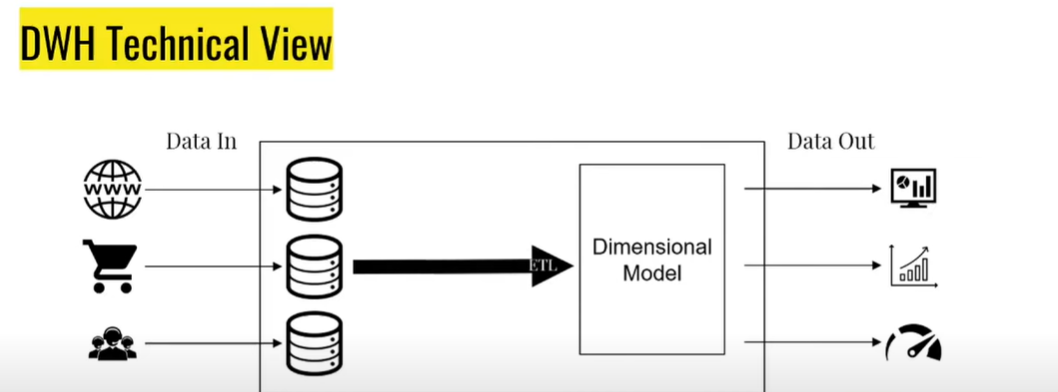
**What is Data Warehouse - Technical View**

A data warehouse is a copy of transaction data specifically structured for query and analysis.

Also

In computing, a data warehouse, also known as an enterprise data warehouse, is a system used for reporting and data analysis and is considered a core component of business intelligence. DWs are central repositories of integrated data from one or more disparate sources.



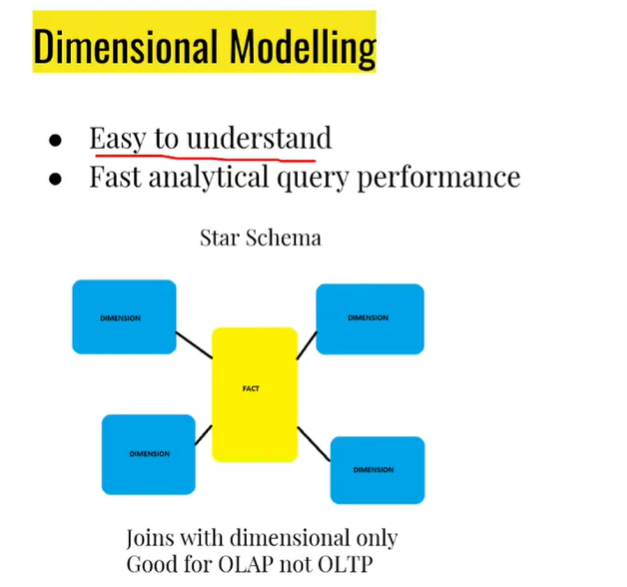
The dimensional model is designed to

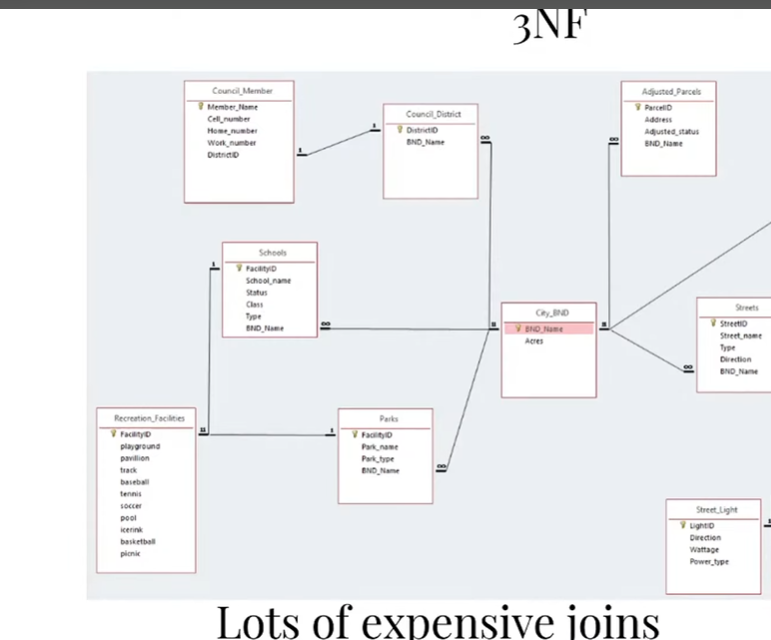
1) Make it easy for business users to work with the data.

2) Improve analytical queries performance.

The technologies used for storing dimensional models(DWH) are different than traditional technologies.

Business-users-facing application such as visualization/dashboards needs clear visuals,aka Business Intelligence (BI) apps(at DATA OUT Stage)





Dimension tables are optimum to use bcoz relation tables can have n no. of joins which can be expensive and not an efficient way of retrieving data.

**Facts & Dimension**

• Fact tables

* Record business events, like an order, a phone call, a book review
* Fact tables columns record events recorded in quantifiable metrics like quantity of an item, duration of a call, book rating

• Dimension tables

* Record the context of the business events, e.g. who, what, where, why, etc...
* Dimension tables columns contain attributes like the store at which an item is purchased, or the customer who made the call, etc...

**How to Identify Fact or Dimension Columns**

For facts, If you're unsure if a column is fact or dimension then simple rule is

that fact is usually: Numeric & Additive

● Examples facts:

* A comment on an article represents an event but we can not easily make a statistic out of its content per se (Not a good fact).
* Invoice number is numeric but adding it does not make sense (Not a good fact)
* Total amount of an invoice could be added to computer total sales (A good fact)

• Example Dimensions:

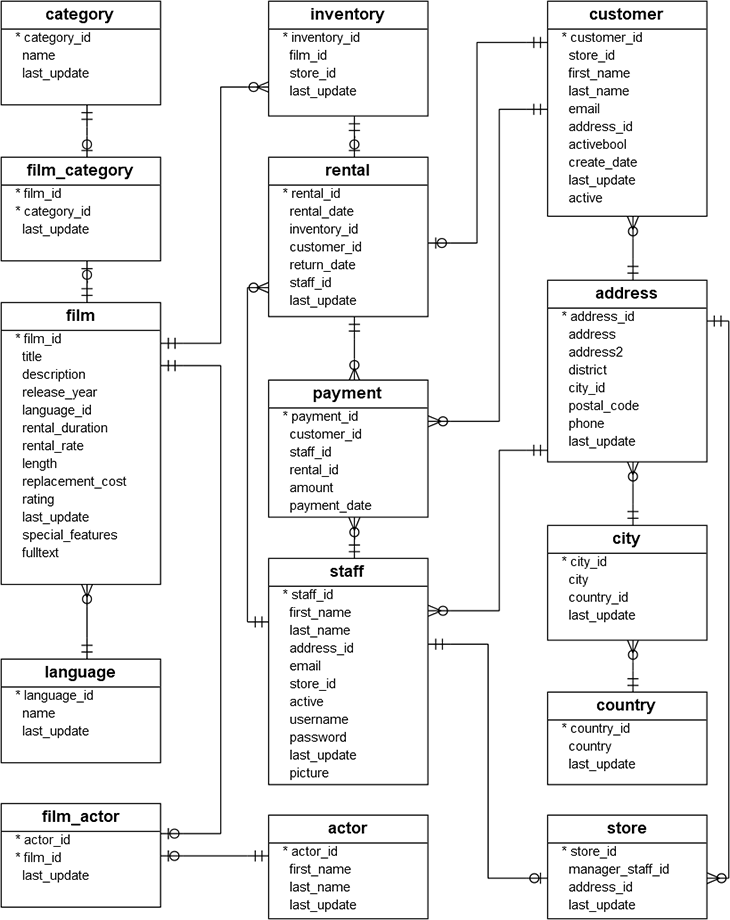
* Data & time are always dimension
* Physical location and their attributes
* Human Roles like customer and staff
* Product sold always good candidates for dimensions

**DVD RENTAL DATA**

**There are 15 tables in the DVD Rental database:**

* actor – stores actors data including first name and last name.
* film – stores film data such as title, release year, length, rating, etc.
* film\_actor – stores the relationships between films and actors.
* category – stores film’s categories data.
* film\_category- stores the relationships between films and categories.
* store – contains the store data including manager staff and address.
* inventory – stores inventory data.
* rental – stores rental data.
* payment – stores customer’s payments.
* staff – stores staff data.
* customer – stores customer data.
* address – stores address data for staff and customers
* city – stores city names.
* country – stores country names.

s



Q1: How easy is it for you to understand this database schema?

A1: Not the complicated, since it is small, but not simple for business users.

Q2: Can you spot candidates for a fact table?

A2: rentals? payments?

Q3: Can you spot candidates for a dimension table?

A3: Customer, Store, Staff, film...

