Real-world Insights from Mining Retail E-Commerce Data

Ronny Kohavi, Ph.D.

**Vice President, Business Intelligence
Blue Martini Software
San Mateo, CA**

http://www.kohavi.com
http://www.bluemartini.com/bi

BLUE MARTINI
Elevating the Customer Experience™

May 22, 2003

Goals

- **Give you a feel for what e-commerce data looks like**

- **Show interesting insights with fun teasers from Blue Martini customers' data**

- **Show things that worked well for us, including architecture and powerful visualizations**

- **Next week: share more detailed data mining lessons and challenges (Rajesh Parekh)**

Agenda

- **Overview of architecture**

- **Usability**

- **Web site traffic**

- **Timeout**

- **Searches, referrers**

- **Micro-conversions and utilizing real-estate**

- **E-mail campaigns**

- **Multi-channel analysis**

- **Cross-sells / Associations**

- **Classification**

- **Summary**

The Vision in 1998

- **In July of 1998, I gave an invited talk at ICML titled Crossing the Chasm: From Academic Machine Learning to Commercial Data Mining** http://robotics.stanford.edu/users/ronnyk/chasm.pdf

- **Most talks have one key slide (some have zero ☺)**

- **The key slide was the following slide, which guided the design of the data mining architecture at Blue Martini software**

BLUE MARTINI

General Suggestions for the Entrepreneurs

- ◆ Team with business people that understand marketing, sales, distribution channels.
- ◆ Build a vertical solution (e.g., manufacturing, warranty DB) and use the end user's vocabulary. Build an application that maximizes profit instead of minimizing RMS.
- ◆ Build a GUI based on the task.
- ◆ Develop a scalable architecture (e.g., client/server).
- ◆ Integrate with other pieces of the process: databases, cleansing tools, reporting tools, post processing, deployment.

Our CEO did this once before

Vertical: e-commerce retail

Ronny Kohavi   ICML 1998
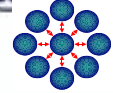
Integrated Architecture
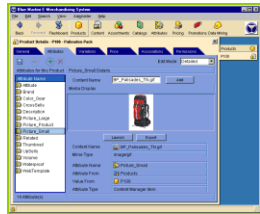
Business Data Definition
(Enterprise Desktop, Remote Desktop)

Stage Data

Customer Interactions
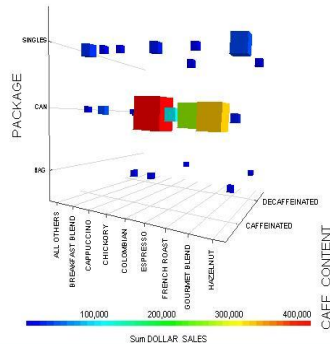(Web, campaigns, Call Center, Wireless, POS)

Deploy Results

Build Data Warehouse (DSSGen)

Analysis
(Reporting, Analytics, Visualizations, OLAP)

marketplace

BLUE MARTINI

Advantages of Architecture

- **It is well documented that "80% of the time spent in knowledge discovery is spent on data preparation"**

- **Our architecture shares enough meta data and there is enough domain knowledge to cut that dramatically**

- **Clickstreams**
    - **Store from the application server layer to the DB (no need to load from flat files on multiple web servers, conflate, and sessionize)**
    - **Collect additional information (screen resolution, local time)**
    - **Tie all activities (registrations, orders) to sessions**
    - **Log high level "Business Events," including cart activities, search information, form errors**

- **More information in *Integrating E-Commerce and Data Mining: Architecture and Challenges*, ICDM 01**
Available at http://robotics.stanford.edu/users/ronnyk

Usability – Form Errors

This was the Bluefly home page

Looking at form errors logged by our architecture, we saw thousands of errors every day on this page

Any guesses?

Improved Home Page

This is the new Bluefly home page

- **Search box added**

- **E-mail box clearly marked as email**

- **As with many insights, hindsight is 20/20**

- **The hard part is collecting the right information and reporting on it**

Bot Detection

- **Bots are automated programs, sometimes called crawlers/robots**
  Examples: search engines, shopping bots, performance monitors

- **Significant traffic may be generated by bots**

- **Can you guess what percentage of sessions generated by bots?**

    23% at MEC (outdoor gear)

    40% at Debenhams

- **Without bot removal, your metrics will be inaccurate**

- **We find about 150 different bot families on most sites.**
  Very challenging problem!
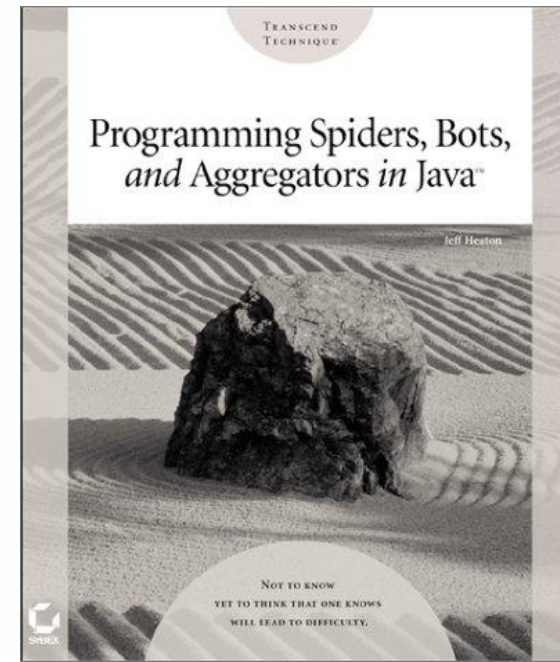
Example: Web Traffic
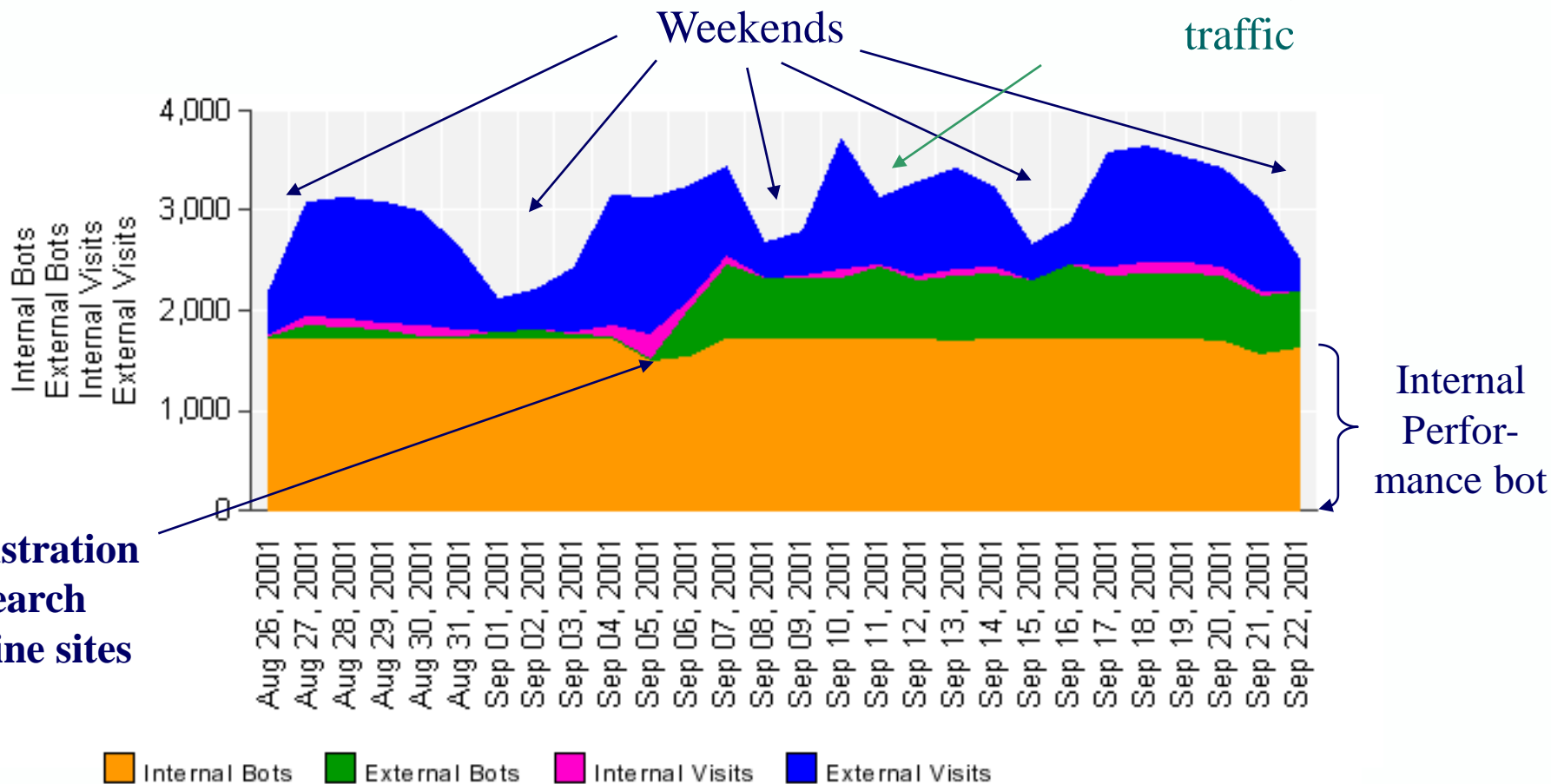
Heat Maps for Day-of-Week (Same Data)

- **Use color to show an additional dimension**
  - Green is low traffic
  - Yellow is medium traffic
  - Red is high traffic
- **The power of visualizations**
  - Weekends are very slow
  - Friday is slow
  - Patterns
    - Sept 11 in green
    - Reduced traffic after Sept 11 (yellow above Sept 11)
    - Sept 3 Labor day in green

Day of week: Sun, Sat, Fri, Thu, Wed, Tue, Mon

Week Jul 30, 2001 · Week Aug 06, 2001 · Week Aug 13, 2001 · Week Aug 20, 2001 · Week Aug 27, 2001 · Week Sep 03, 2001 · Week Sep 10, 2001 · Week Sep 17, 2001

Date (Week Starts on Monday)

Color : Sessions   Size : <None>

50,000.0   100,000.0   150,000.0
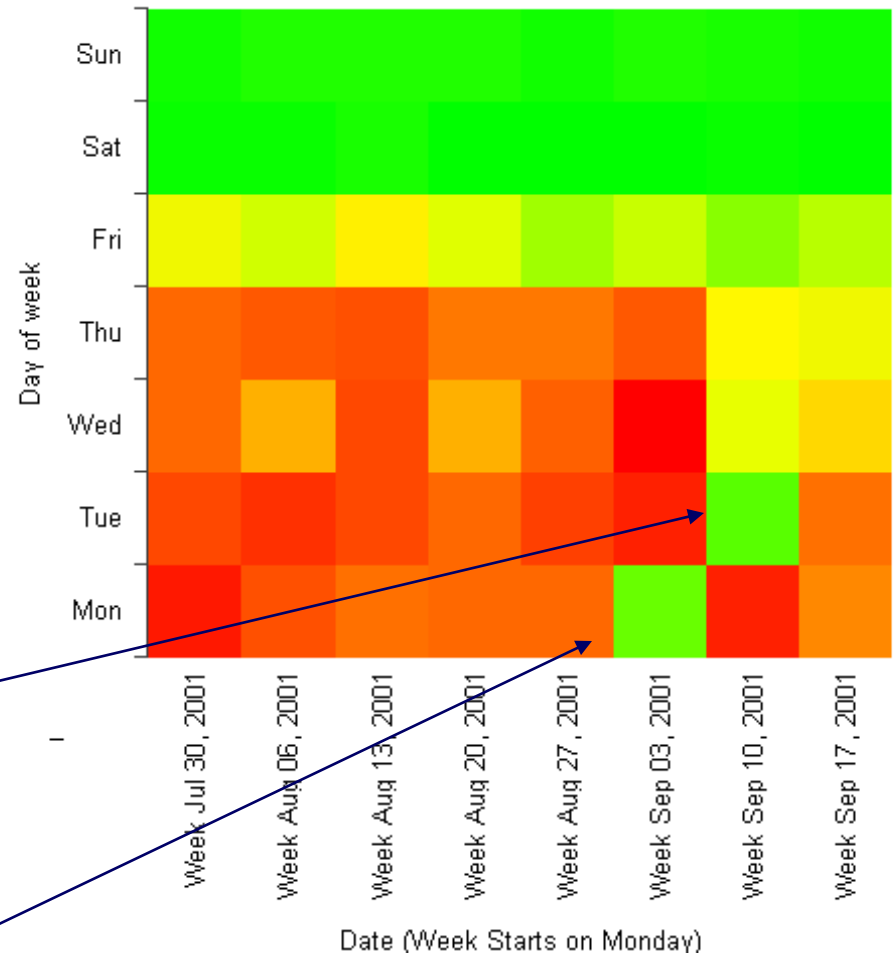
Browsing hours

- **Traffic by hour (server time) Lines show two consecutive weeks**

- **What do you think it looks like?**

- **How stable is it across domains/geographies?**

EST

CST

GMT

Tokyo
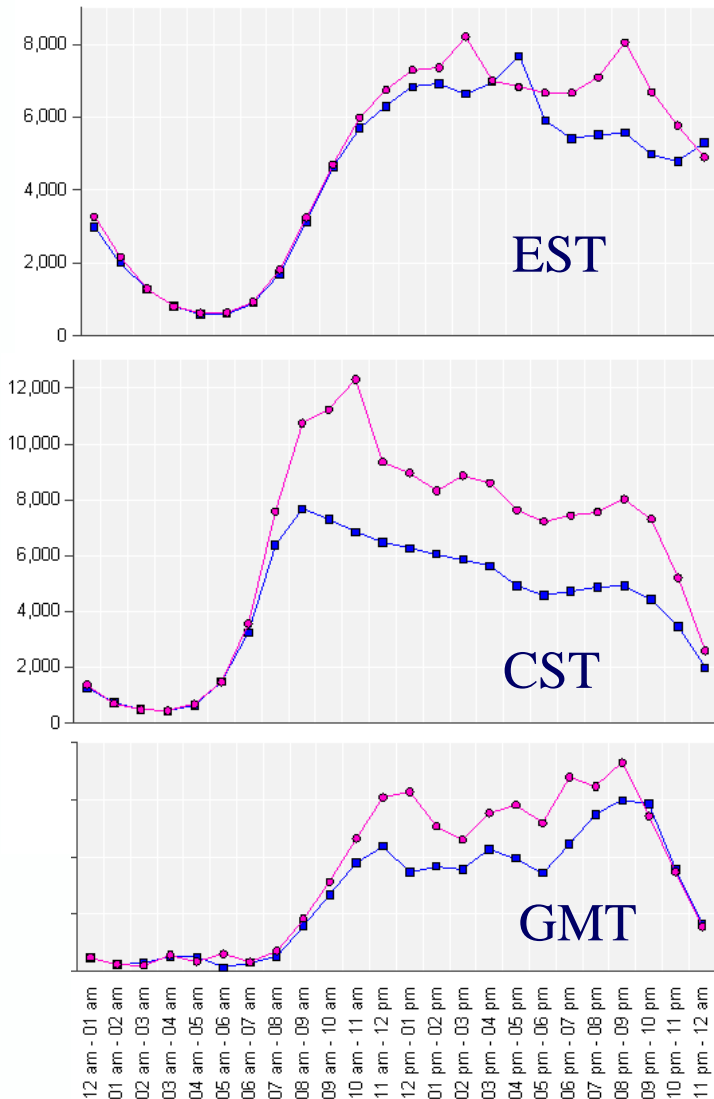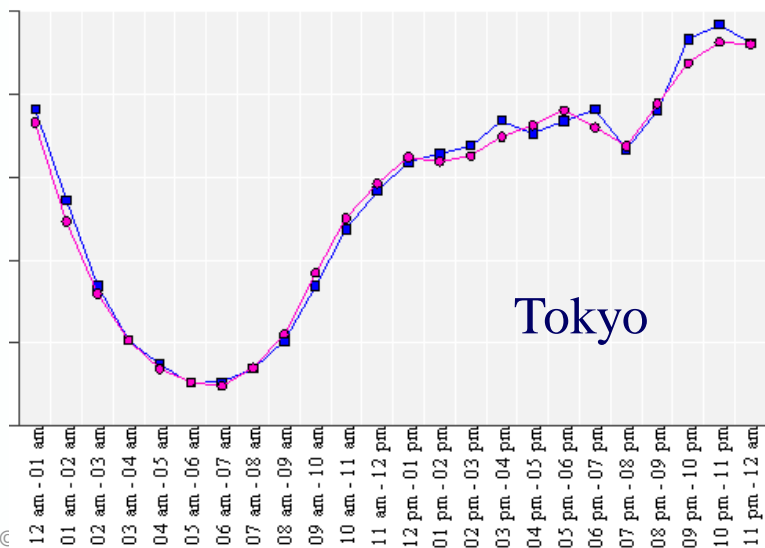
Drill-Down to Hour

- **Same heat map idea applies to hourly patterns**

- **In this case hourly traffic to a web site**

- **Note Sept 11 effect and its effect for rest of week**

Site down at critical hour
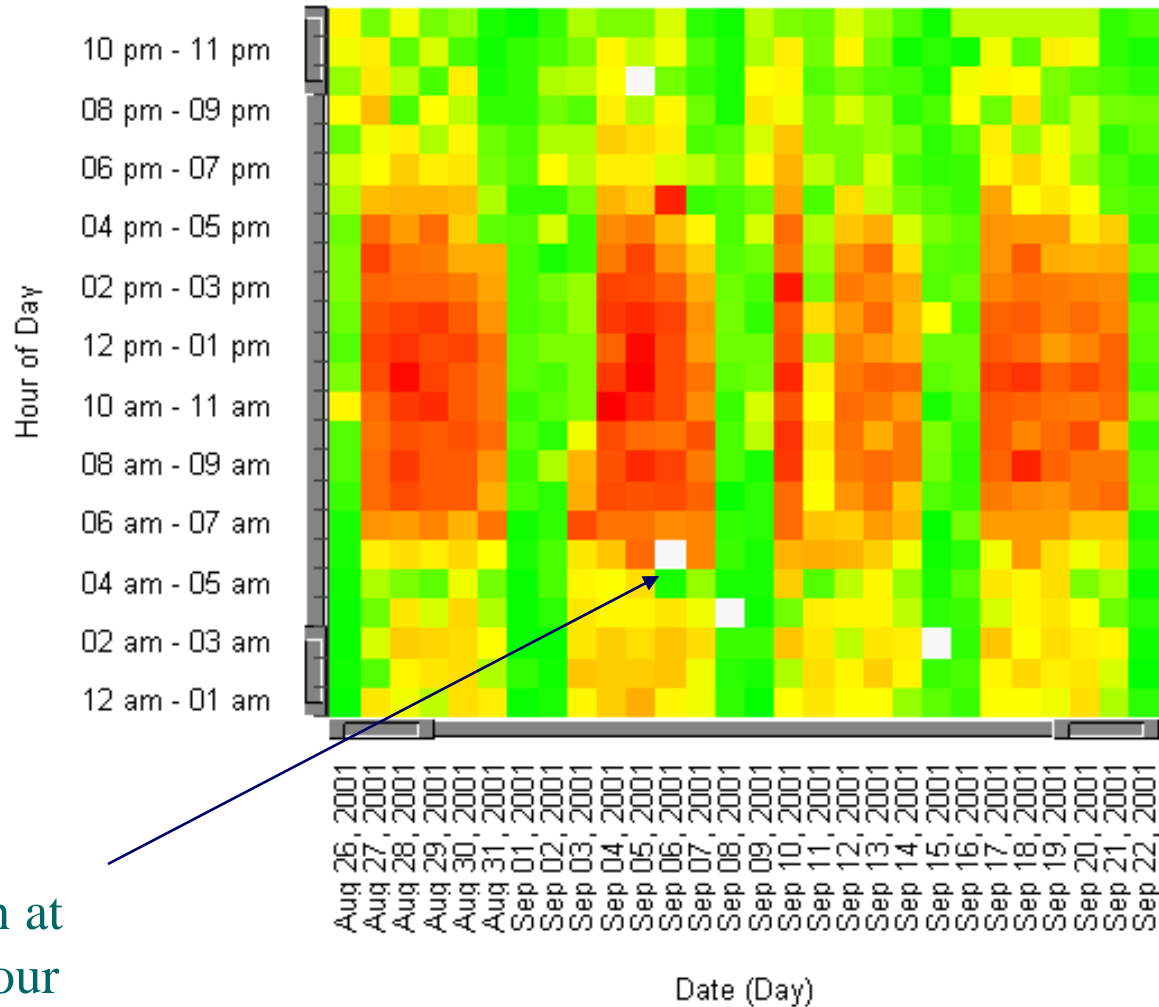
Teaser

- **Here is a similar heatmap**

- **Interestingly, the white square (no traffic) appeared on many sites**

- **But not in Phoenix, AZ servers**

- **Why?**
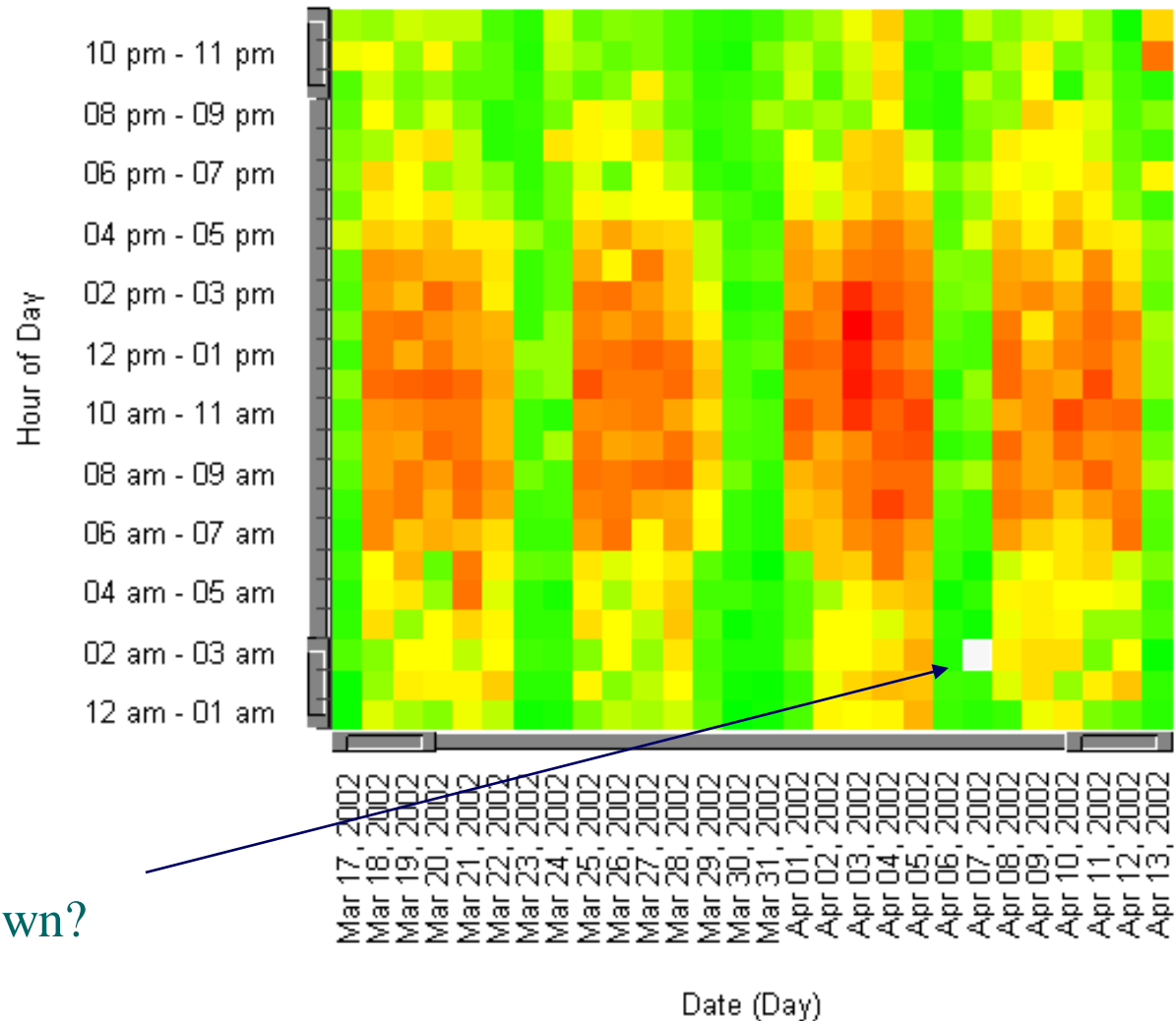
Site down?

Teaser

- **We found that people purchase hours after visiting the site**
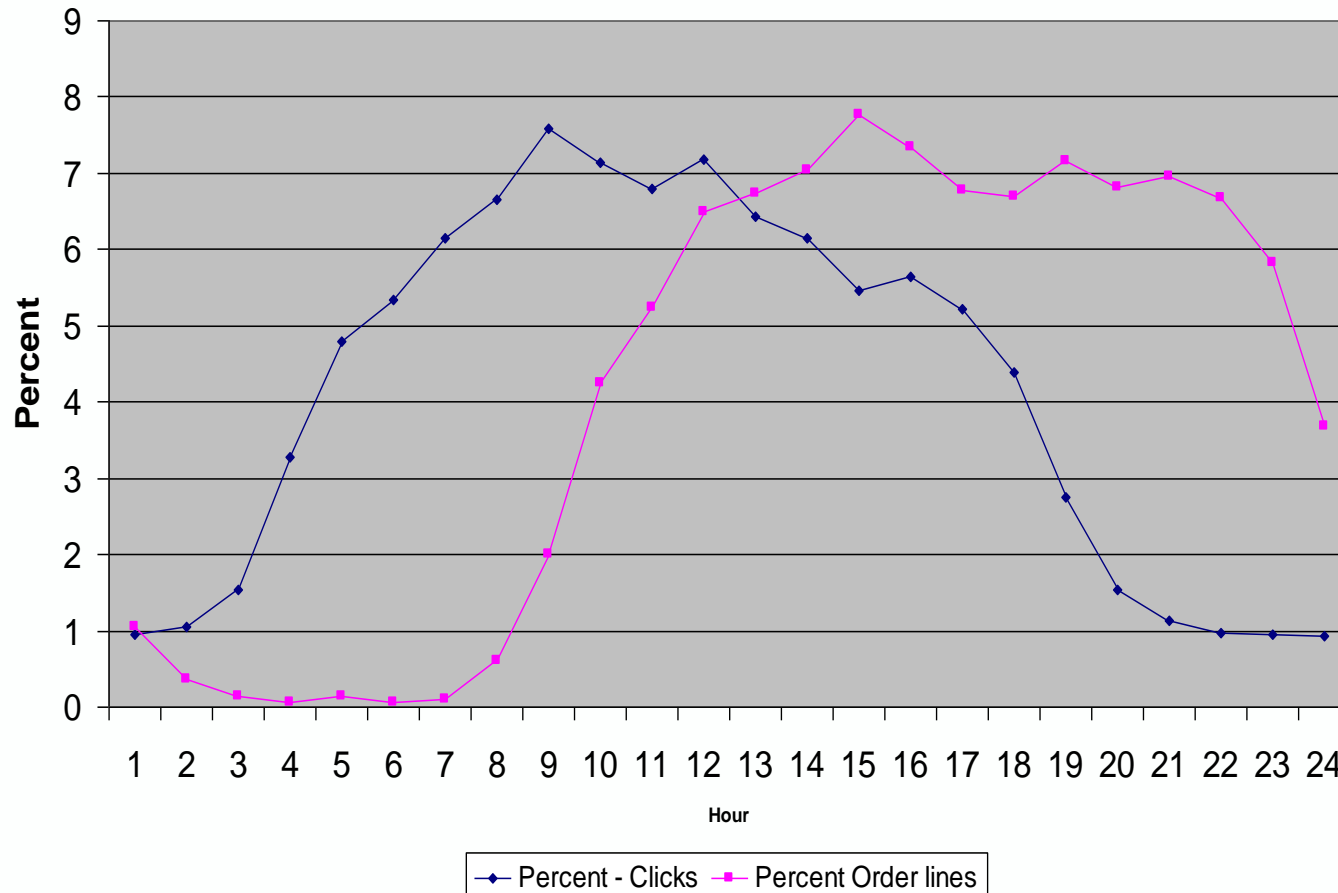
Session Timeout

- **Catledge and Pitkow in a well referenced paper determined that the "optimal" session timeout for analysis should be 25.5 minutes**

- **How many visitors at Debenhams**
  - **Added product to shopping cart**
  - **Waited over 25.5 minutes**
  - **Came back to the site in the next 3 hours?**

  5% (right axis)

  That's a lot of lost shopping carts

Session Timeout Threshold (mins)

% of Sessions with Add To Cart that experience a timeout
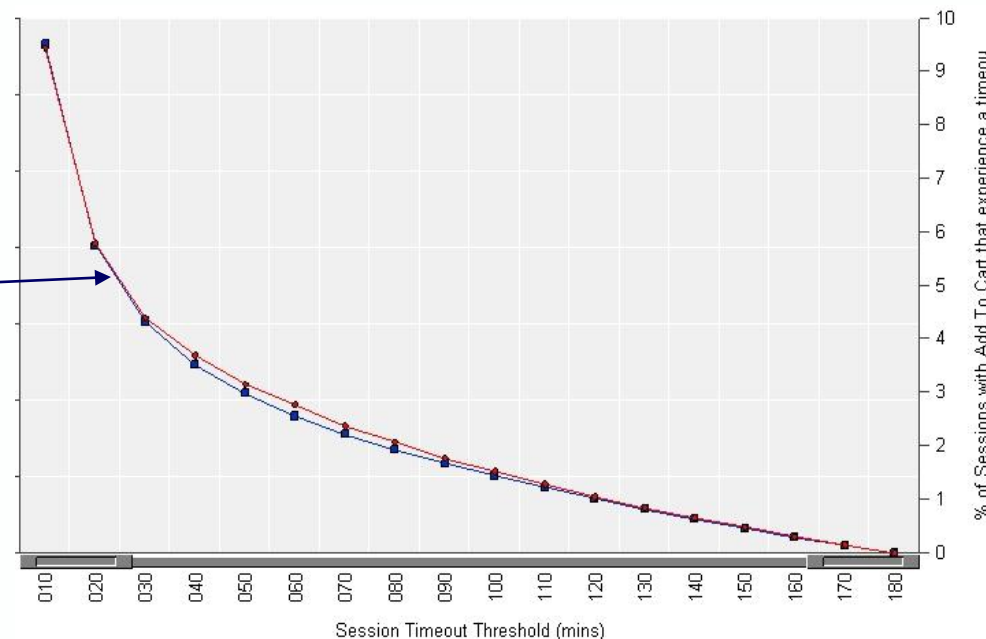
Searches

- **Architecture records every search and the number of results**

- **Top searched keywords (percent of searches)**
    - **Empty search string (3.9%) returns over 160 results**
    - **GPS (1.2%)**
    - **sunglasses (0.8%)**

> *Recommendation:*
>
> **- Do not allow empty search**
>
> **- Create custom pages for often searched keywords**

- **Top failed keywords in the product category (percent of failed searches)**
    - **gift certificate(s) (0.98%) (already implemented since study)**
    - **arc'teryx (0.44%)**
    - **bear spray (0.44%)**
    - **pedometer (0.37%)**
    - **stroller(s) (0.36%)**

> *Recommendation:*
>
> **- Parse search string to remove special characters**
>
> **- Build extensive thesaurus**
>
> **- Consider carrying products**

Synonyms

- **At Publix, an online grocer in the southeast, 'Bath Tissue' was among the top selling assortments**

- **Top failed search?**

Toilet Paper

Search Effectiveness at MEC

- **Customers that search are worth two times as much as customers that do not search**

- **Failed searches hurt sales**

Visit

90%

10%

No Search

Avg sale per visit: $X

Search
(64% successful)

Avg sale per visit: 2.2X

70%

30%

Last Search Succeeded

Avg sale per visit: 2.8X

Last Search Failed

Avg sale per visit: 0.9X
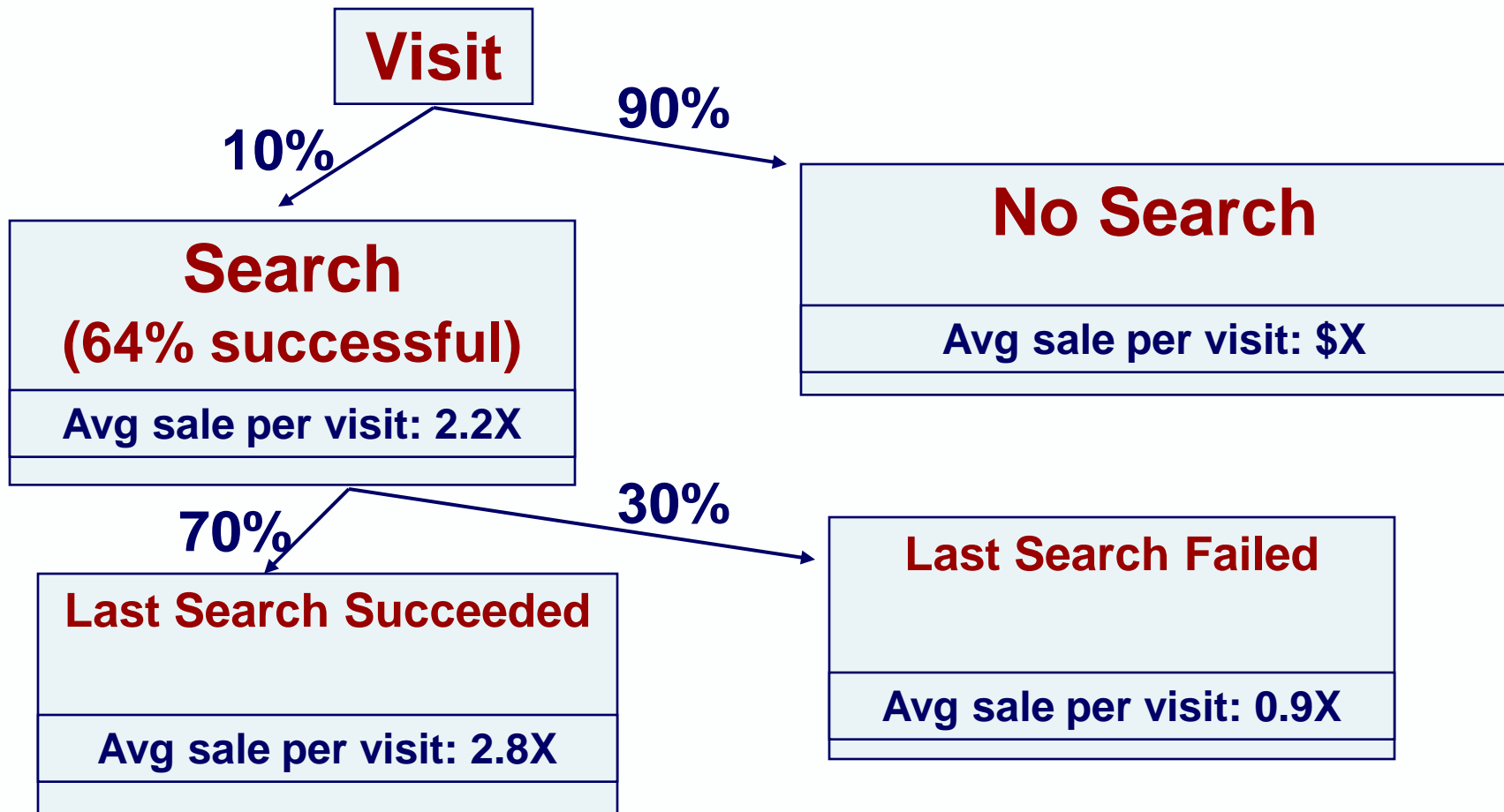
Referrers at Debenhams

- **Top Referrers**
    - **MSN (including search and shopping)**
        - **Average purchase per visit = X**
    - **Google**
        - **Average purchase per visit = 1.8X**
    - **AOL search**
        - **Average purchase per visit = 4.8X**

Micro-Conversion Rates at Debenhams

- Understand abandonment and conversions

- Not just visitor to purchaser, but also the micro-conversions

- Shopping Cart Abandonment 62% =55% + 45% * 17%

- Excellent opportunity to identify problematic steps in processes and improve

- Also a good way to identify abandoned products, send targeted e-mails if those products are on sale

BLUE MARTINI

2.0%

Visitor → Registered Visitor → Purchasing Visitor

7.7% **25%**

Visit → Purchase Visit

2.3%

Add To Cart → Initiate Checkout → Purchase

6% **45%** **83%**

17%

55%

Abandon Cart
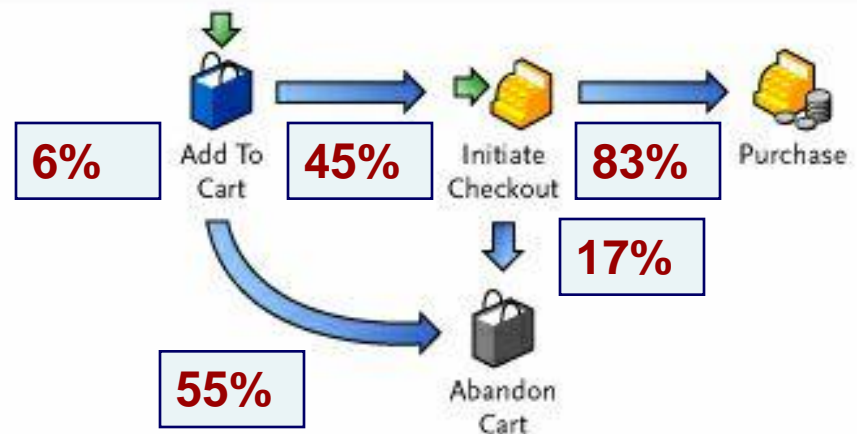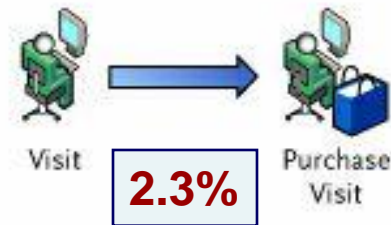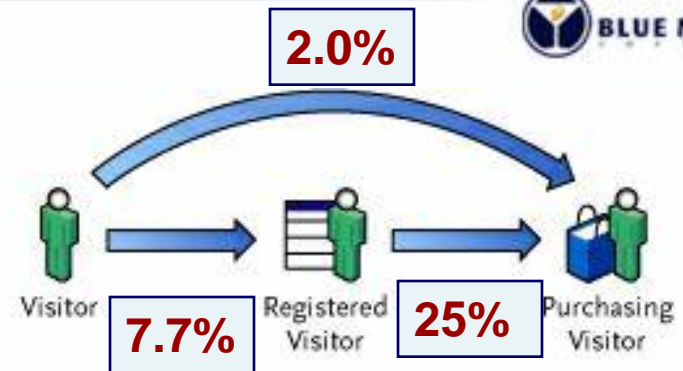
Page Effectiveness
Percentage of visits clicking on different links

14% 3% 2% 8% 2% 13% 9% 0.6% Top Menu 6%

BLUE MARTINI

DEBENHAMS
BRITAIN'S FAVOURITE DEPARTMENT STORE

- welcome - customer services - job opportunities - about us - our stores - account card

Enter Your Word | All Departments | SEARCH | SHOPPING BASKET
0 items

WOMENS | MENS | KIDS | HOMEWARE | GIFTS | WEDDINGS | FLOWERS | MY DETAILS
CATALOGUE | HEALTH & BEAUTY | EXPERIENCES | SALE | LOG - IN

i You are not logged in

3%

2% 2% 0.3% 2%

new year new you

WHICH? WEB TRADER
trust UK

DEBENHA
BIGGEST
HALF PRICE*
SALE

*SELECTED LINES ONLY &
SUBJECT TO AVAILABILITY

Don't let anything come between you and
the Debenhams Biggest sale! Debenhams
Sale

NEWYEAR
NEWYOU

Don't miss our round-up of the best for your
own home gym. From fitness equipment,
solariums to clothing you'll find everything
you need for a fitter new you.
Home fitness

Exclusive
preview
womenswear spring 2003

last few days

EXCLUSIVE
PREVIEW

Preview an exclusive selection of
our fashionable spring womenswear
ranges. And don't miss out on your
chance to save 15% on orders
placed by Monday 13 January.
Exclusive preview

beautiful
gifts
Big beauty
brands
online

Introducing
Nectar
A more
rewarding way
to shop

Happy New
Year
Say it with
flowers

COPYRIGHT© DEBENHAMS RETAIL PLC 2001, all rights reserved. Legal Disclaimer

SITE HELP   SECURITY   DELIVERY   PRIVACY   RETURNS

18% of visits exit at the welcome page

Any product link 7%

Top Links followed from the Welcome Page:
Revenue per session associated with visits

1.4X **2.3X** **X** **2.3X** **1.3X** **4.2X** **1.4X** **5X**

Top Menu 0.2X

10.2X **1.2X** **1.7X** **3.3X** **10X**

Note how effective physical catalog item #s are

Product Links 2.1X

Teaser - High Conversion Rates

BLUE MARTINI

- **Product Conversion Rate is the ratio of product purchases to product views**

- **High can conversion rates be over 100%**

  Conversion rates are high because

  - **Call Center (orders but no views)**

  - **Automatic reordering (send me the medicine every month)**

  - **Bundles (you view X, you get Y for free)**

  - **Wizard (at Virgin Wines, they mix you a case; most people don't even look at the details)**

  - **Quantities over 1 (question of exact definition of conversion)**

Teaser - Privacy

- 92% of Americans are concerned (67% very concerned) about the misuse of their personal information on the Internet.

  - FTC Report, May 2000

- 86% of executives don't know how many customers view their privacy policies.

  - Forrester Report, November 2000

- Q: What percentage of visitors read the privacy statement?
- **A: Less than 0.3%**

Direct Mail Campaigns (Why Spam)

BLUE MARTINI

- **Assumptions:**

  - **Response rate: 3%**
    (This is 0.6% for credit-card solicitations now, but we're going to send a wonderful offer for our Widget and get 3% response)

  - **Average revenue per response: $100**

  - **Profit margin: 20%**
    (after all costs, including handling returns, shipping, etc.)

- **To breakeven, how much should the offer cost per person?**

  - **Think about: creative design costs, letter, brochure, outer envelope, reply envelope, stamp, per-person cost when purchasing list**

Cost should be less than 60 cents!
   3%*$100*20% = $0.60

Obviously, it's not an easy business
That's why e-mail spam are so "cost effective"

Campaign Analysis - Debenhams

- **Analyze the effectiveness of campaigns**

Campaign	Emails Sent	Opens	Click-throughs	Orders
Campaign 1	100% (4.8p/email)	22% (22.3p/email)	9.3% (52p/email)	0.07%
Campaign 2	100% (0.5p/email)	11% (4.8p/email)	3% (17.9p/email)	0.01%
Campaign 3	100% (0.8p/email)	22% (3.6p/email)	5.3% (15.3p/email)	0.01%

Multi Channel Analysis

BLUE MARTINI

Multichannel customers spend 72% more per year than single channel customers

-- State of Retailing Online, shop.org

- **If we define a multi channel customer to have shopped on the web and at a store**

- **How much more do multi channel customers spend at <client> over single channel customers?**

- **More than twice as much for customers with two or more purchases (you can't be multi-channel if you haven't shopped twice).**

Channels by Num Purchases

BLUE MARTINI

The following graph shows that for each known number of purchases, the web-channel-only customer is better

Therefore, our intuition tells us that the web channel is the best channel, right?

Wrong!

Chart: Customer Average Spending vs. Number of purchases (1, 2, 3, 4, 5, >5), comparing Multi-channel and Web-channel only.
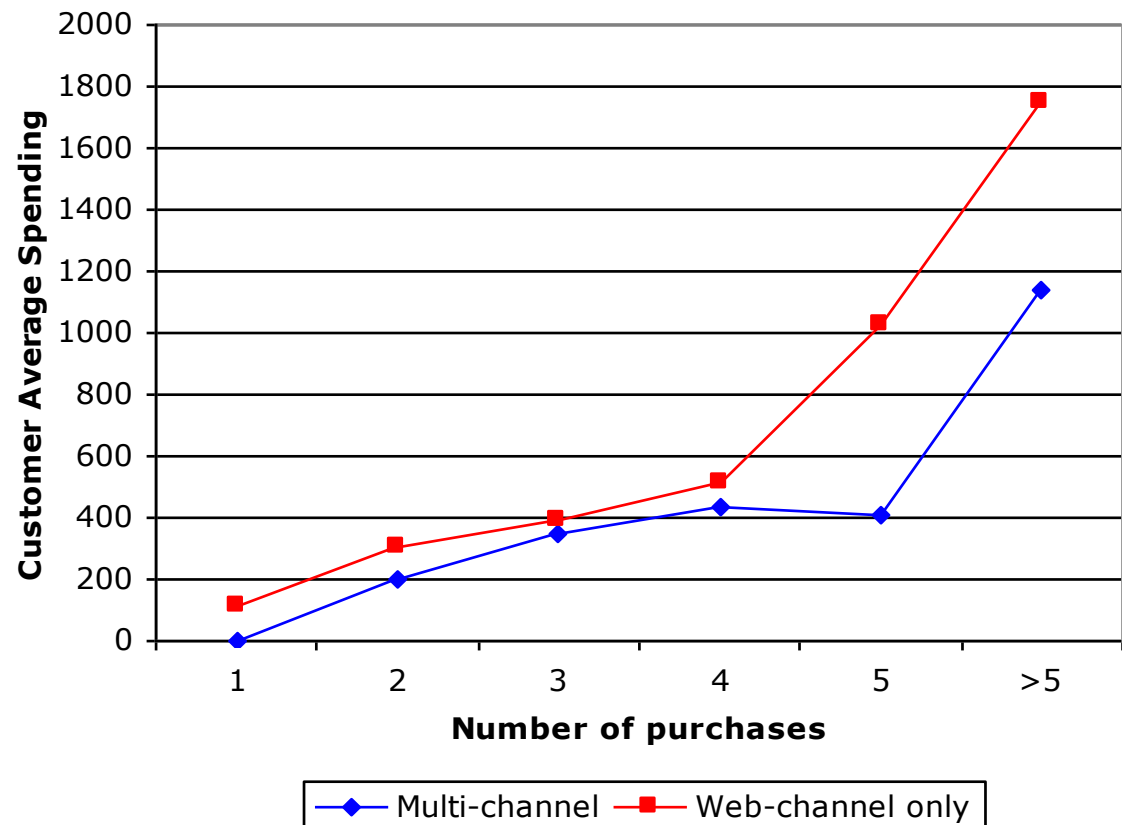
Bug?

- **Multi-channel customers have higher total spending**

- **This is an example of Simpson's paradox**
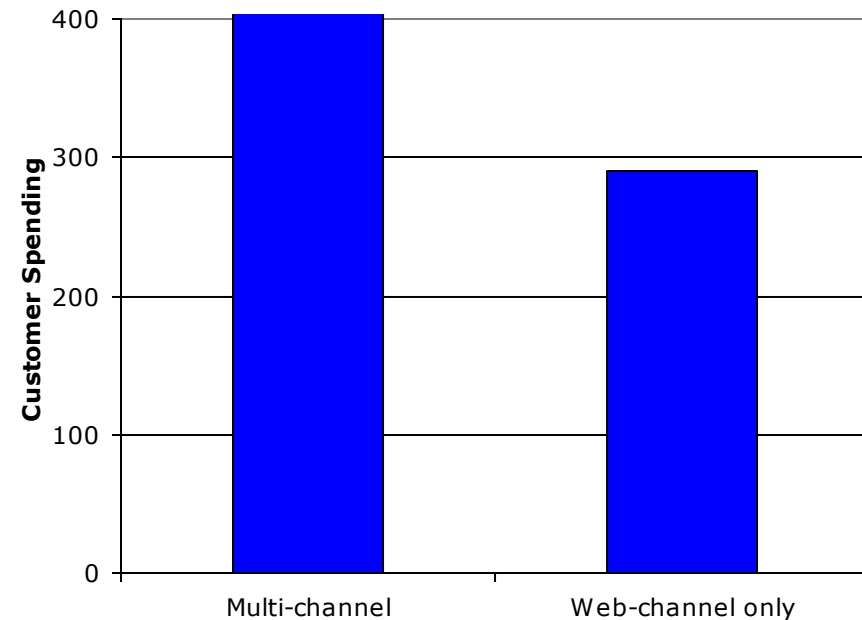
Simpson's Paradox

- **A woman sues Stanford for sex bias**

- **She shows that the school admits 70% of males but only 56% of females**

- **Stanford agrees with these percentages**

- **Shows that in *every* department they accept a higher percentage of females than males**

- **What is amazing is that this can happen**

- **What is more amazing is that it happened in practice**

Subtle Difference in Conversation

- **Alice to Bob: I'm applying to Stanford next year**

- **Bob to Alice: Sorry to hear that; I know they're accepting more males than females**

VS

- **Alice to Bob: I'm applying for department X at Stanford next year**

- **Bob to Alice: Lucky you, I know they're accepting more females than males in department X**

And it doesn't matter what X is!

Here is a Simplified Version

30 customers

Total spending

300 customers

Average

Average

100 customers

200 customers

Blue – web channel
Green – multi channel with web

2   **>5**
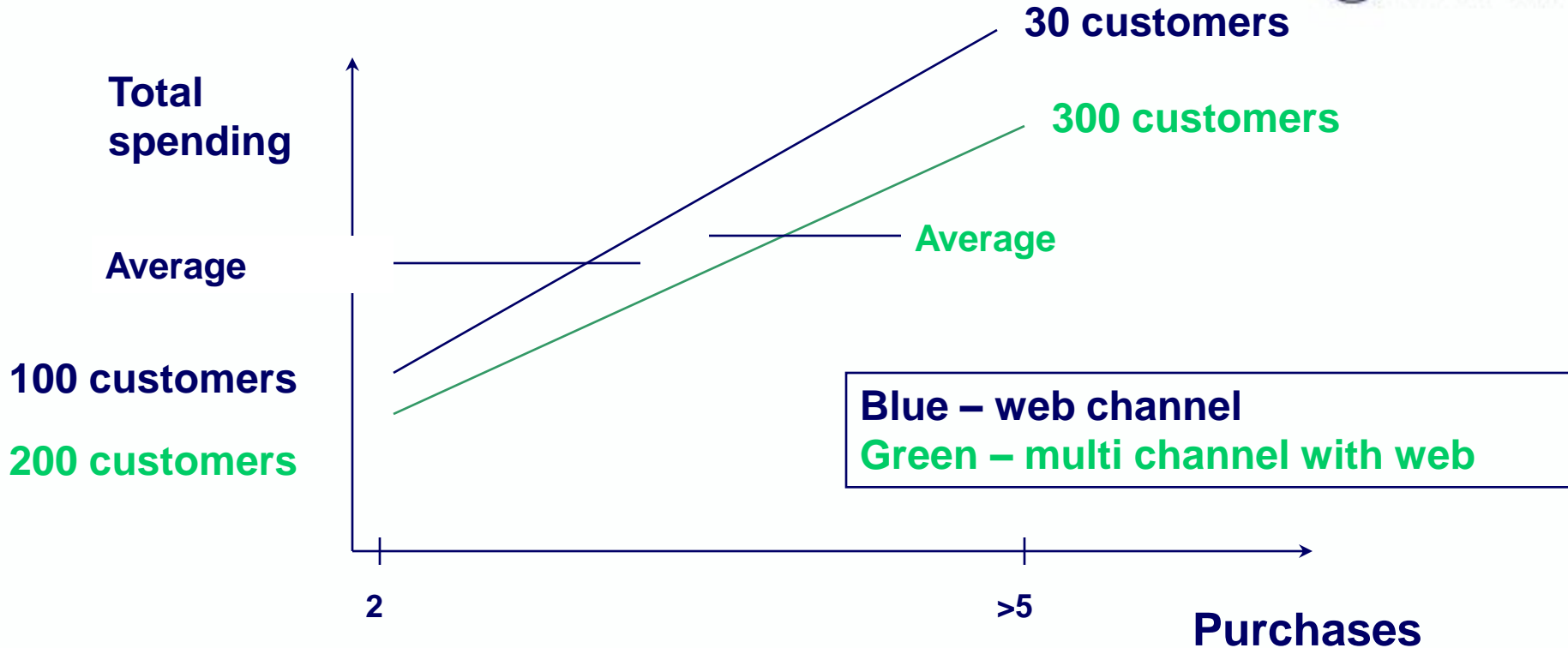
Purchases

The web channel dominates the multi-channel with web
in both 2-purchases and >5 purchases

Product Affinities

- **Which products sell well together**

- **Discovered using the association algorithm**

- **For closing the loop, associations can be used to make cross-sell recommendations at the website**

Product Affinities at MEC
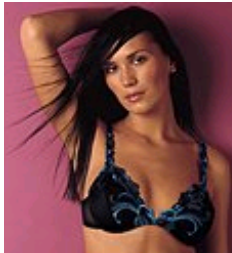
Product		Association		Lift	Confidence	Website Recommended Products
Orbit Sleeping Pad		Orbit Stuff Sack		222	37%	Cygnet Sleeping Bag · Aladdin 2 Backpack · Primus Stove
Bambini Tights Children's		Bambini Crewneck Sweater Children's		195	52%	Yeti Crew Neck Pullover Children's · Beneficial T's Organic Long Sleeve T-Shirt Kids'
Silk Crew Women's		Silk Long Johns Women's		304	73%	Micro Check Vee Sweater · Volant Pants · Composite Jacket
Cascade Entrant Overmitts		Polartec 300 Double Mitts		51	48%	Volant Pants · Windstopper Alpine Hat · Tremblant 575 Vest Women's

- Minimum support for the associations is 80 customers
- Confidence: 37% of people who purchased Orbit Sleeping Pad also purchased Orbit Stuff Sack
- Lift: People who purchased Orbit Sleeping Pad were 222 times more likely to purchase the Orbit Stuff Sack compared to the general population

Product Affinities at Debenhams

Product		Association		Lift	Confidence	Website Recommended Products	
Fully Reversible Mats		Egyptian Cotton Towels		456	41%	J Jasper Towels	
White Cotton T-Shirt Bra		Plunge T-Shirt Bra		246	25%	Black embroidered underwired bra	

Confidence 1.4%

Confidence 1%

- Minimum support for the associations is 50 customers
- Confidence: 41% of people who purchased Fully Reversible Mats also purchased Egyptian Cotton Towels
- Lift: People who purchased Fully Reversible Mats were 456 times more likely to purchase the Egyptian Cotton Towels compared to the general population
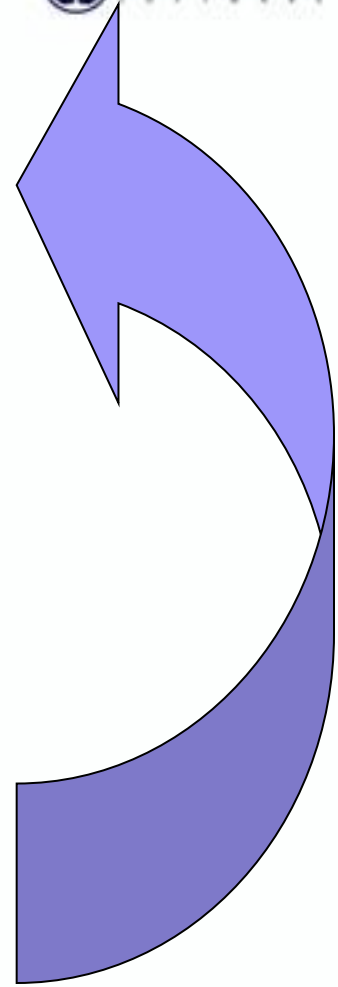
Building The Customer Signature

- **Building a customer signature is a significant effort, but well worth the effort**

- **A signature summarizes customer or visitor behavior across hundreds of attributes, many which are specific to the site**

- **Once a signature is built, it can be used to answer many questions.**

- **The mining algorithms will pick the most important attributes for each question**

- **Example attributes computed:**
    - **Total Visits and Sales**
    - **Revenue by Product Family**
    - **Revenue by Month**
    - **Customer State and Country**
    - **Recency, Frequency, Monetary**
    - **Latitude/Longitude from the Customer's Postal Code**

Migration Study - MEC

- **Customers who migrated from low spenders in one 6 month period to high spenders in the following 6 month period**

Oct 2001 – Mar 2002

| Spent over $200 |
| Spent $1 to $200 |

Migrators

Apr 2002 – Sep 2002

| Spent over $200 (5.5%) |
| Spent under $200 (94.5%) |
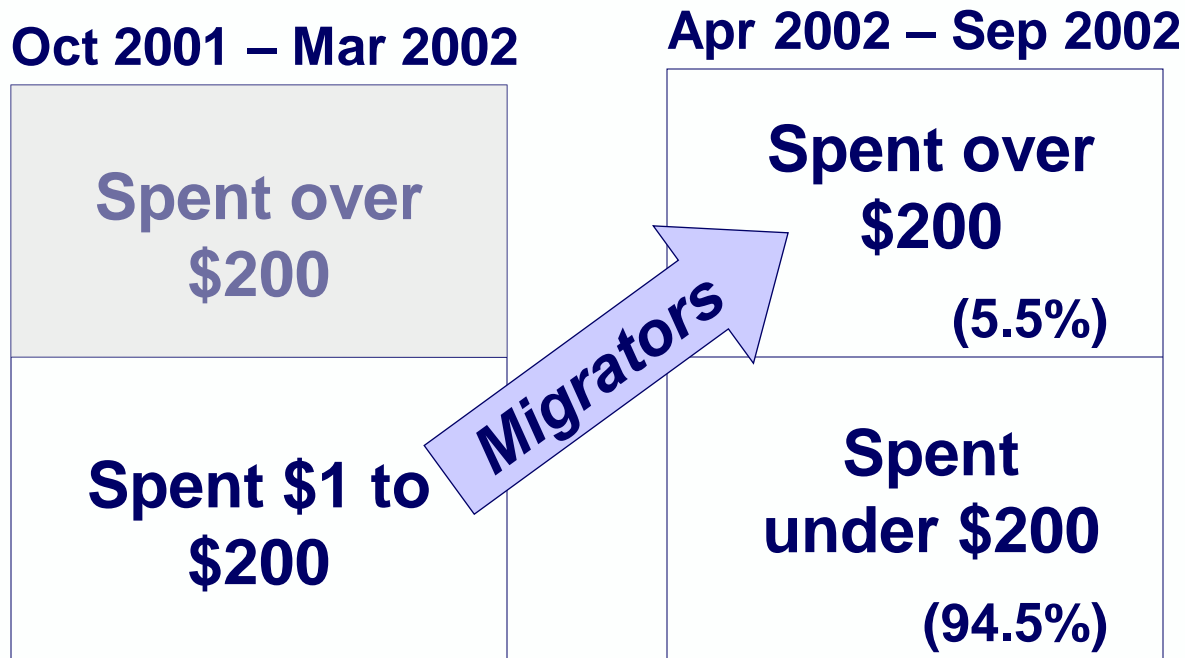
Key Characteristics of *Migrators* at MEC

BLUE MARTINI

- **During October 2001 – March 2002 (Initial 6 months)**

  - **Purchased at least $70 of merchandise**

  - **Purchased at least twice**

  - **Largest single order was at least $40**

  - **Used free shipping, not express shipping**

  - **Live over 60 aerial kilometers from an MEC retail store**

  - **Bought from these product families, such as socks, t-shirts, and accessories**

  - **Customers who purchased shoulder bags and child carriers were LESS LIKELY to migrate**
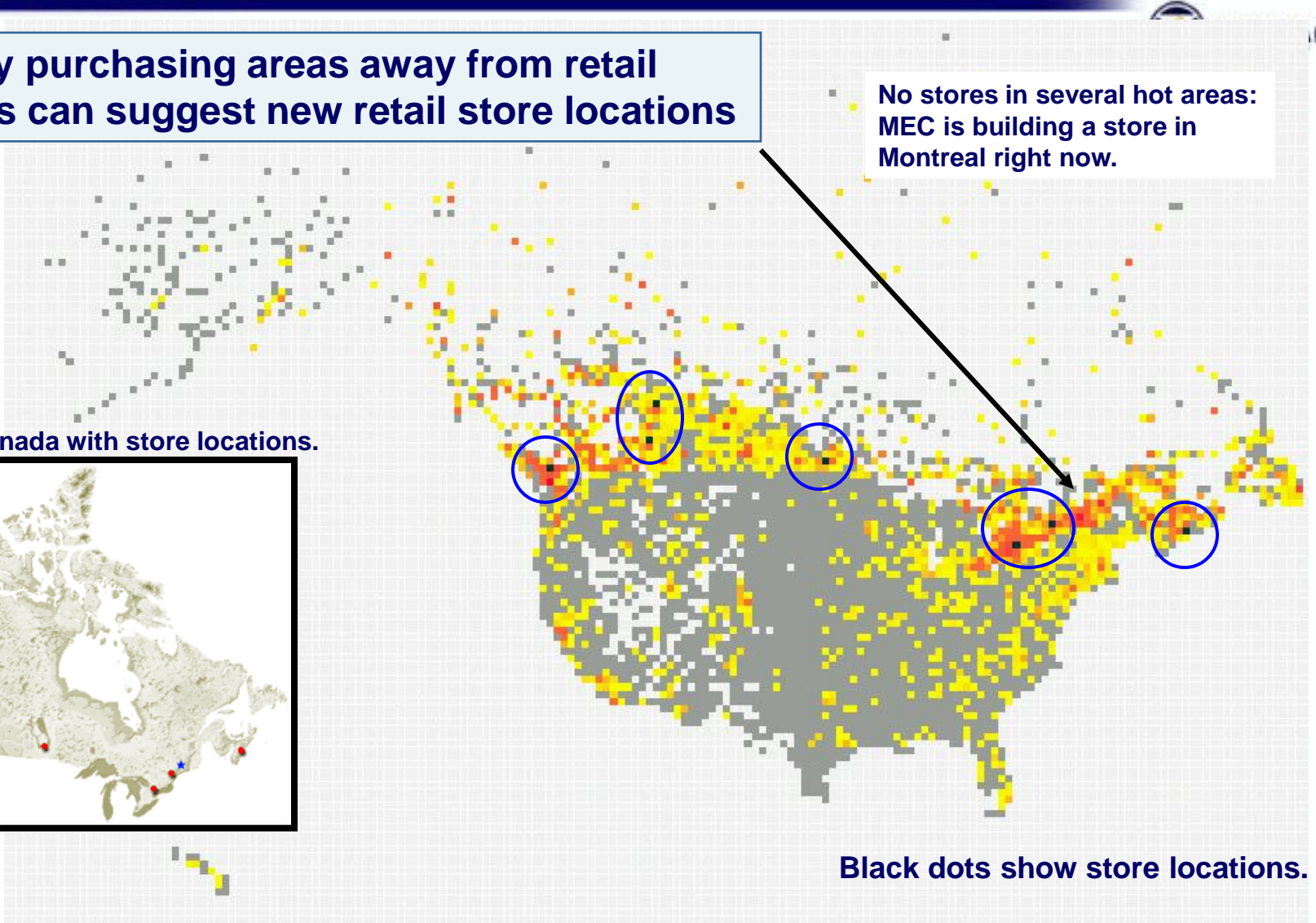
> *Recommendation:*
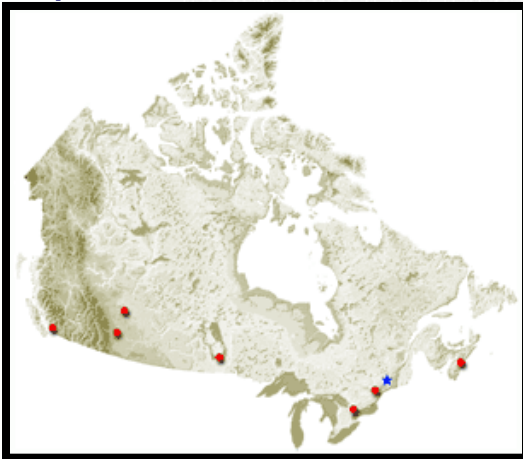> **Score light spending customers based on their likelihood of migrating and market to high scorers.**

Customer Locations Relative to Retail Stores

Heavy purchasing areas away from retail stores can suggest new retail store locations

No stores in several hot areas: MEC is building a store in Montreal right now.

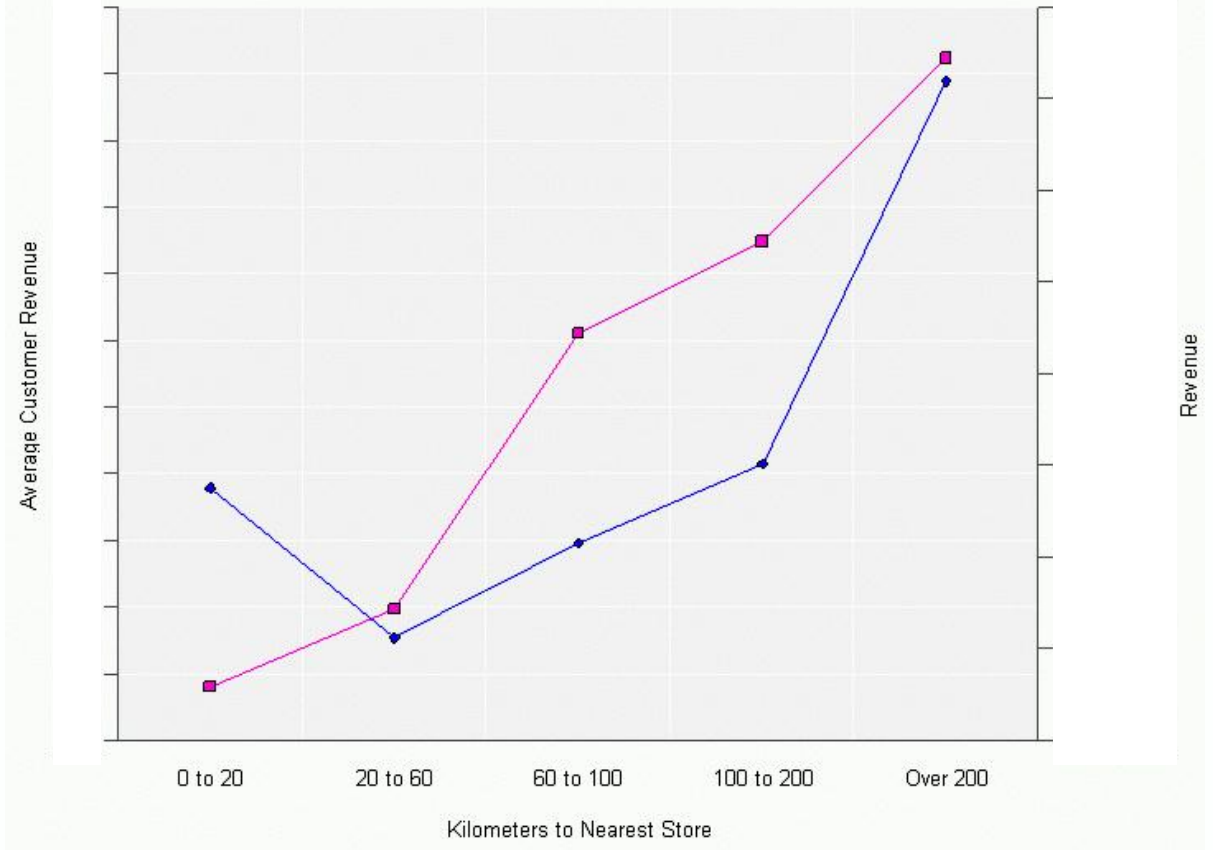**Map of Canada with store locations.**

Black dots show store locations.

Distance From Nearest Store (MEC)

- **People farther away from retail stores**

  - **spend more on average**

  - **Account for most of the revenues**

Other Results at MEC (See Appendix)

BLUE MARTINI

- **Free shipping changed to flat-fee (C$6 flat charge)**
  - **Orders - down 9.5%**
  - **Total sales - up 6.5%**

- **Gear Swap (buy/sell used gear)**
  - **Visit-to-Purchase very low: 0.34% vs. 2.1% for non gear-swap**
  - **However, these visitors converted to purchasing customers (over multiple visits) at a rate 62% higher than visitors who never visited gear swap!**

- **Visits where an FYI page (For-Your-Information) page was viewed had a Visit-to-Purchase conversion of 7.1%**

Other Results at Debenhams (See Appendix)

- **People who got the timeout page for a high percentage of their sessions are less likely to migrate (to heavy spenders)**

- **Revenue due to wedding list item purchases is clearly affected by summer weather**

  – **Weddings are more common in the summer in the UK**

  – **In June/July, 65% of revenues were generated through the wedding list**
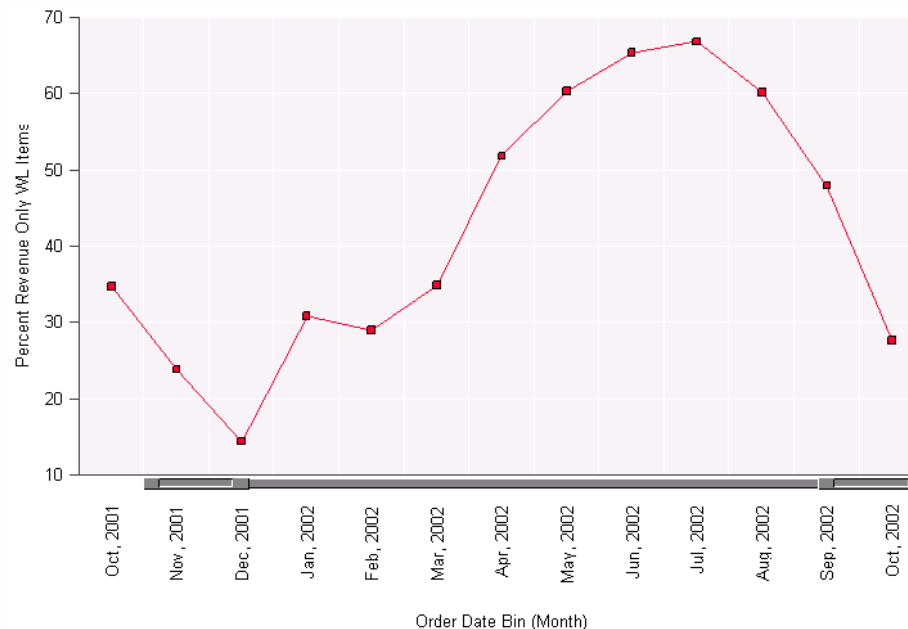
Summary (I)

BLUE MARTINI

- **E-commerce matches the needs of data mining**

  - **Huge datasets (both rows and columns)**

  - **Clean data (collected electronically)**

  - **Very actionable (easy to do controlled experiments)**

  - **Easy to measure return-on-investment**

- **Having a unified architecture (collection, transformation, analysis) saves much of the transformations needed (the 80% factor) and provides access to more data**

- **Customers need to crawl before they walk before they run. Must have simple reports**

Summary (2)

- **Focused on specific vertical – e-commerce retail**
  - **Enabled us to write out-of-the box reports**
    - **Easy for clients to get initial metrics and insights**
    - **Encapsulate our expertise in this domain**
  - **Focuses sales force, easier to demo with right vocabulary**

- **Provide visualization to show patterns (not discussed, but useful: interactive visualization)**

- **Many lessons, both at the business level and at the more data mining technical level to be reviewed by Rajesh Parekh**

Resources

- **WEBKDD workshops**

- **http://www.kohavi.com**

  - **Mining E-commerce Data, the Good, the Bad, and the Ugly, invited talk at KDD 2001 industrial track**

  - **Mining Customer Data, Etail CRM Summit, 2002**

  - **Integrating E-Commerce and Data Mining: Architecture and Challenges, ICDM 2001**

  - **E-metrics Study providing stats for multiple sites, Dec 2001**

  - **Applications of Data Mining to Electronic Commerce, special issue of Data Mining and Knowledge Discovery journal**

  - **Real World Performance of Association Rule Algorithms, KDD2001**

- **http://www.bluemartini.com/bi - case studies, live demo**

Appendix

- **Here are additional slides with some interesting insights**

RFM Analysis

- ## RFM – Recency, Frequency, Monetary **Example**

- ## Insights from Debenhams

  - ### Anonymous purchasers have lower average order amount

  - ### Customers who have opted out [of e-mail] tend to have higher average order amount

  - ### People in the age range 30-40 and 40-50 spend more on average

RFM Analysis (Debenhams)

BLUE MARTINI

- **Recency, Frequency, and Monetary calculations are used extensively in retail for customer segmentation**

- **Implemented the Arthur-Hughes RFM Cube**

  - **R, F, and M scores are binned into 5 equal sized bins**

  - **Each dimension is labeled 1 (best) – 5 (worst)**

- **Interactive visualization using Filter Charts**

- **Look at charts instead of cell-tables**

Complete RFM

RFM Heat Map - Monetary Average

RFM Heatmap - Number of Customers
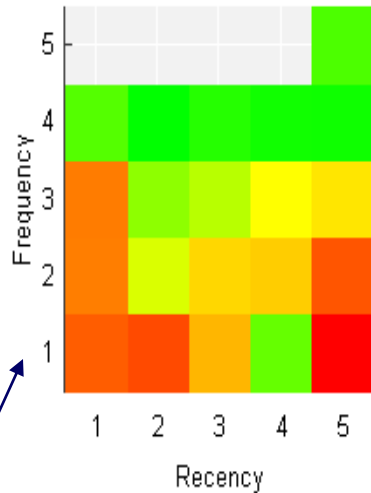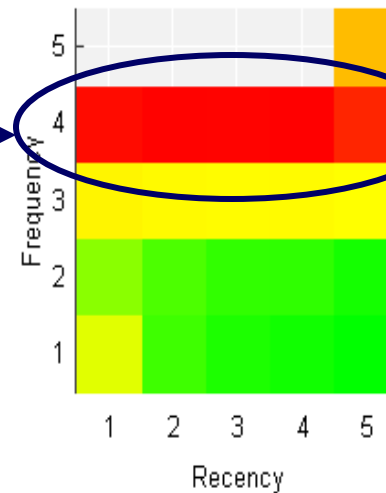
Majority of customers have purchased once

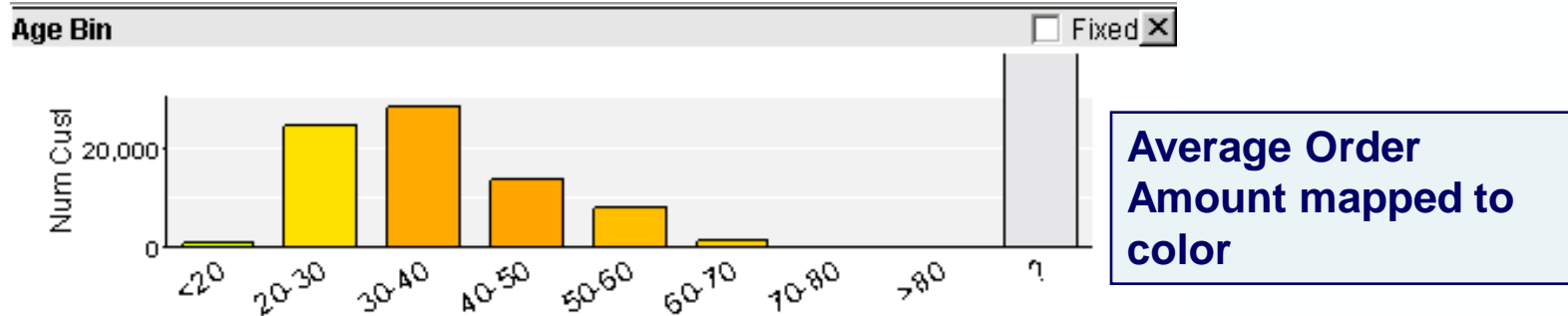**More frequent customers have higher average order amount**

Recommendation

Targeted marketing campaigns to convert people to repeat purchasers, assuming they did not opt-out of e-mails

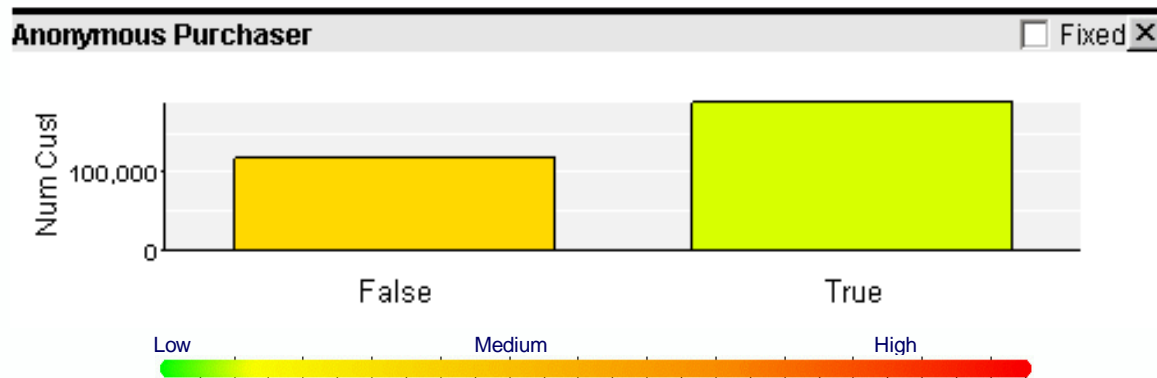# Interacting with the RFM visualization

- ## Explore sub-segments with filter charts

Average Order Amount mapped to color

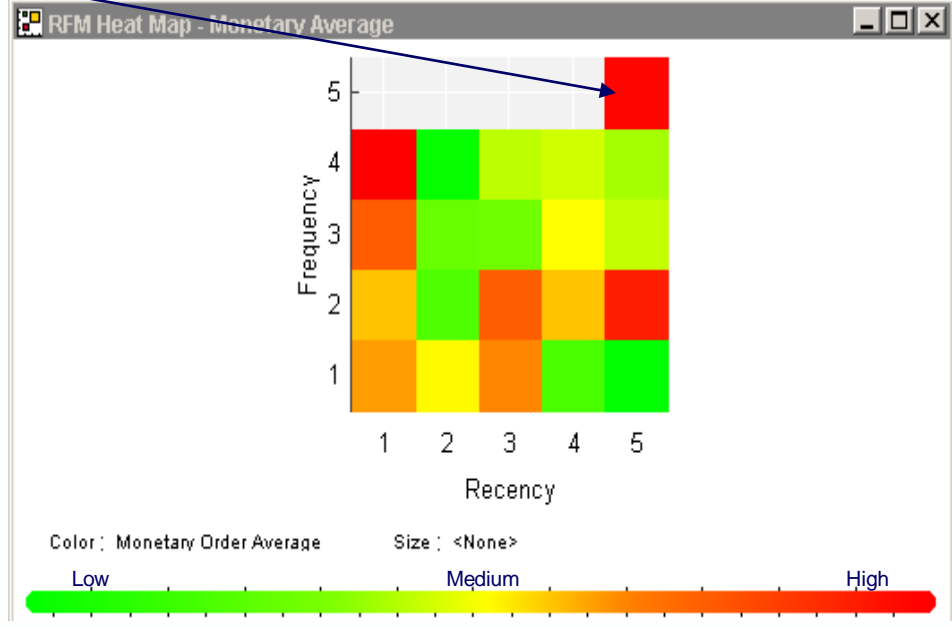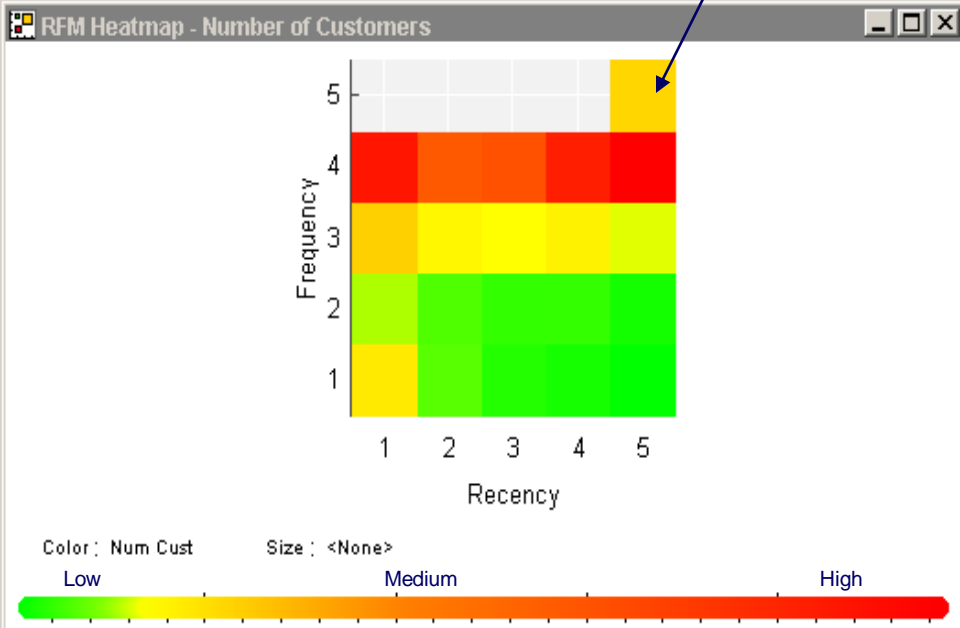- **People in the age range 30-40 and 40-50 spend more on average**

- **Anonymous purchasers have lower average order amount**

RFM for Debenhams Card Owners

Debenhams card owners
Large group (> 1000)
High average order amount
Purchased once (F = 5)
Not purchased recently (R=5)
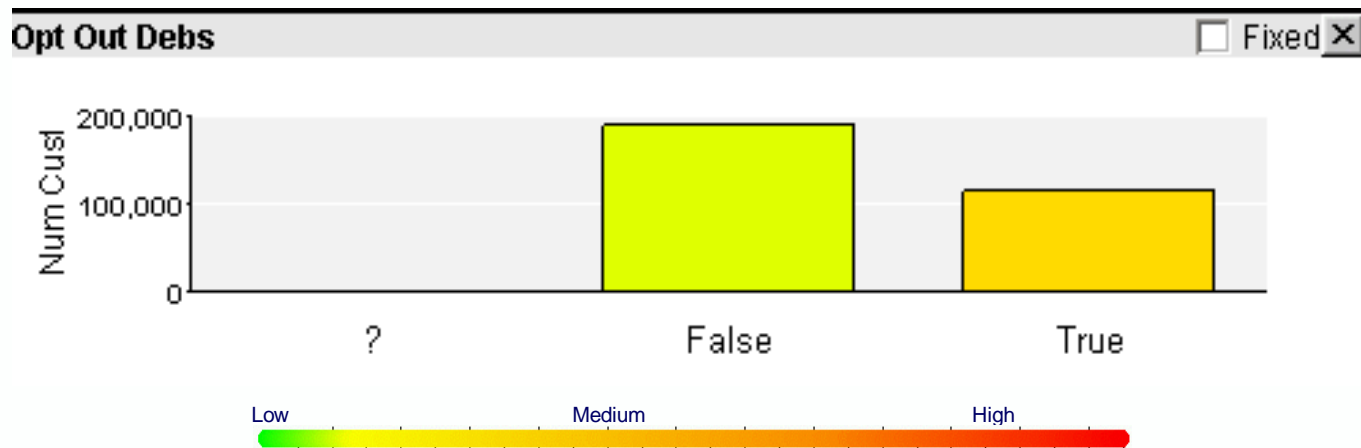
Recommendation

Send targeted email campaign since these are Debenham's customers. Try to "awaken" them!

Customers who have Opted Out

- **Customers who have opted out tend to have higher average order amount**

Recommendation

Send targeted emails to prevent email fatigue

Recommendation

Log changes to opt out settings and track unsubscribes to identify email fatigue
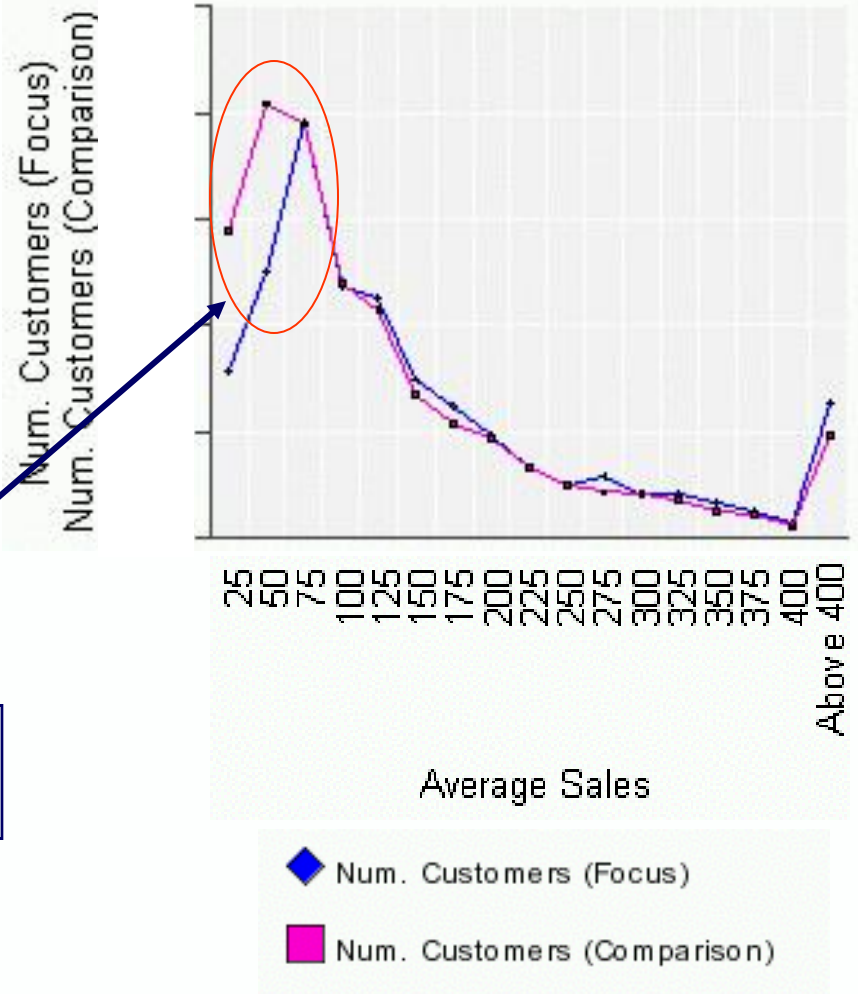
Free Shipping Offer (MEC)

- **Free shipping stopped on Aug 14, 2002**

- **A flat $6 Canadian Dollars shipping charge introduced**

- **Express shipping at higher charge continues**

- **Observations**
  - **Total sales - ↑ up 6.5%**
  - **Revenue (excluding shipping and tax) - ↑ up 2.8%**
  - **Orders - ↓ down 9.5%**
  - **Average Sales per Order – ↑ up 18%**

Free Shipping Offer (Cont.)

BLUE MARTINI

- The distribution shows fewer orders from low spenders (probably a good thing)

- No impact on rest of buyers

Fewer low spenders (<= $50)
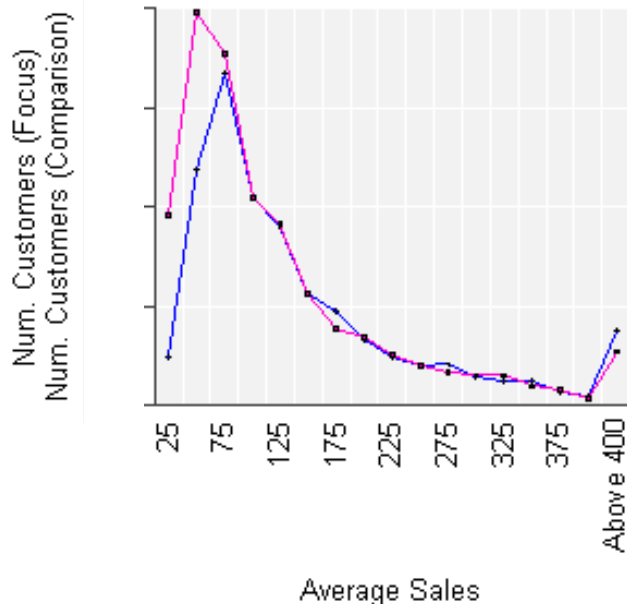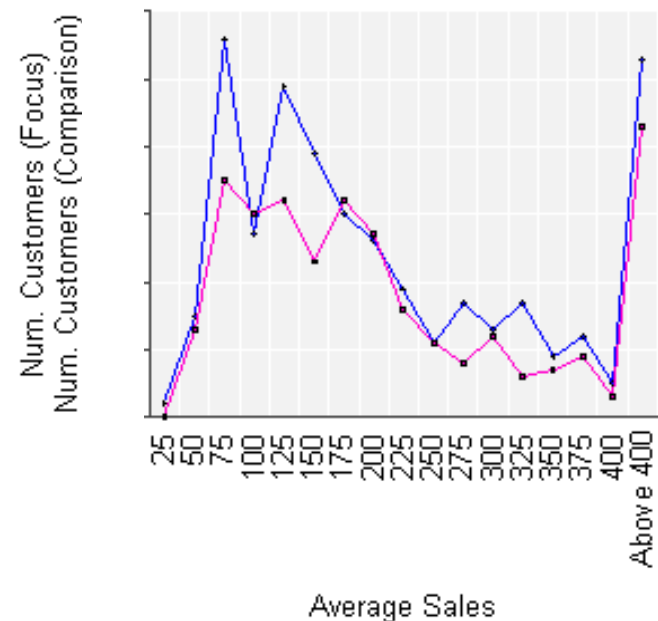
Free Shipping Offer (Cont.)

- **Breakdown of orders by shipping method**

- **More people used express shipping, probably because the delta to ship express wasn't as large (from C$6 instead of from C$0)**

Free/Standard Shipping

Express Shipping

Gear Swap Pages (Cont.)

BLUE MARTINI

> *Recommendation*:
> **Link back to MEC Shopping from Gear Swap**

Shop MEC Cycling

Gear Swap Pages (Cont.)

MOUNTAIN EQUIPMENT CO-OP

Search >

Home   Shop   MEC Community   Help   Login

View Shopping Cart
0 items in your cart.

MEC Community: Outdoor Gear Swap > Cycling

Outdoor Gear Swap

Men's Clothing
Women's Clothing
Packs
Tents
Sleeping Bags
Camping Gear (General)
Climbing

Cycling

Paddling
Snowsports
Kids
Footwear (General)
Miscellaneous Gear
Travel & Trip Partners

GEAR SWAP - Cycling

Looking for a used bike to take on your next outdoor adventure? MEC's OutdoorGearSwap.com is the place to buy, sell, and trade used bikes and other cycling gear.

add new posting >

Hide Details

229 Items Found Show [20 ▼] Items Pages: 1 2 3 4 5 6 7 8 9 10 11 12 ▶

Showing 1-20 of 229 items: **New Filter**

Sort Listing By:	Title	Region	Type	Price	Date

| **Kona Lavadome** | BC | For Sale | 550.00 OBO cad |

Description: 15" green, great bike, like new. Perfect for petite lady or teen wanting to get into mountain biking. Rock Shocks front suspension, Shimano LX components, new brake pads, 1 1/2 inch travel on front tire.

Date: Feb 10
Name: Nicky
Contact Info:
604-904-9456
Email:
toybrink@hotmail.com

| **2002 TIME ATAC CARBON PEDALS** | ON | For Sale | 150 cad |

Shop mec.ca for Cycling Gear

Men's Cycling Clothing
Women's Cycling Clothing
Cycling Footwear
Hydration
Packs, Bags & Panniers
Helmets
Tires, Tubes & Pumps
Bike Parts
Tools
Computers, Lights & Locks
Racks, Fenders & Trailers
Bells, Mirrors & Reflectors
Glasses, Gloves & Masks
Car Racks
Cycling Books

← **Done**

59

Definitions for Gear Swap Analysis

BLUE MARTINI

- **A visitor is defined as someone who is registered (MEC member) or is identified by a cookie**

  - **Note that in the Blue Martini system a registered user will have all of his/her cookies combined into a single visitor ID**

- **Comparing visitors who viewed gear swap with those who did not**

  - **Several non-bot sessions have 1 request that just visited the MEC homepage (Main/home.jsp)**

  - **To get to the Gear Swap section you have to click at least twice**

  - **To make a fair comparison we have excluded all 1 request sessions that just visited the MEC homepage (Main/home.jsp) from the following analysis**

Distribution of Gear Swap Visitors

- **Visitors who viewed Gear Swap pages had a 62% higher visitor to purchaser conversion ratio as compared to those who did not view Gear Swap**

Overall

Visitors:	X
MEC members:	Y
Purchasing Customers:	Z

Visitors who ever viewed Gear Swap

Visitors:	14.3% of X
MEC members:	20.8% of Y
Purchasing Customers:	21.1% of Z

Visitors who never viewed Gear Swap

Visitors:	85.7% of X
MEC members:	79.2% of Y
Purchasing Customers:	78.9% of Z
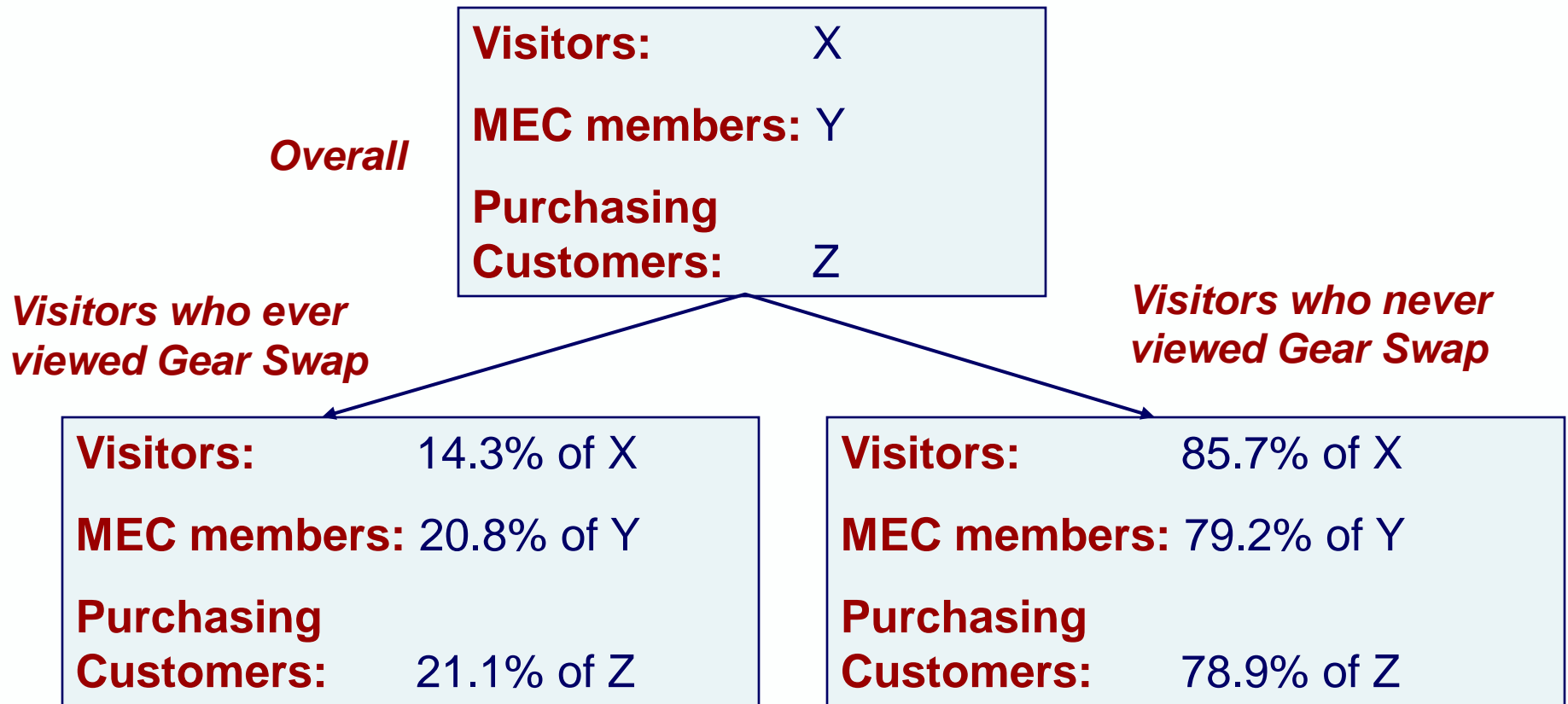
Distribution of Orders (the real ROI)

BLUE MARTINI

Overall

Orders: X

Average Basket Value: $Y

Visitors who ever viewed Gear Swap

Visitors who never viewed Gear Swap

Orders: X

Average Basket Value: 1.05 * Y

Orders: 3,875 (78.3%)

Average Basket Value: 0.98*Y

Distribution of Visits

- **Although, Gear Swap visitors have lower visit-to-purchase conversion than non Gear Swap visitors, they visit more often and their overall visitor-to-purchase conversion is higher**

Overall

Visits:	X

Visitors who ever viewed Gear Swap

Visitors who never viewed Gear Swap

Visits:　　24.8% of X

Visit to Purchase Conversion:　　1.94%

Visits:　　75.2% of X

Visit to Purchase Conversion:　　2.3%
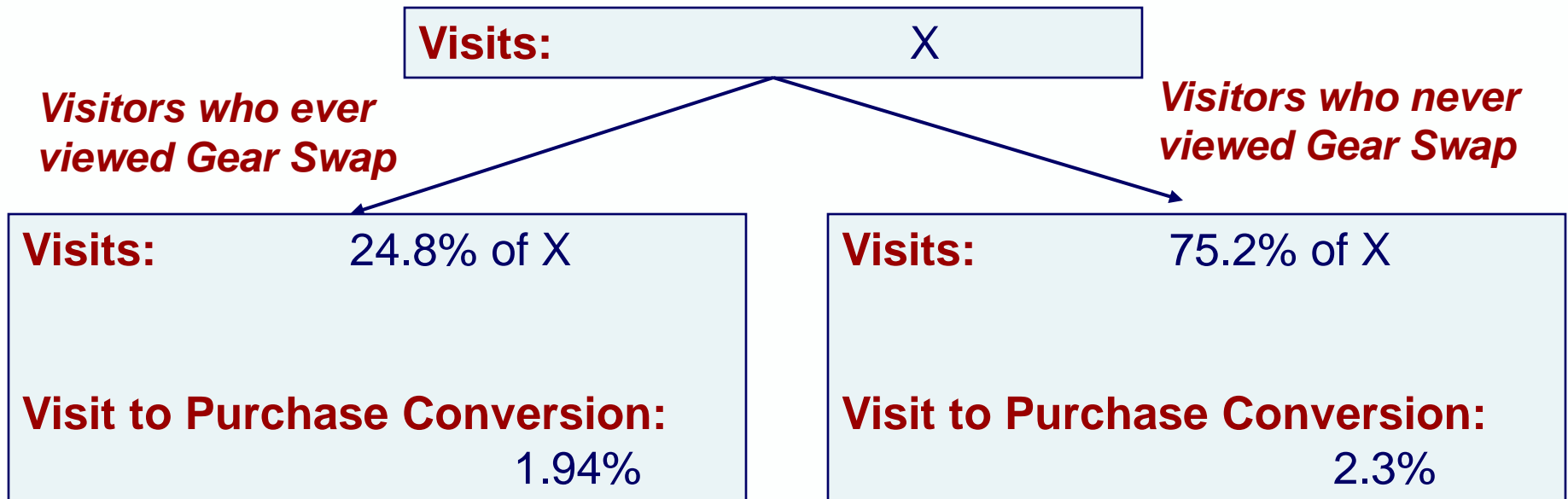
Effectiveness of FYI Pages

- **People viewing FYIs are more likely to purchase**

Viewed FYI

    Visits: 6.2% of all

    Purchases: 23% of all

    Visit-to-Purchase: 7.1%

    Avg. Sales per Visit: 6.1X

Did Not View FYI

    Visits: 93.8% of all

    Purchases: 77% of all

    Visit-to-Purchase: 1.2%

    Avg. Sales per Visit: $X

Recommendation:
Controlled experiment to study the effect of FYIs

FYIs (Cont.)

- **Setting up controlled experiments to study the cause-effect relationship of FYI**

  - Select a handful of products (say 6) for introducing FYIs

  - Randomly show the new FYIs to 50% of the visitors viewing these products and don't show the FYIs to the other 50% of the visitors

  - At the end of the trial period (say 2-3 weeks) measure the visit-to-purchase conversion of the two groups

  - Determine if there is a significant difference in the visit-to-purchase conversion of the two groups
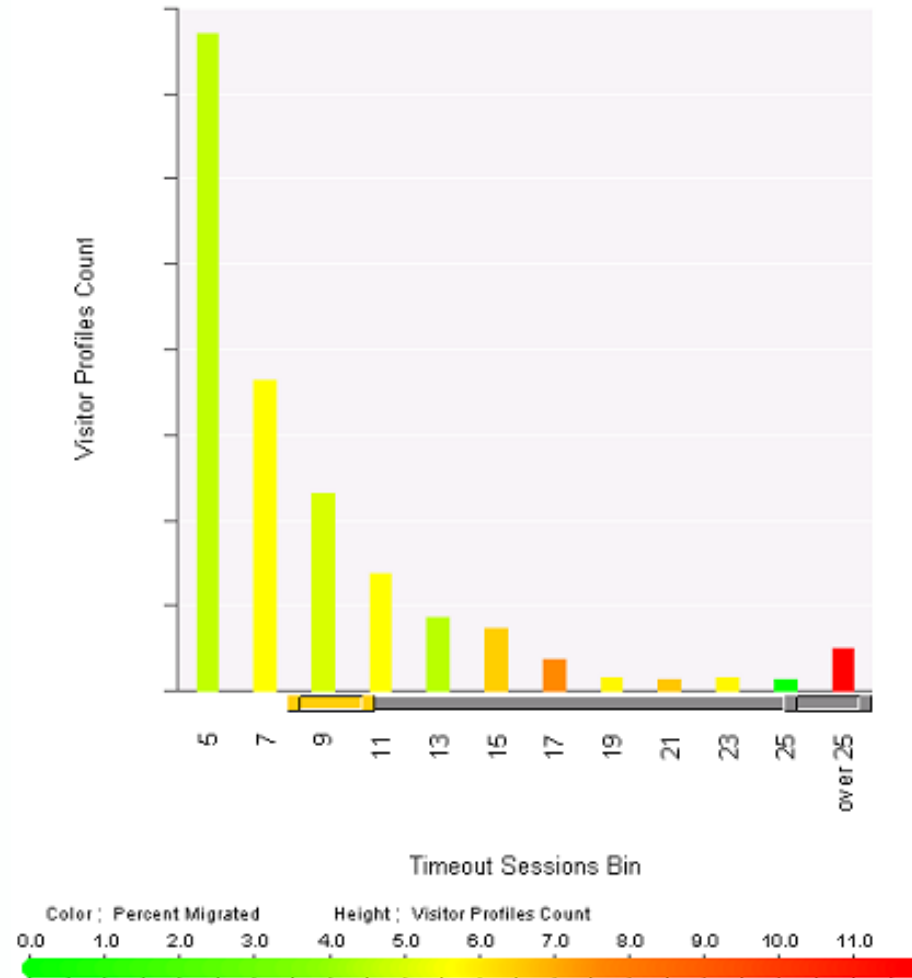
Debenhams Migrators: Timeout

Some attributes are more useful when combined with other attributes

For each visitor we computed the number of sessions which went to the page timeout.jsp

This was binned as shown on the X axis of the chart

The height shows the number of visitors in each bin and color shows the percentage of those visitors who migrated

Just looking at this variable alone it is difficult to tell what the pattern is
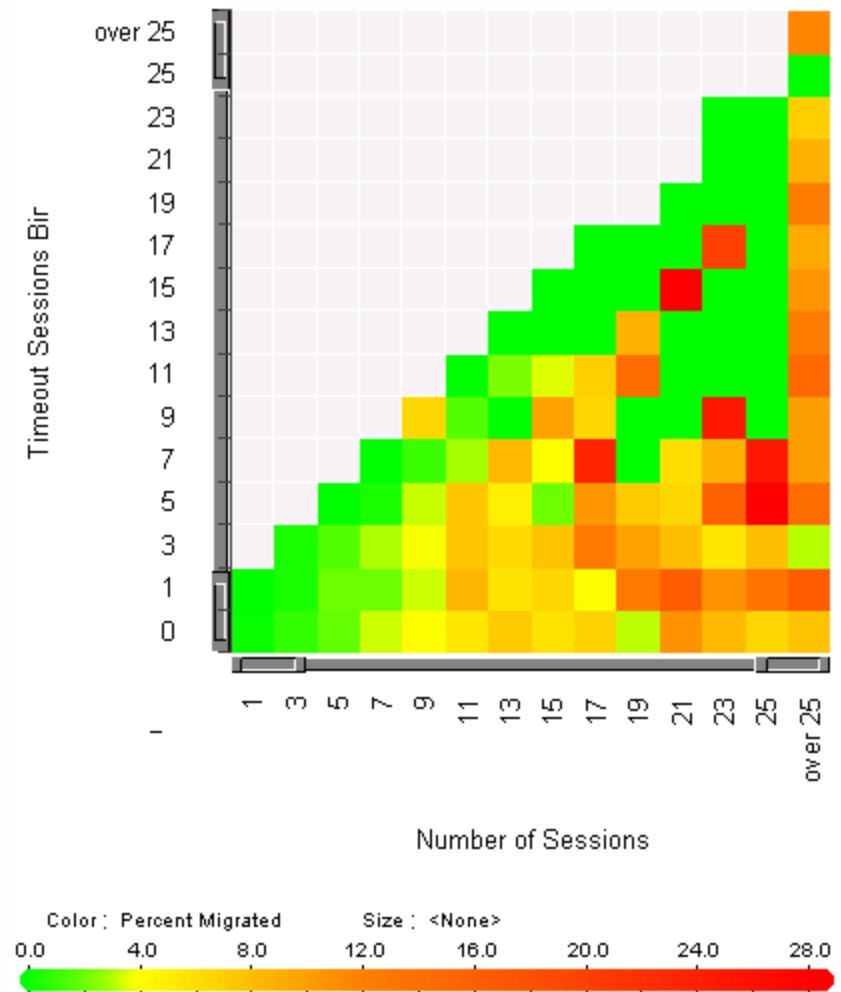

Timeout Sessions Bin

Color : Percent Migrated   Height : Visitor Profiles Count

0.0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0 9.0 10.0 11.0

Migrators: Timeout

By combining the number of timeout sessions with the total number of sessions for each visitor a pattern emerges

In this heatmap the X axis shows the total number of sessions, the Y axis shows the number of timeout sessions, and color shows the percentage of migrators at each pair of values

The green along the diagonal shows that people who got the timeout page for a high percentage of their sessions are less likely to migrate
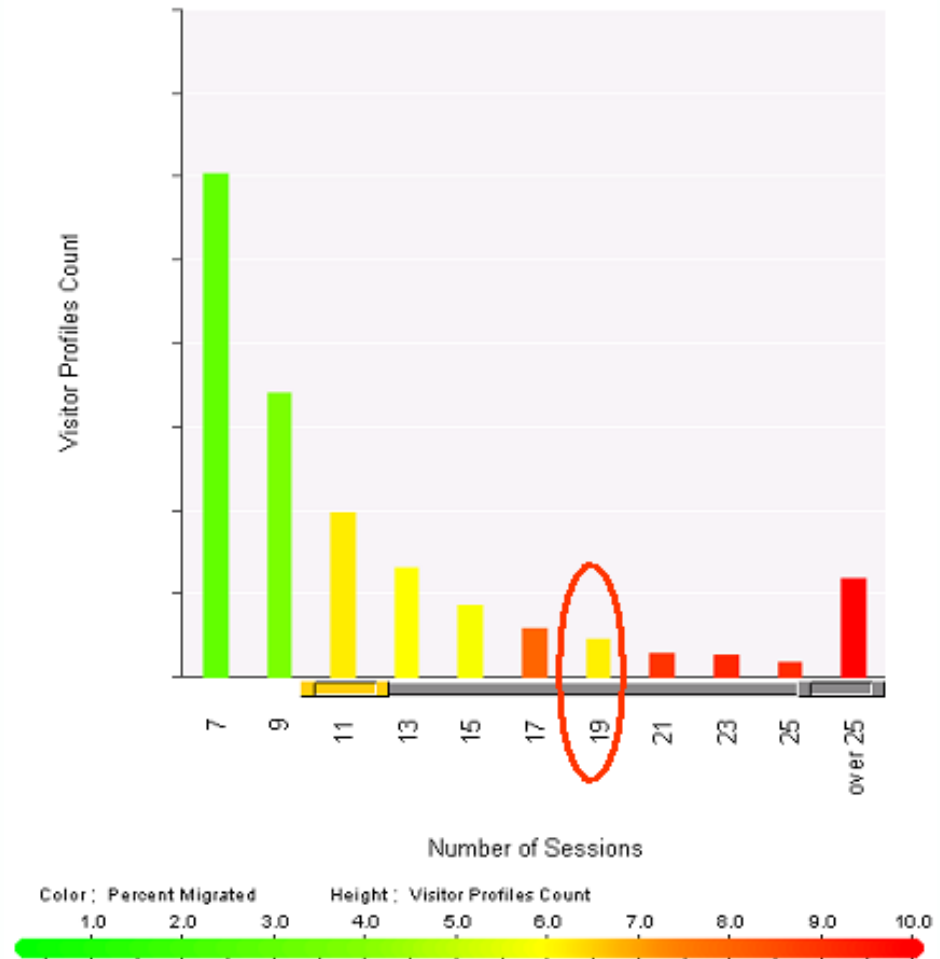
Migrators: Timeout

The number of sessions a visitor has is a good indicator of whether or not they will migrate

However there are some inconsistencies that are apparent. For example, why does the percent of visitors who migrate drop at 19 sessions?

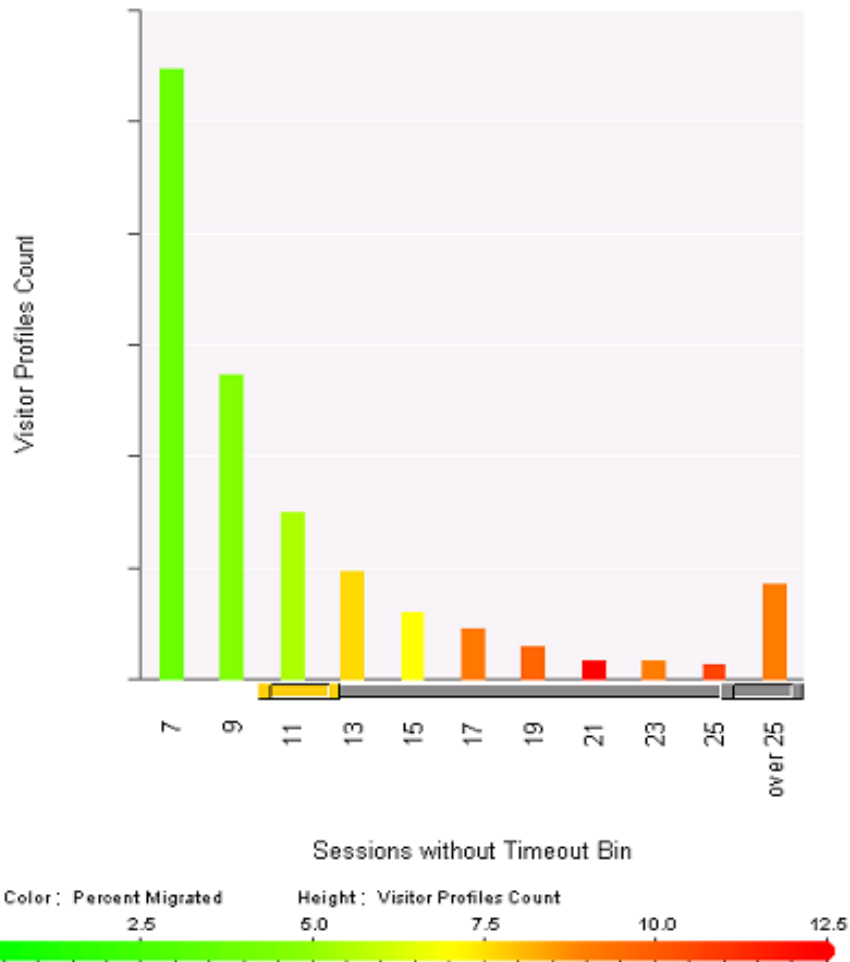We can construct new attributes based on the relationship we saw between the number of timeouts and the number of sessions

Two more attributes can be created:

• Number of sessions that did not time out
• Percentage of sessions that did not time out

Number of Sessions

Color : Percent Migrated   Height : Visitor Profiles Count
1.0   2.0   3.0   4.0   5.0   6.0   7.0   8.0   9.0   10.0

Migrators: Timeout

Number of sessions without timeout is a good predictor of migration
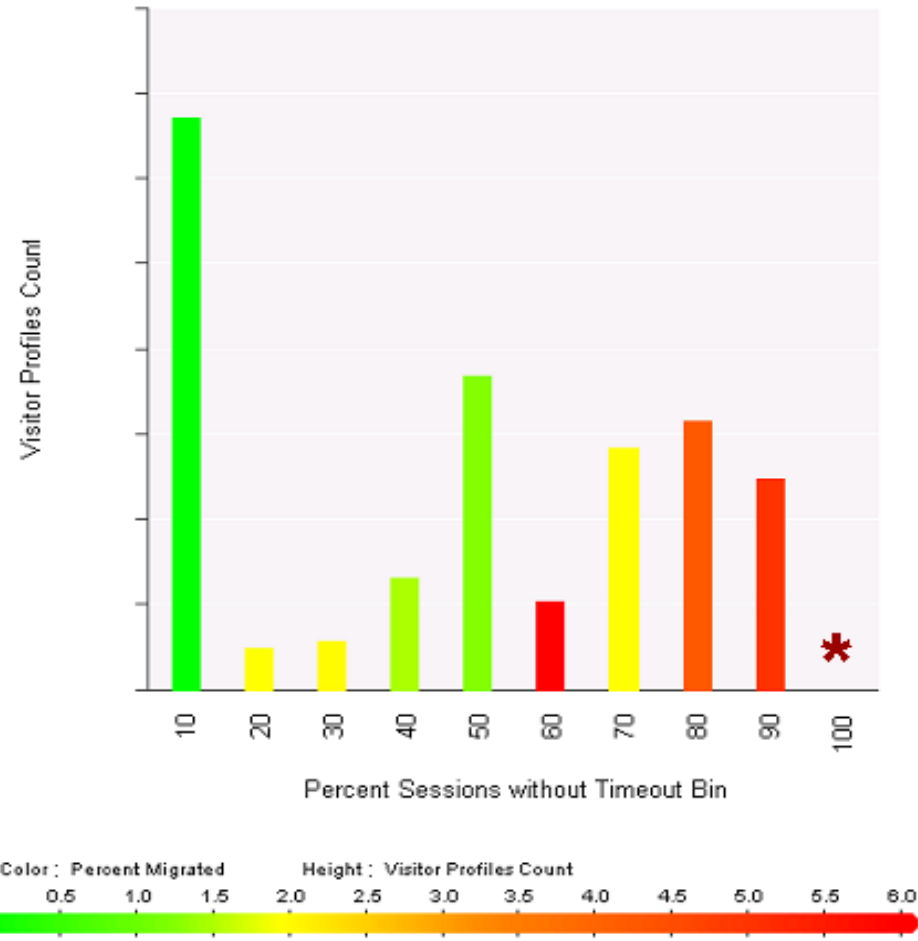
Percentage of sessions without timeout is also a good indicator of migration

* **68,000 visitors with no timeout sessions have been filtered out**

Distribution of Wedding Purchases over Time

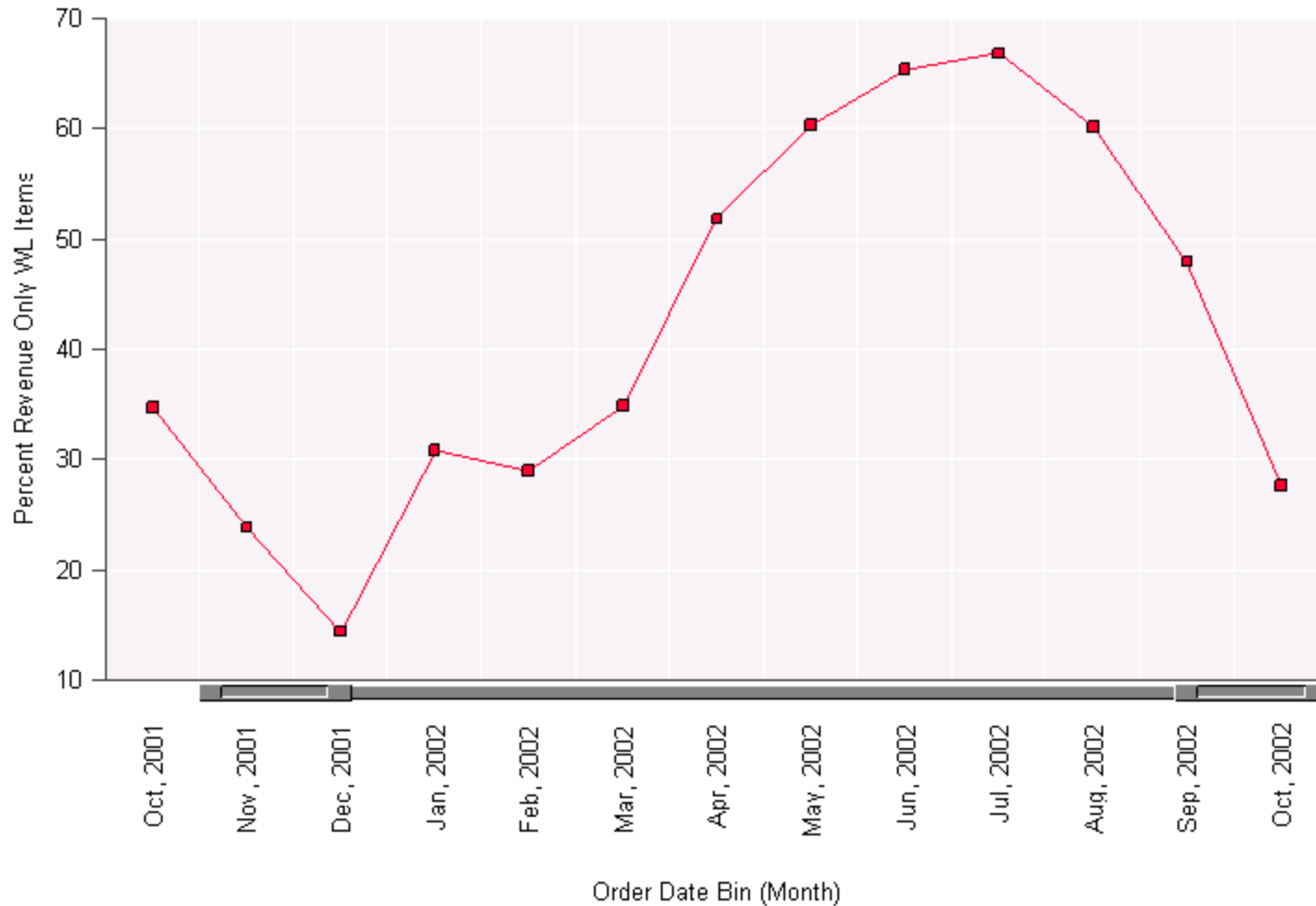**Revenue due to wedding list item purchases clearly affected by summer weather, when weddings are more common in the UK**

Acxiom

- **BMS supports ADN – Acxiom Data Network**

- **Seamless integration: get username/password
  Note: Acxiom recently changed their interface, so you will
  need a patch**

- **Comprehensive collection of US consumer and telephone
  data available via the internet**

  - **Multi-sourced database**

  - **Demographic, socioeconomic, and lifestyle information.**

  - **Information on most U.S. households**

  - **Contributors' files refreshed a minimum of 3-12 times per year.**

  - **Data sources include:** County Real Estate Property Records, U.S.
    Telephone Directories, Public Information, Motor Vehicle Registrations,
    Census Directories, Credit Grantors, Public Records and Consumer
    Data, Driver's Licenses, Voter Registrations, Product Registration
    Questionnaires, Catalogers, Magazines, Specialty Retailers, Packaged
    Goods Manufacturers, Accounts Receivable Files, Warranty Cards

Example - Income

BLUE MARTINI

- **Graph showing incomes for a company that targets high-end customers based on POS purchases**

- **Income of their customers in blue**

- **The US population in red**
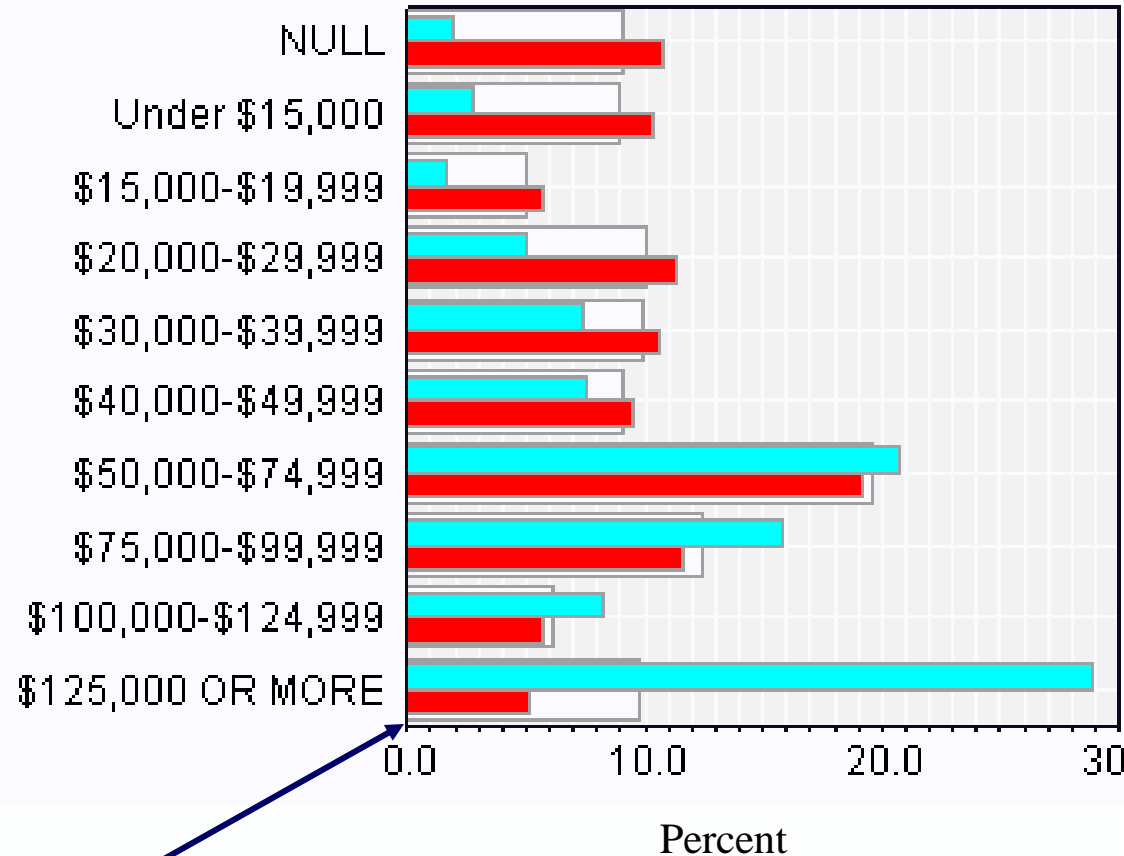
Note highest bracket (30% vs. 5% for US)
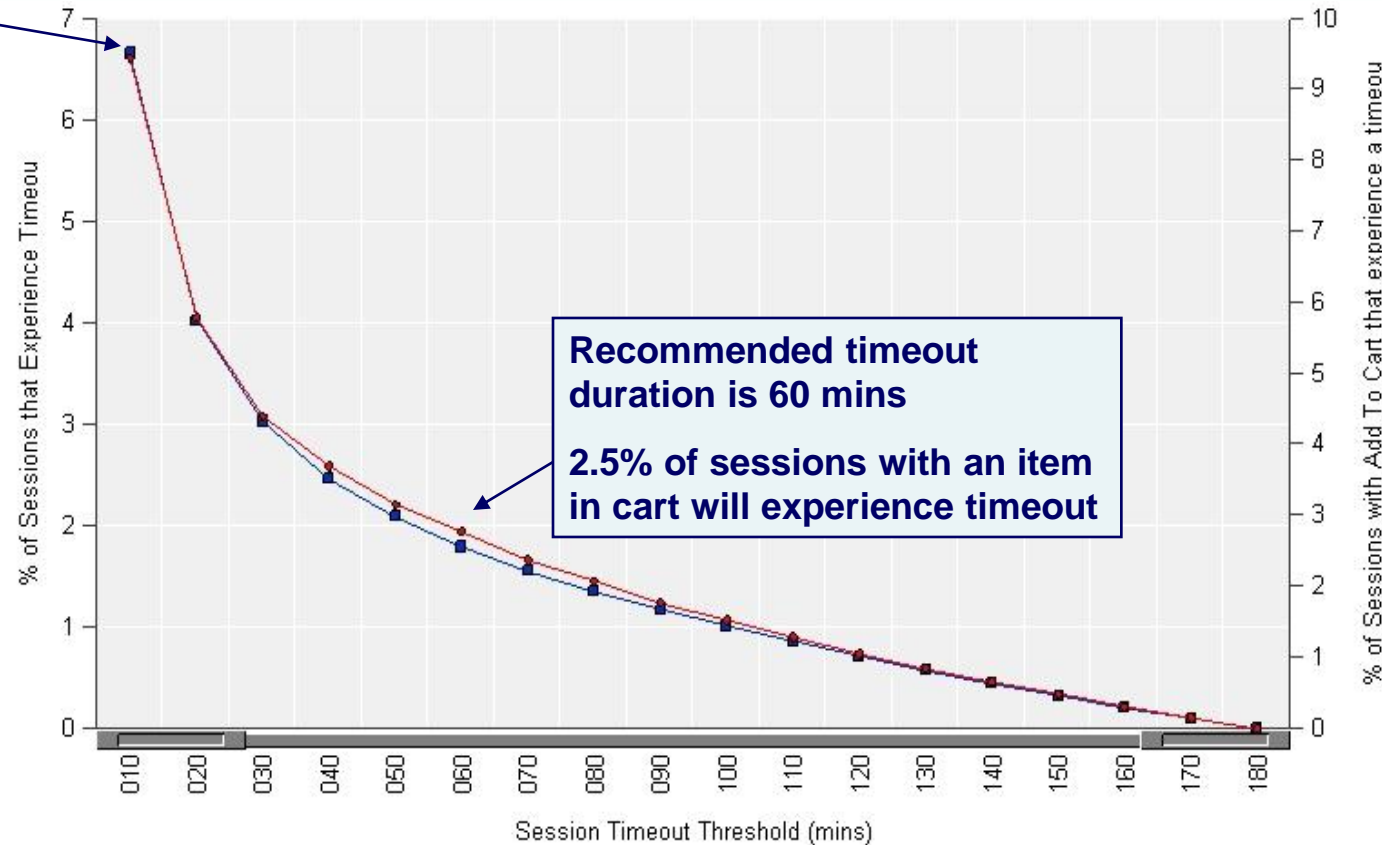
Setting Session Timeout (Debenhams)

Debenhams set session timeout to 10 minutes to reduce memory footprint.

9.5% of visitors with an item in the cart lost it when they came back within 3 hours

Look for upcoming article by us on developer summarizing this

RFE filed to automatically extend sessions with carts

RFE filed to remove bot sessions (one-click) immediately to reduce memory footprint

Recommended timeout duration is 60 mins

2.5% of sessions with an item in cart will experience timeout

Session Timeout Threshold (mins)

Left axis	Right axis
■ % of Sessions that Experience Timeout	◆ % of Sessions with Add To Cart that experience a timeout

World Wide Revenue Detail

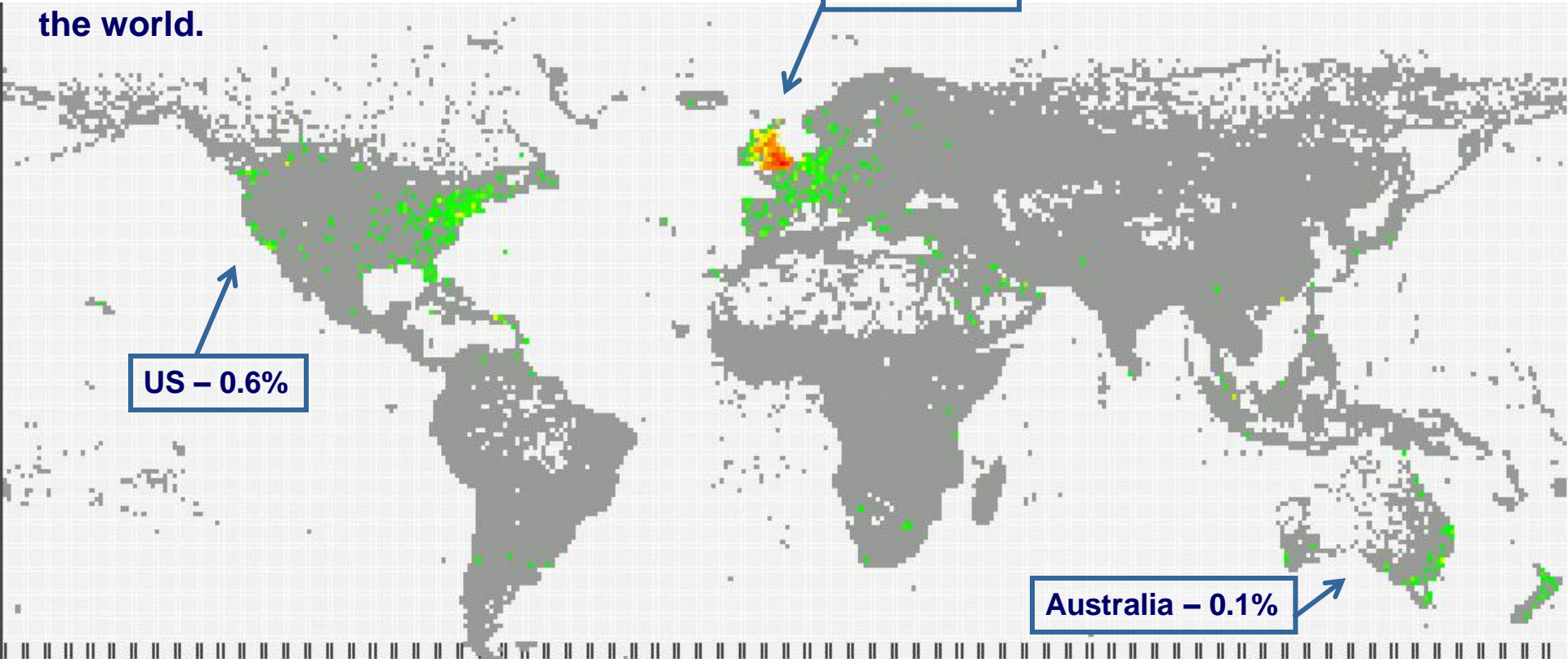**Although Debenhams online site only ships in the UK, we see some revenue from the rest of the world.**

BLUE MARTINI

UK – 98.8%

US – 0.6%

Australia – 0.1%

Low   Medium   High

NOTE: About 50% of the non-UK orders are wedding list purchases

Acxiom Integration

- **Web behavior is one axis**

- **Demographic information is another**

- **Blue Martini provides an extremely tight integration with Acxiom:**

  – **Sign an agreement to get a password**

  – **DSSGen will pull information from Acxiom over the internet as a part of building the data warehouse or as an option for an existing warehouse**

  – **ZERO effort. No tapes, no customizations needed**

Consumer Demographics

- **Using Acxiom, we compared online shoppers to a sample of the population**

    - **People who have a Travel and Entertainment credit card are 48% more likely to be online shoppers (27% for people with premium credit card)**

    - **People whose home was built after 1990 are 45% more likely to be online shoppers**

    - **Households with income over $100K are 31% more likely to be online shoppers**

    - **People under the age of 45 are 17% more likely to be online shoppers**

Demographics - Income

- **A higher household income means you are more likely to be an online shopper**

BLUE MARTINI
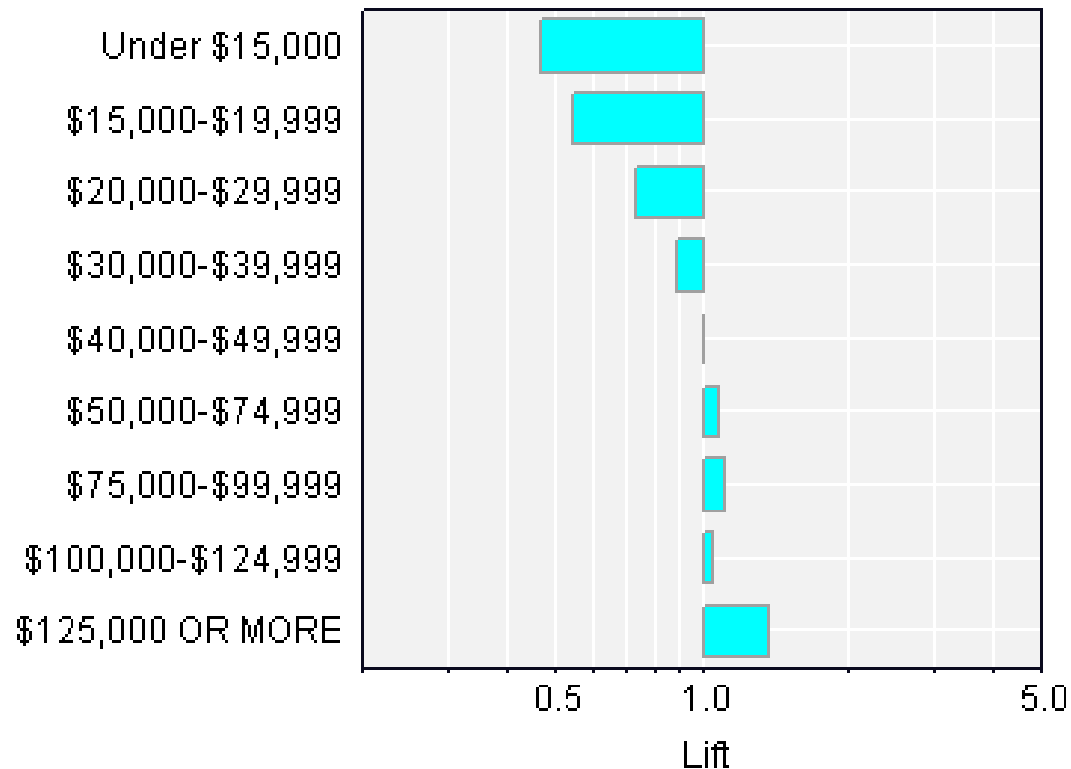
- **The more credit cards, the more likely you are to be an online shopper**