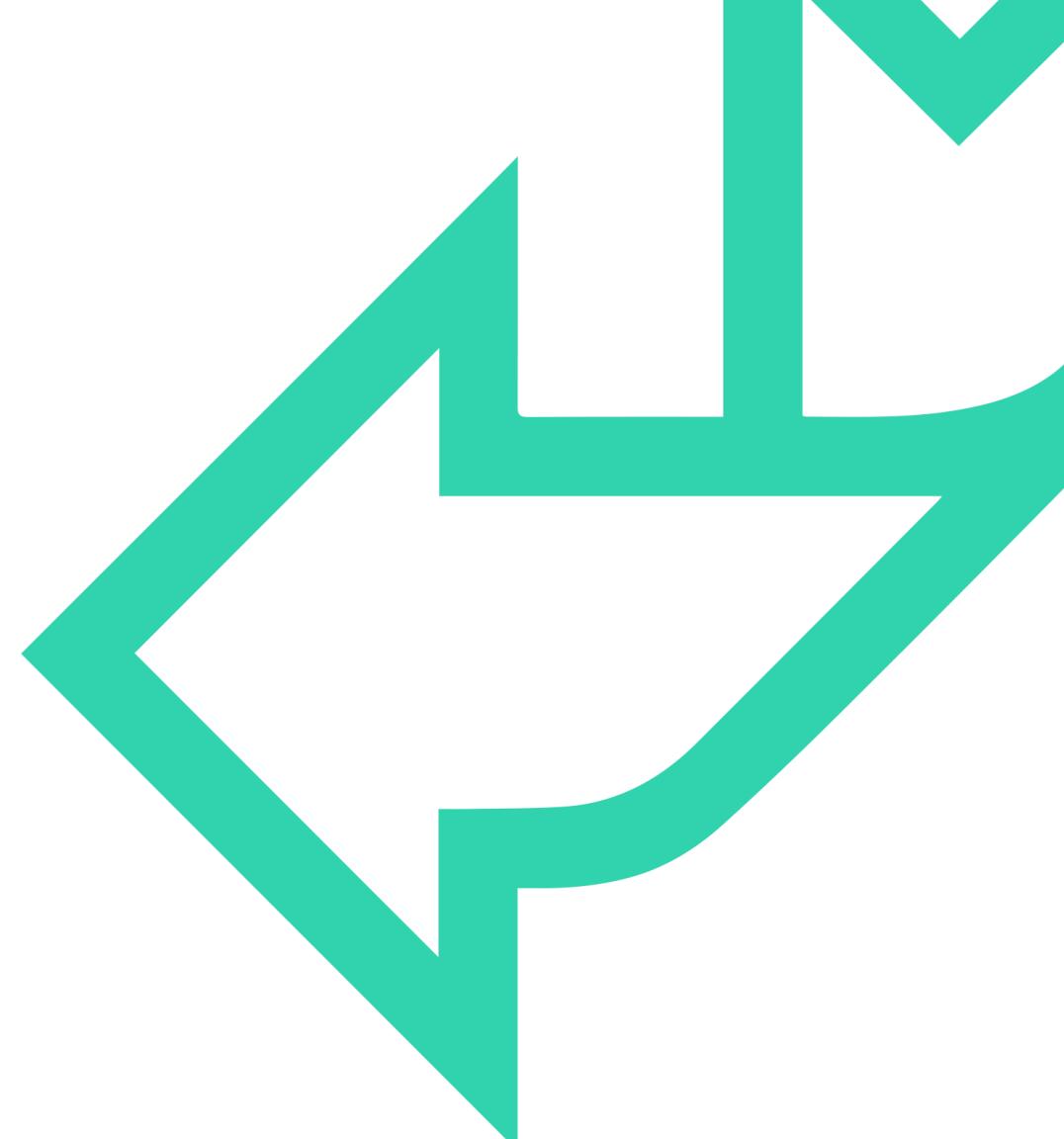




Databases and SQL

Tutor: Rob Menzies
Email: rob.menzies@qa.com



Introduction to Data Modeling

Timings

→ Start 9.30 am

→ Mid morning break 11.00am – 11.15am.

→ Lunch 12.30pm – 1.30pm

→ Mid afternoon break 3.00pm – 3.15pm.

→ Finish 4.00pm

→ QA 4.00pm - 4.30pm

WHAT IS DATA?



UNITED ELECTRICITY

Mr and Mrs Smith
21 Eastern Ave
Ealing
London
W13 100

Bill period from 29 Dec 2007 to 3 Apr 2008
Tax point date 10 April 2008
Your customer account number 26361 234096

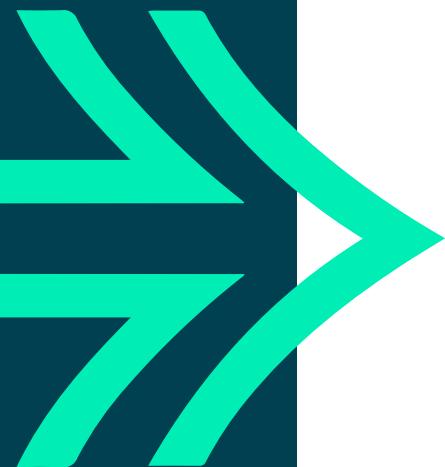
3 Meter number(s)
F03C29801
General Domestic

4 12224 13423 E
5 C - Customer reading
E - Estimated reading
No code - Company reading

6 For any enquiry, please phone 0811 744 000

Reading last time	Reading this time	Tariff	Units	Price of each unit in pence	Amount £p
7 Unrestricted units 1286					
Charges 29 Dec 2007 - 31 Dec 2007					
Standard energy	22	9.02	198		
Discounted energy	19	8.55	162		
Standard energy at 13.730p for 3 day(s) 0.41					
Charges 1 Jan 2008 - 3 Apr 2008					
Standard energy	705	9.90	69.80		
Discounted Energy	540	9.38	50.65		
Standard charge at 15.070p for 93 day(s) 14.02					
1 Monthly Direct Debit Discount VAT at 5.00% on charges of £131.04 6.55					
2 Total this invoice 137.69					
Balance from previous bill Payment received 20 January 2008 40.17 Payment received 20 February 2008 38.00CR Payment received 20 March 2008 38.00CR					
3 DEBIT BUDGET ACCOUNT BALANCE CARRIED FORWARD 63.78					

DATA ANALYSIS IS...



... The process of examining an organisation's (or business system's) data independently of individual applications, so that a data structure can be developed which is capable of supporting all the organisation's (or business system's) func



WHY DO WE NEED TO DO DATA ANALYSIS?

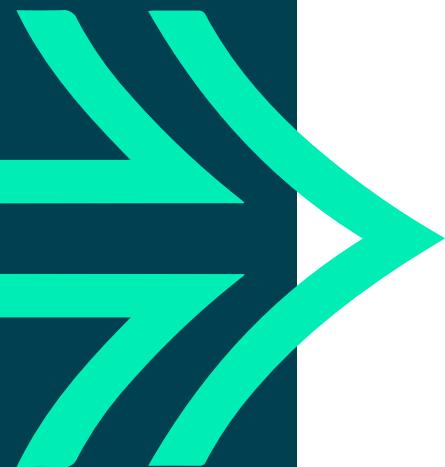


To determine the data requirements of the system
(or organisation !)

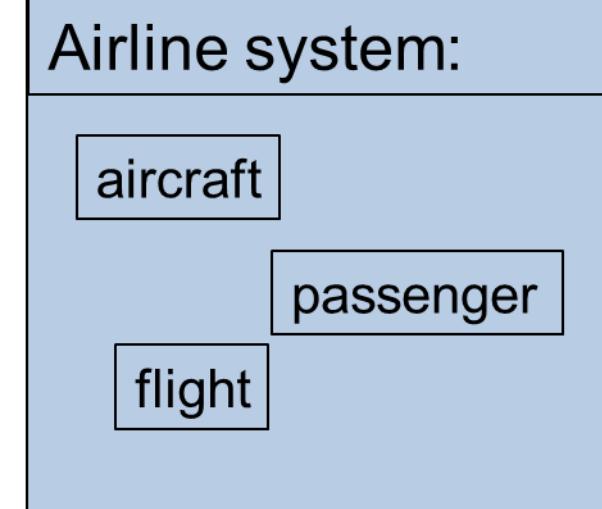
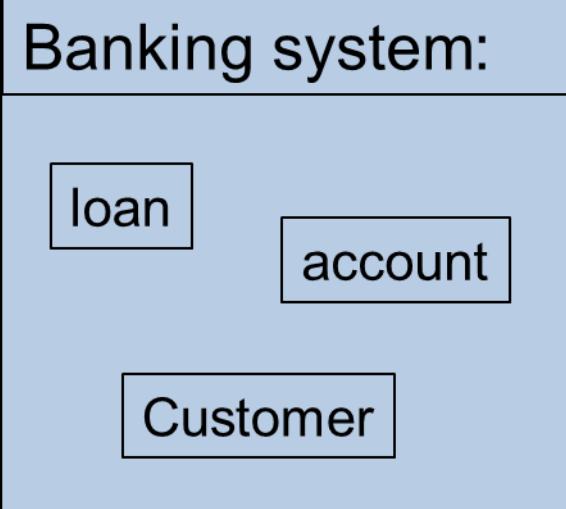
To eliminate (or control) redundancy, so data is
collected
recorded
and stored . . . once and once only

To enable high-level performance planning (at design
stage)

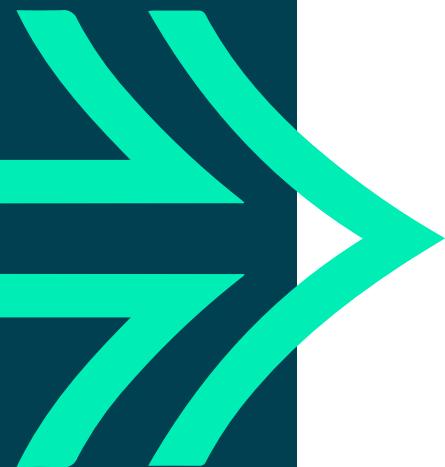
DATA MODELLING



Identify Entities (Tables) and their relationships
Entities can be physical things but also conceptual
Customer (Physical)
Transaction (Conceptual)
Always named in the singular
Draw diagrams that show the system



THE RELATIONAL DATABASE MODEL

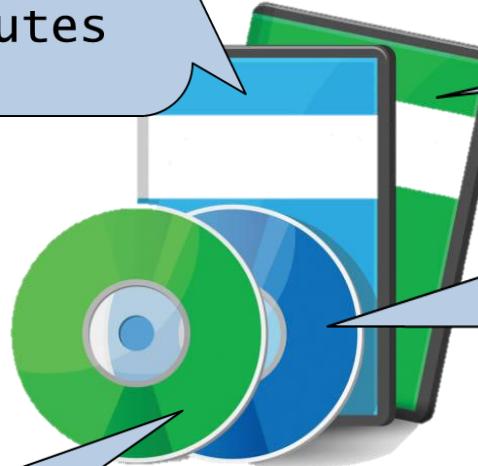


Rank	Title	Director	Worldwide gross	Year	Duration
1	Avatar	James Cameron	\$2,783,165,628	2009	162
2	Titanic	James Cameron	\$1,843,201,268	1997	194
3	The Lord of the Rings: The Return of the King	Peter Jackson	\$1,119,110,941	2003	201
4	Pirates of the Caribbean: Dead Man's Chest	Gore Verbinski	\$1,066,179,725	2006	151
5	Toy Story 3	Lee Unkrich	\$1,063,165,731	2010	103
6	Alice in Wonderland	Tim Burton	\$1,024,299,801	2010	108
7	The Dark Knight	Christopher Nolan	\$1,001,921,825	2008	152
8	Harry Potter and the Philosopher's Stone	Chris Columbus	\$974,733,550	2001	152
9	Pirates of the Caribbean: At World's End	Gore Verbinski	\$960,996,492	2007	169
10	Harry Potter and the Deathly Hallows: Part 1	David Yates	\$954,501,070	2010	146

DVD COLLECTION - QUESTIONS



Give me a list of all films with runtime greater than 120 minutes



what films has Johnny Depp appeared in?

Give me a list of all the film titles in my collection

who directed Alice in wonderland?

DATABASE EXAMPLE TABLE 1



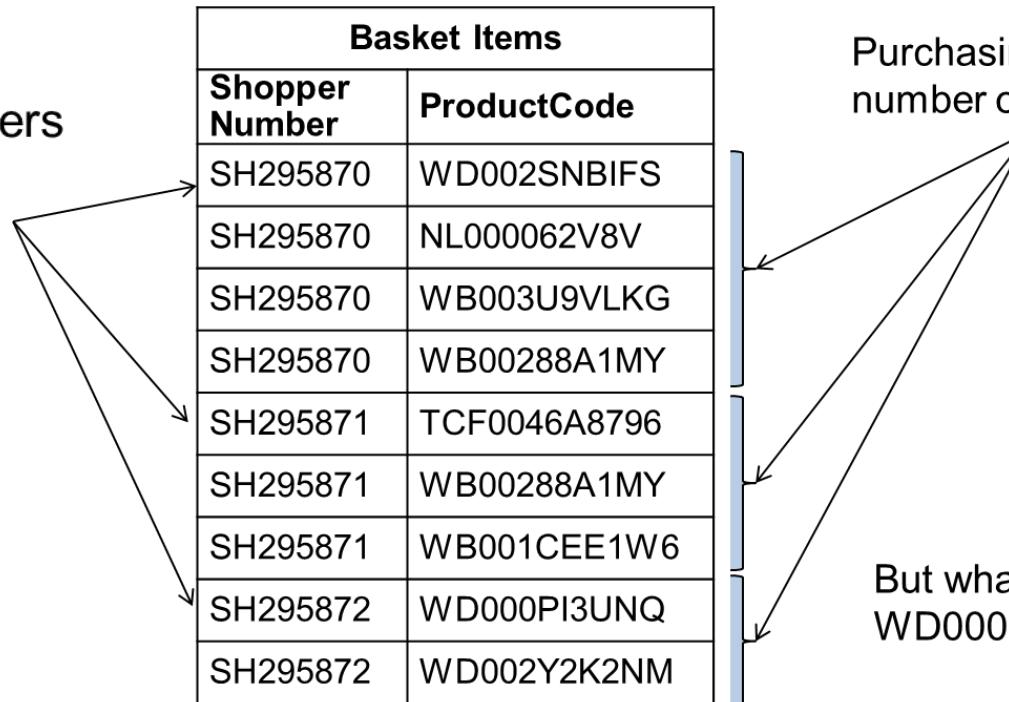
To An Online Film store: Films table

Films			
ProductCode	Title	Director	Price
TCF0046A8796	Avatar	James Cameron	£11.93
TCF00004CZDS	Titanic	James Cameron	£19.99
NL000062V8V	The Lord of the Rings: The Return of the King	Peter Jackson	£6.79
WD000PI3UNQ	Pirates of the Caribbean: Dead Man's Chest	Gore Verbinski	£3.99
WD002SNBIFS	Toy Story 3	Lee Unkrich	£12.99
WD0033WRUN2	Alice in Wonderland	Tim Burton	£6.00
WB001CEE1W6	The Dark Knight	Christopher Nolan	£6.99
WB00288A1MY	Harry Potter and the Philosopher's Stone	Chris Columbus	£4.48
WD002Y2K2NM	Pirates of the Caribbean: At World's End	Gore Verbinski	£3.99
WB003U9VLKG	Harry Potter and the Deathly Hallows: Part 1	David Yates	£14.99

DATABASE EXAMPLE TABLE 2

An on-line film store: BasketItems table

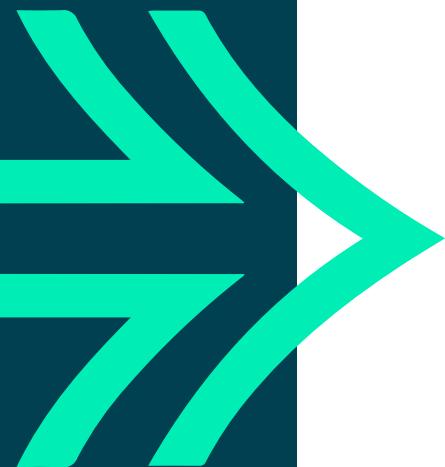
3 Customers



Purchasing Different number of items

But what is an WD000PI3UNQ?

DATA MODELLING PROCESS



An on-line film store: BasketItems table

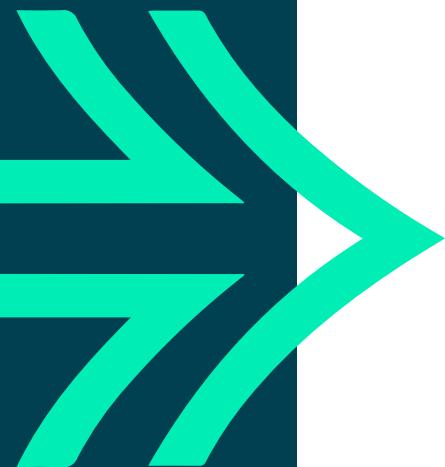
Produce the order for customer SH295872

Basket
Items

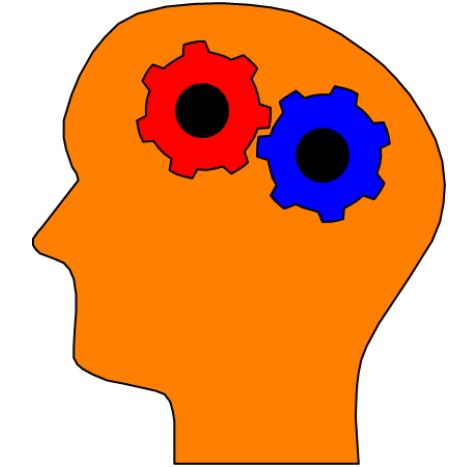
Films

Pirates of the Caribbean: Dead Man's Chest £3.99
Pirates of the Caribbean: At World's End £3.99

DATA MODELLING PROCESS



- Requirements gathering
- Analysis of enterprise information
- Identifying business entities
- Identifying business attributes
- Identifying business relationships



The 3 Layer Model

User (or External) Views

Logical (or Conceptual) View

Physical (or Internal) View

SOURCES OF INFORMATION FOR DATA MODELLING



**Avoid computer system documentation (initially!)
Use a Top-down approach:**

- Management reports
- Front-end input

A TYPICAL SOURCE

Order
Customer

Sales Order Record					
Order No:	Order Date:	Order Value:	Priority:	Salesperson Name:	
Customer No:	Customer Name:				
Invoice Address:		Consignee Address:			
ORDER DETAILS:					
Item No.	Product Code	Description	Qty	Price	Delivery Date
01					
02					
03					
04					
05					
06					
07					
08					
09					
10					
Delivery Instructions:			Special Instructions:		

Sales Person

Product

Order Item

FROM ANALYSIS TO DATA MODEL...

