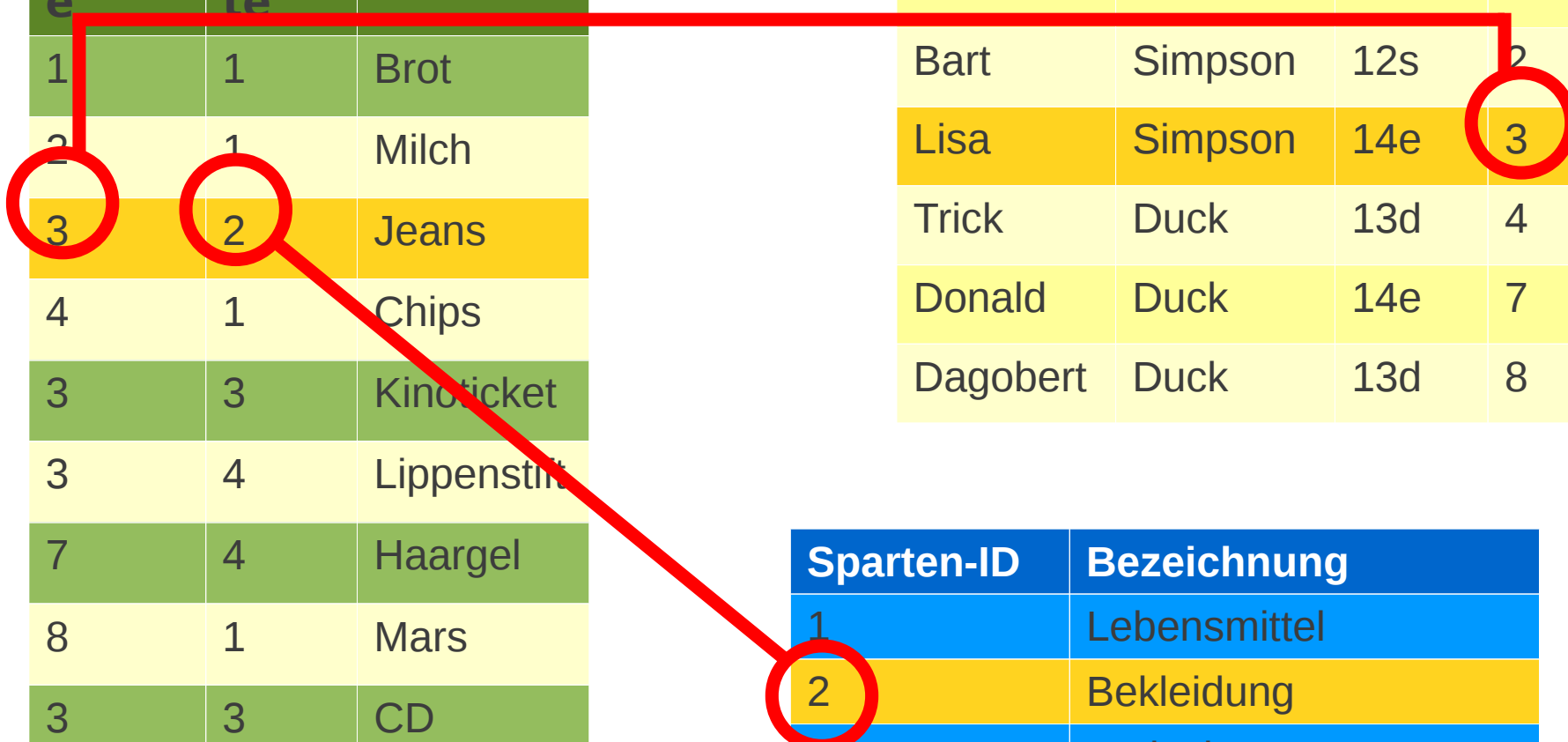
Fluggesellschaft	Abflugs-ort	Ziel	Flug-#	Abflugs-datum	Abflugs-zeit	Ankunfts-zeit
Swiss	ZRH	JFK	LX-16	23.3.2012	9:55	13:00

Verknüpfung von Tabellen

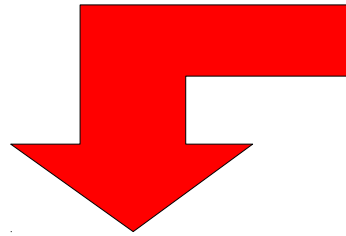
Kunde	Sparte	Artikel
1	1	Brot
2	1	Milch
3	2	Jeans
4	1	Chips
3	3	Kinoticket
3	4	Lippenstift
7	4	Haargel
8	1	Mars
3	3	CD
3	1	M&Ms

Vorname	Nachname	Klasse	Nr
Tick	Duck	12s	1
Bart	Simpson	12s	2
Lisa	Simpson	14e	3
Trick	Duck	13d	4
Donald	Duck	14e	7
Dagobert	Duck	13d	8

Sparten-ID	Bezeichnung
1	Lebensmittel
2	Bekleidung
3	Freizeit
4	Kosmetik



Welche Felder eignen sich für Verknüpfungen?



Vorname	Nachname	Klasse	Nr
Tick	Duck	12s	1
Bart	Simpson	12s	2
Lisa	Simpson	14e	3
Trick	Duck	13d	4
Donald	Duck	14e	7
Dagobert	Duck	13d	8

Der Wert des Feldes muss den Datensatz unter allen *möglichen* Datensätzen eindeutig bestimmen können!

Solche Felder nennt man
Schlüsselkandidaten

Finde Schlüsselkandidaten!

Hinweis:

Schlüsselkandidaten dürfen mehrere Felder umfassen.

Vorname	Nachname	Adresse	PLZ	Ort	Telefon	Geburtsdatum
Pascal	Schuppli	Oberdorfstrasse 18	3054	Schüpfen	031 862 10 07	2. Feb. 1977

Autor	Titel	Seiten	Verlag	Publikationsort	Publikationsjahr	ISBN
King, Stephen	It	1104	Signet	New York City	1987	978-0451169518

Fluggesellschaft	Abflugsort	Ziel	Flug-#	Abflugsdatum	Abflugszeit	Ankunftszeit
Swiss	ZRH	JFK	LX-16	23.3.2012	9:55	13:00

Primär- und Fremdschlüssel

Fremdschlüssel

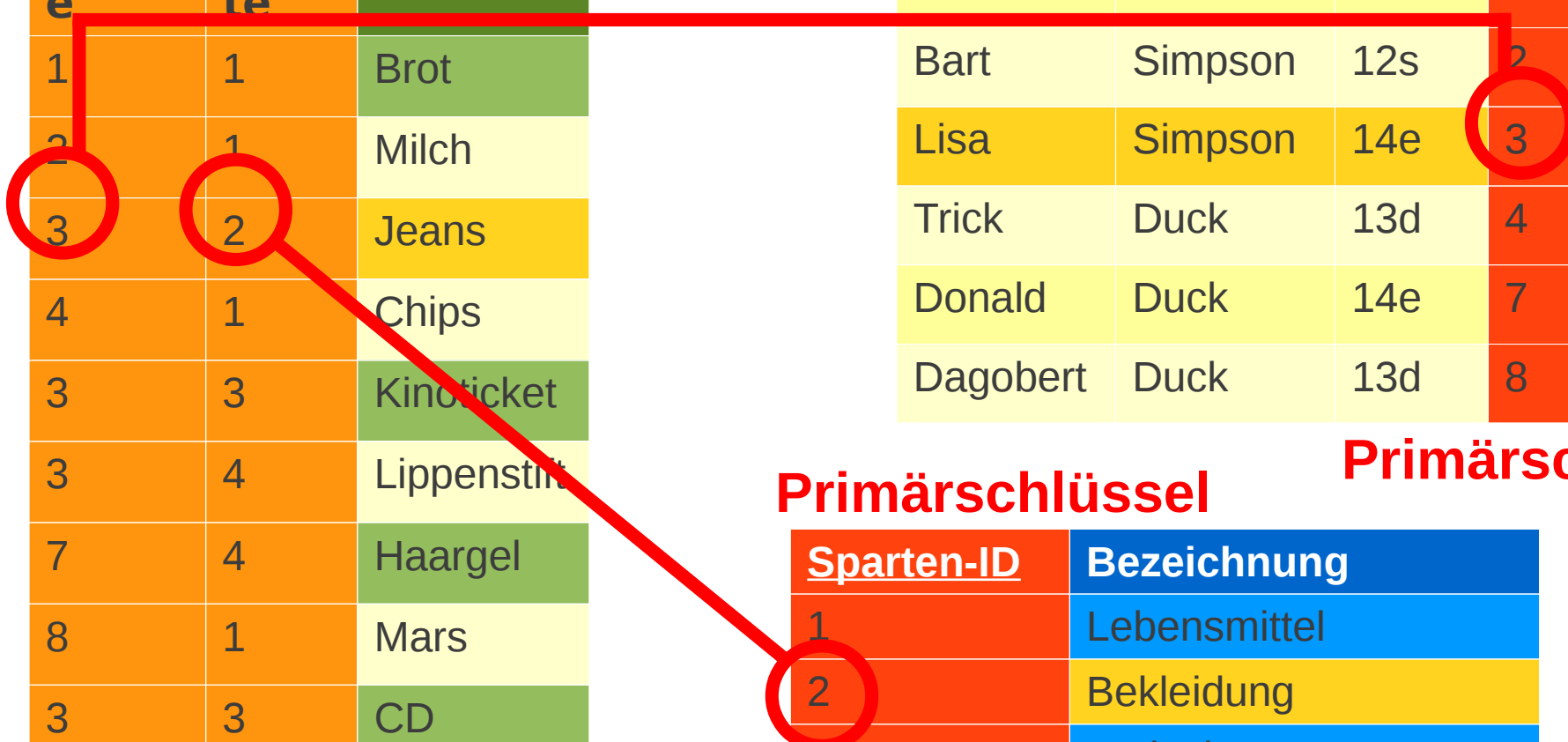
Kunde	Sparte	Artikel
1	1	Brot
2	1	Milch
3	2	Jeans
4	1	Chips
3	3	Kinoticket
3	4	Lippenstift
7	4	Haargel
8	1	Mars
3	3	CD
3	1	M&Ms

Vorname	Nachname	Klasse	Nr
Tick	Duck	12s	1
Bart	Simpson	12s	2
Lisa	Simpson	14e	3
Trick	Duck	13d	4
Donald	Duck	14e	7
Dagobert	Duck	13d	8

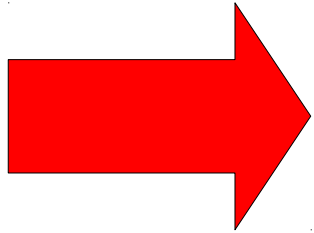
Primärschlüssel

<u>Sparten-ID</u>	Bezeichnung
1	Lebensmittel
2	Bekleidung
3	Freizeit
4	Kosmetik

Primärschlüssel



Normalisieren



Ziel: Eindeutigkeit und keine Redundanz!

„ Wieviele Tabellen muss ich erstellen?

Welche Felder muss ich machen und in welcher Tabelle?

“

Welche Schlüssel wähle ich aus?

1. Normalform:

Sind alle Felder atomar?

Kunde	Sparte	Artikel	Preis	Laden	Ort
1	1	Brot	2.50	Coop	Bahnhof Biel
2	1	Milch	1.90	Migros	Innenstadt Biel
3	2	Jeans	49.50	Migros	Bahnhof Bern
4	1	Chips	3.70	Coop	Biel
5	3	Kinoticket	17.90	Alhambra	Bern
6	4	Lippenstift	14.95	Migros	Innenstadt Biel
7	4	Haargel	12.80	Coop	Bahnhof Biel
8	1	Mars	1.20	Migros	Bahnhof Bern
9	3	CD	19.90	City-Disk	Biel
10	1	M&Ms	4.20	Coop	Bern

1. Normalform:

Verbesserungsvorschläge?

Kunde	Sparte	Artikel	Preis	Laden	Ort
1	1	Brot	2.50	Coop	Bahnhof Biel
2	1	Milch	1.90	Migros	Innenstadt Biel
3	2	Jeans	49.50	Migros	Bahnhof Bern
4	1	Chips	3.70	Coop	Biel
5	3	Kinoticket	17.90	Alhambra	Bern
6	4	Lippenstift	14.95	Migros	Innenstadt Biel
7	4	Haargel	12.80	Coop	Bahnhof Biel
8	1	Mars	1.20	Migros	Bahnhof Bern
9	3	CD	19.90	City-Disk	Biel
10	1	M&Ms	4.20	Coop	Bern


2. Normalform: Keine Abhängigkeiten zu Teilmengen von Schlüsselkandidaten?

<u>CD- Id</u>	Album	Interpret	Gründ ung	<u>Trac k</u>	Lied
10	Not that Kind	Anastacia	1999	1	Not that Kind
10	Not that Kind	Anastacia	1999	2	I'm Outta Love
10	Not that Kind	Anastacia	1999	3	Cowboys & Kisses
11	Wish You Were Here	Pink Floyd	1964	1	Shine On Your Crazy Diamond
12	Freak of Nature	Anastacia	1999	1	Paid My Dues

Quelle. Wikipedia, Artikel „Normalisierung (Datenbanken)“

2. Normalform: Verbesserungsvorschläge?

Funktional abhängig von CD-Id,
aber nicht von Track!



<u>CD-Id</u>	Album	Interpret	Gründung	<u>Track</u>	Lied
10	Not that Kind	Anastacia	1999	1	Not that Kind
10	Not that Kind	Anastacia	1999	2	I'm Outta Love
10	Not that Kind	Anastacia	1999	3	Cowboys & Kisses
11	Wish You Were Here	Pink Floyd	1964	1	Shine On Your Crazy Diamond
12	Freak of Nature	Anastacia	1999	1	Paid My Dues

2. Normalform oder nicht?

Hersteller	Modell	<u>Modellname</u>	Hersteller land
Forte	X-Prime	Forte X-Prime	Italien
Forte	Ultraclean	Forte Ultraclean	Italien
Dent-o-Fresh	EZ Brush	Dent-o-Fresh EZ Brush	USA
Kobayashi	St-60	Kobayashi St-60	Japan
Hoch	Tooth Master	Hoch Tooth Master	Deutschland
Hoch	X-Prime	Hoch X-Prime	Deutschland

Quelle. Wikipedia, Artikel „Second normal form“

Nicht 2. NF wegen Schlüsselkandidat!

Hersteller	Modell	<u>Modellname</u>	Hersteller land
Forte	X-Prime	Forte X-Prime	Italien
Forte	Ultraclean	Forte Ultraclean	Italien
Dent-o-Fresh	EZ Brush	Dent-o-Fresh EZ Brush	USA
Kobayashi	St-60	Kobayashi St-60	Japan
Hoch	Zahnmeister	Hoch Zahnmeister	Deutschland
Hoch	X-Prime	Hoch X-Prime	Deutschland



Quelle: Wikipedia, Artikel „Second normal form“

3. Normalform: Hängen Felder nur von Schlüssel- feldern ab?

<u>CD- Id</u>	Album	Interpret	Gründ ung
10	Not that Kind	Anastacia	1999
11	Wish You Were Here	Pink Floyd	1964
12	Freak of Nature	Anastacia	1999

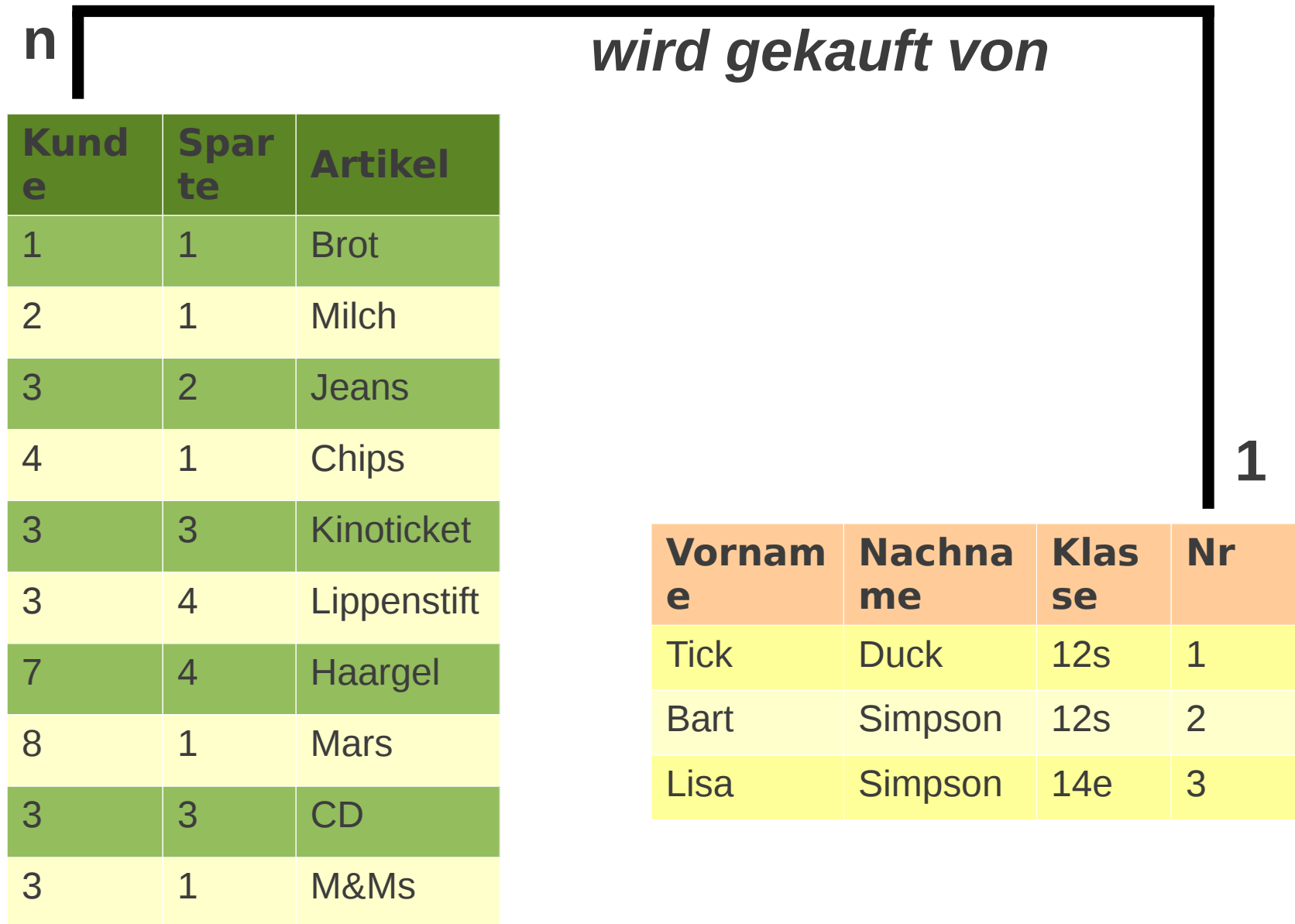
Keine 3. NF, denn: transitive Abhängigkeiten!



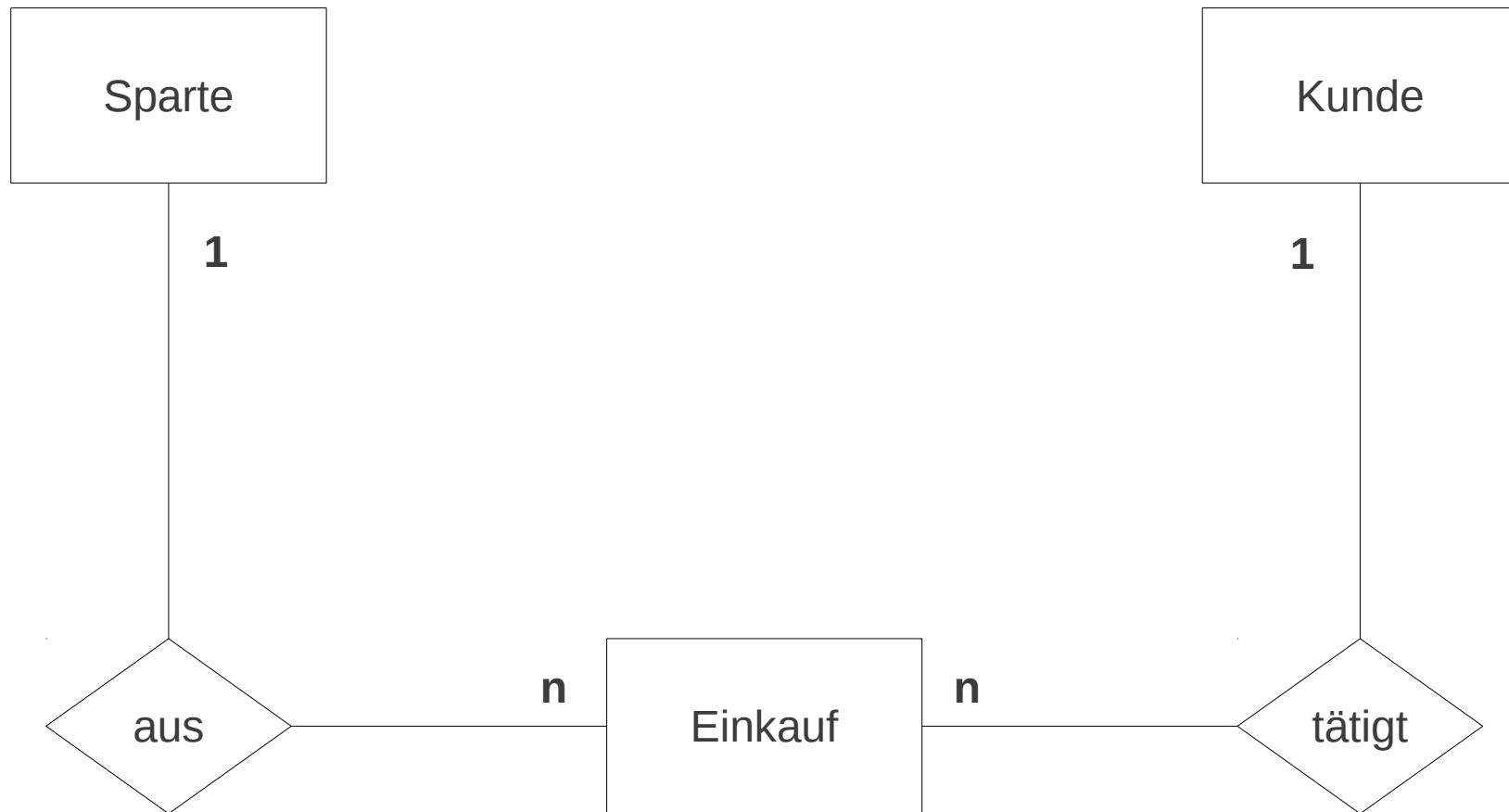
<u>CD-Id</u>	Album	Interpret	Gründung
10	Not that Kind	Anastacia	1999
11	Wish You Were Here	Pink Floyd	1964
12	Freak of Nature	Anastacia	1999

Kardinalität

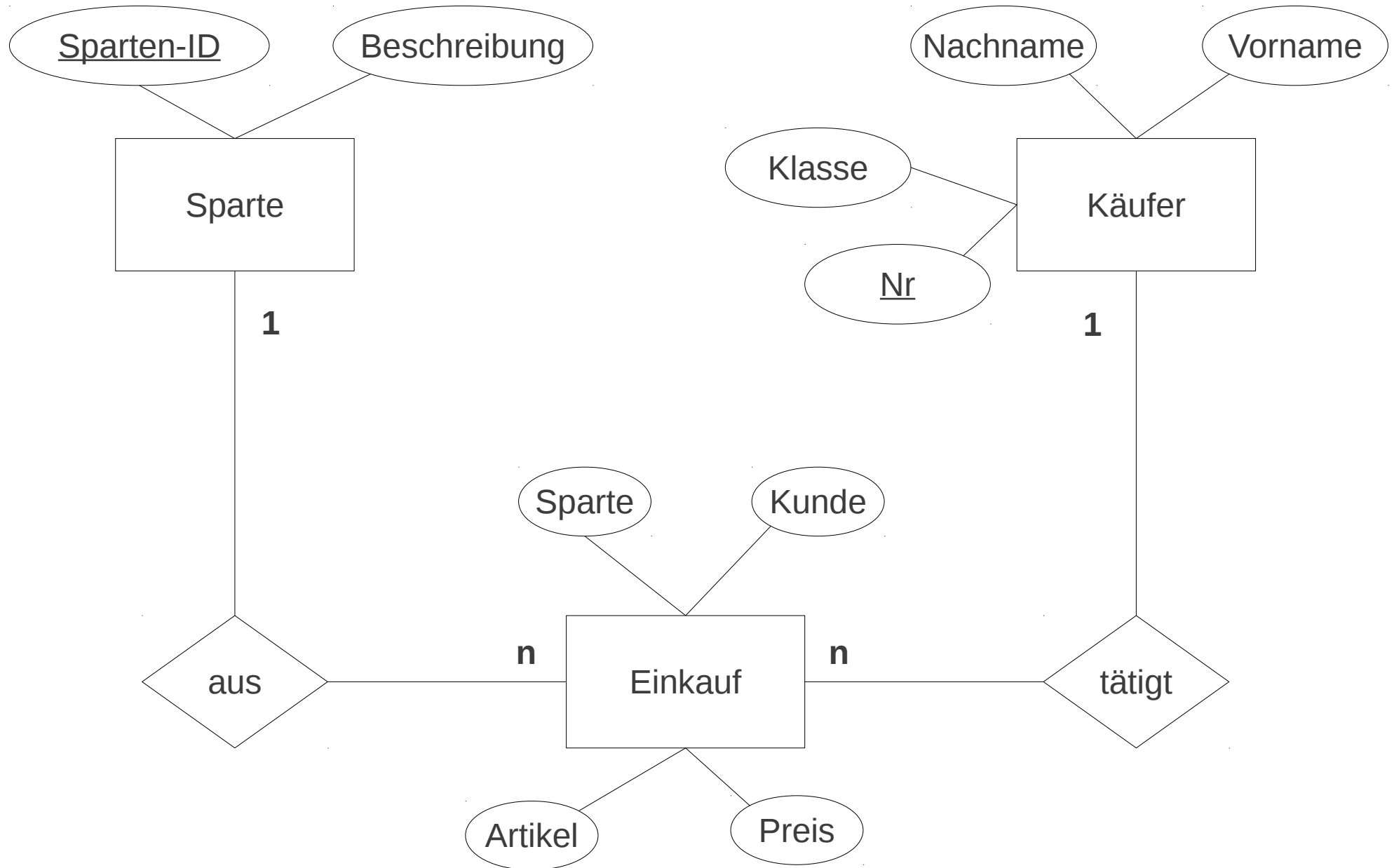
1:1, 1:n, m:n



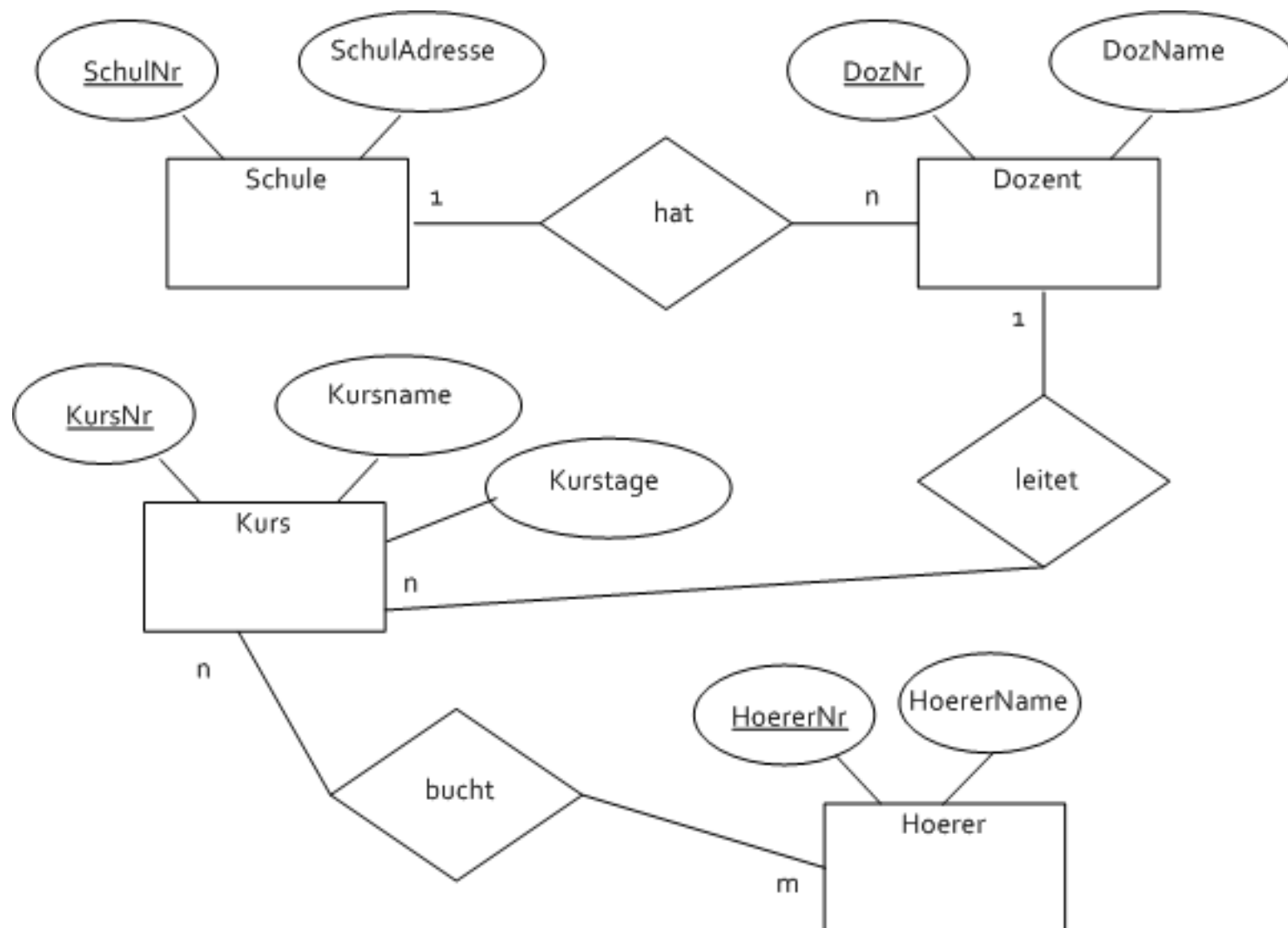
ER-Diagramme



ER-Diagramme mit Feldern

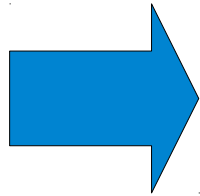


ER-Diagramm für die Schule

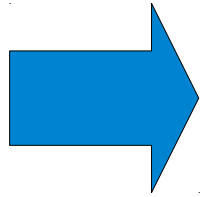


Tabellen aus ER-Diagrammen ableiten

Faustregel:



Eine Tabelle für jede Entity



Eine Tabelle für jede Relation

Welche Tabellen?

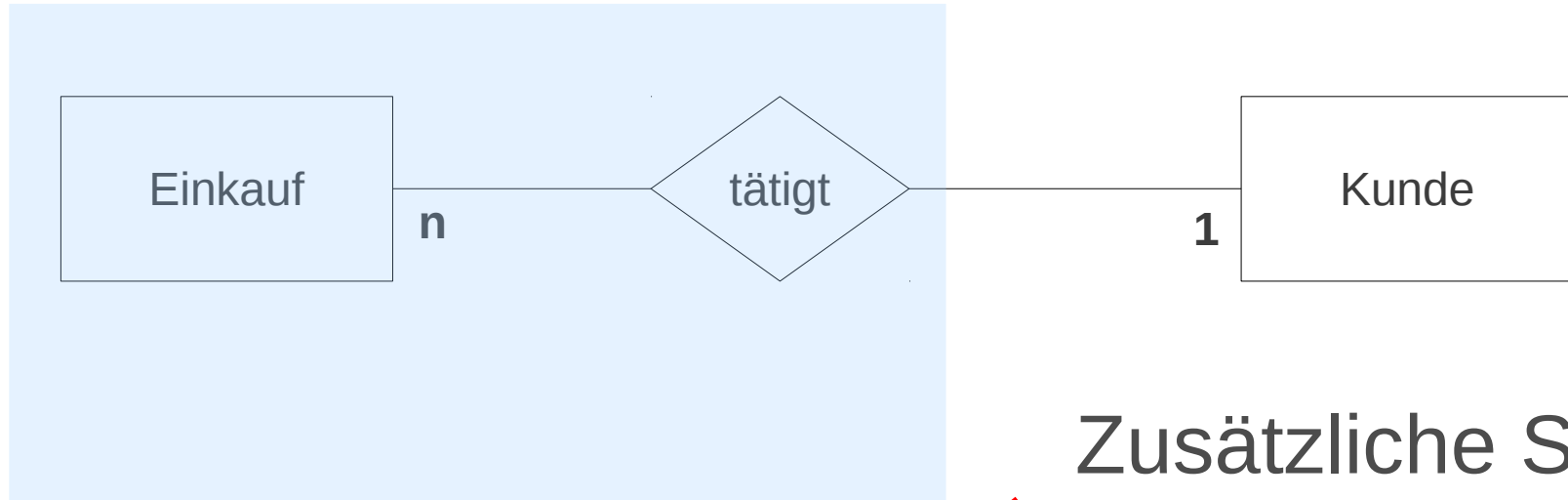


ENr	KNr
1	2
2	2
3	3
4	4
5	3
6	3

ENr	Klasse	Sparte	Artikel
1	12s	1	Brot
2	12s	1	Milch
3	14e	2	Jeans
4	13d	1	Chips
5	14e	3	Kinoticket
6	14e	4	Lippenstift

Vorname	Nachname	Klasse	KNr
Tick	Duck	12s	1
Bart	Simpson	12s	2
Lisa	Simpson	14e	3

1:n-Relations: Nur 2 Tabellen!

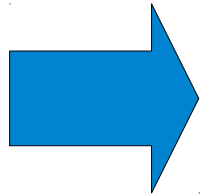


Zusätzliche Spalte
mit Fremd-
schlüssel

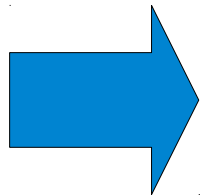
ENr	Klasse	Sparte	Artikel	KNr
1	12s	1	Brot	2
2	12s	1	Milch	2
3	14e	2	Jeans	3
4	13d	1	Chips	4
5	14e	3	Kinoticket	3
6	14e	4	Lippenstift	3

Vorname	Name	Klasse	KNr
Tick	Duck	12s	1
Bart	Simpson	12s	2
Lisa	Simpson	14e	3

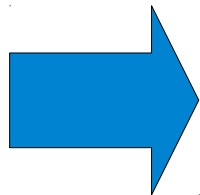
Tabellen aus ER-Diagrammen ableiten



Eine Tabelle für jede Entity
Attribute sind Felder/Spalten



Eine Tabelle für jede **m:n**-
Relation mit Schlüsseln der
beiden involvierten Entities



1:n: Entity auf der n-Seite um
Fremdschlüssel-Attribut für
die andere Entity erweitern