



**David Eccles
School of Business**

THE UNIVERSITY OF UTAH

Association Rule Mining

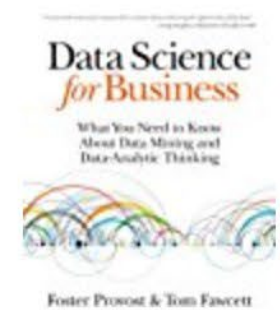
Data Mining for Business Intelligence

AGENDA

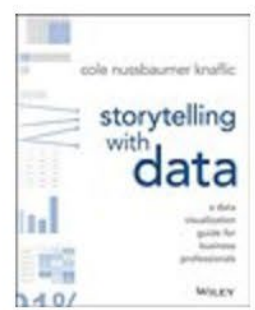
- ▶ Motivating examples and main approach
- ▶ Set basics and important metrics
 - ▶ Support, confidence and lift
- ▶ Apriori property and Apriori algorithm
- ▶ Data preparation decisions
- ▶ Extension of association rules

FROM “DIAPERS AND BEER ” TO

Frequently Bought Together



+



Total price: **\$60.77**

Add both to Cart

Add both to List

- ☒ **This item:** Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking by Foster Provost Paperback **\$33.27**
- ☒ **Storytelling with Data: A Data Visualization Guide for Business Professionals** by Cole Nussbaumer Knaflic Paperback **\$27.50**

Customers Who Bought This Item Also Bought

Page 1 of 17



Data Smart: Using Data Science to Transform Information into Insight

› John W. Foreman

★★★★☆ 105

#1 Best Seller in Computer Simulation

Paperback

\$31.99 ✓Prime



Storytelling with Data: A Data Visualization Guide for Business Professionals

› Cole Nussbaumer...

★★★★☆ 113

Paperback

\$27.50 ✓Prime



Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die

› Eric Siegel

★★★★☆ 283

Paperback

\$13.07 ✓Prime



Doing Data Science: Straight Talk from the Frontline

› Cathy O'Neil

★★★★☆ 50

Paperback

\$29.59 ✓Prime



Data Science from Scratch: First Principles with Python

› Joel Grus

★★★★☆ 65

#1 Best Seller in Computer Programming...

Paperback

\$31.72 ✓Prime



Big Data: A Revolution That Will Transform How We Live, Work, and Think

› Viktor...

★★★★☆ 385

Paperback

\$9.00 ✓Prime

1 Ref: <http://www.dssresources.com/newsletters/66.php>

Information Systems

Matt Pecsok

ASSOCIATION

Under a given condition (or at an event):

A set of objects \rightarrow (implies) another set of objects



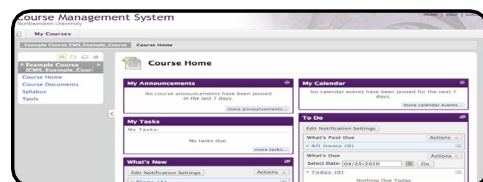
Retail items purchased together



Services subscribed by a customer



Web pages accessed by a user in a session



Courses taken by a student



Medications prescribed by a doctor



Genes expressed at the same level

ASSOCIATION

Customers who bought

A set of items → (tend to buy) another set of items



Retail items purchased together

ASSOCIATION

Customers who access (visit)

A set of pages → (tend to visit) another set of pages



Retail items purchased together



Web pages accessed by a user in a session

ASSOCIATION

From cancer patients' tissue

A set of expressed genes → (imply) another set of genes are expressed



Retail items purchased together



Web pages accessed by a user in a session



Genes expressed at the same level

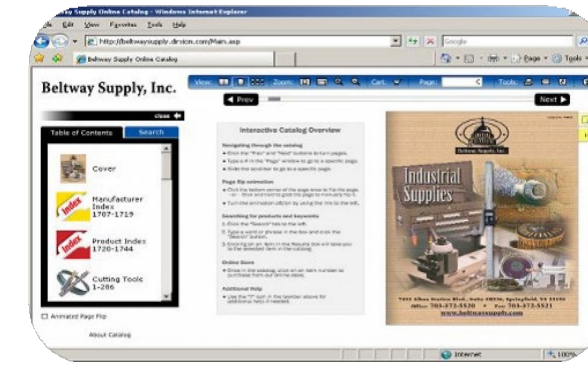
DECISION SUPPORT APPLICATIONS



Customer relationship management



Retail merchandise placement



Online retail catalog design



Website link reorganization



Fraud detection



Gene analysis for cancer prevention

ASSOCIATION RULES

Given: A database of transactions

Example of transactions:



- A customer's purchase at a grocery store



- An online order at an e-tailer such as 'Amazon.com'

ASSOCIATION RULES MINING

- ▶ **Search for:** Association rules (patterns), $X \rightarrow Y$
- ▶ **Strength of association rule**
 - ▶ Correlate the presence (existence) of one set of items (X) with the presence (predict) of another set of items (Y)
- ▶ **Common correlation measure**
 - ▶ Conditional probability $P(Y|X)$
- ▶ **Challenge:** Too many patterns, time consuming searches
- ▶ **Approach**
 - ▶ Find interesting rules with sufficient strengths and presence



David Eccles
School of Business

THE UNIVERSITY OF UTAH