



INFO90002 Database Systems and Information Modelling

Dr Sean Maynard

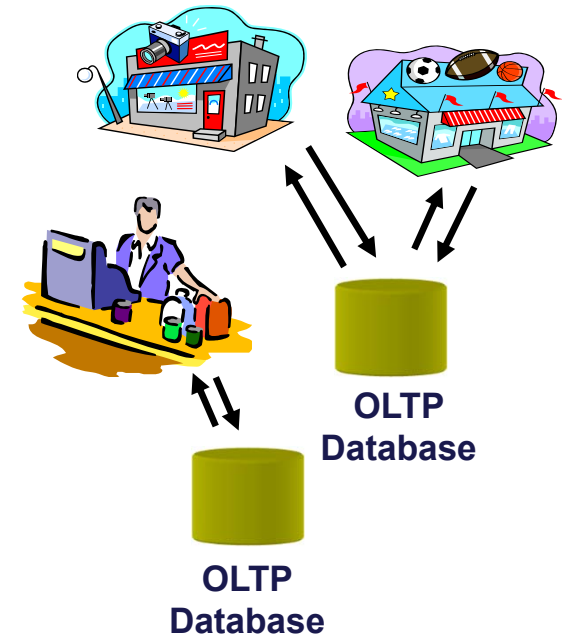
Applications of Database Technology:
Data Warehousing and
Business Intelligence / Analytics



- Different Types of Database Systems
 - Operational
 - Informational
- An Intro to Data Warehousing
 - Data Warehouse architectures
 - Data Marts and Operational Data Stores
- An Intro to Business Intelligence / Analytics

- Why did we get databases in the 1st place
 - Some reasons
 - Reduce duplication of data
 - Increase data sharing
 - Improve data consistency
 - Plus...(not discussed today)
 - Increase data independence
 - Reduce program maintenance

- Used to run day to day business operations - also known as Online Transaction Processing Systems
- Uses OLTP databases to store daily business transactions.
- Characteristics of data:
 - Transaction oriented
 - May be inconsistent and incomplete
 - Volatile – continually changing
 - Current





- Too many of them
 - Everybody wanted one, or two, or more
 - Production, Marketing, Sales, Accounting ...
- Everybody got what was best for them
 - IBM, Oracle, Access, Excel
- Eventually this re-created the problem databases were meant to solve
 - Duplicated data
 - Inaccessible data
 - Inconsistent data



So what's the problem?

- We are doing just fine ...
 - Local optimisation → Global sub-optimisation
 - Organisation not able to capture benefits from its data and technology assets.
 - “data in jail”
- Data are valuable
 - Islands of data, integrated data not available
 - Help to improve processes, improve performance
 - Useful for analysis and decision-making
 - If only we had the tools to analyse the data
- Competitive performance suffers
 - Our competitors are improving and we aren't!



- Need an integrated way of getting the WHOLE organisations data!
- Its really an Informational Database, rather than a Transactional Database
 - A single database that allows all of the organisations data to be stored in a form that can be used to support organisational decision processes and organisational processes.



Difference between Transactional & Informational Systems

Characteristic	Transactional	Informational
Primary Purpose	Run the day to day business	Support decision making
Type of Data	Current data – representing the state of the business	Historical data – snapshots and predictions
Primary Users	Customers, clerks and other employees	Managers, analysts
Scope of Usage	Narrow, planned, fixed interfaces	Broad, ad hoc, complex interfaces
Design Goal	Performance and availability	Flexible use and data accessibility
Volume	Many constant updates and queries on a few tables or rows	Periodic batch updates, complex querying on multiple or all rows



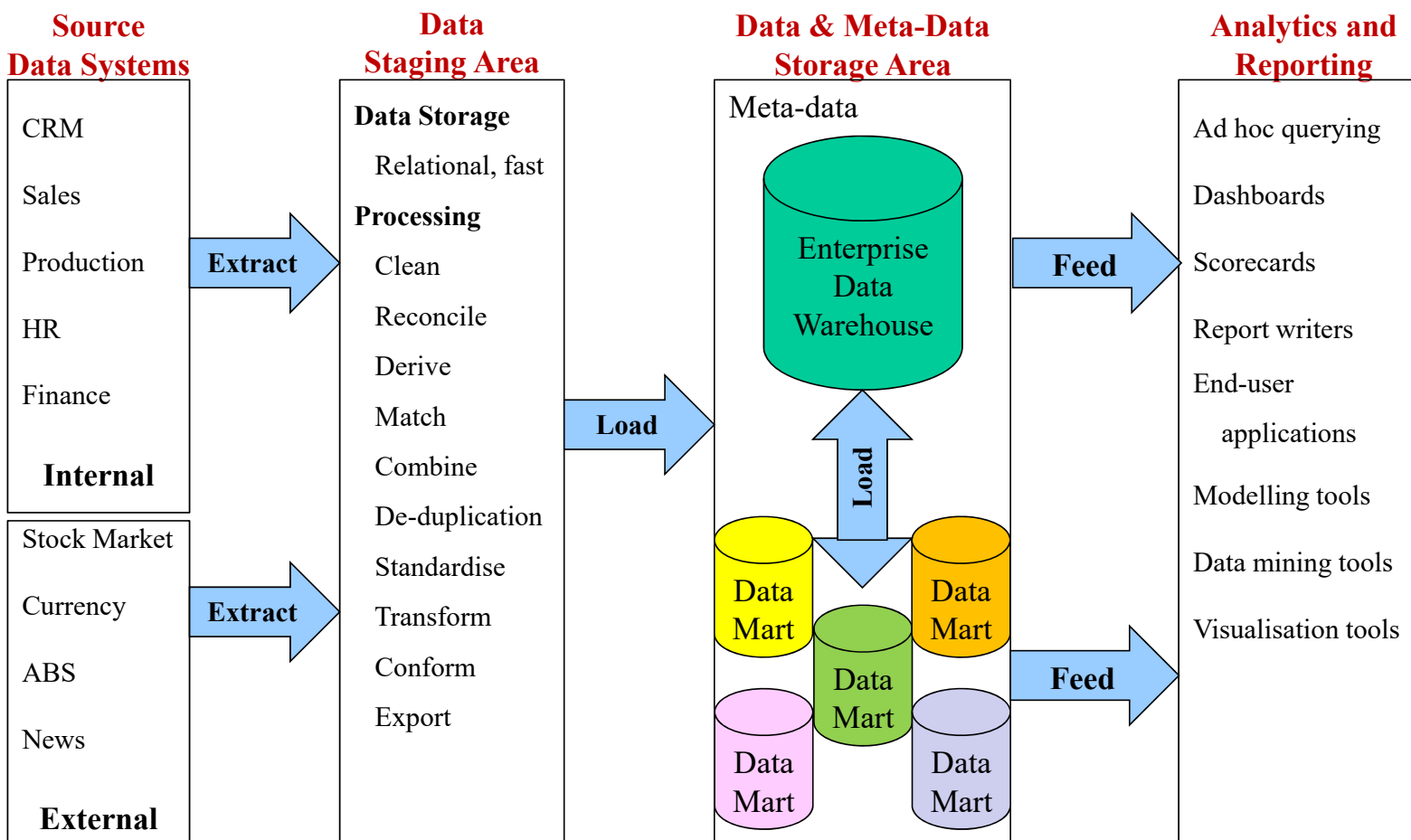
- Data Warehouse
 - A single repository of organisational data
 - Integrates data from multiple sources
 - Extracts data from source systems, transforms, loads into the warehouse
 - “Single version of truth”
 - Makes data available to managers/users
 - Supports analysis and decision-making
- Involve a large data store (often several Terabytes, Petabytes, Exabytes, Zetabytes, Yottabytes of data)



- Data warehouses are used for analysing enterprise-wide data from a range of sources
- Contain both summary and detailed data
- Data need not to be normalised:
 - Data do not need to be updated – no chance of anomalies
 - Speed up query performance in generating reports



A Traditional DW Architecture

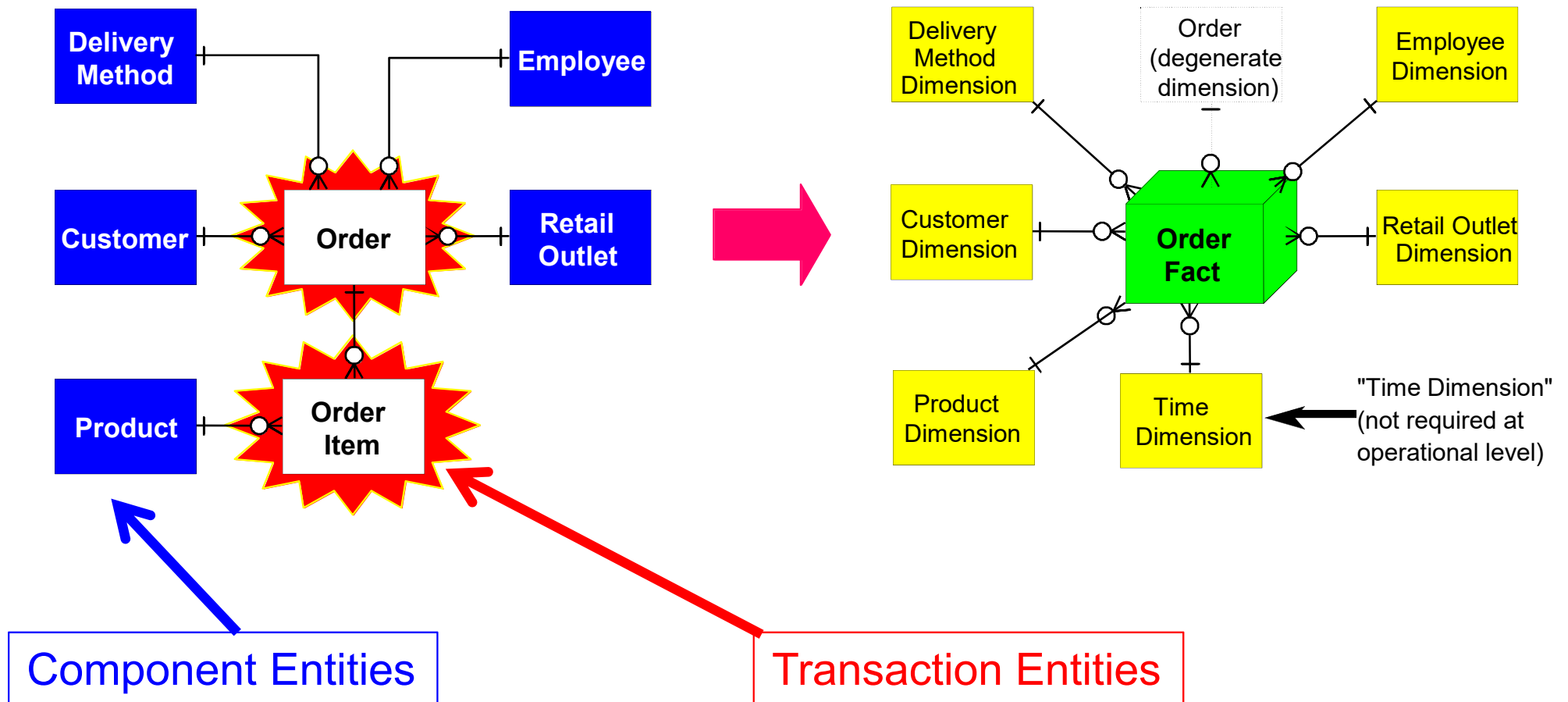




- Subject oriented - organised around particular subjects, integrated across functions (eg sales, customers, products)
- Validated, Integrated data - Data from different systems converted to a common format: allows comparison and consolidation of data from different sources
 - Data from various sources are validated before storing them in a data warehouse.
 - Data quality is crucial to the credibility of the warehouse
- Time variant
 - Trend analysis crucial for decision support: requires historical data
 - Data consist of a series of “snapshots” which are time stamped
- Non-volatile - Users have read access only
- Supports management needs - Used by end users
 - Data warehouses require a simple and easy to navigate structure

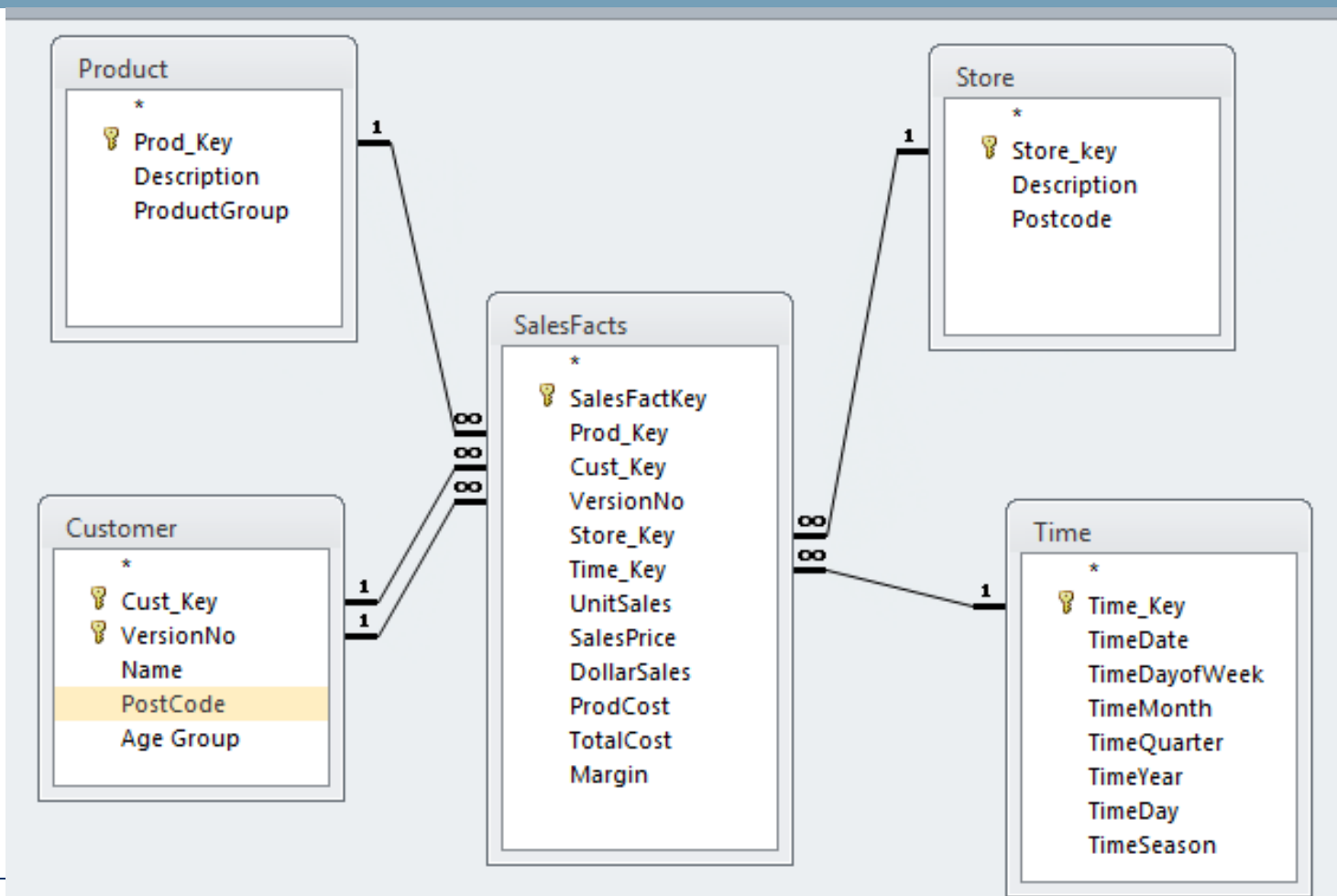


- In a DW rather than focusing on every single entity we focus on... (a rough guide)
 - Facts
 - use a “Fact table”
 - From the ER model these are
 - Transaction entities (business events)
 - » Eg. orders, shipments, payments ...
 - We then determine the components needed to get the facts
 - Generate the Star Schema from these 2 things...





A Simple Example





- What do we do with it?
 - Data warehouses are used for analysing enterprise-wide data from a range of sources
 - Data sources can be from internal or external to an organisation
- Analysis of data
 - Commonly called Analytics, or Business Intelligence



- Forecasting, predictive modeling
 - Extrapolate historical data, trends
 - Statistical techniques: regression, time series
- Simulations
 - ‘What if’ analysis, model
 - Analyse alternative assumptions and scenarios
- Just give me this information
 - Information retrieval and visualisation
 - Probably the most used functionality

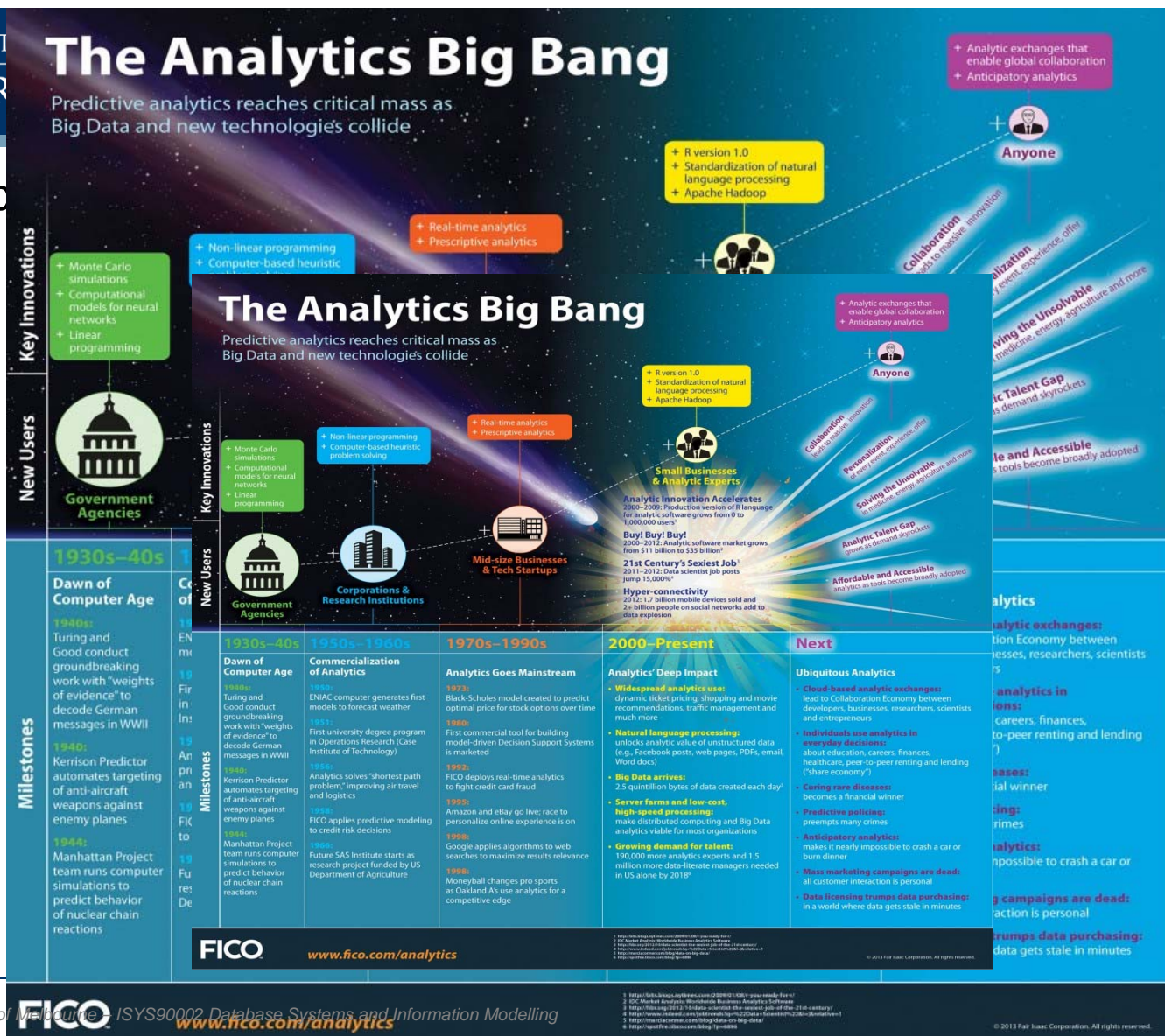


- Basic Query and Reporting
 - Basic question: What happened?
 - Historical focus, limited flexibility
 - Data source : OLTP database, ODS, data mart, data warehouse
- On-Line Analytical Processing (OLAP)
 - Basic question: What happened and why?
 - Historical focus, multi-dimensional - look at data from many points of view, with medium flexibility (slice & dice, drill down)
 - Data Source: Data Mart, Data Warehouse
- Data Mining
 - Basic questions: What is interesting? What might happen?
 - Future focus, high flexibility
 - Extract relationships, patterns and trends, predict future trends.
 - Data source: Data Warehouse



- Harrah's winners information network
 - For your clients, all of whom will lose money
- Marriott: optimal room price in real time
 - Loyalty programs, competitive intelligence,
 - Yield management
- United Postal Service
 - Identification of profit opportunities
 - Transport management
- Australian Tax Office
 - Taxation Fraud Identification and Case Management

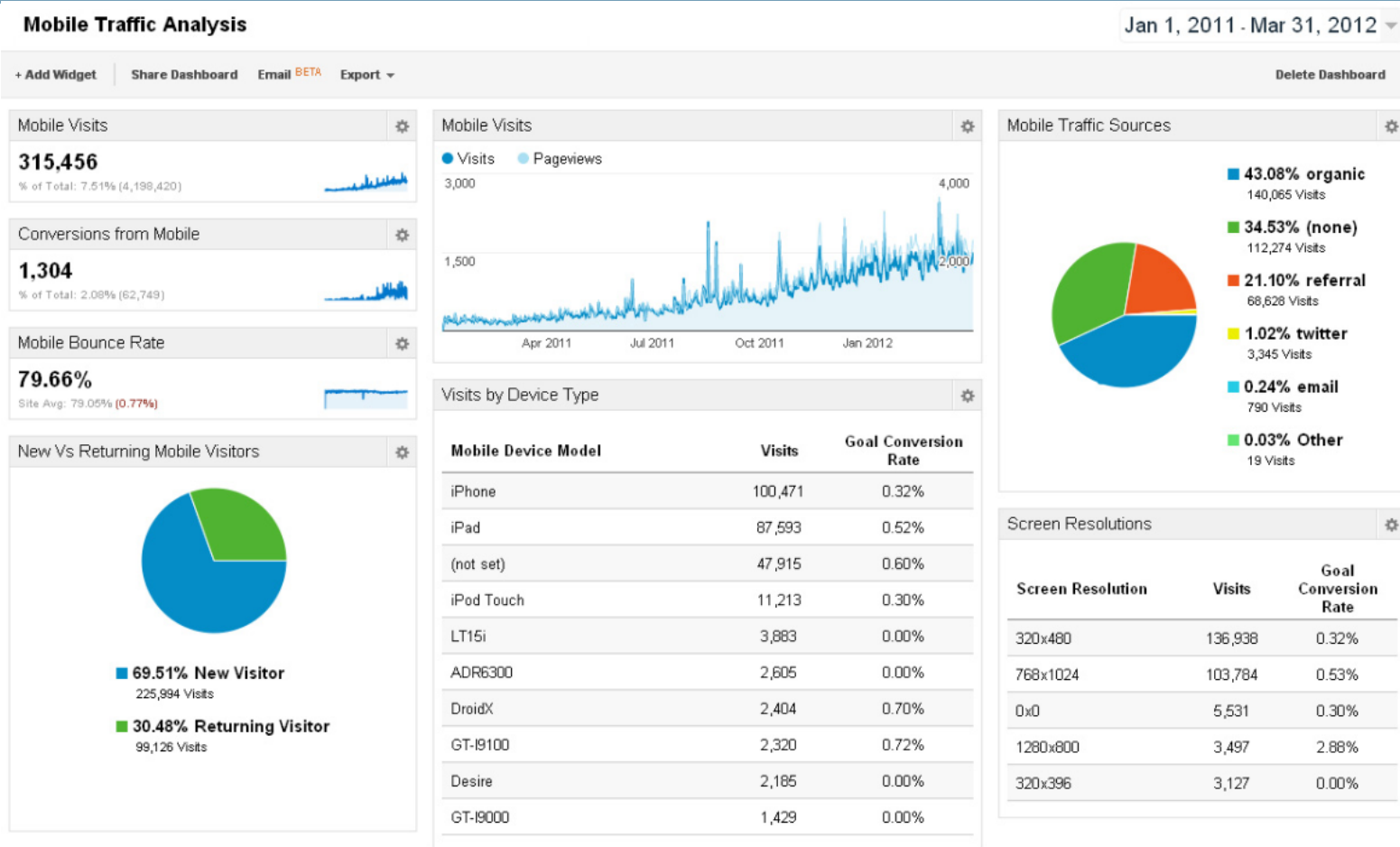
• <http://www.fico.com/analytics>

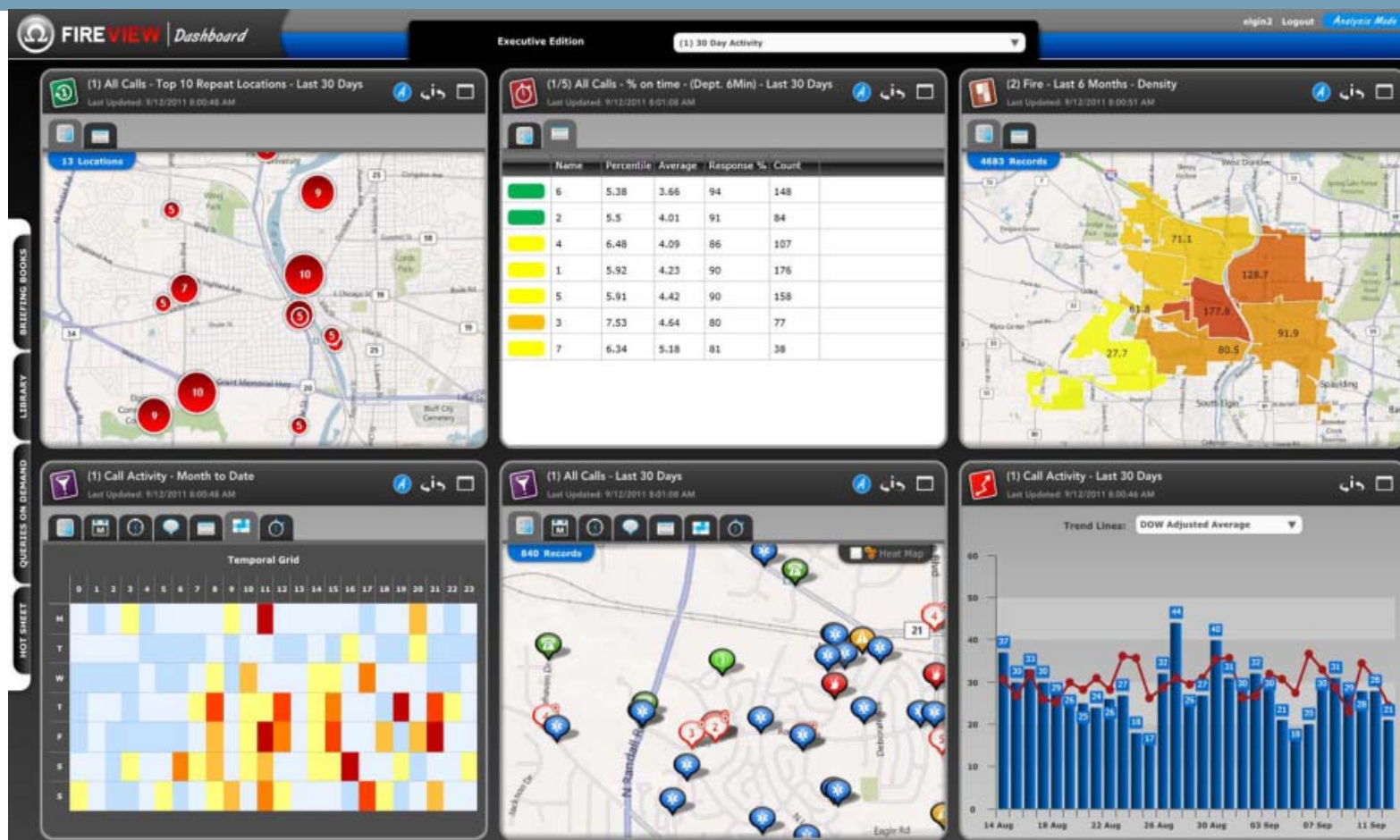


aphic.pdf



- IDC report claims mean 401% ROI
 - median ROI 167%
 - sponsored by DW industry, survey of 62 sites
 - average investment \$2.2m, av payback 2.3 yrs
- Another IDC study (from Negash 2004)
 - 43 North American and European organisations
 - median five-year ROI of 112%
 - mean 457%, range 17% to 2000%
 - \$2 million investment
- Time/cost saving not primary expected benefit
 - “the hope is that a good BI system will lead to a big bang return at some time in the future” (Negash, 2004, p. 185)

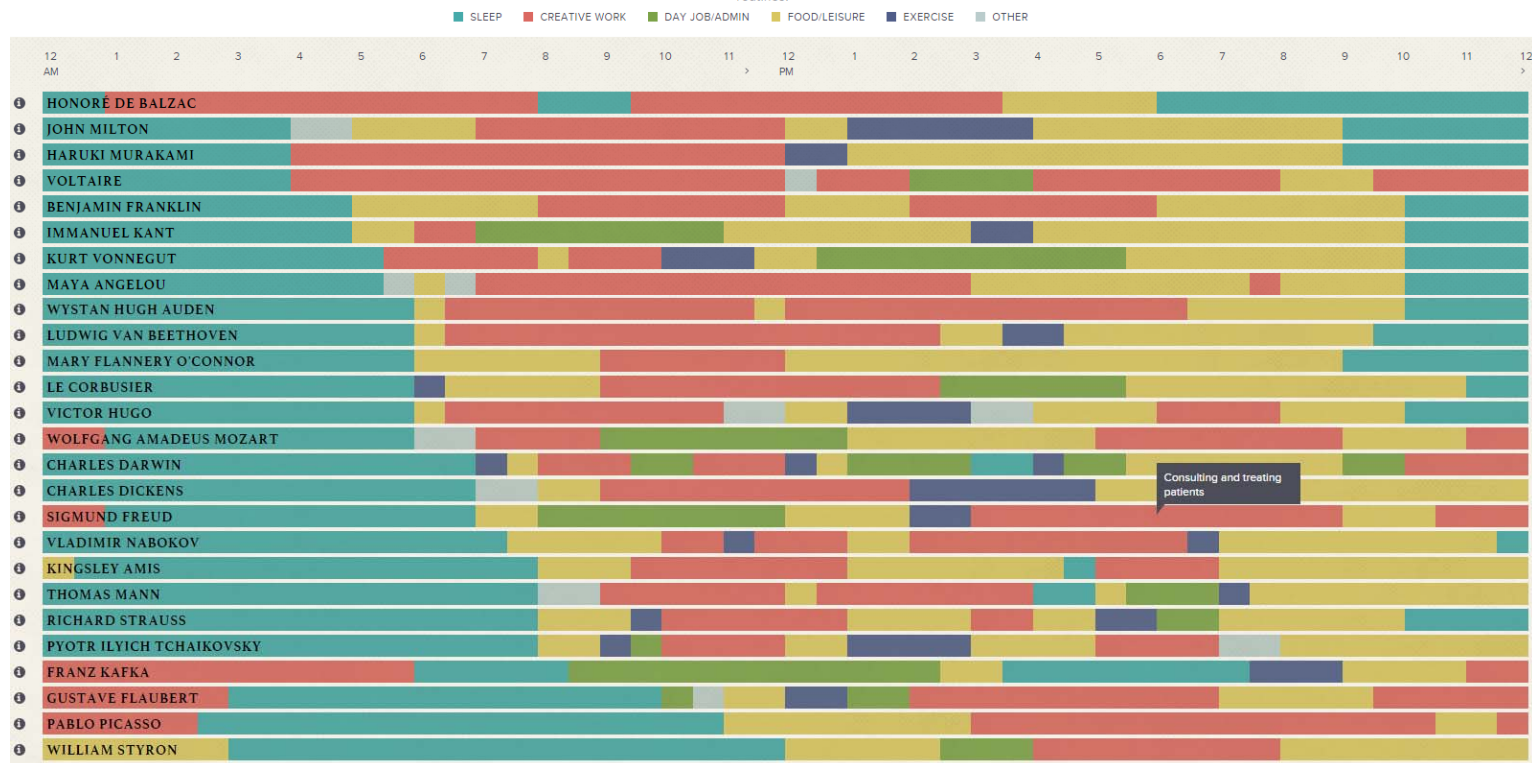


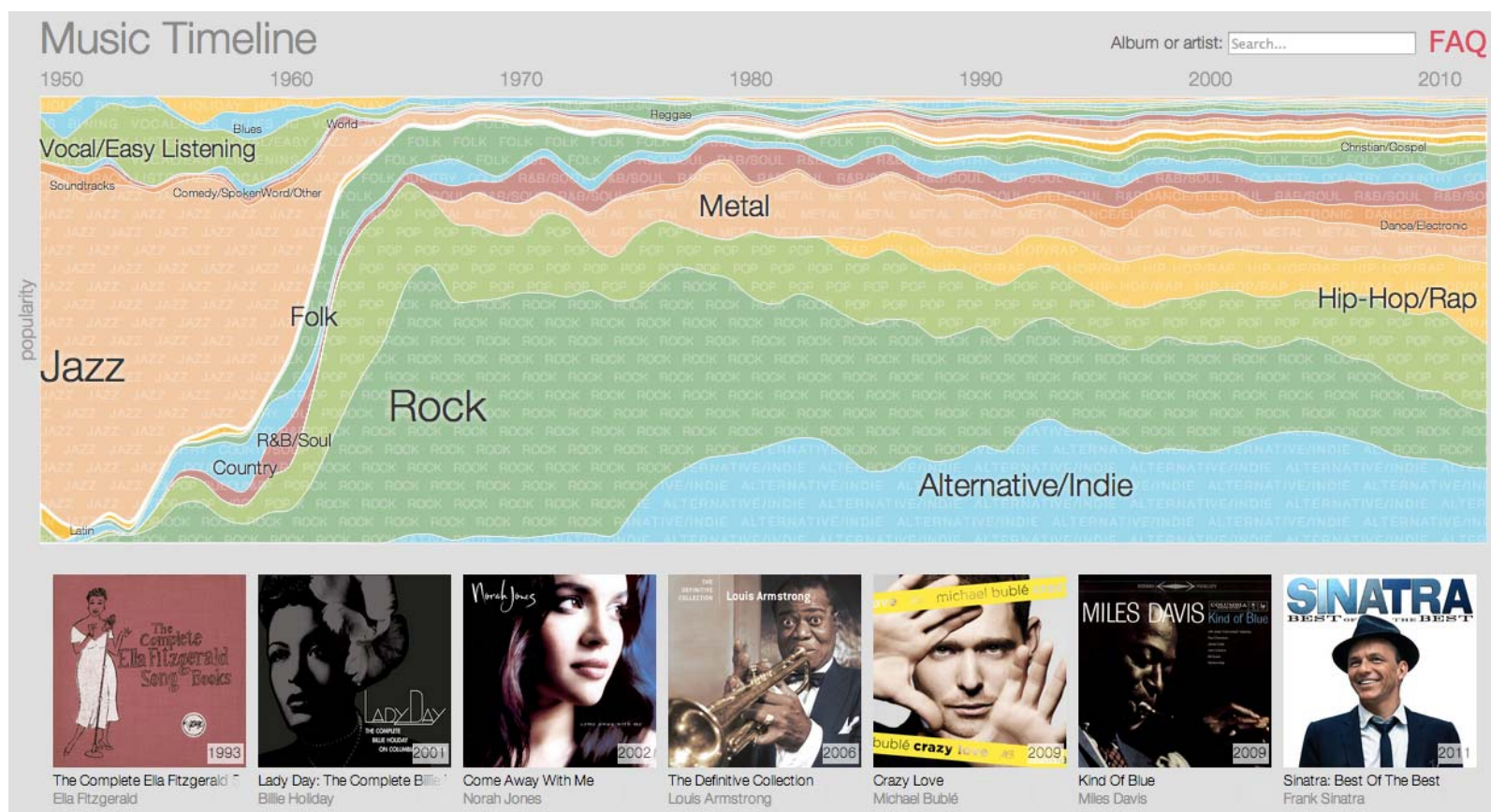




THE DAILY ROUTINES OF FAMOUS CREATIVE PEOPLE

Turns out great minds don't think alike. Discover how some of the world's most original artists, writers and musicians structured their day, based on 'Daily Rituals' by Mason Currey. Filter the different categories by toggling on or off, and hover over the colored bars to learn more about the daily routines.







- <http://homicide.igarape.org.br/>
- <http://benschmidt.org/profGender>
- <http://flowingdata.com/2015/02/04/when-do-americans-leave-for-work/>



- ISYS90086 – Data Warehousing
 - Teaches Data Warehousing
 - Available in most Postgrad Degrees
 - Eg MIT, MIS
 - Taught in the Summer Semester
 - Probably 3 Jan to 7 Feb (plus exam the week after the 5th)
 - TO BE CONFIRMED



- Informational systems are designed to support decision making, while operational systems are designed to support daily business transactions.
- Data warehouse consolidates data from various sources to support decision making.
 - Use Star Schema's
- Data mining tools can provide insights (knowledge discovery) and make predictions based upon historical data