



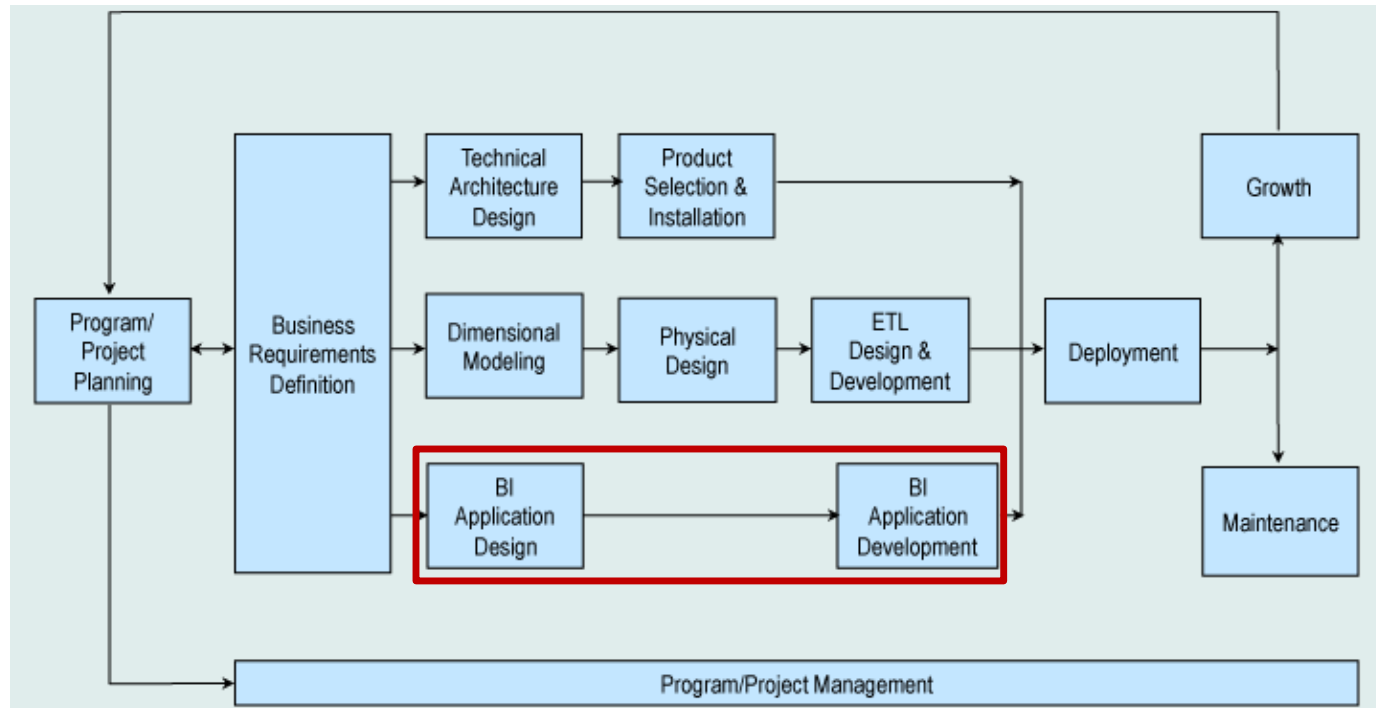
Introduction

School of Information Studies
Syracuse University

Agenda

- Define business intelligence
- Explain the six categories of BI
- Discuss how we measure BI
- Describe the process of developing BI
- Demonstrate how to build OLAP cubes in a multidimensional database

Kimball Lifecycle



BI Defined...

Inmon

“Systems that help make companies understand what makes the wheels of the corporation turn and to help predict the future impact of decisions.”

Kimball

“A generic term to describe leveraging the organization’s internal and external information assets to support improved business decision-making.”

Aren’t they saying the same thing?

DW Is the Foundation for BI

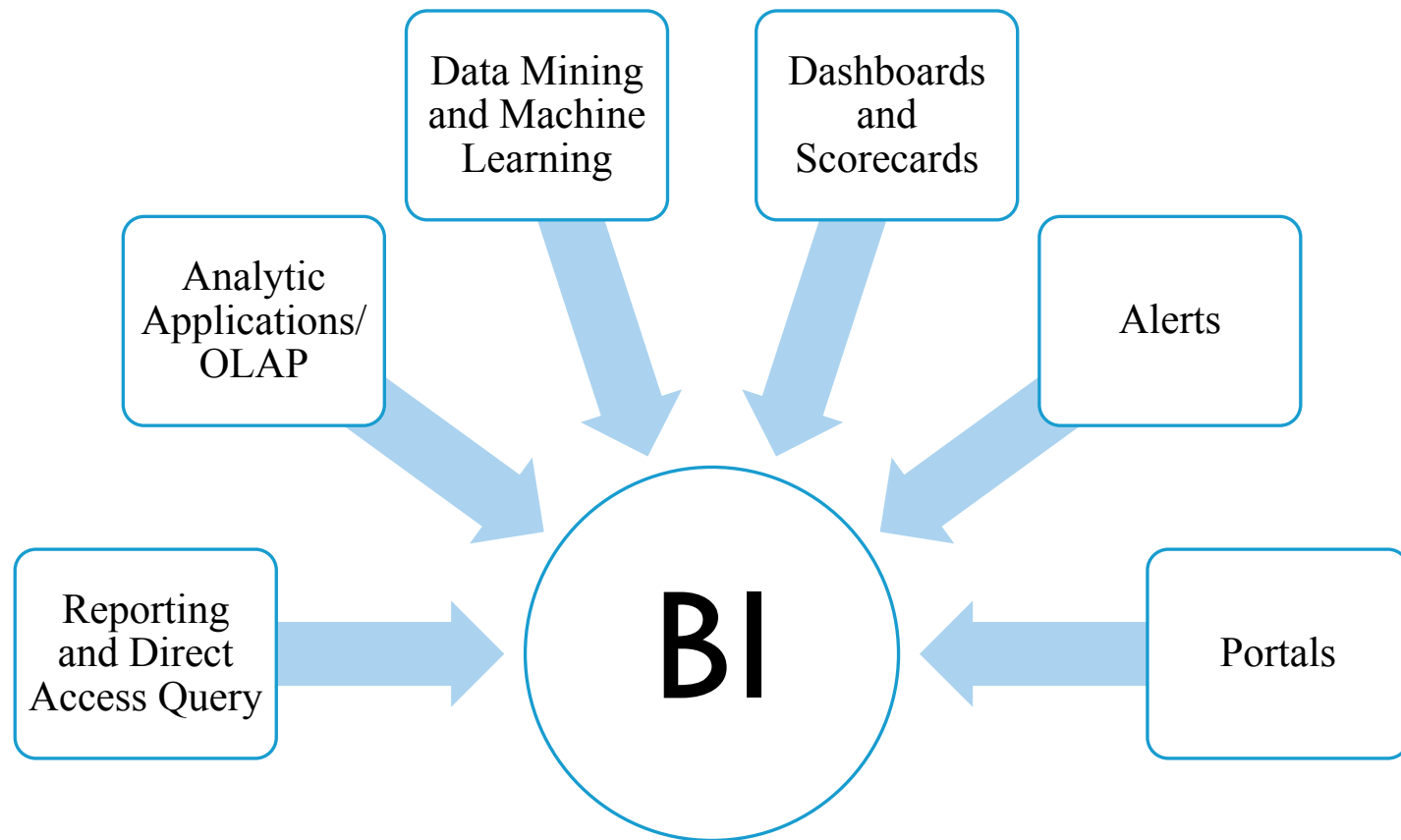




Introduction

School of Information Studies
Syracuse University

Six Categories of BI

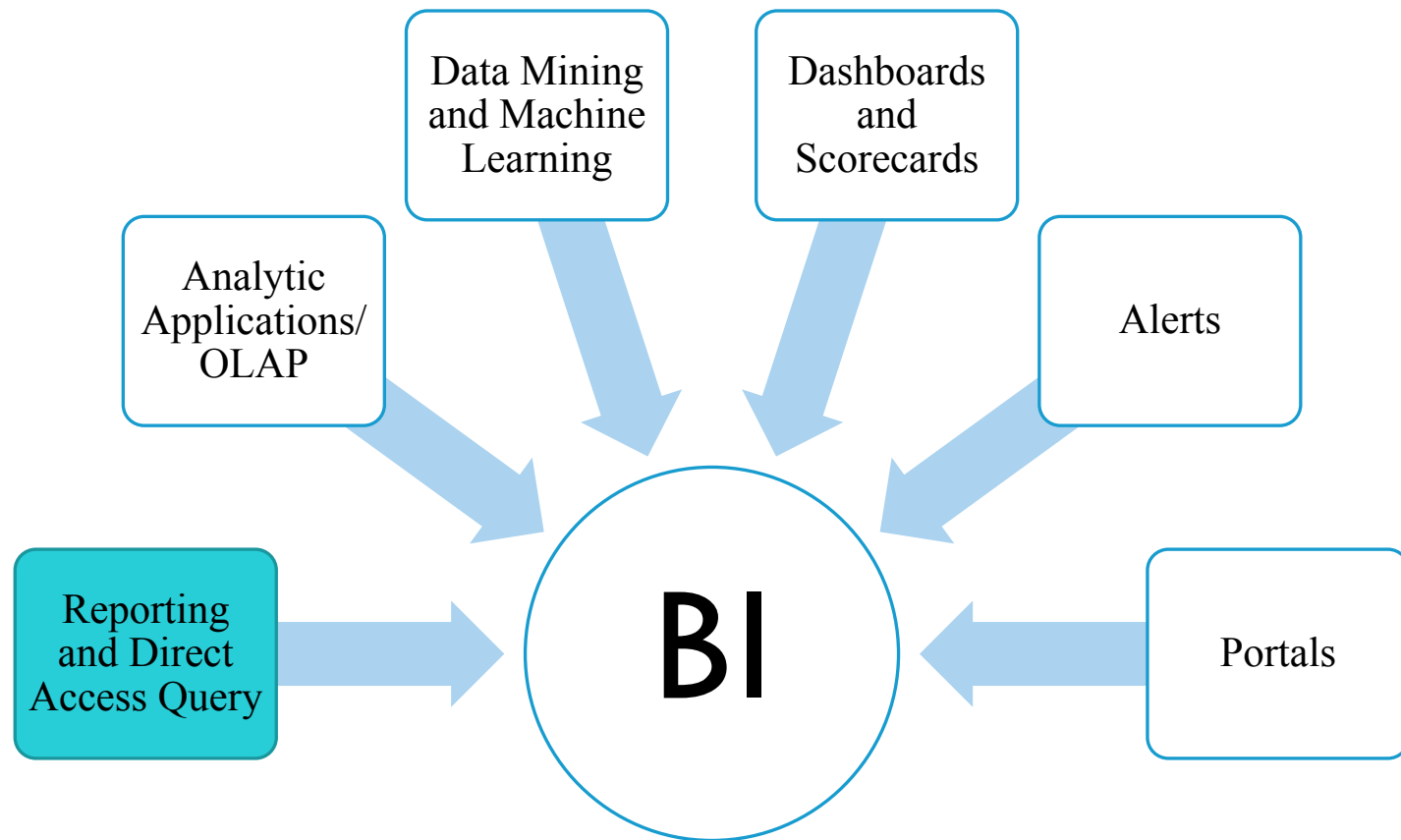




Reporting

School of Information Studies
Syracuse University

Six Categories of BI



Reports

- Information presented in a tabular, matrix format or as a chart.
- Parameters can be added to make them dynamic.
- Typically delivered to business user or available in the portal.
- Pros:
 - Simple to build and execute
 - Easy to understand without training
- Cons:
 - Not flexible
 - Meets only a specific informational need

Example: Reports

Firefox - ClassInstructors - Report Manager

data.ischool.syr.edu/Reports/Pages/Report

Home > Classes > ClassInstructors

Choose Term: 1133 (Summer 2013)

View Report

Find | Next

Term	ClassN	Title	suid	Email	Last Name	First Name	Last Updated
1133	73670	Independent Technology Ed.	100505719	dlnosky@syr.edu	Nosky	Deborah L	7/2/2013 7:45:14 AM
1133	71983	GET Internship	100464719	smbonzi@syr.edu	Bonzi	Susan Monica	7/2/2013 7:45:14 AM
1133	73853	Independent Technology Ed.	100505719	dlnosky@syr.edu	Nosky	Deborah L	7/2/2013 7:45:14 AM
1133	71999	Startup Sandbox	88274	joliddy@syr.edu	Liddy	John DuRoss	7/2/2013 7:45:14 AM
1133	72019	Information Technol	66332	awenner@syr.edu	Wenner	A Randall	7/2/2013 7:45:14 AM
1133	71961	Web Design and Mgt	22792	arperuta@syr.edu	Peruta	Adam ricci	7/2/2013 7:45:14 AM
1133	72015	Blogging for Info Pro	11140	smkeeley@syr.edu	Keeley	Sean Matthew	7/2/2013 7:45:14 AM
1133	71989	Enterprise Technolo	92684	ybmooon@syr.edu	Moon	Young Bai	7/2/2013 7:45:14 AM
1133	72005	Human Comp. Inter	20293	ykim58@syr.edu	Kim	Youngseek	7/2/2013 7:45:14 AM
1133	71941	Internship in Info M	64719	smbonzi@syr.edu	Bonzi	Susan Monica	7/2/2013 7:45:14 AM
1133	71943	Internship in Info M	64719	smbonzi@syr.edu	Bonzi	Susan Monica	7/2/2013 7:45:14 AM
1133	71939	Coop Educ in Info M	64719	smbonzi@syr.edu	Bonzi	Susan Monica	7/2/2013 7:45:14 AM
1133	73669	Coop Educ in Info M	64719	smbonzi@syr.edu	Bonzi	Susan Monica	7/2/2013 7:45:14 AM
1133	71967	Intro to Library & Inf	25436	rdlankes@syr.edu	Lankes	Richard David	7/2/2013 7:45:14 AM
1133	72007	Applied Info Security	16207	mmpollitt@syr.edu	Pollitt	Mark Mollineaux	7/2/2013 7:45:14 AM
1133	71959	Info Arch for Interne	08563	pwfitzgi@syr.edu	Fitzgibbons	Patrick W	7/2/2013 7:45:14 AM
1133	73685	Effective Search Tec	09054	jahurst@syr.edu	Hurst-Wahl	Jill Ann	7/2/2013 7:45:14 AM
1133	71997	Startup Sandbox	88274	joliddy@syr.edu	Liddy	John DuRoss	7/2/2013 7:45:14 AM
1133	73672	Scripting Application	92209	njmccrac@syr.edu	McCracken	Nancy	7/2/2013 7:45:14 AM
1133	73777	Social Informatics	80014	armill05@syr.edu	Miller	Alison Ruth	7/2/2013 7:45:14 AM
1133	72025	Social Networking in	80014	armill05@syr.edu	Miller	Alison Ruth	7/2/2013 7:45:14 AM

Direct Access Query and Reporting Tools

Power users have access to software and the dimensional model for writing their own queries.

Four key functions of these tools:

- **Query formulation:** assist with data queries
- **Analysis and presentation capabilities:** placing the data in “presentation quality” format
- **User experience:** metadata access, easy to use, prevent misuse of data
- **Technical features:** multitasking, scheduling, import/export

Examples: MS Excel, Hyperion

Example: Query Tool

The screenshot displays the Hyperion Interactive Reporting Studio interface for a query named 'ist_classSections_populate'. The interface includes a menu bar (File, Edit, View, Insert, Format, Query, DataModel, Tools, Window, Help) and a toolbar with various icons. The main workspace is divided into several sections:

- Request:** A list of fields including last_update, term, course_num, course_subj, class_section, class_number, course_cr, course_title, facility_id, instructor_name, Name Last First Mid, instructor_role_id, class_m, class_t, class_w, class_h, class_f, class_s, class_u, class_start, class_end, course_component, class_notes, class_campus, instructor_suid, logon_id, instructor_email, and facility_desc.
- Sort:** Fields term, course_num, and class_section are selected.
- Filter:** A series of AND conditions: Acad Org AND Campus AND Class Notes Seq AND Schedule Print Ind AND Institution AND Acad Career AND current_terms.
- Tables:** Four tables are visible, each with a list of fields:
 - V Sis Term:** Institution, Acad Career, Term, Term Desc, Term Sh Desc, Term Start Date, Term End Date.
 - V Crs Class Meeting:** Course Id, Course Offer Nbr, Term, Session Code, Class Section, Class Meeting Number, Class Nbr, Instr Emplid, Instr Assign Seq, Course Topic Id, Course Topic Desc, Course Topic Sh Desc, Description, End Date, Facility Id, Facility Desc, Facility Room Capacity, Facility Sh Desc, Friday Day Ind, Instr Grade Roster Access, Instr Grade Roster Sh Desc, Instructor Contact Minutes, Instructor Load Factor, Instructor Name, Instructor Role, Instructor Role Desc, Instructor Role Sh Desc.
 - V Crs Class:** Subject, Catalog Nbr, Course Id, Course Offer Nbr, Term, Session Code, Class Section, Associated Class, Class Nbr, Institution, Course Component, Acad Career, Acad Career Desc, Acad Career Sh Desc, Acad Group, Acad Group Desc, Acad Group Sh Desc, Acad Org, Acad Org Desc, Acad Org Sh Desc, Acad Progress Ur, Ap Account, Ap Affiliate, Ap Budget, Period, Ap Business Unit, Ap Class Field, Ap Department.
 - V Cc Personal Data:** Emplid, Birth Country, Birth Country Desc, Birth Country Sh Desc, Birth Date, Citizenship Code, Citizenship Desc, Citizenship Sh Desc, Curr Addr Line1, Curr Addr Line2, Curr Addr Line3, Curr Addr Line4, Curr Addr Ctv.

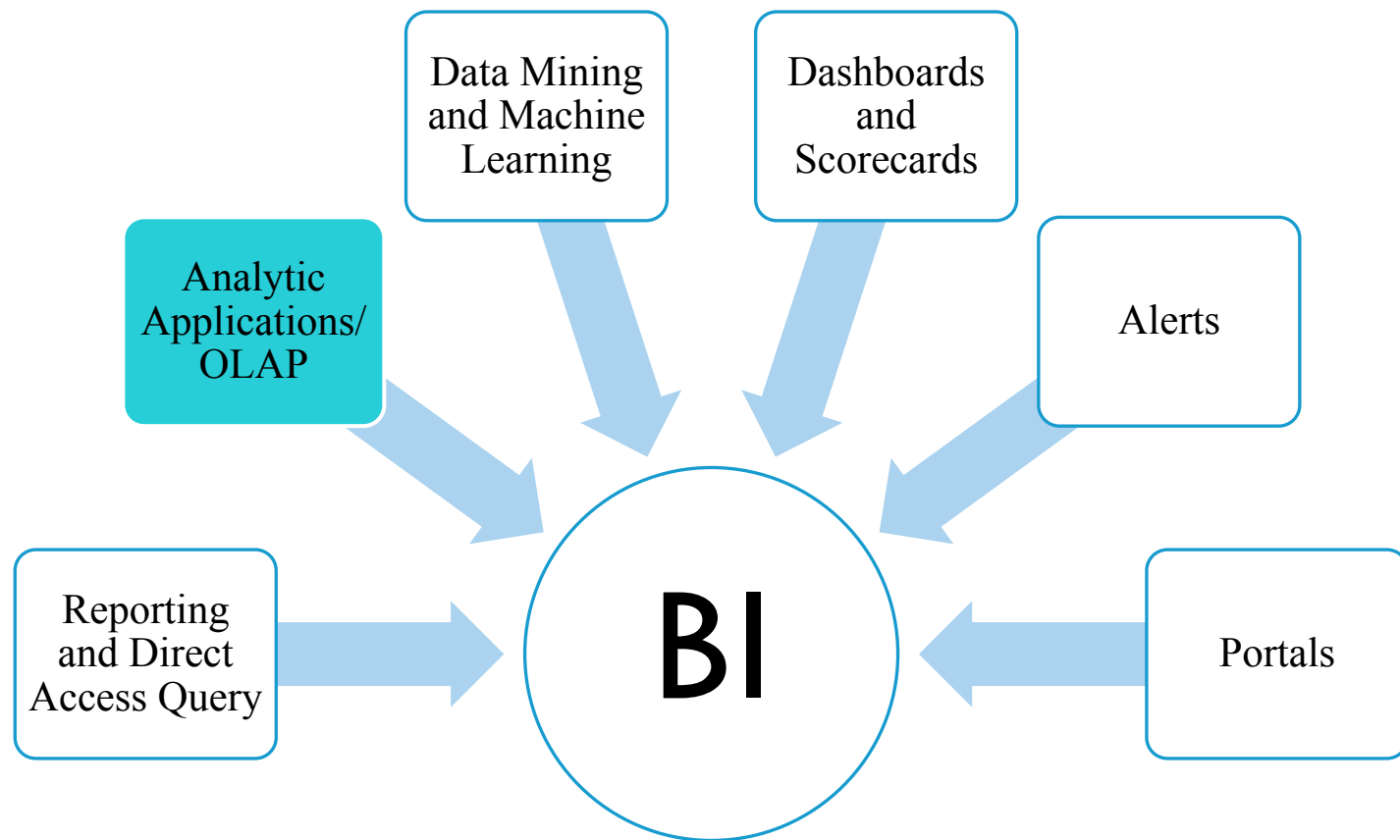
Relationships are indicated by lines with symbols: a single line for 'V Sis Term' to 'V Crs Class Meeting', a double line with a plus sign for 'V Cc Personal Data' to 'V Crs Class Meeting', and single lines for 'V Crs Class Meeting' to 'V Crs Class'.



Analytic Applications

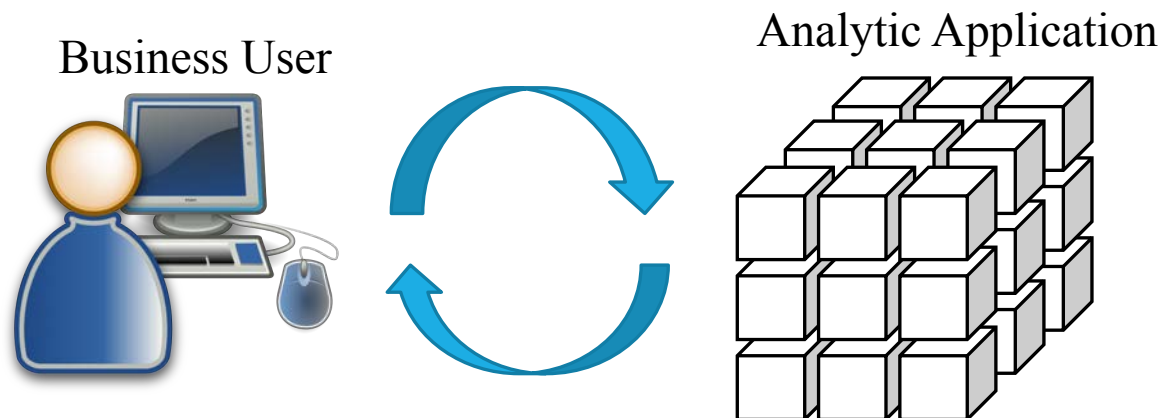
School of Information Studies
Syracuse University

Six Categories of BI



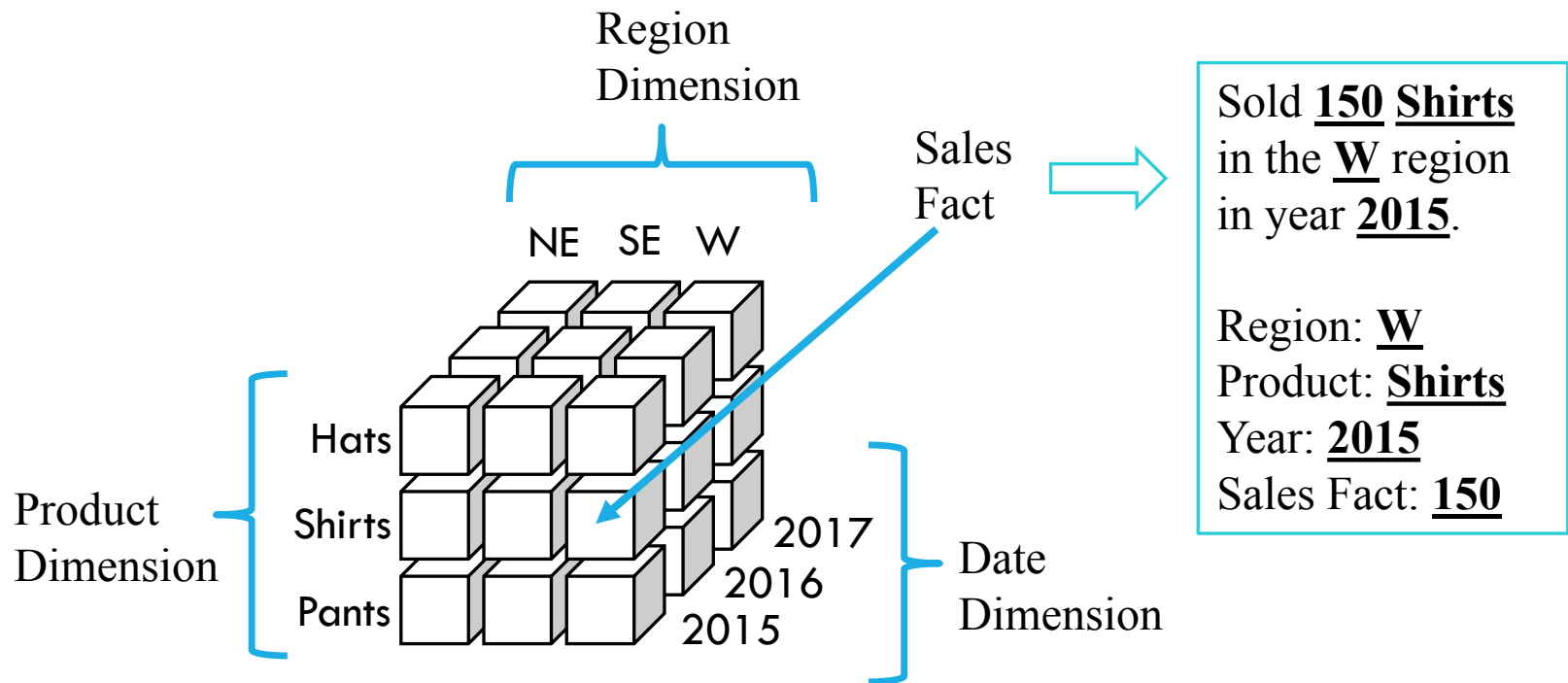
Analytic Applications

- Based on the DDS, use fact tables and dimensions.
- Users interact with the data and use them to answer their own questions.
- OLAP = online analytical processing. Instead of a report, presents an explorable dataset.
- Typically use a multidimensional database.



OLAP Cube

Data structure for representing analytic data.



Example: Analytic Application

The screenshot displays an Excel spreadsheet with a PivotTable summarizing product data. The PivotTable is located in the range A14:C14. The columns are labeled 'Row Labels', 'Gross Profit', and 'Product Gross Profit Margin'. The rows list various product categories and their corresponding financial metrics. The 'Grand Total' row is highlighted in blue.

Row Labels	Gross Profit	Product Gross Profit Margin
Accessories	\$634,467.16	49.88%
Bikes	\$10,515,096.61	11.11%
Clothing	\$368,836.00	17.42%
Bib-Shorts	\$51,256.59	30.74%
Caps	(\$1,203.20)	-2.35%
Gloves	\$83,412.52	34.36%
Jerseys	(\$92,921.27)	-12.35%
Shorts	\$155,945.59	37.71%
Socks	\$12,163.31	40.89%
Tights	\$60,714.84	30.08%
Vests	\$99,467.61	38.33%
Components	\$1,032,966.48	8.75%
Grand Total	\$12,551,366.25	11.43%

The PivotTable Fields task pane on the right shows the following configuration:

- Show fields:** (All)
- Product Categories:** Checked
- Product Model Lines:** Unchecked
- FILTERS:** (Empty)
- COLUMNS:** Σ Values
- ROWS:** Product Categories
- VALUES:** Gross Profit, Product Gross Profit Margin, Product Gross Profit Margin ...

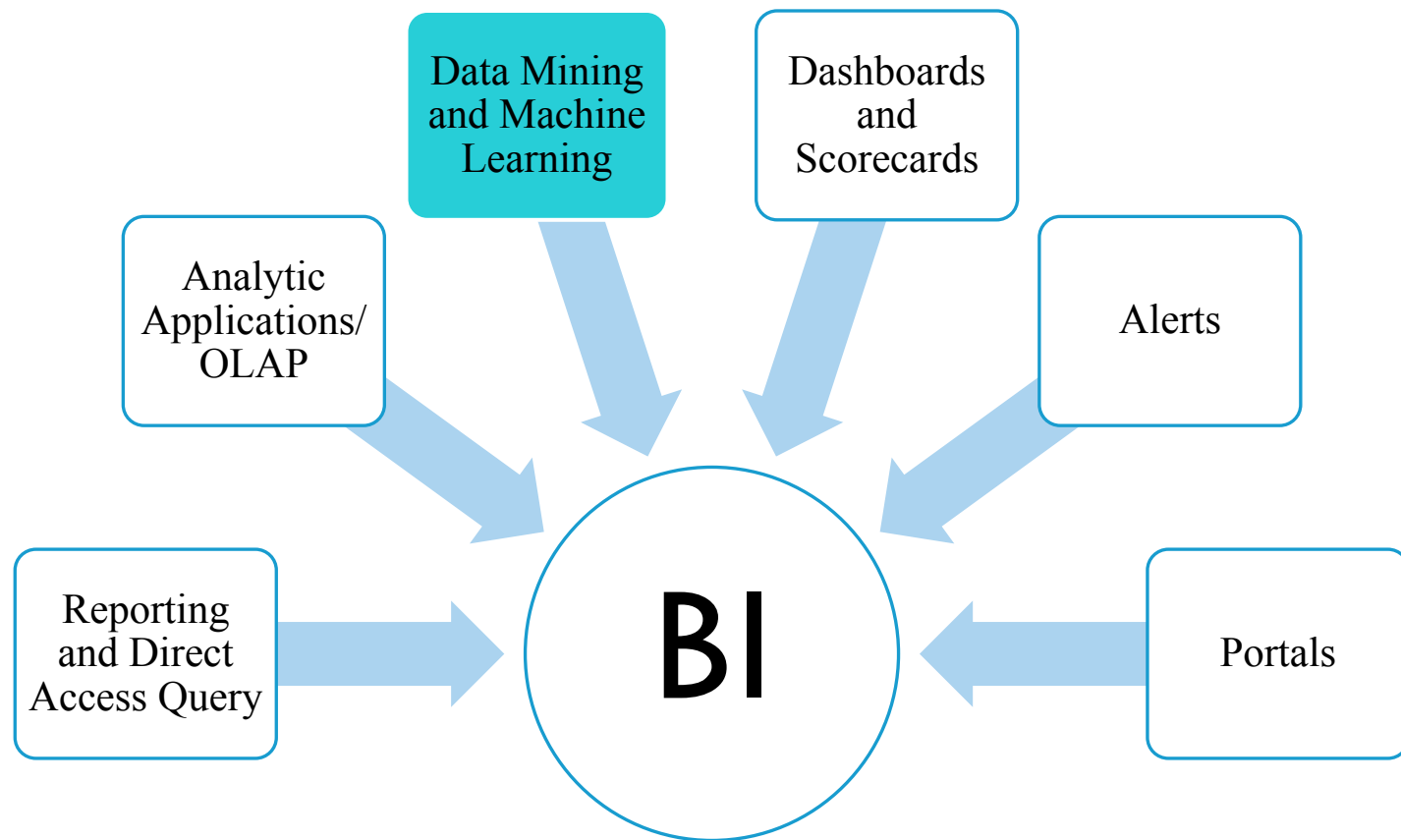
The status bar at the bottom indicates 'READY' and 'Sheet1'.



Data Mining

School of Information Studies
Syracuse University

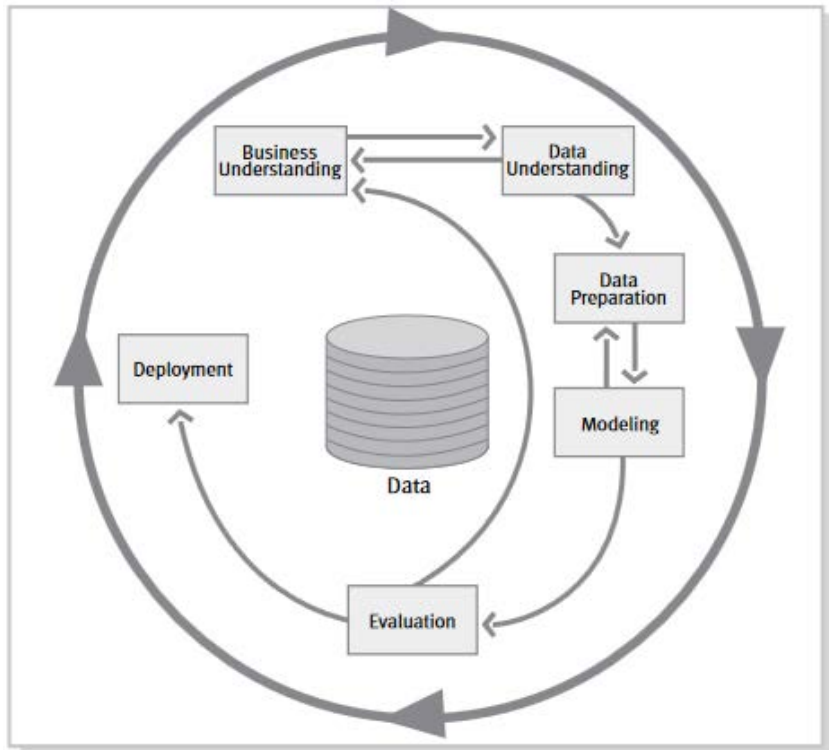
Six Categories of BI



Data Mining

- A process of **data exploration** with the intent to find **patterns, relationships, or insights** of organizational value.
- Data mining is part of **descriptive** and **predictive** analytics.
 - **Clustering and classifying data:** “Customer 1 is Type A, Customer 2 is Type B.”
 - **Estimating and predicting:** “Customer Type A will spend \$N this year.”
 - **Affinity grouping:** “Customers who buy Product X are also likely to buy Product Y.”
 - **Anomaly detection:** fraud detection, unusual patterns, outliers.

CRISP-DM: Cross-Industry Standard Process for Data Mining (2000)



*CRISP-DM Step-By-Step
Guide 1.0*

<https://www.the-modeling-agency.com/crisp-dm.pdf>

Data Mining and Machine Learning

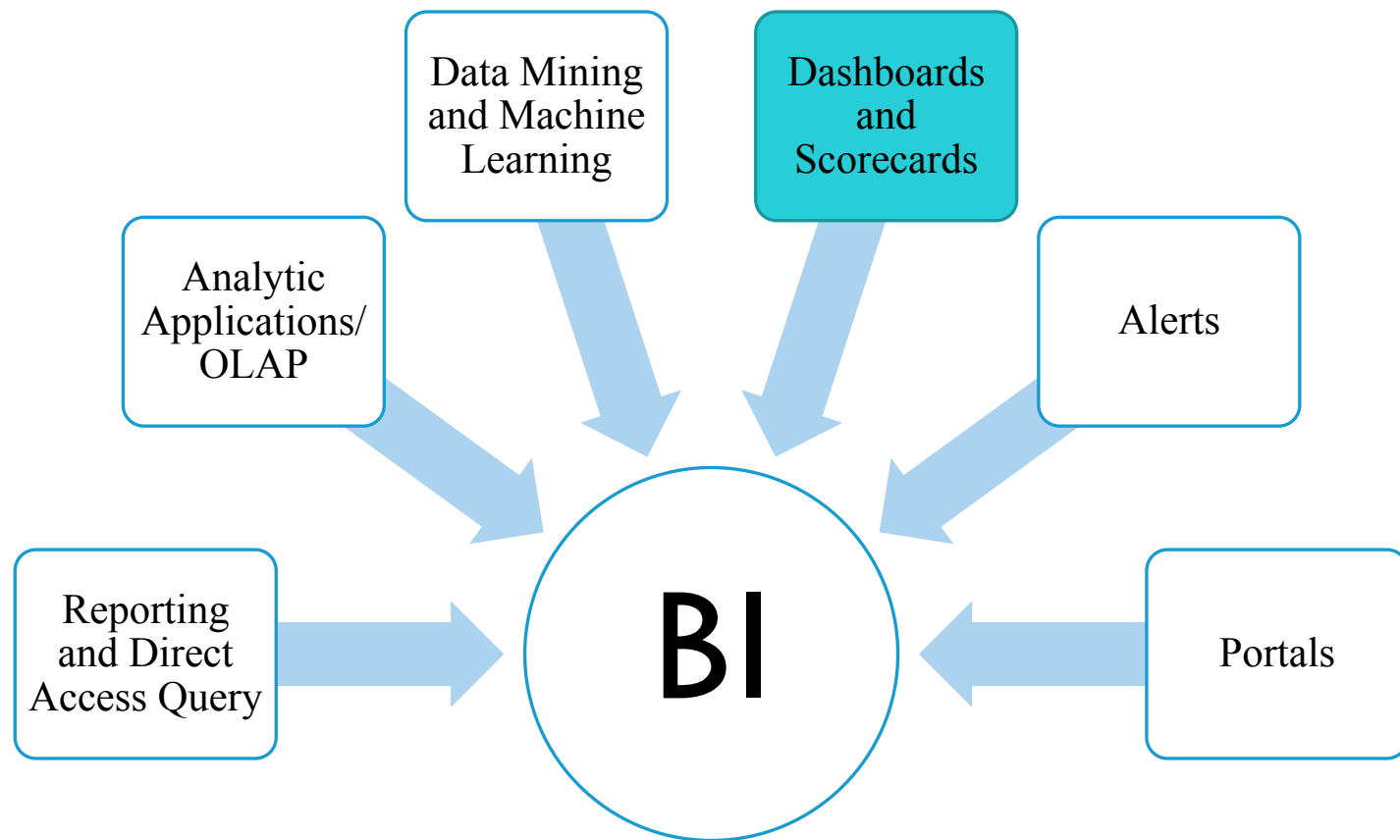
- What is the relationship between these two?
- We use **machine learning** algorithms to perform **data mining**.
- Machine learning is divided into **supervised** and **unsupervised** algorithms.
 - **Supervised** algorithms are trained using a feature set and labeled targets to predict future unknown values.
 - **Unsupervised** algorithms discover patterns or relationships in data from a feature set.



Dashboards and Scorecards

School of Information Studies
Syracuse University

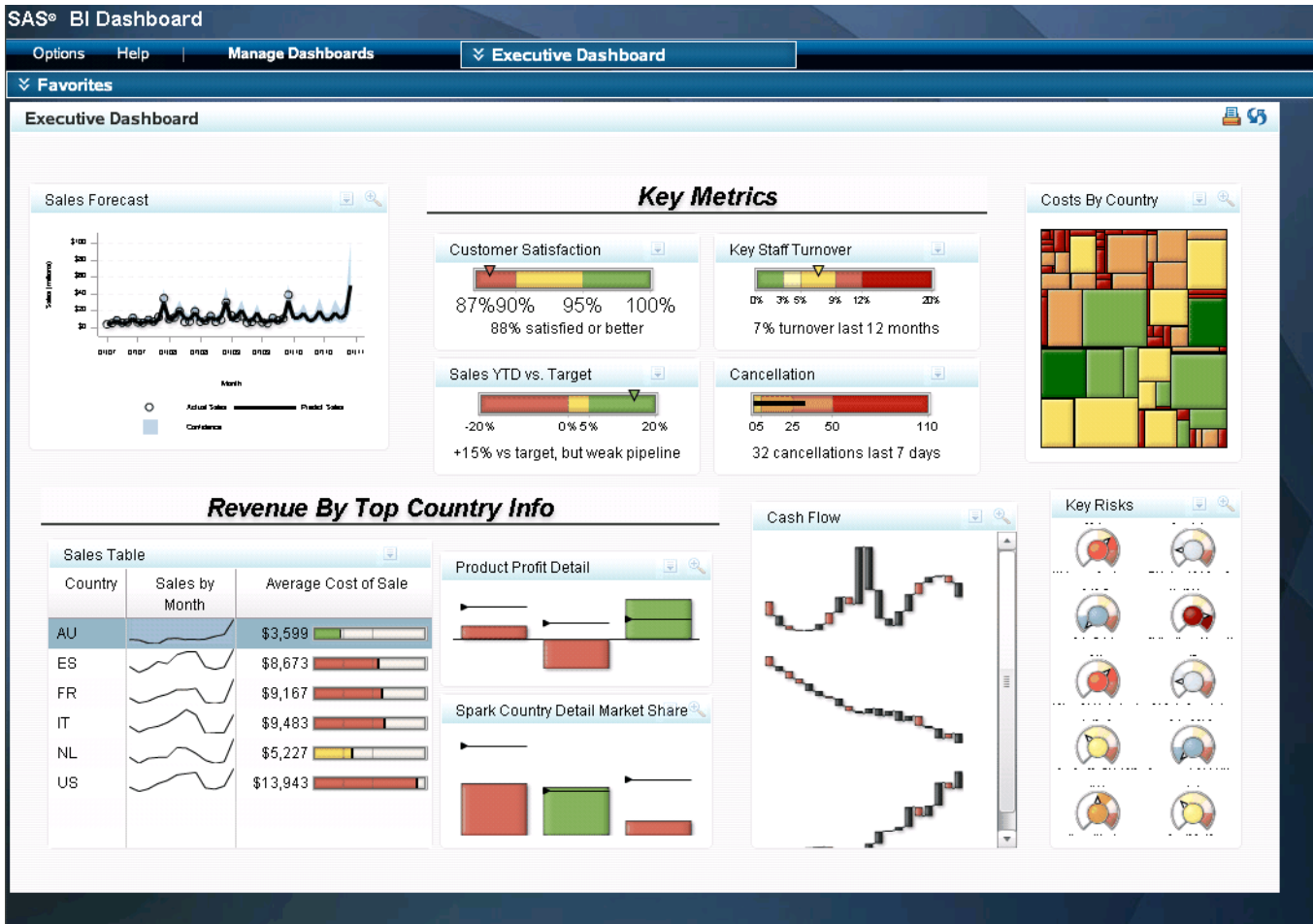
Six Categories of BI



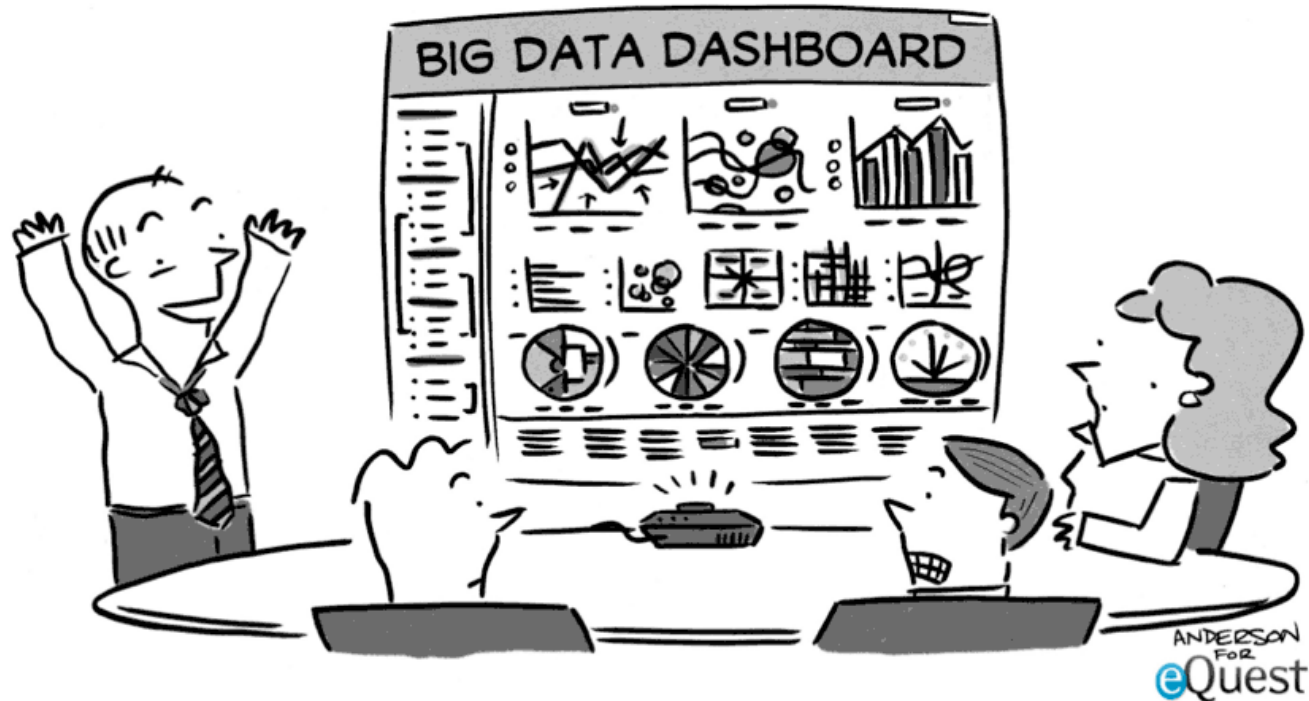
Dashboard and Scorecards

- Originally just executive interfaces, now organization-wide
- Provide a high-density and information-rich visual representation of data
- Usually web-based and highly interactive
- Consist of gauges, charts, and cards
- Contain KPIs (key performance indicators) for measuring goals
- As much of an organizational challenge as it is a technical one

Example: Dashboard



Just Because You Can Doesn't Mean You Should



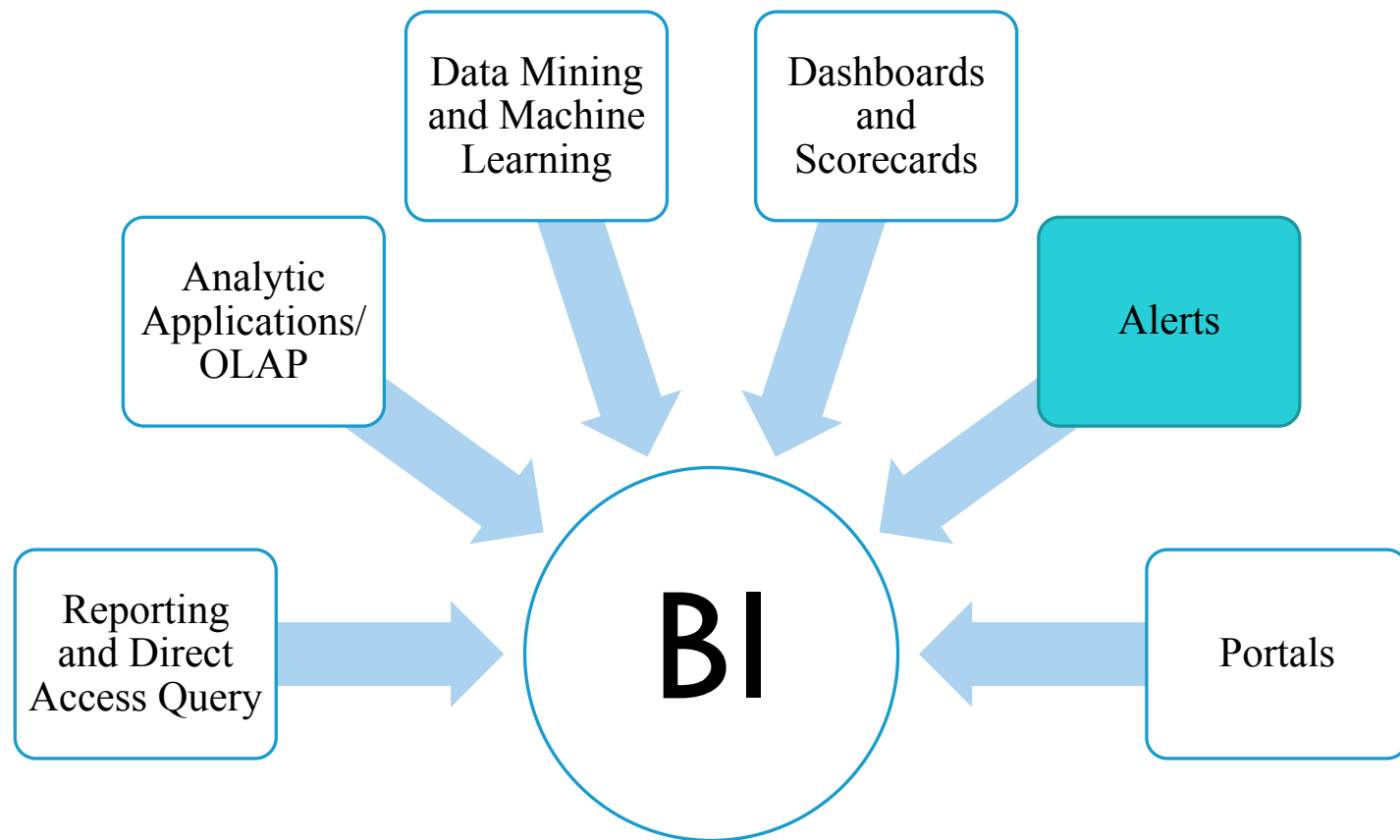
"After careful consideration of all 437 charts, graphs, and metrics, I've decided to throw up my hands, hit the liquor store, and get snookered. Who's with me?!"



Alerts

School of Information Studies
Syracuse University

Six Categories of BI



Alerts

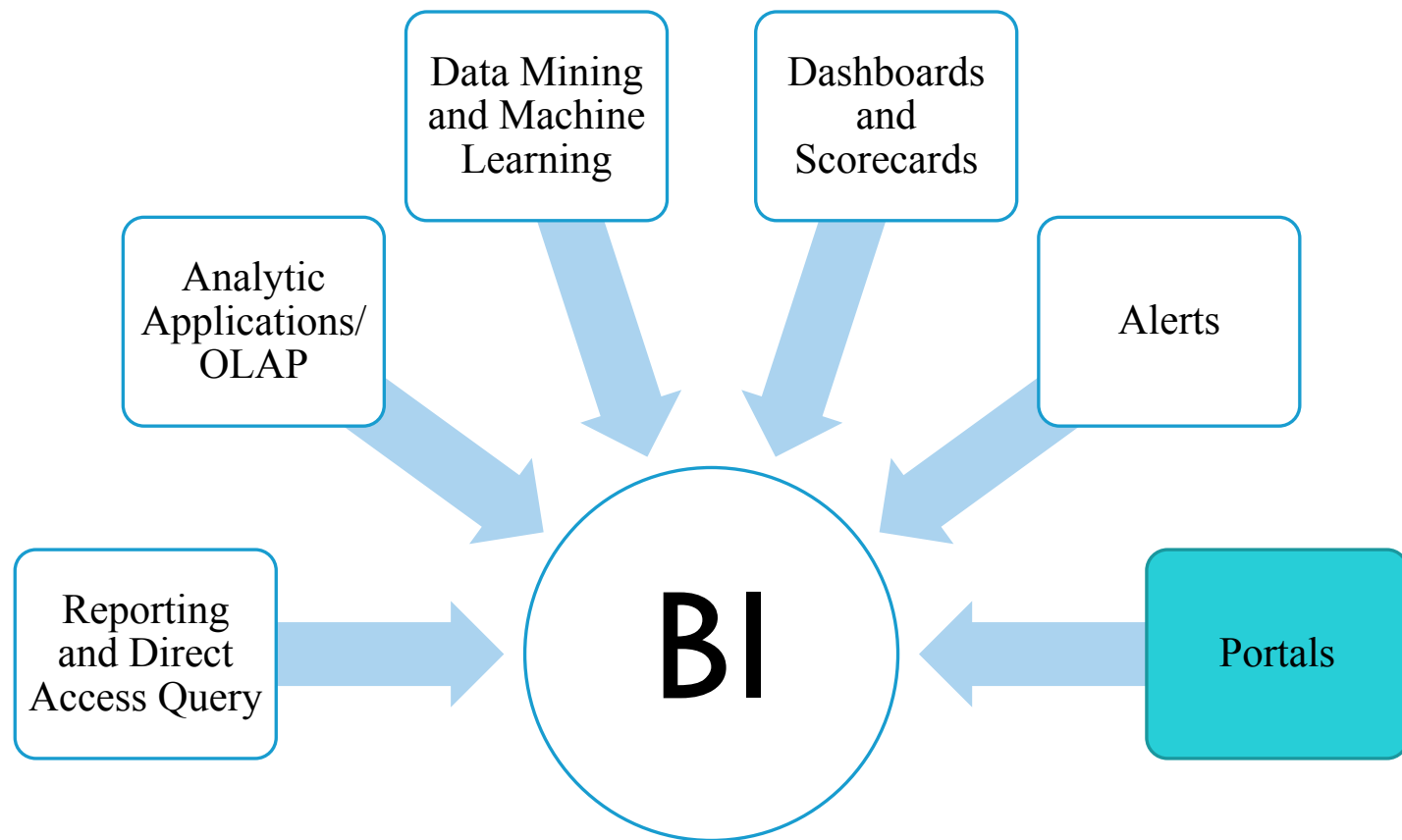
- Notifications to business users that an event or condition has happened
- Allow business users to react or respond
- Examples:
 - An inventory manager receives notification when a stock level drops too low.
 - The customer account representative receives a notification when one of his assigned customers completes the feedback form on the company's website.
 - When the daily sales figure for a retail store drops under a certain percentage of the target.



Portals

School of Information Studies
Syracuse University

Six Categories of BI



The BI Portal

- Provides access to BI applications and data.
- Security can be set up to restrict access.
- Makes it easier to navigate a large BI implementation.
- Requirements of any BI portal:
 - **Useable:** easy to find what you need
 - **Content rich:** information shown in context
 - **Clean:** simple design not overwhelming
 - **Current:** new content often
 - **Interactive:** browse data, customization for relevance
 - **Value-oriented:** users need to see value in it

Example: Portal

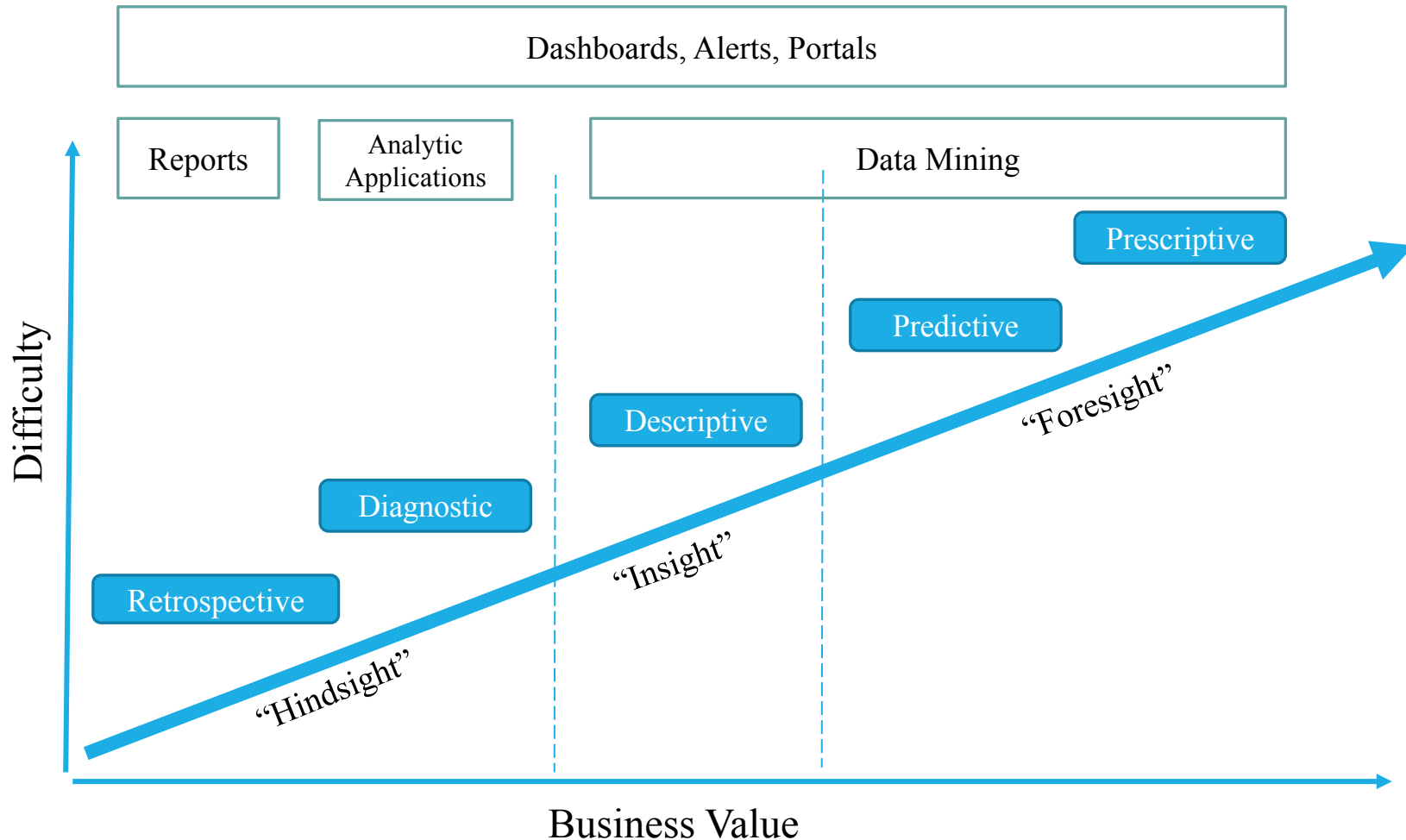




Revisiting the Five Types of Analytics

School of Information Studies
Syracuse University

Six Categories and Analytics





Three Types of BI

School of Information Studies
Syracuse University

Three Types of BI

	Operational BI	Tactical BI	Strategic BI
Business Focus	Manage daily operations; integrate BI with operational systems	Conduct short-term analysis to achieve strategic goals	Achieve long-term organizational goals
Purpose	Make immediate decision	Short-term decision-making	Long-term planning
Primary Users	Analysis, operational users	Executives, managers	Executives, managers
Data	Real-time metrics	Historical metrics	Historical metrics

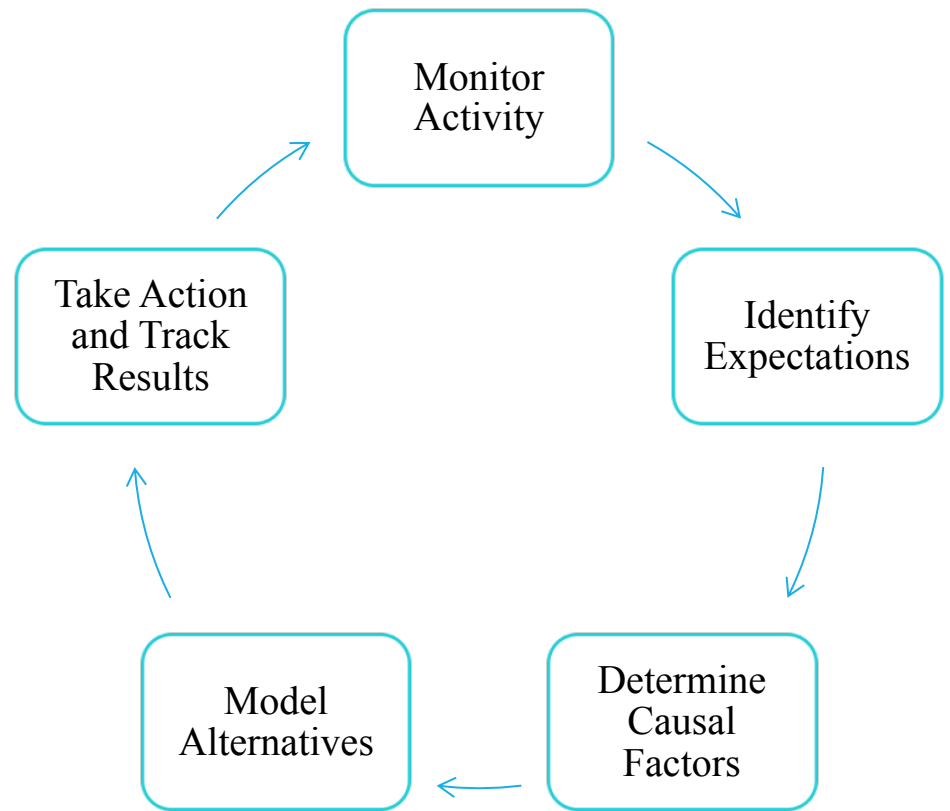


Analytic Cycle for BI

School of Information Studies
Syracuse University

Analytic Cycle for BI Analysis

- A model for BI development
- Improve BI through feedback of business users
- This helps us :
 - Understand how our users will use BI
 - Determine the tools we must provide to make their experience positive and productive





Multidimensional Database Management Systems

School of Information Studies
Syracuse University

What Is a Multidimensional Database?

- A multidimensional database is a storage mechanism for dimensional data. It consists of **dimensions** and **cubes**.
- Facts are stored in cube cells referenced by the atomic values of the dimension, similar to a multidimensional array.
- Aggregations are precalculated at higher levels in each dimension and stored in the cube cells, too.
- The underlying structure is very compact and fast compared to a relational star schema.
- The drawback being it takes longer to add/update data as the time to precalculate each aggregation can be time-intensive.

OLAP, ROLAP, HOLAP, MOLAP

- OLAP, or online analytical processing, is the act of exploring and interacting with dimensional data (facts and dimensions).
- OLAP can be done with relational data, multidimensional data, or both.

OLAP Type	DBMS	Data Stored in:
ROLAP	Relational	Star schema
MOLAP	Multidimensional	Cube
HOLAP	Hybrid (Both)	Data in star schema, aggregates in cube

Multidimensional Database Key Features

- Consolidate multiple data sources into a single cube
- Independent processing of facts and dimensions
- Semantic model:
 - Unknown members
 - Robust attribute properties
 - Define hierarchies
- Predefined aggregates (attribute relationships)
- Calculations/KPIs
- Perspectives

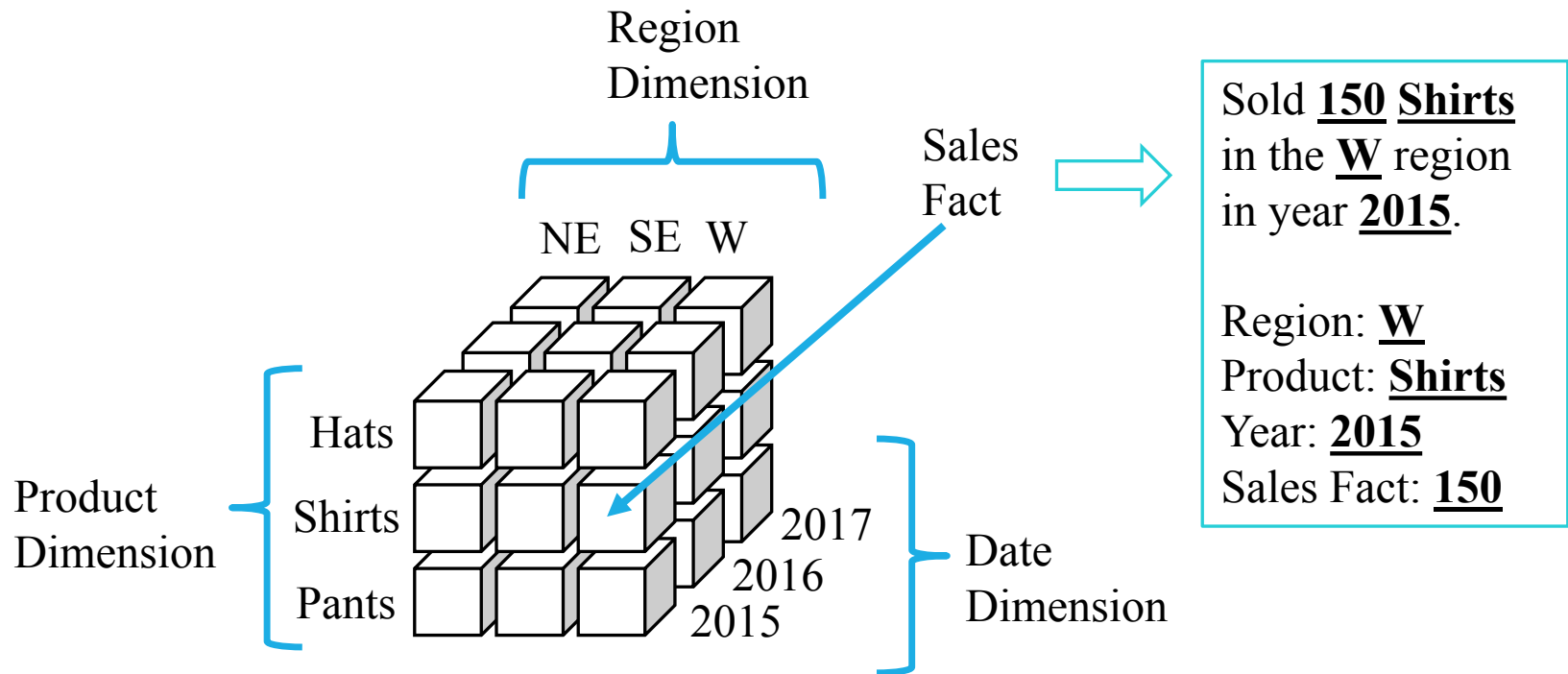


OLAP Terminology

School of Information Studies
Syracuse University

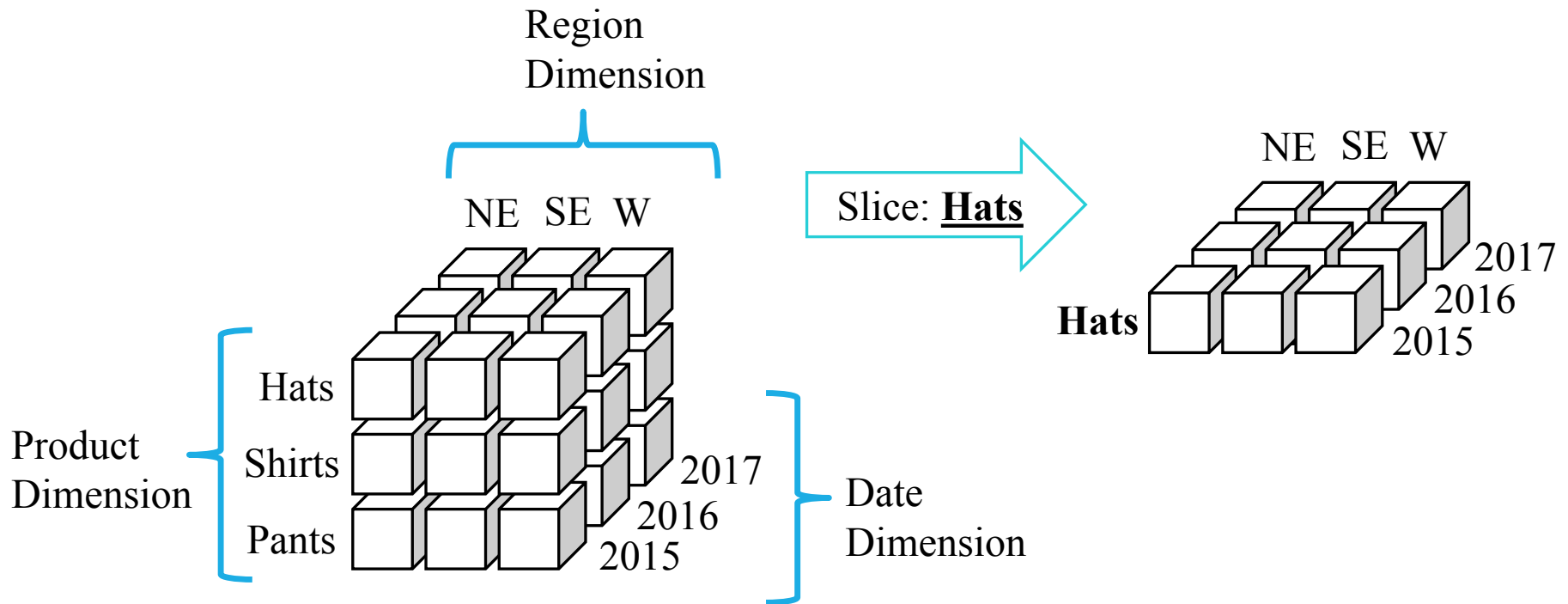
OLAP Cube

Data structure for representing analytic data.



OLAP Terminology: Slice

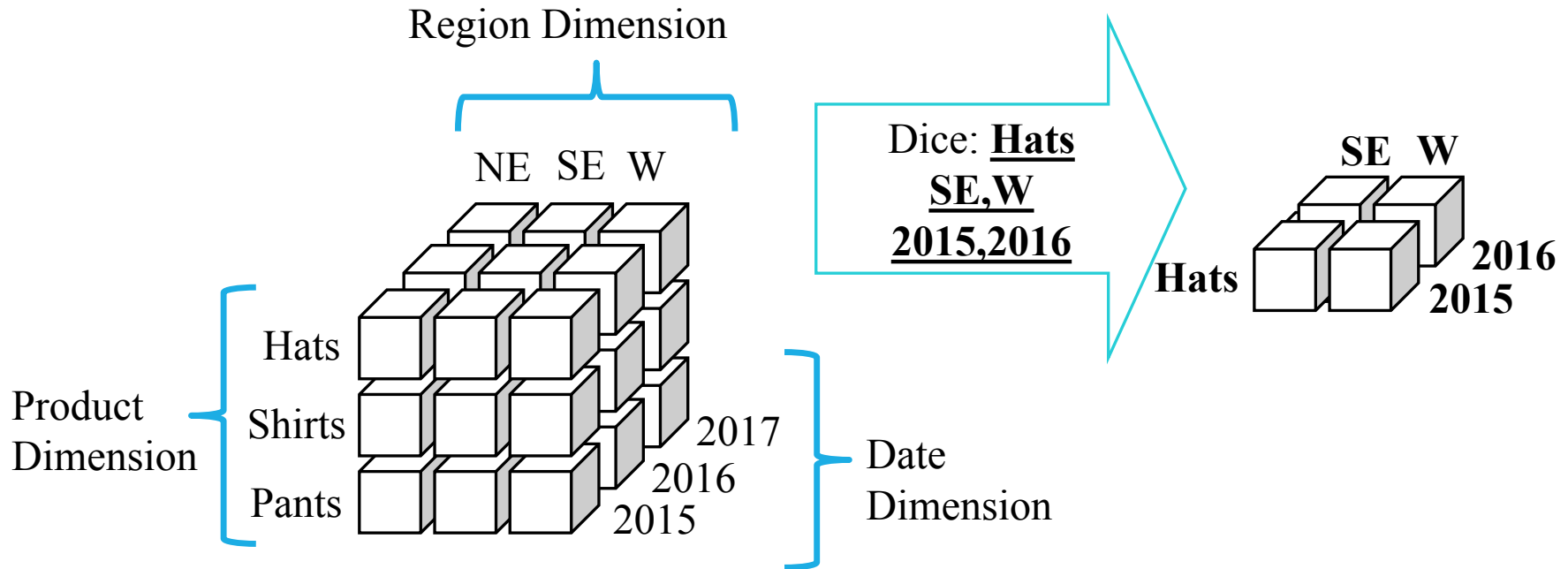
Slice: Filtering on one dimension



OLAP Terminology: Dice

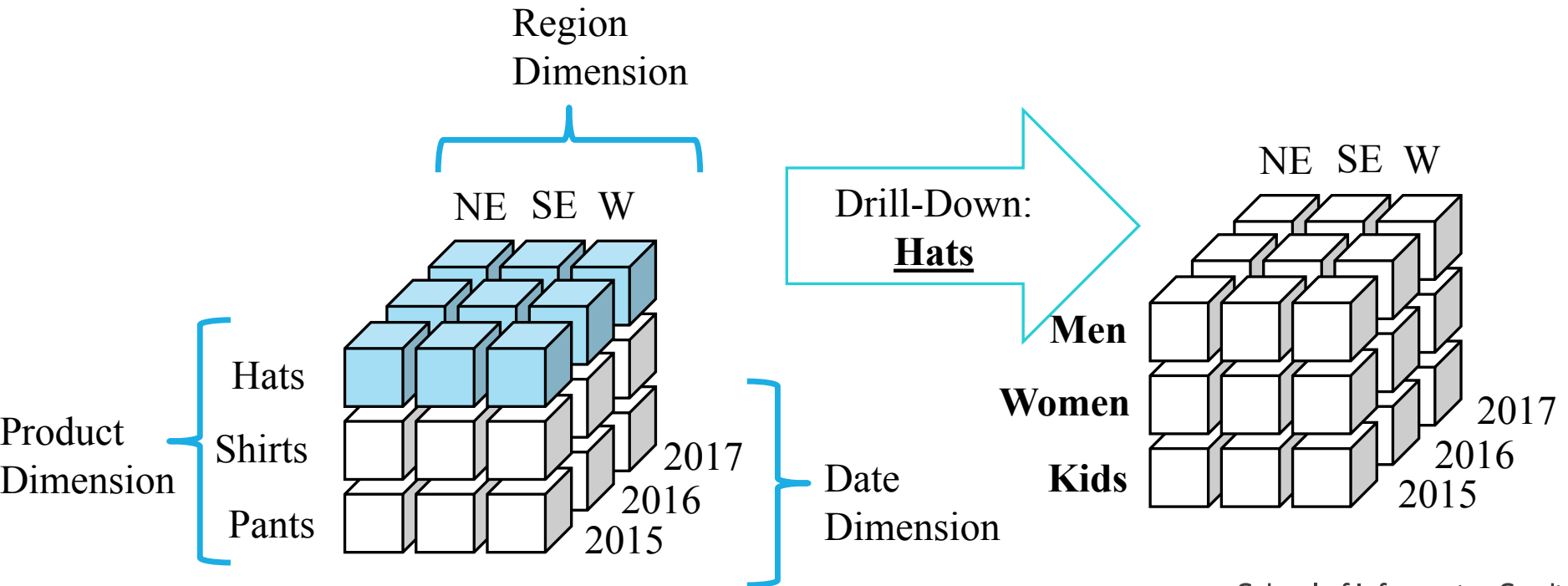
Dice: Filtering on more than one dimension to make a subset of the cube.

Perspective: A named dice of the cube



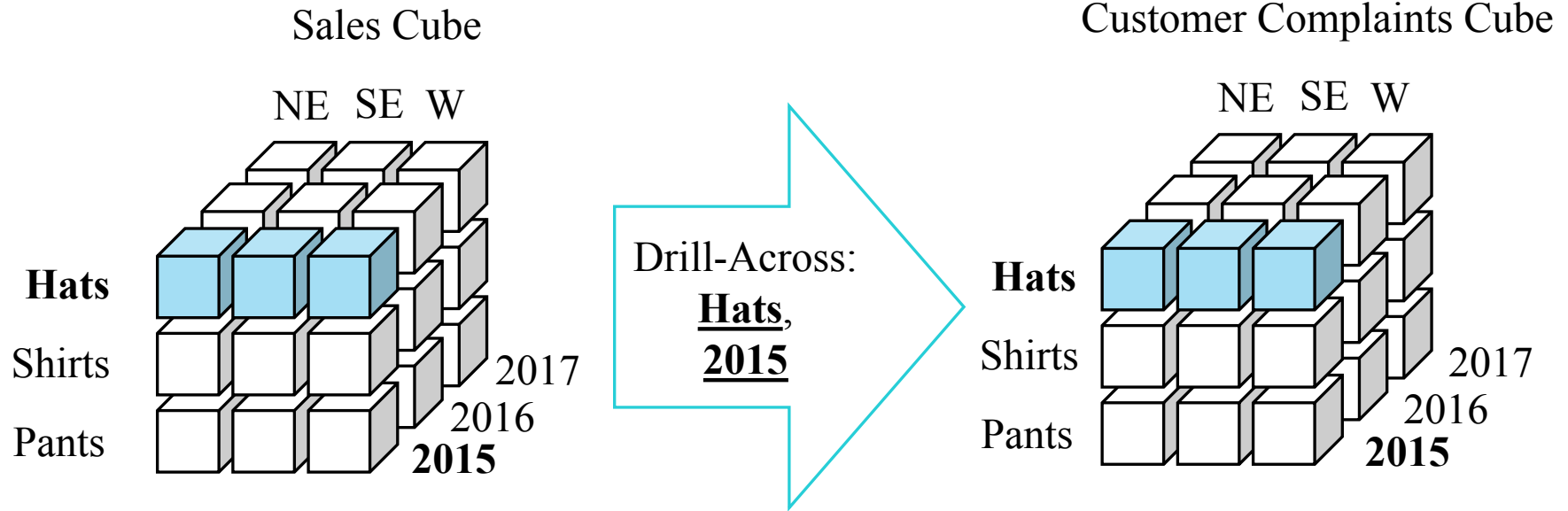
OLAP Terminology: Drill-Down/Up

- **Drill-down:** Moving down a level in a dimension hierarchy to view more detail
- **Drill-up:** Moving up a level in a dimension hierarchy



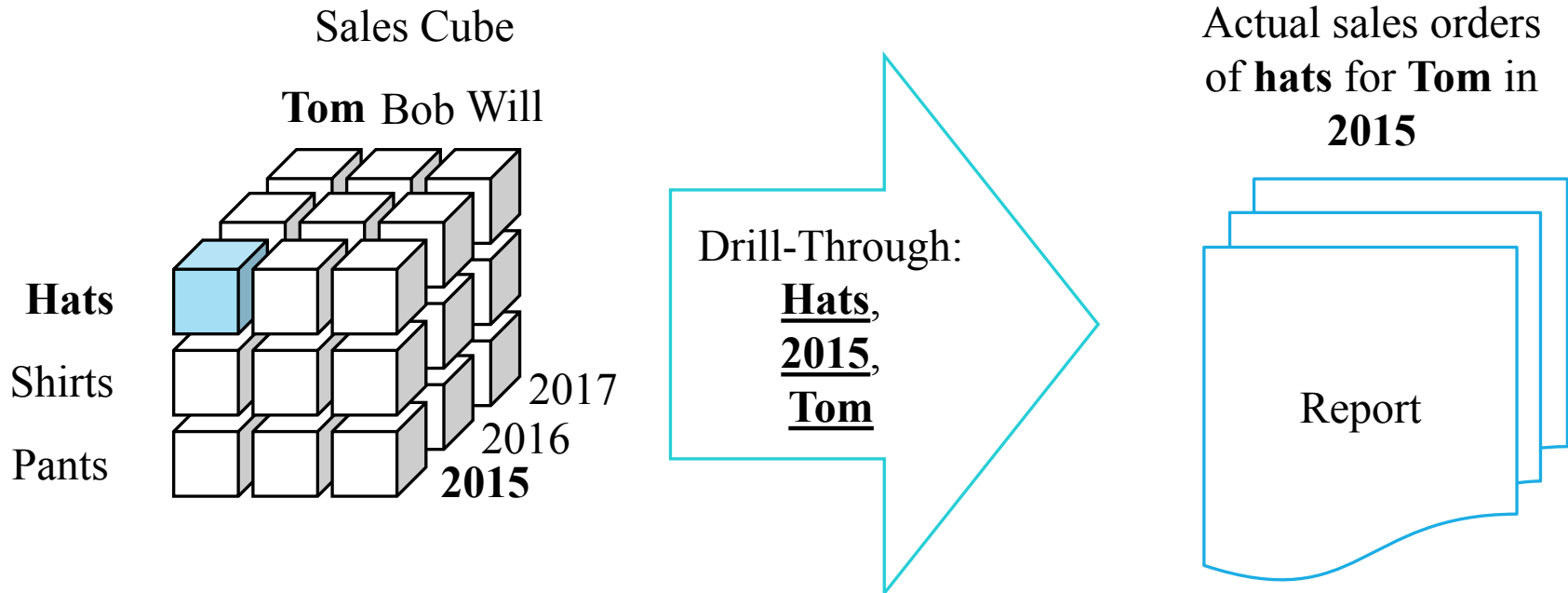
OLAP Terminology: Drill-Across

Drill-across: Use a sliced or diced dimension to view facts in a different cube.



OLAP Terminology: Drill-Through

Drill-through: Retrieve the OLTP detail of the actual event in the cube.

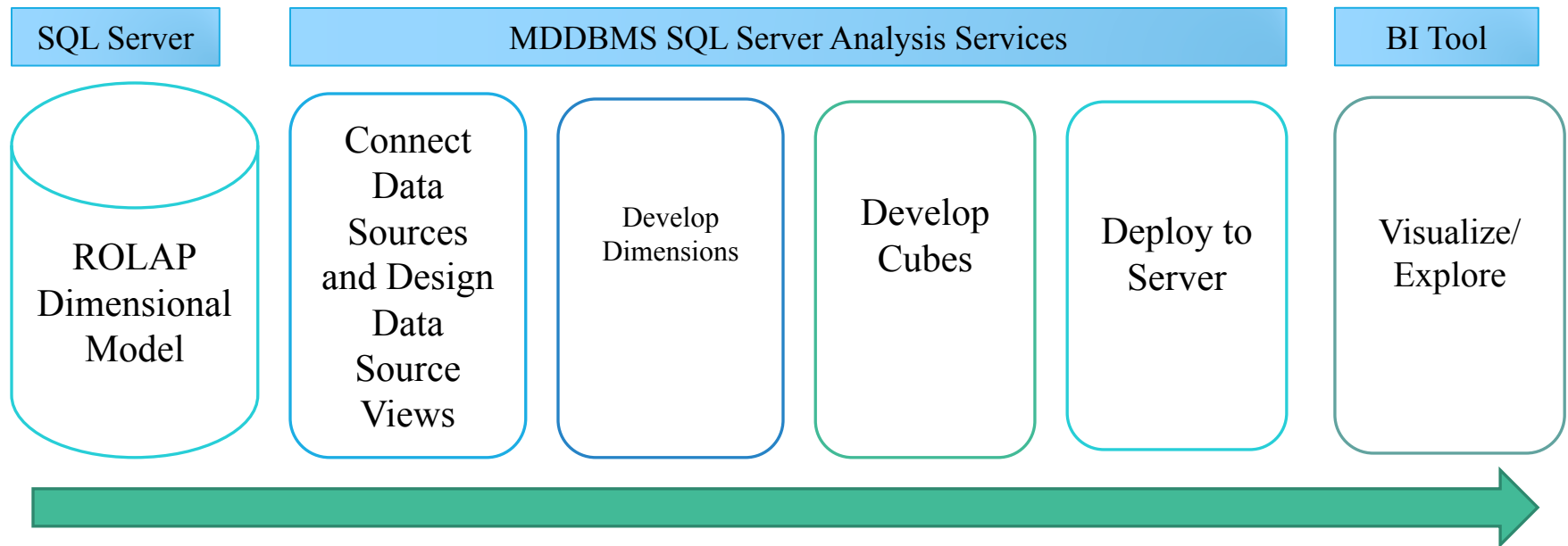




An Overview of the Process

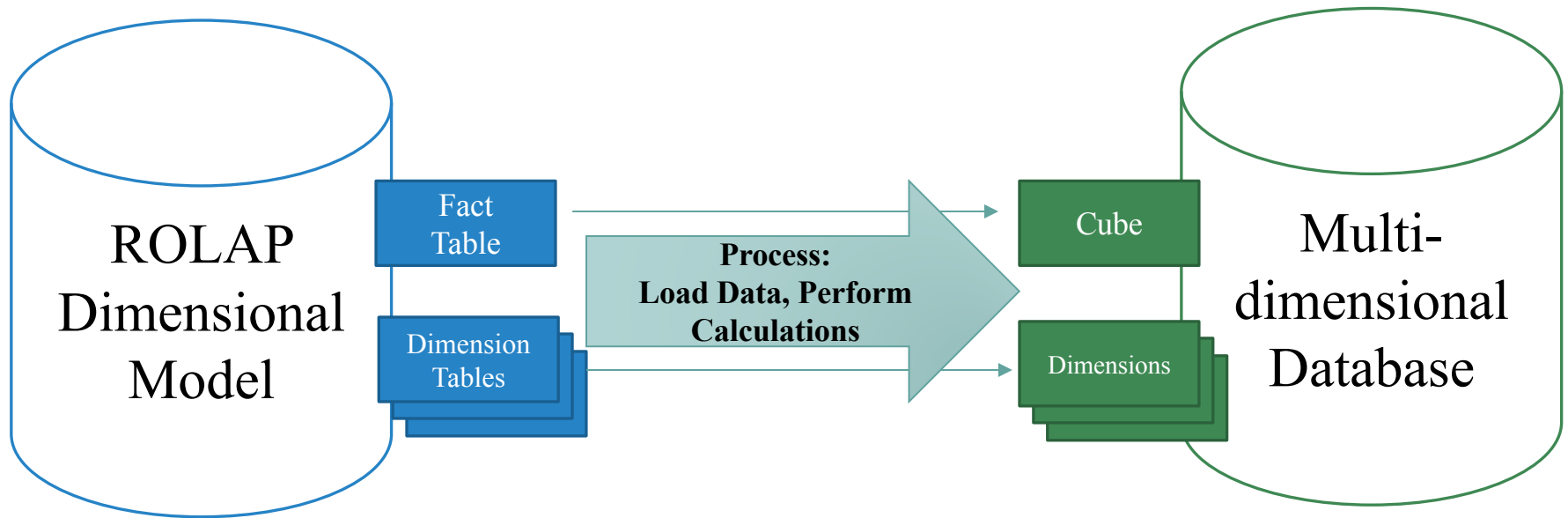
School of Information Studies
Syracuse University

The MOLAP Development Process



MOLAP: Processing

Before you can view changes in a MDDDBMS, you must first **process** the data.



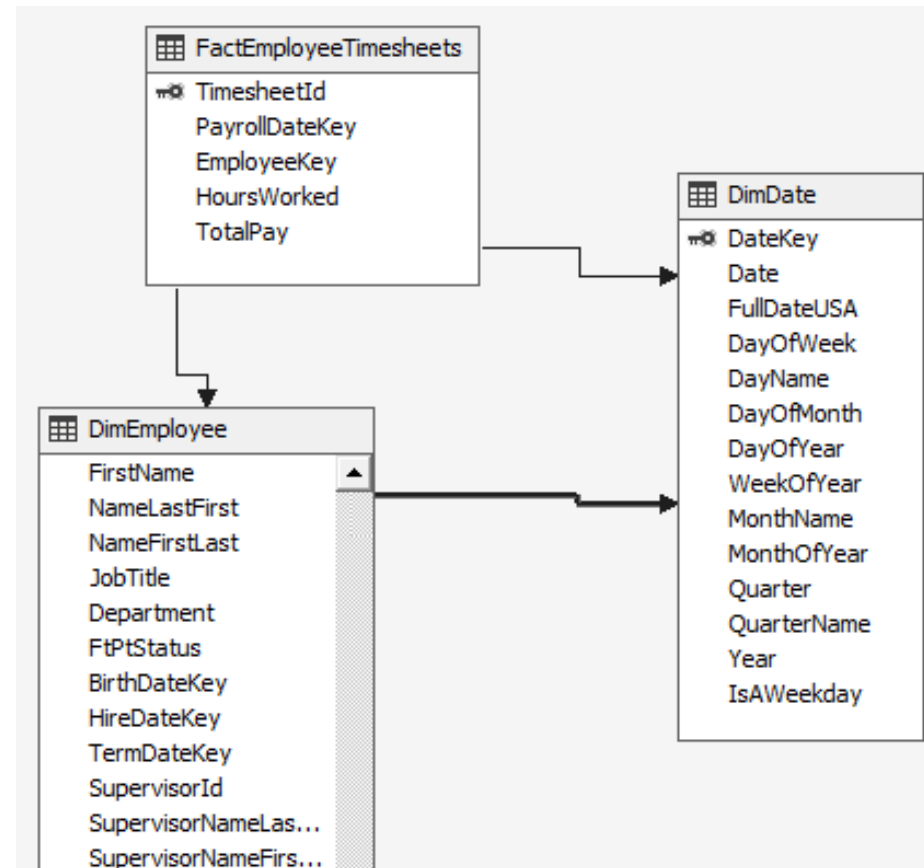


Connecting Data Sources and Creating Views in SSAS

School of Information Studies
Syracuse University

Connecting Data Sources and Views

- Connect to ROLAP sources.
- Create a data source view to combine tables from disparate sources.
- Derive calculated columns for values that do not exist.
- Establish relationships that are not in the ROLAP schema.





Building Dimensions

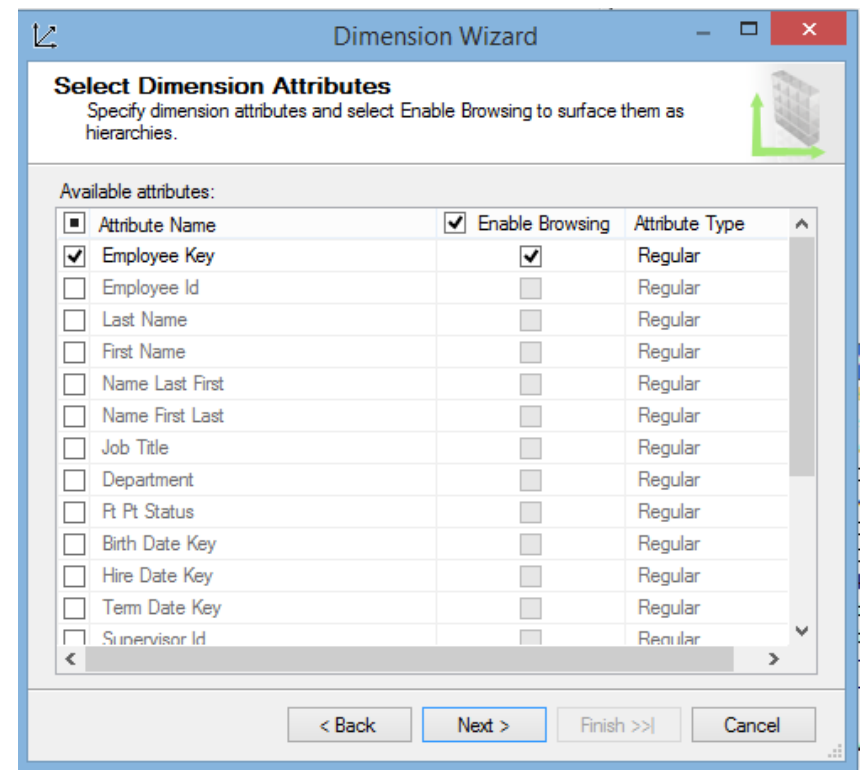
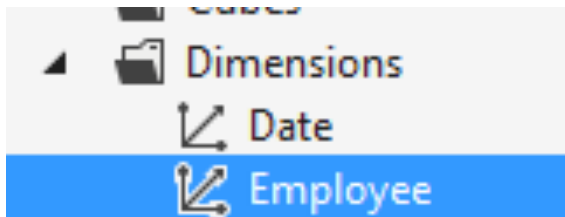
School of Information Studies
Syracuse University

Developing Dimensions: The Steps

1. Create dimension from source.
2. Add and configure attributes.
3. Configure hierarchies.

1. Creating Dimensions From Source

- Choose the ROLAP dimension from the data source view.
- Start with only the key attribute.
- Add the other attributes manually.
- Best practice: Drop “Dim” from the name.



2. Configuring Attributes

- We configure how we want the attributes presented to the user.
- Important properties
 - Name
 - Key column
 - Name column
 - Value column
 - Order by (key or name?)

Example: Month

- Key: month of year
- Value: month name
- Order by: key



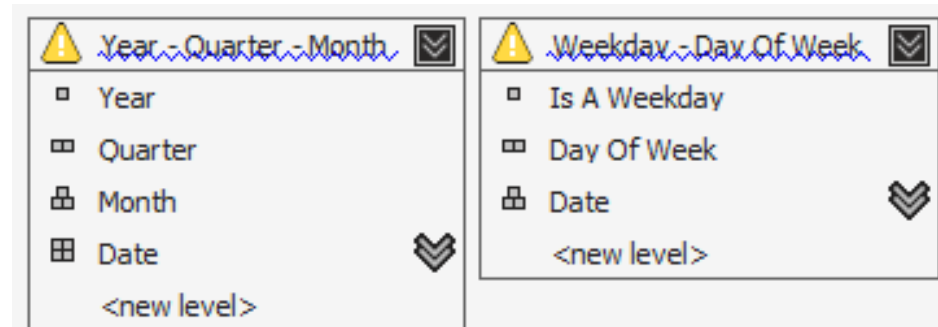
3. Configure Hierarchies

Allow us to drill-down through attributes.

Three types:

1. Natural: 1-M relationship among hierarchies
 - E.g., Year → Month → Day
2. Unnatural: no dependent relationship
 - E.g., Color → Size
3. Parent-child
 - E.g., Employee → Supervisor
 - Automatic when self-join exists in dimension.

Examples:





The Cube Development Process

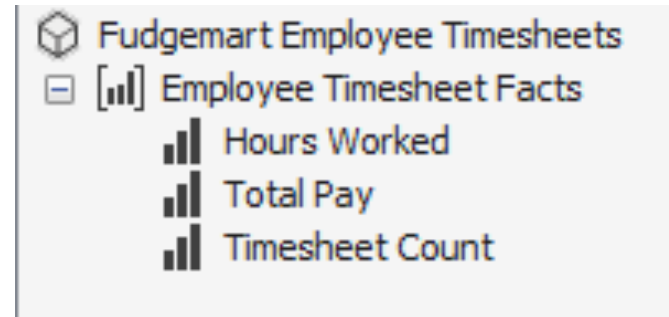
School of Information Studies
Syracuse University

Develop Cubes: The Steps

1. Select fact table: adds measures and determines dimensions.
2. Add and configure each measure (fact) properties.
3. Add calculations: business rules to the cube.
4. Add KPIs: key performance indicators.
5. Configure partitions: summary tables.
6. Configure perspectives: cube views.

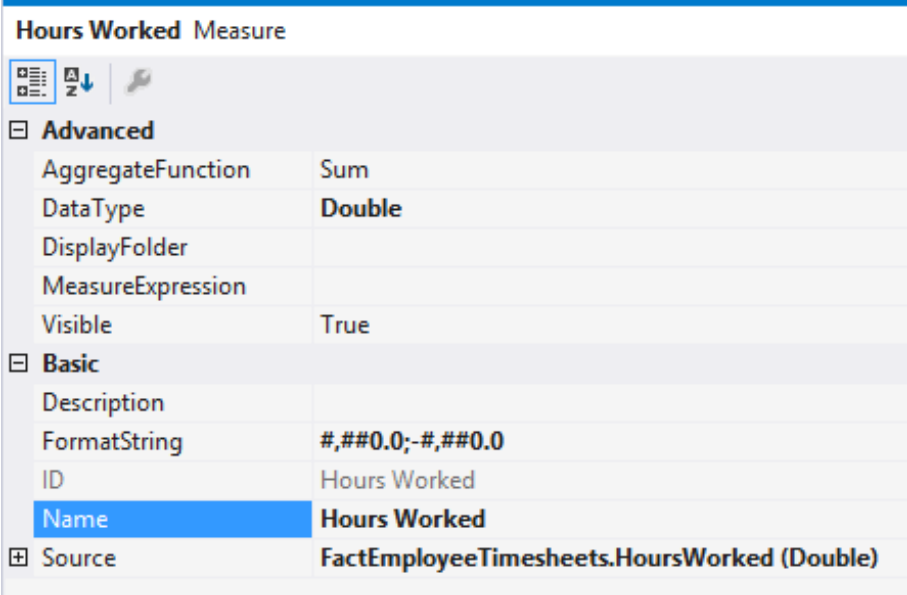
1. Select Fact Table


- To build a cube you must include at least one fact table. This initiates the cube-building process.
- You can have more than one fact table in the cube.
- This allows you to consolidate logic into a single point of access for the user.
- Each fact grain will contain the same measure group.



2. Configuring Measure Properties

- We configure how we want the measures presented to the user.
- Important properties:
 - Name
 - Format string
 - Aggregate function: sum, count, min/max, none
 - Visibility: show/hide
 - Measure group



Hours Worked Measure	
	
Advanced	
AggregateFunction	Sum
DataType	Double
DisplayFolder	
MeasureExpression	
Visible	True
Basic	
Description	
FormatString	###0.0;-###0.0
ID	Hours Worked
Name	Hours Worked
Source	FactEmployeeTimesheets.HoursWorked (Double)

3. Add Calculations

- Additional calculations based on current measures.
- Function builder to help you out.
- These are MDX expressions. MDX is the SQL of multidimensional databases.

Name: [Overtime Hours]

Parent Properties

Parent hierarchy: Measures

Parent member: [Change]

Expression

IIF([Measures].[Hours Worked]>40,[Measures].[Hours Worked]-40,null)

Additional Properties

Format string: "Standard"

Visible: True

Non-empty behavior: []

Associated measure group: Employee Timesheet Facts

Display folder: []

4. KPIs—Key Performance Indicators

KPIs are metrics defined by the organization to measure performance.

You need:

1. Value: What you're measuring.
2. Goal: What you want to achieve.
3. Status: How close is the value to the goal? Range between -1 and 1 .
4. Trend: Status over time periods. Range between -1 and 1 .

Example:

1. Value: overtime hours
2. Goal: 0
3. Status:
 - When overtime = 0 then 1
 - When overtime ≥ 10 then -1
 - Else 0
4. Trend:
 - When overtime < last time period overtime, then 1, etc.

5. Partitions

- Filter cube data rows based on an expression; to make smaller cubes
- Improves cube performance
- Example:
 - Partition by date so current FY is in a single partition
 - Partition by store so each branch has its data

6. Perspectives

- Reduce complexity of cubes by eliminating irrelevant facts and dimension attributes
- Similar to SQL views in relational databases
- Example:
 - Single physical cube with sales and projections can be simplified into two perspectives sales and projects



Visualizing Your Cube

School of Information Studies
Syracuse University

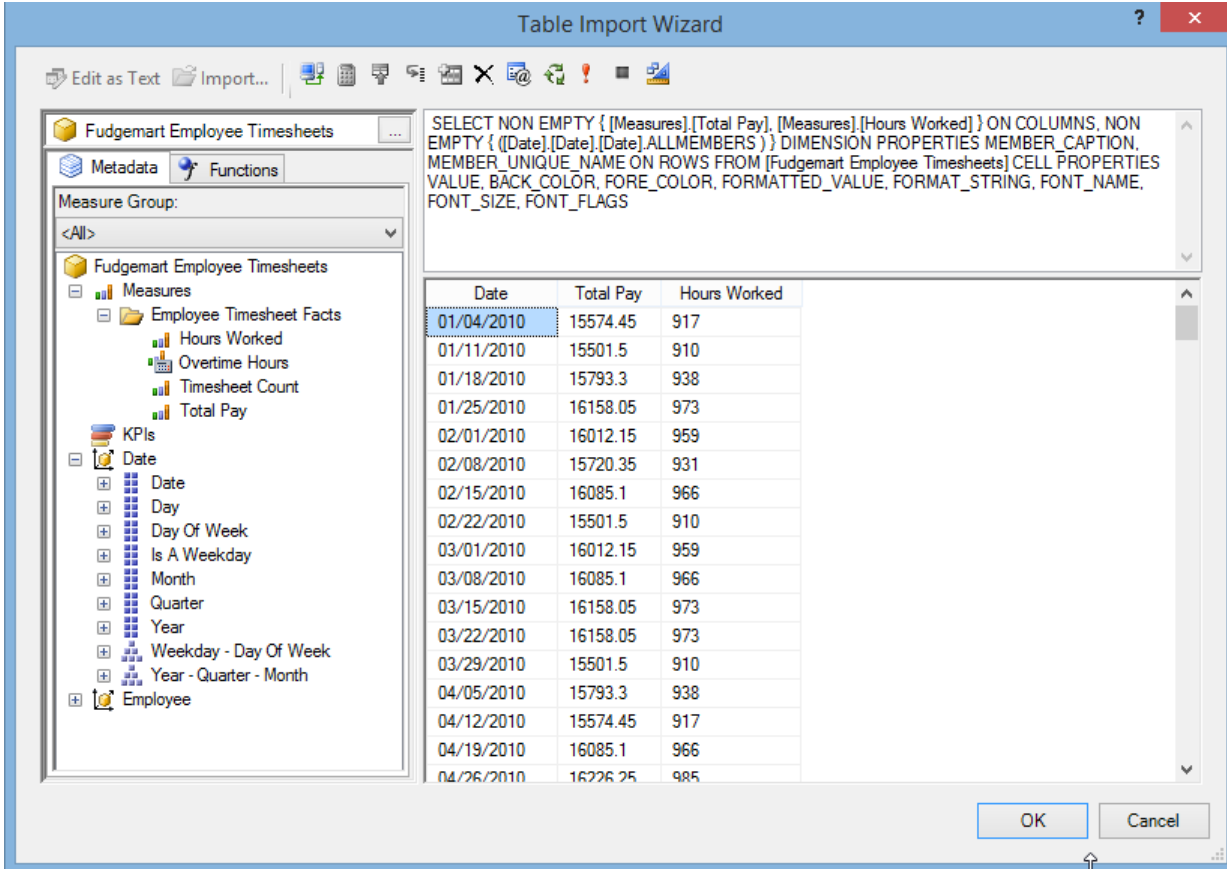
Visualizing Cubes

- SSAS Cube Browser
 - During cube development
- Excel Pivot Table
 - For power users and cube testing
- Power BI or other BI tool
 - For business users and BI testing

MDX Query

MDX is the SQL for MOLAP databases.

Allows you to query your cube and retrieve data in tabular form.



The screenshot shows the 'Table Import Wizard' window. The title bar is 'Table Import Wizard'. The window has a toolbar with icons for 'Edit as Text', 'Import...', and other functions. The main area is divided into two panes. The left pane shows a tree view of the 'Fudgemart Employee Timesheets' cube. The right pane shows the MDX query and a preview of the data.

MDX Query:

```
SELECT NON EMPTY { [Measures].[Total Pay], [Measures].[Hours Worked] } ON COLUMNS, NON  
EMPTY { ([Date].[Date].[Date].ALLMEMBERS) } DIMENSION PROPERTIES MEMBER_CAPTION,  
MEMBER_UNIQUE_NAME ON ROWS FROM [Fudgemart Employee Timesheets] CELL PROPERTIES  
VALUE, BACK_COLOR, FORE_COLOR, FORMATTED_VALUE, FORMAT_STRING, FONT_NAME,  
FONT_SIZE, FONT_FLAGS
```

Data Table:

Date	Total Pay	Hours Worked
01/04/2010	15574.45	917
01/11/2010	15501.5	910
01/18/2010	15793.3	938
01/25/2010	16158.05	973
02/01/2010	16012.15	959
02/08/2010	15720.35	931
02/15/2010	16085.1	966
02/22/2010	15501.5	910
03/01/2010	16012.15	959
03/08/2010	16085.1	966
03/15/2010	16158.05	973
03/22/2010	16158.05	973
03/29/2010	15501.5	910
04/05/2010	15793.3	938
04/12/2010	15574.45	917
04/19/2010	16085.1	966
04/26/2010	16226.25	985

Example: Cube Browser

Northwind Daily In...y Snapshot [Online] Northwind Sales [Online] X

Cube Structure Dimension Usage Calculations KPIs Actions Partitions Aggregations Perspectives Translations **Browser**

Language: Default X

Edit as Text Import...

Northwind Sales

Metadata

<All>

Northwind Sales

- Measures
 - Sales
 - Discount Amount
 - Extended Price Amount
 - Order To Shipped Lag In Days
 - Quantity
 - Sales Count
 - Sold Amount
- KPIs
 - Customer
 - Employee
 - Order Date
 - Product
 - Shipped Date

Dimension	Hierarchy	Operator
Customer	Contacts	Equal
<Select dimension>		

Day Of Week	Sold Amount	Discount Amount	Extended Price Amount
Monday	54027.59	2871.5	56899.09
Tuesday	48357.6199	3316.02	51673.64
Wednesday	38088.285	2965.295	41053.58
Thursday	22190.8	2273	24463.8
Friday	66682.1625	2487.4475	69169.61

Pivot Tables

Use SSAS as a data source.

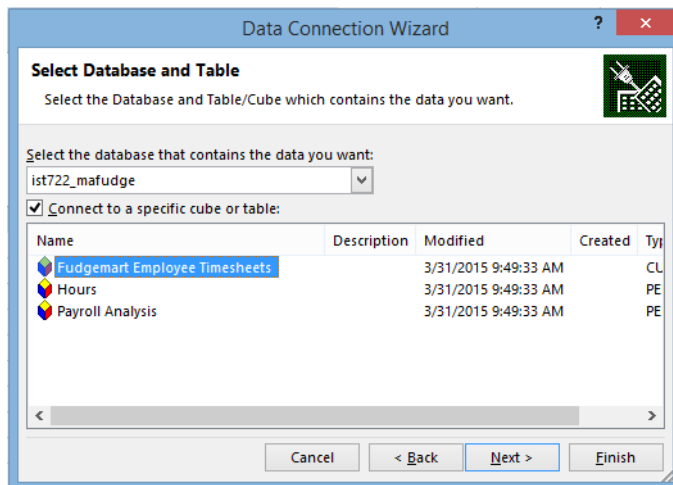
BI semantic model:

- Understands measures, data types, and hierarchies

PivotTable Fields

Choose fields to add to report:

- ☒ **Employee Timesheet Facts**
 - ☐ Hours Worked
 - ☐ Overtime Hours
 - ☐ Timesheet Count
 - ☒ **Total Pay**
- ☒ **Date**
 - ☐ Weekday - Day Of Week
 - ☒ **Year - Quarter - Month**
 - ☐ More Fields
- ☒ **Employee**
 - ☒ **Department - Employee**
 - ☐ Job Title - Employee
 - ☐ Status - Employee
 - ☐ Supervisor - Employee
 - ☐ More Fields



Total Pay	Column Labels				
Row Labels	2010	2011	2012	2013	Grand Total
Clothing	\$86,783.60	\$94,996.43	\$100,468.44	\$99,744.19	\$381,992.66
Customer Service	\$302,224.00	\$305,382.90	\$320,597.21	\$315,682.45	\$1,243,886.56
Electronics	\$127,192.00	\$134,451.83	\$148,067.41	\$146,442.40	\$556,153.64
Hardware	\$95,295.05	\$114,229.65	\$119,920.56	\$118,082.34	\$447,527.59
Housewares	\$113,744.00	\$115,833.11	\$122,052.11	\$120,675.38	\$472,304.60
Sporting Goods	\$111,047.50	\$109,174.16	\$132,897.30	\$132,777.07	\$485,896.03
Grand Total	\$836,286.15	\$874,068.06	\$944,003.02	\$933,403.84	\$3,587,761.08