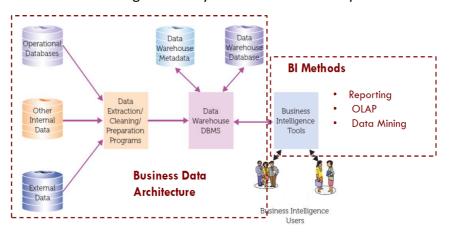# *Business intelligence*

Operational database vs Data warehouse

- **Operational database:**
  Used by various information systems across the organisation on a daily basis. Whenever we talked about databases thus far we talked about operational databases.

- **Data warehouse:**
  Large storage container which collects <u>historical data</u> from operational databases. During this process data needs to be cleaned (data cleansing) so its in the correct format for business intelligence analysis. Unclean data will produce unclean results.

Business Intelligence Methods (3 methods)

1. **Reporting**
   Visualisation tool that integrates raw data and **deliver reports** by:
   - **Filtering** data
   - **Sorting** data
   - **Grouping** data
   - And making simple **calculations** on the data

2. **Online Analytical Processing** [OLAP] →Exam Q (what does OLAP stand for?)
   Data visualisation tool (analysis is manual) with two functions
   1. **Drilling down:** looking at data in increasing depth, depending on their preferences (i.e. zoom in from region, to district to each store)
   2. **Multi-dimensional**: allows users to compare information in different ways and analyse database information from multiple databases at a time.
      OLAP cube: rotate the cube→bring different information to surface based on what you want to see. i.e. which colour car is selling the best for which car? Define the X (car) and Y axis (colour) to see. Simply change the X or Y or Z axis based on what you want to see

3. Data Mining (most sophisticated)

- Use of algorithms to anticipate events or predict future outcomes (could be unethical, target example)

NB EXAM! Trend analysis, fraud detection + **DEFINE CUSTOMER CHURN**

Applications	Description
Consumer clustering	Identify the common characteristics of customers who tend to buy the same products/ services
Customer churn	Identify the reason customers switch to competitors; predict which customers are likely to do it
Fraud detection	Identify characteristics of transactions that are most likely to be fraudulent
Direct marketing	Identify which prospective clients should be included in mailing lists to obtain the highest response rate
Interactive marketing	Predict what each individual accessing a website is most likely to be interested in seeing
Market basket analysis	Understand what products or services are commonly purchased together, and on what days of the week
Trend analysis	Reveal the difference between a typical customer this month and a typical customer last month

MBA Calculations:

1. Probability	Likelihood that a customer will buy item B $$P(B)= B/ \text{total products}$$	
2. Support	Probability customer buys two items together, **read off the table** $$P(A\&B)= \text{read off the table } A\&B/ \text{total items}$$	
3. Confidence	Probability that a customer buys B given that they already bought A $$P(A|B)=\text{support}/\text{probability B}$$	
4. Lift	**NB DEFINE LIFT: ********The factor by which the probability that someone buys product A increases, if they are already buying a product B $$\text{Lift}= \text{Confidence}/ \text{Probability A the base}$$ **> 1 is GOOD, <1 is BAD.** Below 1 means after buying product A, he/she is less likely to buy B. Might happen if two things are substitutes i.e. if you buy a mobile phone, less likely to buy a GPS device. **Which products to cross-sell? Look for highest lift score!**	