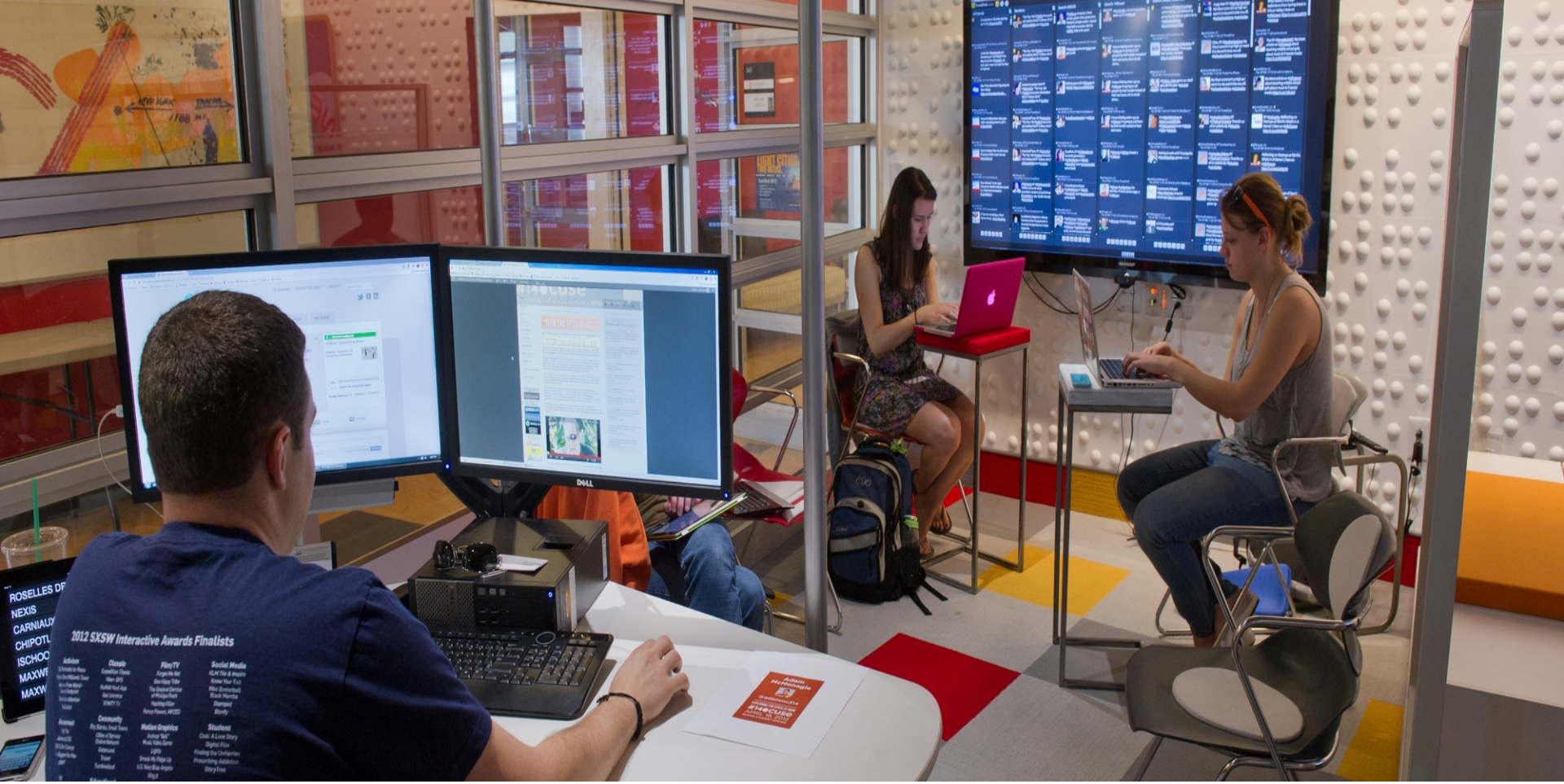


ASSOCIATION RULE MINING

SYRACUSE UNIVERSITY
School of Information Studies

ASSOCIATION RULE (AR) MINING

Given a set of transactions, find:

Items that co-occur frequently

Rules such as "if a customer bought *x*, he or she would buy *y*, too"

TID	*Items*
1	Bread, Milk
2	Bread, Diaper, Beer, Eggs
3	Milk, Diaper, Beer, Coke
4	Bread, Milk, Diaper, Beer
5	Bread, Milk, Diaper, Coke

Strong rules

{Milk} --> {Coke}

{Diaper, Milk} --> {Beer}

FREQUENT ITEMSETS

Itemset:

A collection of one or more items

k-itemset contains k items

1-itemset:

{A}:3, {B}:3, {C}:2, {D}:4, {E}:3, {F}:2

2-itemset:

{A,B}:1, {A,D}:3

3-itemset:

{A,B,C}:0, {B,E,F}:2

Transaction ID	Items Bought
10	A, B, D
20	A, C, D
30	A, D, E
40	B, E, F
50	B, C, D, E, F

Frequently Bought Together

This item: The Manga Guide to Database

The Manga Guide to Statistics by Shin Taka

The Manga Guide to Linear Algebra by Shi

ASSOCIATION RULES

Association rule:

An implication of the form X → Y, where X and Y are itemsets

E.g., {E, F} → {B}

LHS: Left-Hand Side →

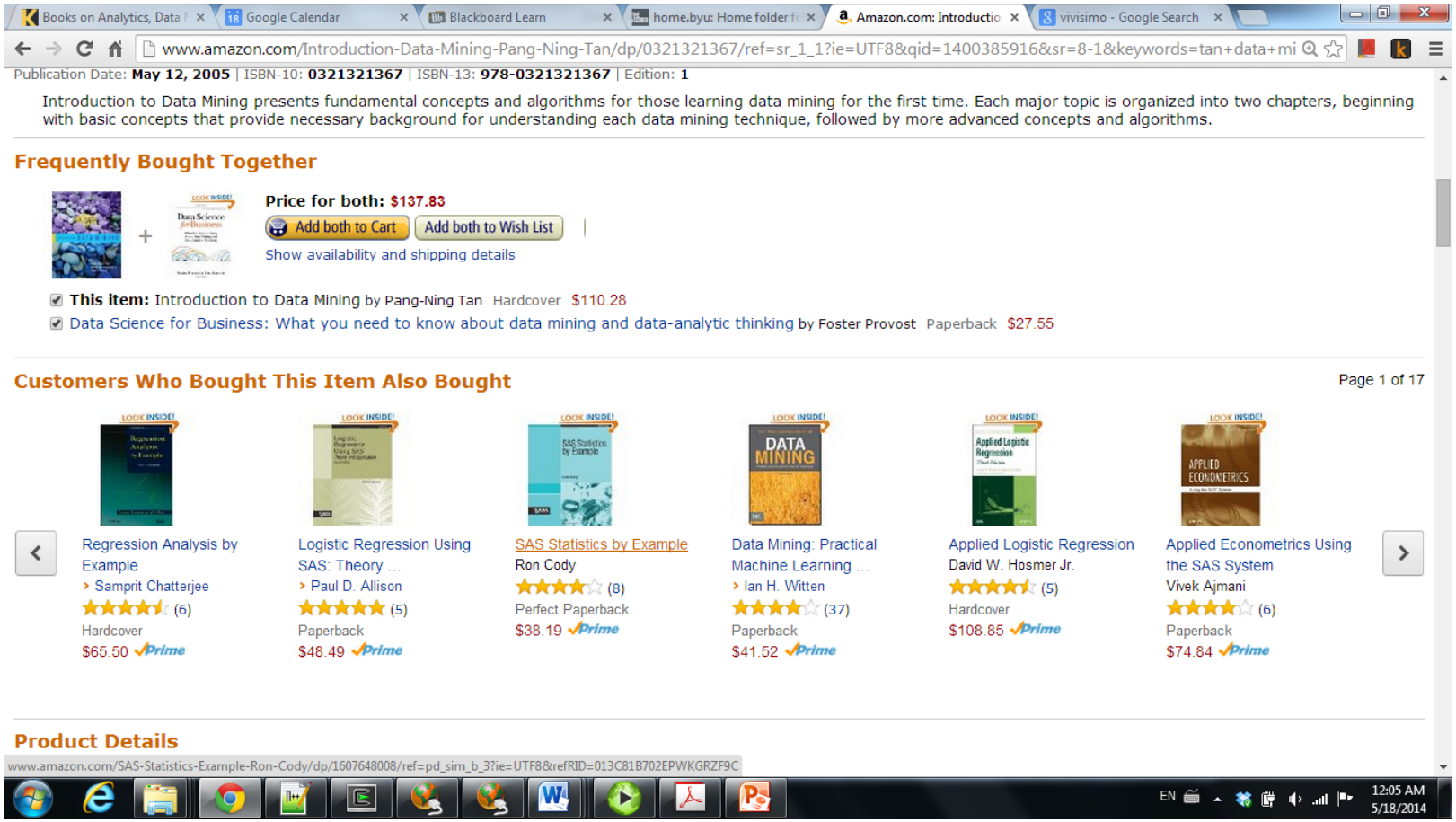
Example Rules:

{B, E} → {F}
{E, F} → {B}
{B, F} → {E}
{B} → {E, F}
{E} → {B, F}
{F} → {B, E}

← RHS: Right-Hand Side

Transaction ID	Items Bought
10	A, B, D
20	A, C, D
30	A, D, E
40	B, E, F
50	B, C, D, E, F

AR MINING APPLICATION 1: MARKETING AND SALES PROMOTION



AR MINING APPLICATION 2: SHELF MANAGEMENT

Supermarket shelf management

Goal: To identify items that are bought together by sufficiently many customers

Approach: Process the point-of-sale data collected with barcode scanners to find dependencies among items.

A classic rule:

If a customer buys diapers and milk, then he is very likely to buy beer.

So don't be surprised if you find six-packs stacked next to diapers!

AR MINING APPLICATION 3: INVENTORY MANAGEMENT

Inventory management

Goal: A consumer-appliance repair company wants to anticipate the nature of repairs on its consumer products and keep the service vehicles equipped with right parts to reduce the number of visits to consumer households.

Approach: Process the data on tools and parts required in previous repairs at different consumer locations and discover the co-occurrence patterns.