



LWTN Data Modeling

Mitchell Pearson

Agenda

Dimensional Modeling

- Facts and Dimensions

- Star Schema

Creating a Data Model in Power BI

Multiple Fact Tables

Role Playing Tables

Data Analyst of the Future with Microsoft Fabric



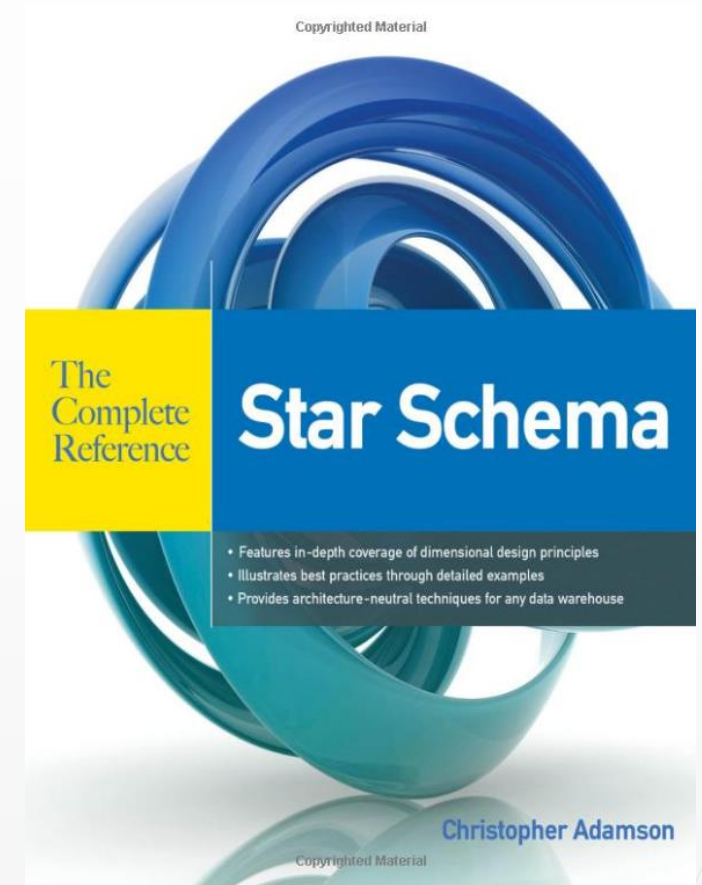
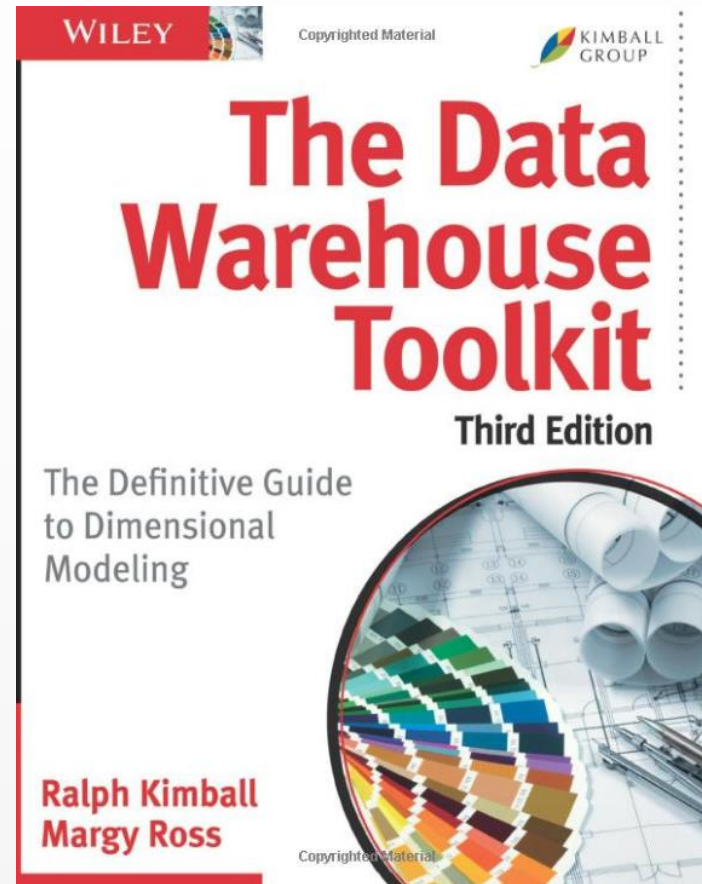
Recommended Books

Data Warehouse Toolkit

- Ralph Kimball

Star Schema

- Christopher Adamson





Dimensional Modeling

What is a Data Model?

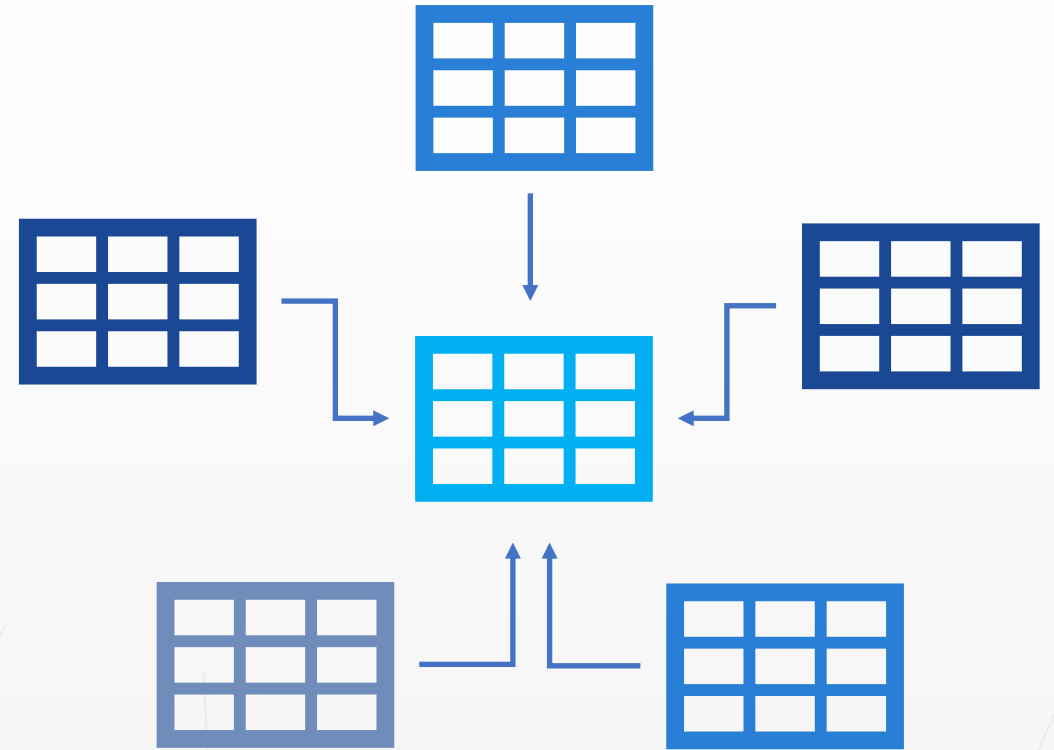
Attributes of a Good Data Model

Can be easily understood and consumed

Large data changes are scalable

Provides **predictable performance**

Is **flexible and adaptable**, but not at the expense of the other attributes



Things to consider....

- 1) What are you measuring?
- 2) What types of business problems are you trying to solve?
- 3) How much data are you working with?
- 4) What are your data sources?

What is easier with a good data model?

Managing Storage constraints

Performance Tuning

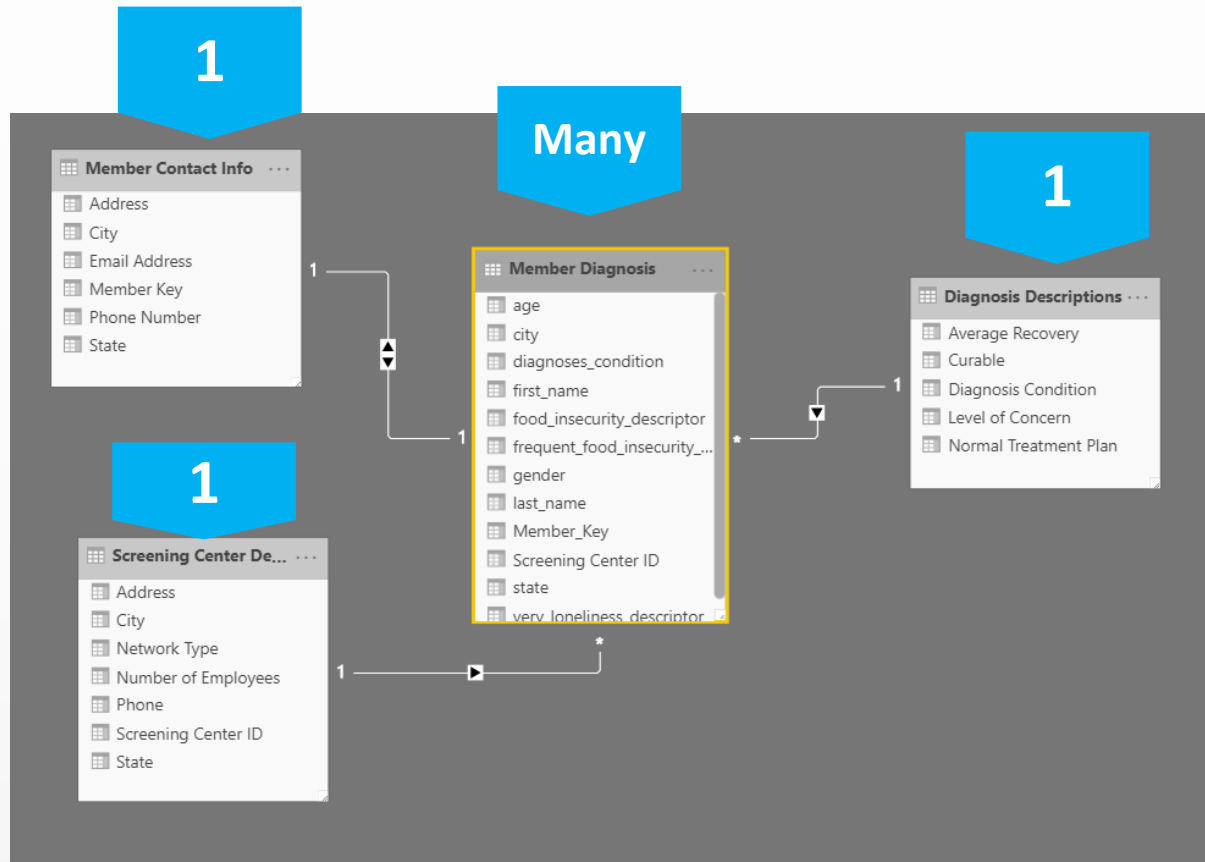
Managing Row Level Security

Authoring DAX

Everything!



Star Schema



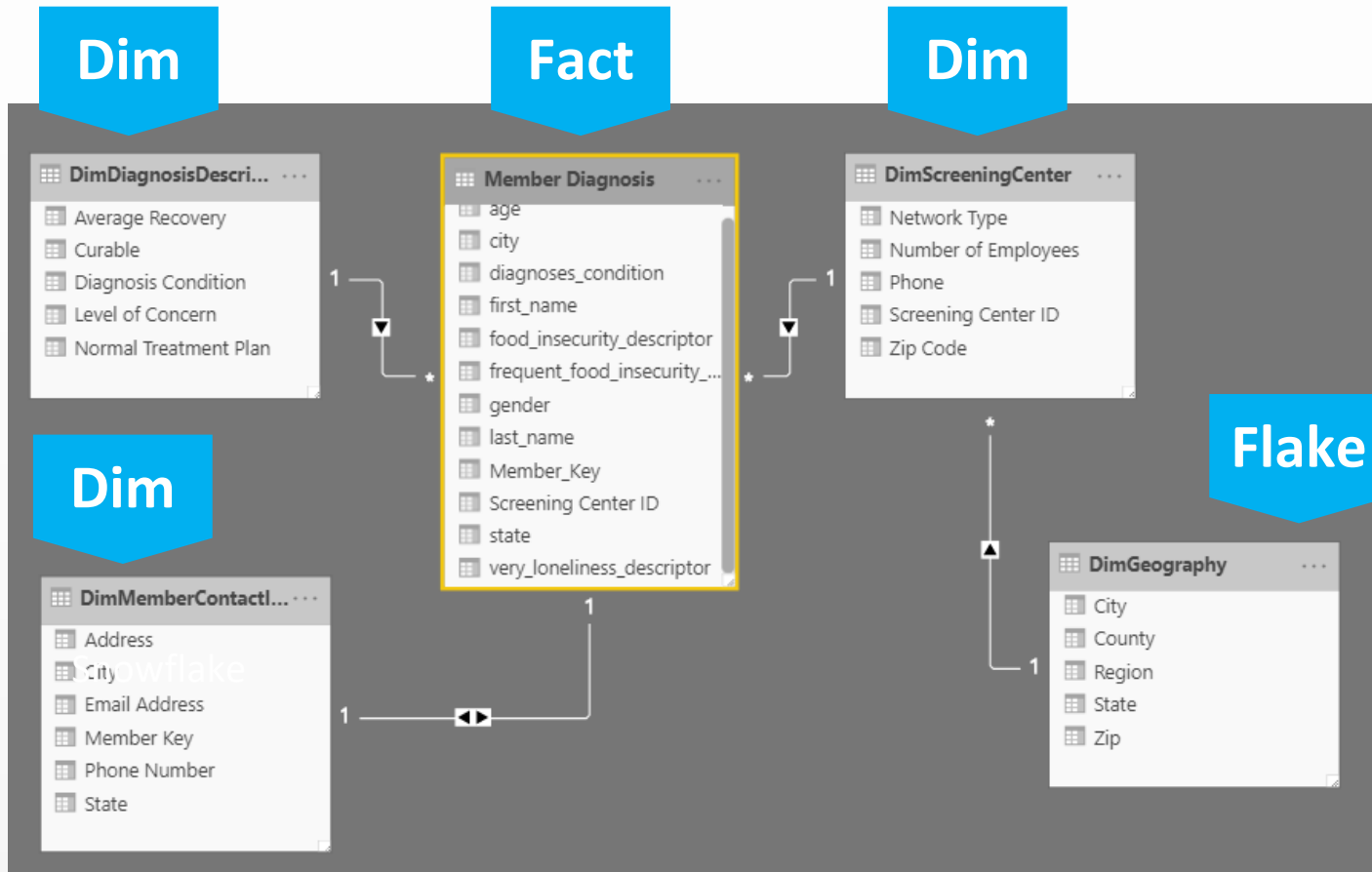
Fact table in the middle

Surrounded by Dims

Looks like a 'Star'

Fact table is the "Many" side of the (one to many) relationship

Snowflake Schema



Center is a Star schema

Fact table in middle

Surrounded by Dims

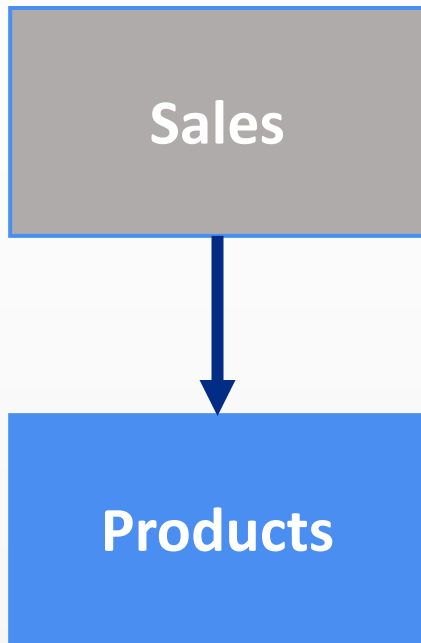
Dims “snowflake” off of other Dims

If you have many, it looks like a ‘Snowflake’

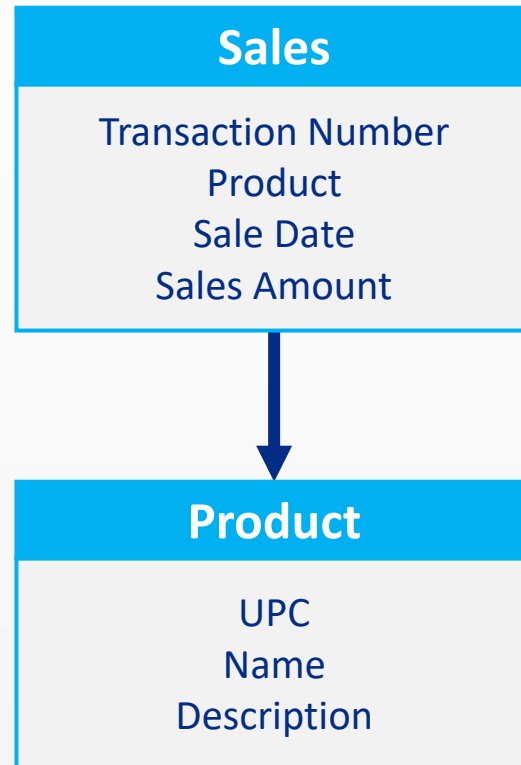
Dim or Fact tables can be the “Many” side of the relationship

Model Types

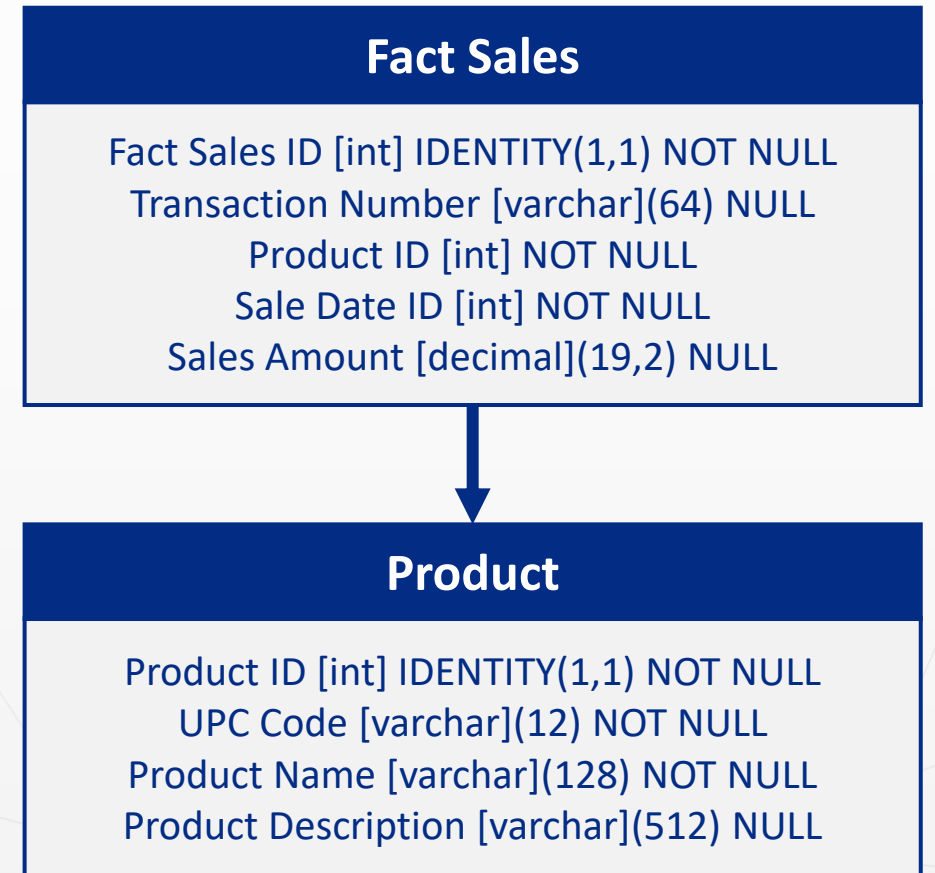
Conceptual



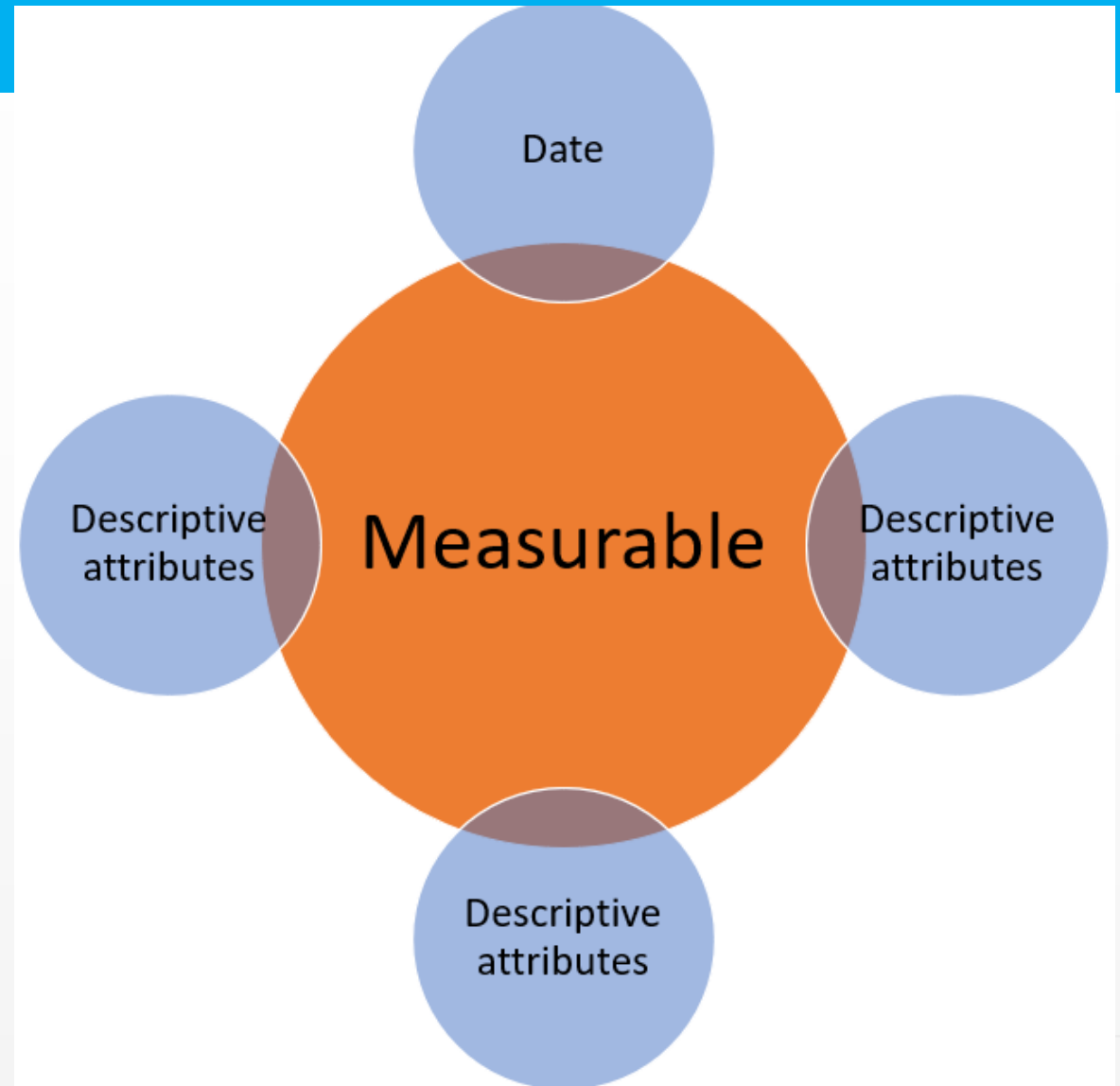
Logical



Physical



Conceptual Model





Dimensional Model Design

Dimensional Model Basics | Model Types

Dimensional Model – Terminology

Dimensional Model

Organizes the data so it is easy to retrieve for reporting purposes

Fact Table

A fact is an event that may or may not include measures.

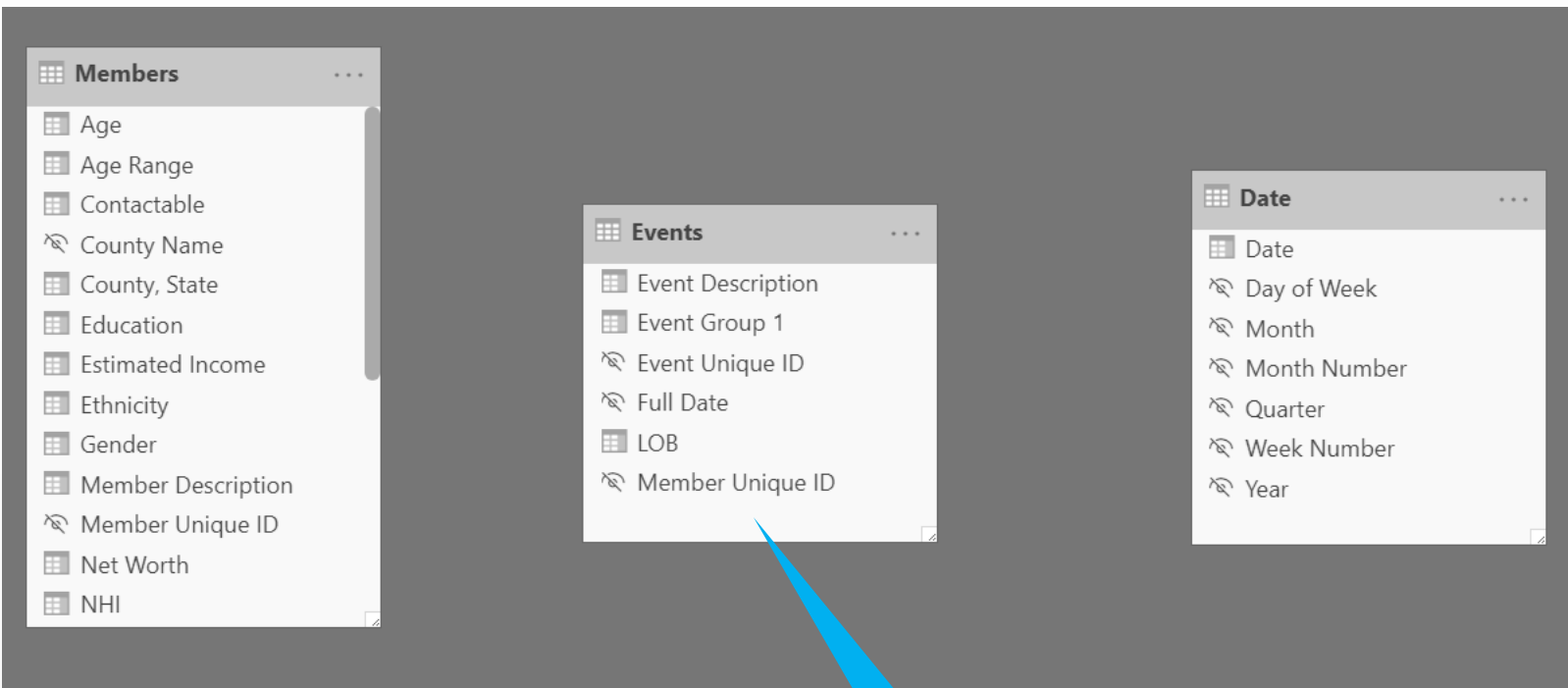
Dimension Table

Category of information, or a noun, descriptive

Attribute (*column in dimension table*)

Descriptor of the object

Fact Tables



Fact

A fact is an **event** that may or may not include measure

Fact Table

Contains Measures
(or items to be aggregated)
of a business process

Examples

Claim Amount, Screenings, Total Claims, Cost

Measures

Usually sliceable

Examples:

By Month, By Member

Granularity

Lowest level of information that will be stored in the fact table, or the values that would make the row distinct compared to all other rows

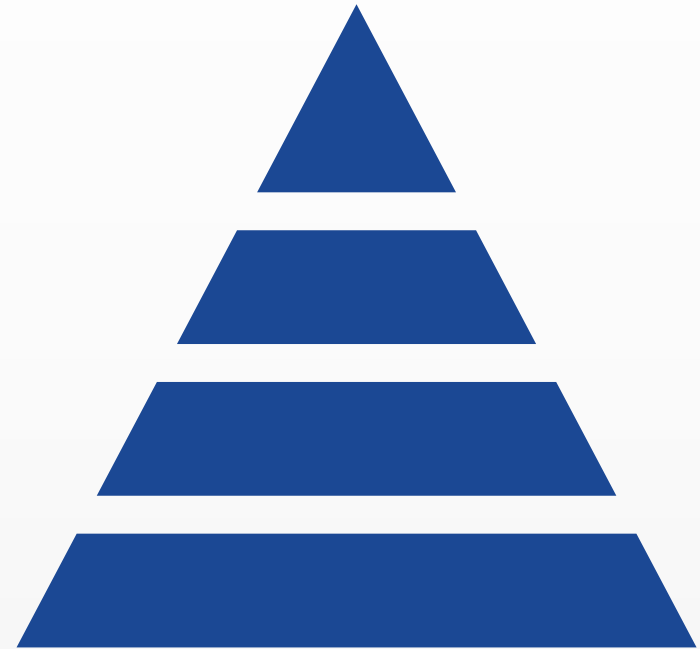
Fact Tables

Fact

A fact is an **event** that may or may not include measures

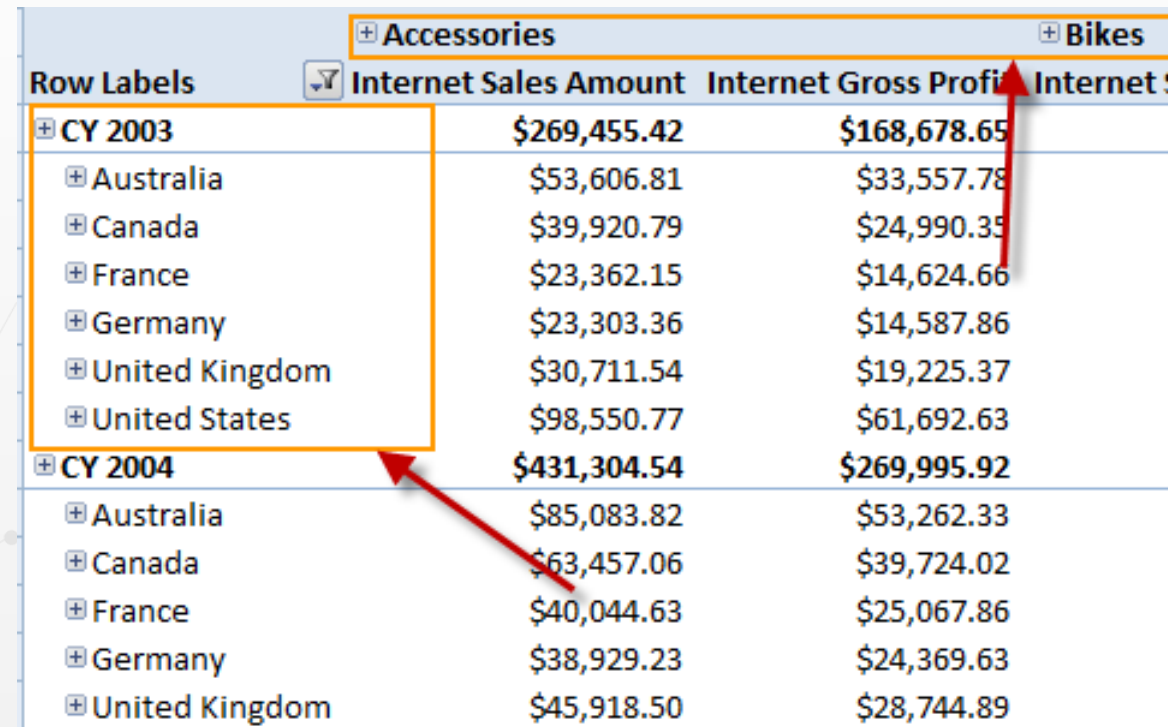
Granularity

Lowest level of information that will be stored in the fact table, or the values that would make the row distinct compared to all other rows



Defining Dimension

- A Dim (or Dimension) table contains descriptive attributes that define how a fact should roll up
- *“Dimensions provide the “who, what, where, when, why, and how” context surrounding a business process event.*
- - *Ralph Kimball*



The screenshot shows a data table with dimensions. The 'Row Labels' column is highlighted with an orange box, and the 'Internet Sales Amount' and 'Internet Gross Profit' columns are also highlighted with orange boxes. Red arrows point from the 'CY 2003' and 'United States' rows to the 'Internet Sales Amount' and 'Internet Gross Profit' columns, respectively.

Row Labels	Internet Sales Amount	Internet Gross Profit	Internet Sales Amount
+ CY 2003	\$269,455.42	\$168,678.65	
+ Australia	\$53,606.81	\$33,557.78	
+ Canada	\$39,920.79	\$24,990.35	
+ France	\$23,362.15	\$14,624.66	
+ Germany	\$23,303.36	\$14,587.86	
+ United Kingdom	\$30,711.54	\$19,225.37	
+ United States	\$98,550.77	\$61,692.63	
+ CY 2004	\$431,304.54	\$269,995.92	
+ Australia	\$85,083.82	\$53,262.33	
+ Canada	\$63,457.06	\$39,724.02	
+ France	\$40,044.63	\$25,067.86	
+ Germany	\$38,929.23	\$24,369.63	
+ United Kingdom	\$45,918.50	\$28,744.89	

Dimension Architecture



- Wide Table
- Surrogate Key (Unique ID)
- Natural Key
- Best Attributes are Descriptive
- Start Date / End Date
- Flags

DEMO TIME!!