Dimensional modeling

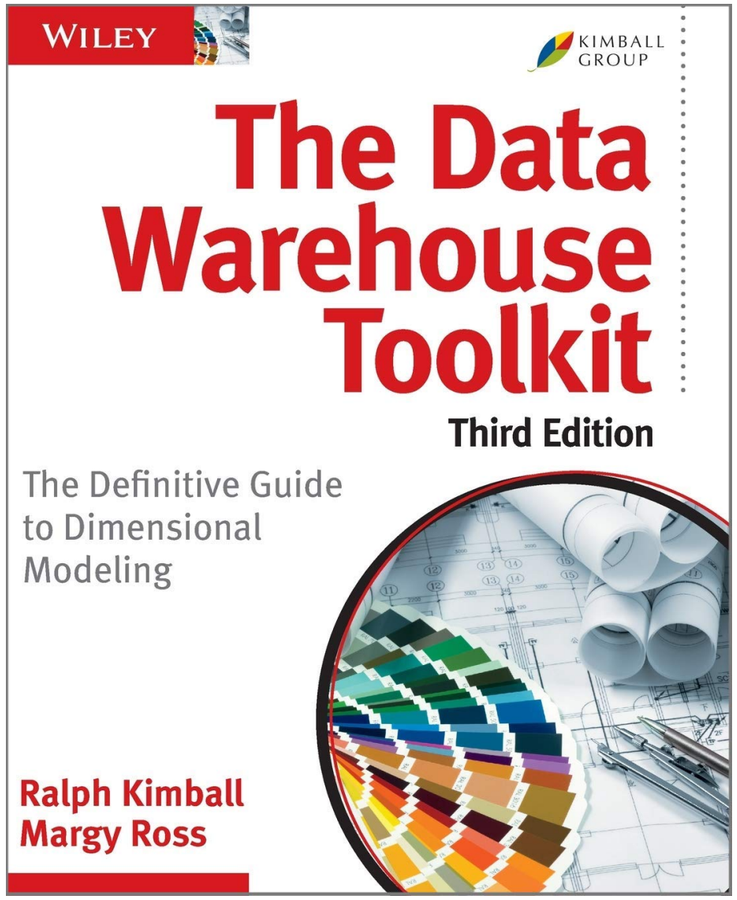
DATA MODELING IN POWER BI

Sara Billen
Curriculum Manager at DataCamp

datacamp

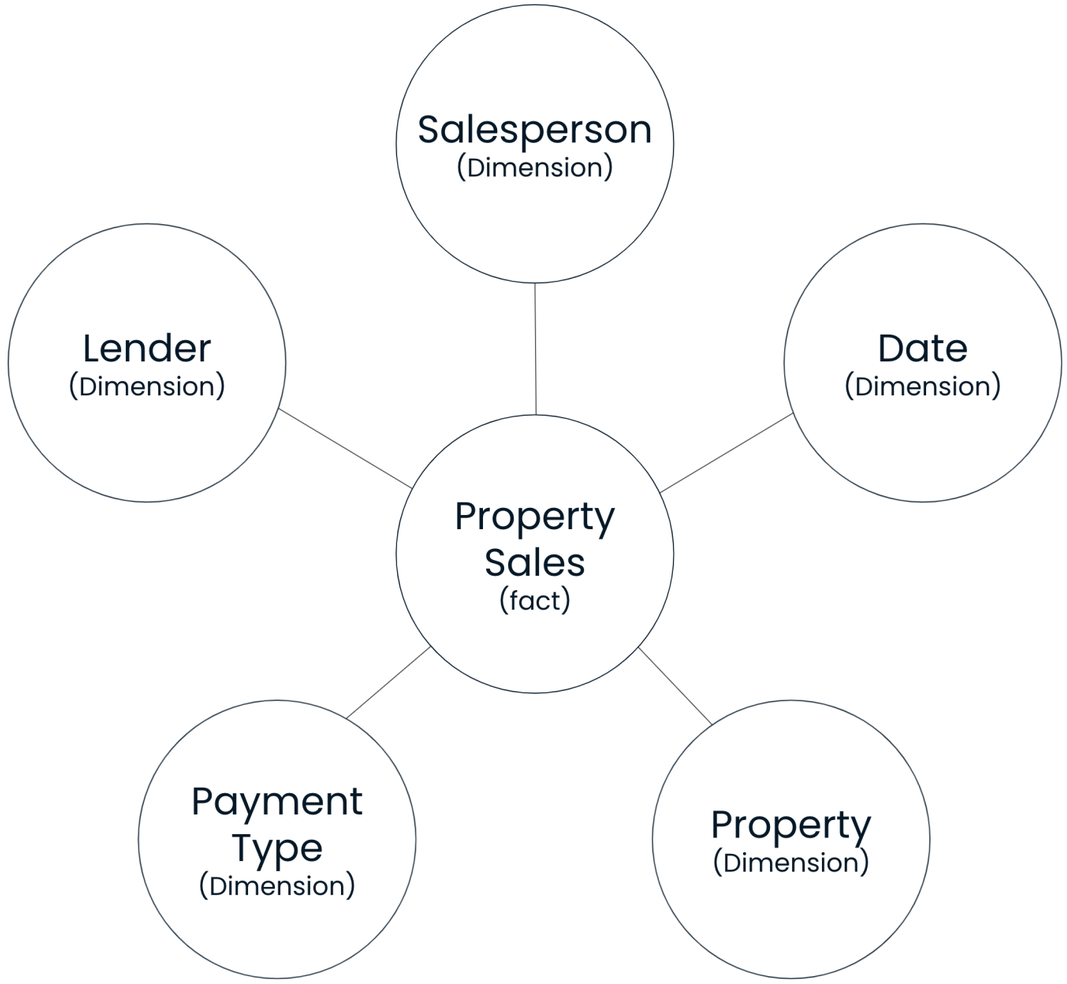
The Kimball Model

The dimensional model



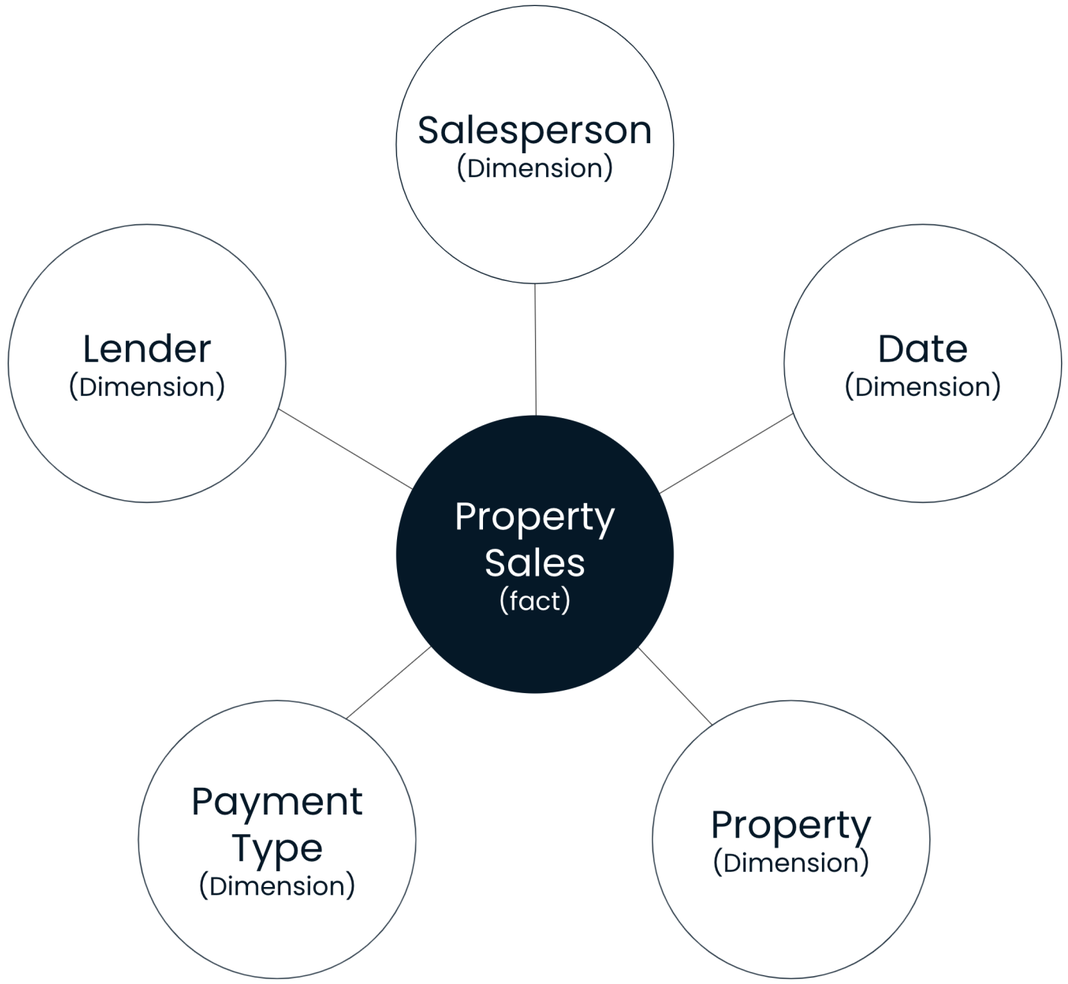
The Kimball Model

- Key concepts
  - **Facts:** metrics from a business process
  - **Dimensions:** context surrounding a business process
  - Combine to form a **star schema**
- Star schemas are used in data warehouses
- Power BI is optimized for star schemas



Fact tables

- Made up of
  - **Facts (measures)**
    - Measurements or metrics from your business process
  - **Keys**
    - Used to establish relationships between fact and dimension tables
- Fact tables are long and narrow
  - Lots of rows
  - Fewer columns



Fact tables: an example

Property Sales table

LenderID	StartDateID	PropertyID	PaymentTypeID	SalesPersonID	Rent	Duration
CO76	20200624	PG14	P2	SA9	750	24
CO56	20200907	PG4	P4	SA12	1250	12
CO62	20201201	PG16	P3	SA5	3000	36
CO43	20200201	PG6	P3	SA6	500	24
CO76	20200530	PG20	P2	SA6	5000	12
CO76	20200115	PG11	P2	SA2	2000	24
CO32	20201201	PG15	P2	SA9	450	36
...	...	...	...	...	...	...

Fact tables: an example

Keys: establish relationships between tables

LenderID	StartDateID	PropertyID	PaymentTypeID	SalesPersonID	Rent	Duration
CO76	20200624	PG14	P2	SA9	750	24
CO56	20200907	PG4	P4	SA12	1250	12
CO62	20201201	PG16	P3	SA5	3000	36
CO43	20200201	PG6	P3	SA6	500	24
CO76	20200430	PG20	P2	SA9	5000	12
CO76	20200115	PG11	P2	SA2	2000	24
CO32	20201201	PG15	P2	SA9	450	36
...	...	...	...	...	...	...

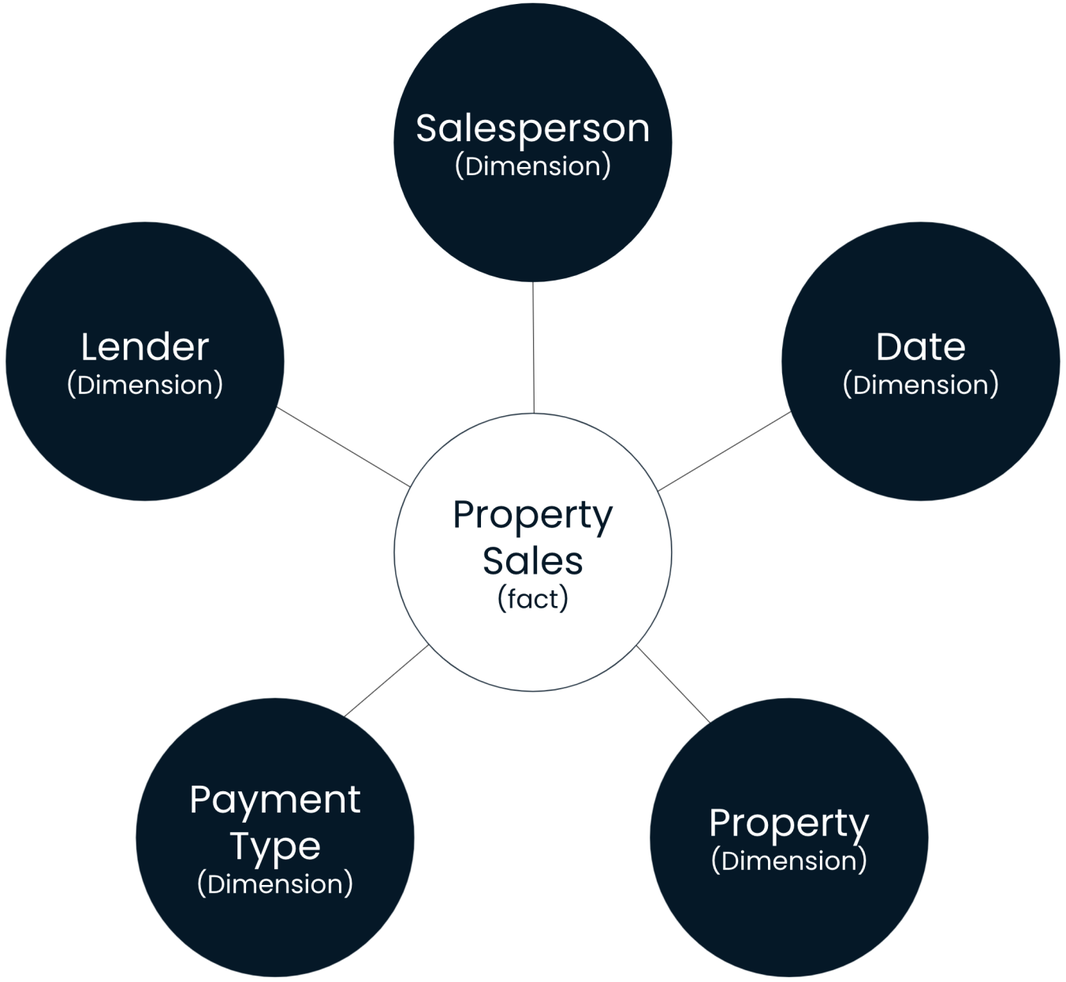
Fact tables: an example

Measures: metrics from the business process

LenderID	StartDateID	PropertyID	PaymentTypeID	SalesPersonID	Rent	Duration
CO76	20200624	PG14	P2	SA9	750	24
CO56	20200907	PG4	P4	SA12	1250	12
CO62	20201201	PG16	P3	SA5	3000	36
CO43	20200201	PG6	P3	SA6	500	24
CO76	20200430	PG20	P2	SA9	5000	12
CO76	20200115	PG11	P2	SA2	2000	24
CO32	20201201	PG15	P2	SA9	450	36
...	...	...	...	...	...	...

Dimension tables

- Provide context
  - Who, what, when, where, why?
- Shared business concepts
  - E.g., person, employee, customer, vendor
- Contain static or "slowly changing" data
  - E.g., name, date of birth, height
- Dimension tables are short and wide
  - Few rows
  - Lots of columns



Dimension tables: an example

Salesperson table

SalesPersonID	FirstName	LastName	DateOfBirth	Salary
SA9	Mary	Howe	1990-02-19	24000
SA12	David	Ford	1978-03-24	18000
SA5	Ann	Beech	1980-11-10	12000
SA6	Julie	Lee	1985-06-13	30000
SA9	John	White	1965-10-01	9000
...	...		...	

Dimension tables: an example

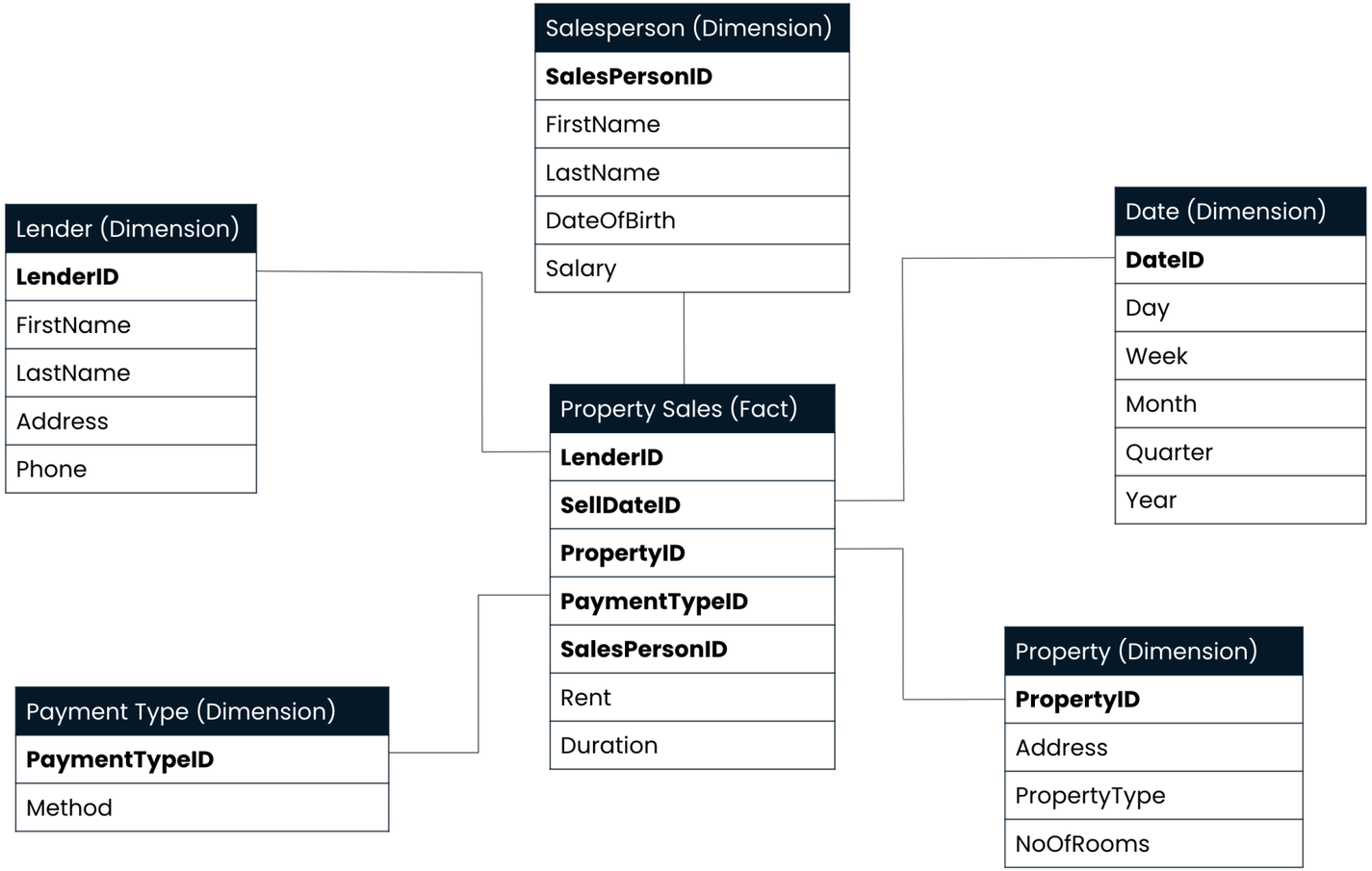
Key: establish relationship with fact table

SalesPersonID	FirstName	LastName	DateOfBirth	Salary
SA9	Mary	Howe	1990-02-19	24000
SA12	David	Ford	1978-03-24	18000
SA5	Ann	Beech	1980-11-10	12000
SA6	Julie	Lee	1985-06-13	30000
SA9	John	White	1965-10-01	9000
...	...		...	

Dimension tables: an example

Attributes: various characteristics of the dimension

SalesPersonID	FirstName	LastName	DateOfBirth	Salary
SA9	Mary	Howe	1990-02-19	24000
SA12	David	Ford	1978-03-24	18000
SA5	Ann	Beech	1980-11-10	12000
SA6	Julie	Lee	1985-06-13	30000
SA9	John	White	1965-10-01	9000
...	...	...	...	...



- Dimensions are used in multiple facts
- Dimensions do not link to other dimensions

The dataset

Fact

- **Establishment Survey:** number of employees, number of firms, ...

Dimensions

- **Industry:** NAICS code, industry group, subsector, sector
- **Time:** year, decade, century
- **Age:** establishment age
- **Geography:** country, state

Let's practice!

DATA MODELING IN POWER BI

Creating a star schema

DATA MODELING IN POWER BI

Sara Billen
Curriculum Manager at DataCamp

datacamp

Let's practice!

DATA MODELING IN POWER BI