**Definition**

Dimensional models are the models that look like the way business people think about their operation.

They are optimized to query and easier to retrieve data. Mostly used in reporting.

Another name for Dimensional model is Star Schema.

**Fact table**

Stores the most basic unit of measurement of a business process.

e.g. a coffee sale includes name of coffee, amount, size, store location, store attendee etc.

stores quantitative data.

**Dimension table**

Who, What, When and Where of business processes. Dimensions are the object or context. Dimensions are the things about which something is being spoken.

e.g. coffee names, list of stores, list of employees etc

stores qualitative data or categorical data as they represent distinct categories rather than numbers. We cant do mathematical operations on these.

**Lookup table**

lookup table for the Quarter attribute would include a list of all of the quarters available in the data warehouse. Each row (each quarter) may have several fields, one for the unique ID that identifies the quarter, and one or more additional fields that specifies how that particular quarter is represented on a report (for example, first quarter of 2001 may be represented as "Q1 2001" or "2001 Q1").

**Hierarchy**

Relationship between different attributes within a dimension.

e.g. Time dimension: Year -> Quarter -> Month -> Day

Quarter can be a dimension with other attributes e.g. name “Q1 2018”

**Benefits of Dimensional Modeling**

**Simpler Data Navigation and answers to data questions**

How many vanilla chai lattes were sold to customers age 18-25 in December 2015? What were the top 5 stores in California for iced coffee sales last week?

**Faster Database Performance**

Fewer joins, minimized data redundancy and operations on numbers instead of text which is more efficient use of CPU and memory

**Flexible to business changes**

**Tips to get started with modeling**

**Identify the business process you want to track**

e.g. tracking sales in credit card transactions

**Choose the granularity of the fact data**

Depends on total volume of data. Usually finest grain of data.

e.g. each transaction

**Find out dimensions**

Identify the attributes involved in each transaction and create separate dimension tables for each of them.

**Identify the attributes or properties of dimensions**

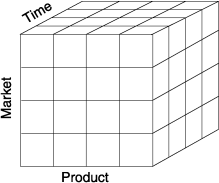
e.g. store location, city, store open date etc.

**Consolidate the facts**

Quantity and sales amount are your measures that you will store in a fact table along with related dimensions (keys).

Suppose your business sells products in different markets and you want to evaluate the performance over time. It is easy to conceive of this business process as a cube of data, which contains dimensions for time, products, and markets. The following figure shows this dimensional model. The various intersections along the lines of the cube would contain the *measures* of the business. The measures correspond to a particular combination: product, market, and time data.

*Figure 1. A dimensional model of a business that has time, product, and market dimensions*



Dimension tables only have a single join that attaches them to the fact table

Qualitative Data

Qualitative data are also called categorical data as they represent distinct categories rather than numbers. In case of dimensional modeling, they are often termed as "dimension". Mathematical operations such as addition or subtraction do not make any sense on that data.

Qualitative data can be further classified into below classes:

NOMINAL :

Nominal data represents data where order of the data does not represent any meaningful information. Consider your passport number. There is no information as such if your passport number is greater or lesser than some one else's passport number.

#### ORDINAL :

Order of the data is important for ordinal data. Consider height of people - tall, medium, short. Although they are qualitative but the order of the attributes does matter

Quantitative data

Quantitative data are also called numeric data as they represent numbers. In case of dimensional data modeling approach, these data are termed as "Measure".

Example of quantitative data is, height of a person, amount of goods sold, revenue etc.

Quantitative attributes can be further classified as below.

INTERVAL :

Interval classification is used where there is no true zero point in the data and division operation does not make sense. Bank balance, temperature in Celsius scale, GRE score etc. are the examples of interval class data. Dividing one GRE score with another GRE score will not make any sense. In dimensional modeling this is synonymous to semi-additive facts.

RATIO :

Ratio class is applied on the data that has a true "zero" and where division does make sense. Consider revenue, length of time etc. These measures are generally additive.