



Gaining shopper insights using market basket analysis

Innar Liiv, Ph.D

Tallinn University of Technology, Estonia

innar.liiv@ttu.ee

World of **contradictions**

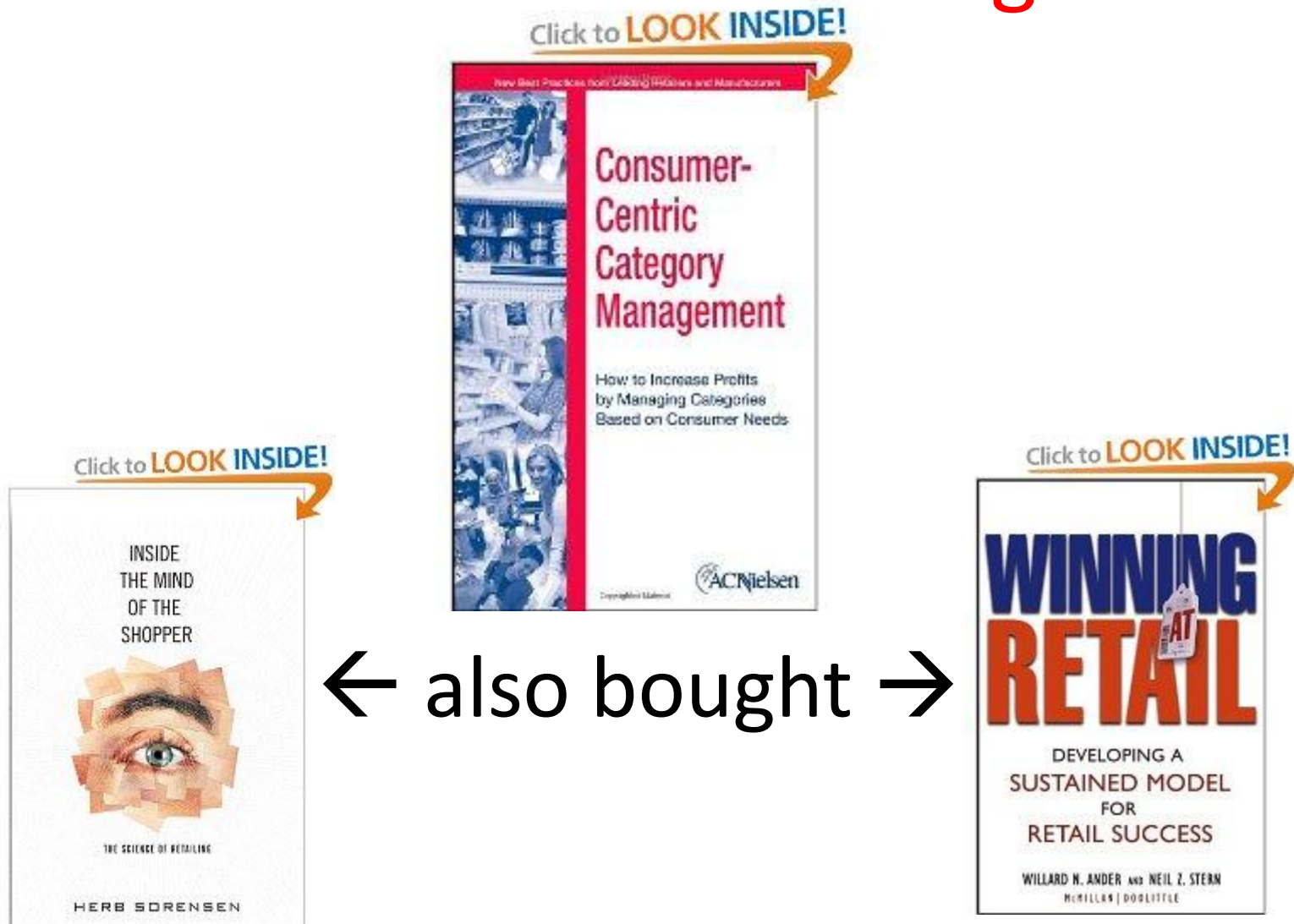
“Surely, if knowledge is power, then both, **retailers and manufacturers could increase their power and their profits** by pooling their knowledge, **sharing information, coordinating marketing** and so on – instead of fighting as usual.” Source: ECR Europe POS Data Management Group

- Yes, in jail (Article 82 EC Treaty)

Agenda

- What is Market Basket Analysis?
- How does it fit in to general category management framework and traditional customer behavior analysis?
- Why should retailers and suppliers care?
- What kind of important business questions does it allow to answer?
- What kind of data is needed?
- Why now? What made and makes it hard to implement in practice?

Customers who bought



(AFFINITY ANALYSIS)

Motivation

- Amazon's revenue in 2009: \$24.5B
- ~\$5B came from product **recommended** using Market Basket Analysis (affinity analysis)
- This works in online shops, but how would it look like in the offline context?
- **Walmart's** RetailLink, Dunnhumby (**TESCO**), APT Market Basket Analyzer (US, **Family dollar** w/ 6600+ shops), IntelliStats Market Basket Analyzer (Castle Rock, CO), Oracle Retail (Retek), NetPerceptions (KD1) Market Basket Analyzer, deployed at **Walgreens, Lowes** and **Meijer**


The Profitable Product Death Spiral

- Using product profitability to drive product decisions can lead to a Profitable Product Death Spiral in which more and more products become unprofitable
- Profitability comes from customers, not products, product decisions should be based on the assortment of products a customer buys
- Keeping some unprofitable products may be necessary to retain some profitable customers
- **Understanding interdependencies between products** is the key factor avoiding that!

Basket composition is the **context!**

- Daily sales data for “Kamajahu”:**Qty:14 / €49**
- Is it a good result?
- **Classical approach:** can't tell, trend matters
- Actually: trend **is not enough to show** that an average basket value including this product was **€200**.

Trends in customer analytics

- 
- confirmatory data analysis
 - Carefully planned
 - Often indirect
 - Samples, generalizations ...
 - exploratory data analysis
 - Data-driven
 - Secondary data
 - We have all the data
 - Attention deficit
 - social data revolution

Category Management

- Early 1990s by The Partnering Group (TPG):
 1. Category definition
 2. Category role
 3. Category assessment
 4. Category scorecard
 5. Category strategies
 6. Category tactics
 7. Plan implementation
 8. Category review

1) **Definition and grouping** by factual shopping behaviour

2) **Basket analysis:** What do people buy during the same shopping trip? How do you build bigger baskets through stronger cross-category connections?

3) Which **categories** are related? What else is in the buyer's basket?

New questions enabled

- How does the shopping behaviour correlate with the category (mgmt) tree?
- What are the **typical** baskets, where my product **X** is in? (category mix %, avg price)
- How does a campaign for my product **X** influence other products, categories?
- How do sales/campaigns in other categories influence my product **X**?

New questions enabled (2)

- What % of customers cherry-picked this product?
- What products are most cherry-picked?
- What is the real proportion of cherry-pickers?
- What categories are often bought together?
- Is it similar in different regions, shops?

New questions enabled (3)

- Are there groups categories, which are profitable, but weakly connected in actual shopping behaviour? (exit, rebranding?)
- Efficient Product Introduction: in what kind of baskets the new product started to appear first? (quick adoption)
- What kind of baskets demonstrated a strongest lift due to the campaign?

What **kind of data** is needed?

- Mandatory Data Profile (SLSRPT) + **Category** flag from Optional Data Profile
(ECR Europe POS Data Management Project Group)
- **Row-level** information from POS receipts:
 - Product
 - Qty
 - (Price)
- Product Category Hierarchy
- Product list with links to category/hierarchy

	A	B	C	D	E	F	G
1	Categories:	7512	7533	7610	7620	8960	8972
2	Basket1	1	1	1	0	0	0
3	Basket2	1	1	1	0	0	0
4	Basket3	1	1	1	0	0	0
5	Basket4	1	1	1	0	0	0
6	Basket5	0	0	0	1	0	0
7	Basket6	0	0	0	1	0	0
8	Basket7	0	0	0	0	1	1
9	Basket8	0	0	0	0	1	1
10	Basket9	0	0	0	0	1	1
11	Basket10	0	0	0	0	1	1

Why now? What **makes it hard**?

- Regardless what service providers say, real-time POS data analytics is still not a reality;
- Legal issues (privacy, competition law);
- Unified product category hierarchy not feasible – business models are not uniform!
- Data quality issues (product & hierarchies)
- Handling changing categories
- Direct, measurable and predictable ROI

Conclusions

- Market basket data is one of the actionable potentials enabled by collaborative POS data management;
- Summarized sales information is not enough, basket-level POS data has to be stored in the data warehouse;
- Can answer lots of interesting and important business questions!

Thank you!

Thoughts, questions?

innar.liiv@ttu.ee
+3725200552