

Marketing Analytics

Aula 7

Professora:

Marina Muradian

Agenda

1. Marketing Analítico no contexto atual
2. Pesquisa em Marketing (dados primários)
3. Pesquisa Experimental (testes A/B)
4. Oferta: produto e serviço (Conjoint)
5. Marca e posicionamento (MDS)
6. Segmentação (Cluster)
7. Propensão a compra (Logit) + Basket Analysis

Apresentação do trabalho e aplicação
das técnicas (exercício final)

Análise de Cesta de Compras

Basket Analysis

Porque o comportamento do Shopper diz mais do que sua idade ou seu nível sócio econômico



Quais os dados?

		 	
SELL OUT	100%	100%	100%
DADOS DE CLIENTES			
CLIENTES	+/- 1,6 MM	+/- 3,7 MM	+/- 14,4 MM

Segmentação
Hábito de
Compra



Fazer o que é certo para seu
consumidor é a chave



Se o GPA perder **1 cliente fiel**, ele precisará de
**14 clientes descompromissados para
compensar**



dunnhumby

Hábitos de Compra - Método

		Padrão de Visita							
		Todo dia	2 X semana	Semanal	Intermitente	De vez em quando	Quase nunca	Em abandono	Abandono
Gastos	Alto	Premium	Valioso					Em abandono	Inativos
	Médio	Valioso	Potenciais						
	Baixo		Descomprometidos						

Clientes são classificados em função de sua frequência, recencia e seus gastos (modelo RFV)

O mesmo cliente pode participar de diferentes missões e se comportar de forma diferente

Missões de Compra



Abastecimento



Reposição



Refeição para hoje



On the Go



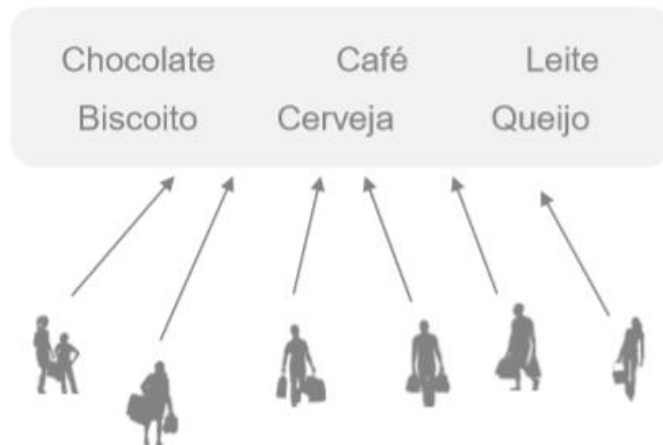
Eventos especiais/Presentes



Não Alimentos

Promoção de Produtos?

Targeting Padrão



Targeting Padrão

A partir de um conjunto de ofertas, identifica o melhor grupo de clientes

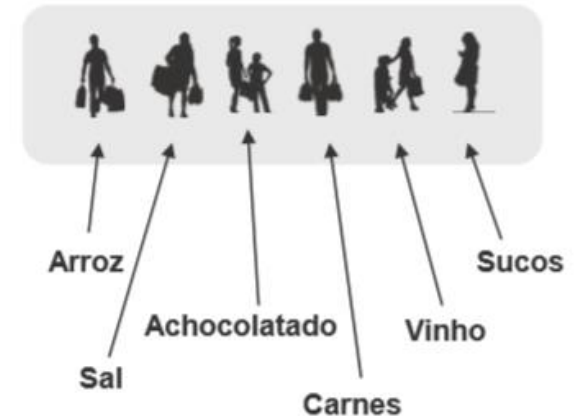
Ofertas Temáticas

Estimular vendas cruzadas e trade-up

Ofertas Exclusivas

Incentivar experimentação e educar clientes

Alocação por Otimização



Alocação por Otimização

A partir do cliente, identifica o melhor conjunto de ofertas

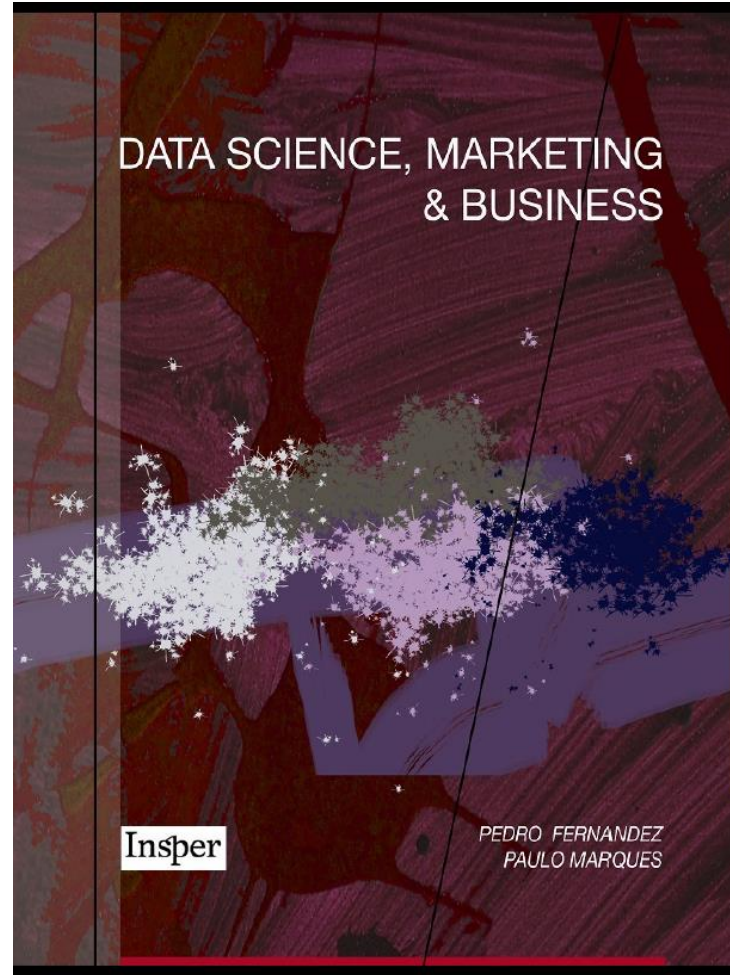
Ofertas Personalizadas

Aumentar e manter os clientes fiéis

Resultado



Aplicação – pag. 27



<https://datascience.insper.edu.br/datascience.pdf>

Sacola de Compras

Suponha que $\Omega = \{L, C, S, R, V\}$ é um conjunto formado pelos itens:

- (L) leite
- (C) café
- (S) suco
- (R) refrigerante
- (V) vinho

Uma **sacola de compras** B é um subconjunto (não vazio) de Ω que representa os itens adquiridos por um consumidor em determinada compra.

Nossos dados consistem em um conjunto de sacolas de compras:

$$B_1 = \{L, C, R\}$$

$$B_4 = \{L, C, S\}$$

$$B_2 = \{L, C, R, V\}$$

$$B_5 = \{C, S\}$$

$$B_3 = \{L, R\}$$

$$B_6 = \{S, R\}$$



Insper

Suporte

Definição: para um conjunto de itens $X \subset \Omega$, o **suporte** de X ($supp(X)$) é a fração de sacolas de compras que contém os itens de X

$$X_1 = \{L\} \Rightarrow supp(X_1) = \frac{4}{6}$$

$$X_2 = \{C\} \Rightarrow supp(X_2) = \frac{4}{6}$$

$$X_3 = \{S\} \Rightarrow supp(X_3) = \frac{3}{6}$$

$$X_4 = \{R\} \Rightarrow supp(X_4) = \frac{4}{6}$$

$$X_5 = \{L, C\} \Rightarrow supp(X_5) = \frac{3}{6}$$

$$X_6 = \{L, R\} \Rightarrow supp(X_6) = \frac{3}{6}$$

$$X_7 = \{C, S\} \Rightarrow supp(X_7) = \frac{2}{6}$$

$$X_8 = \{C, R\} \Rightarrow supp(X_8) = \frac{2}{6}$$

$$X_9 = \{L, C, R\} \Rightarrow supp(X_9) = \frac{2}{6}$$

Ou, em termos de interpretação probabilística
 $supp(X_1) = \Pr(X_1) = \Pr(\{L\})$

Quanto mais itens houver em X , menor o suporte, pois seria menos provável selecionarmos uma sacola de compras em que todos os itens estivessem presentes



Insper

Regra de Associação

Dados dois conjuntos de itens, X e Y , tais que $X \cap Y = \emptyset$, uma regra de associação é uma expressão simbólica:

$$X \Rightarrow Y$$

Em que X é o antecedente e Y o consequente.

Suporte da Regra de Associação

$$\text{supp}(X \Rightarrow Y) = \text{supp}(X \cup Y)$$

Ou seja, a fração de cestas de compra que **contém todos os itens de X e Y** .

Exemplo:

$$\text{supp}(\{L, C\} \Rightarrow \{R\}) = \text{supp}(\{L, C\} \cup \{R\}) = \text{supp}(\{L, C, R\}) = \frac{2}{6}$$

Regra de Associação

Confiança da Regra de Associação

A confiança de uma regra de associação $X \Rightarrow Y$ é definida:

$$conf(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X)}$$

Ou, em termos de interpretação probabilística:

$$conf(X \Rightarrow Y) = \Pr(Y|X)$$

Exemplo:

$$conf(\{L, C\} \Rightarrow \{R\}) = \frac{supp(\{L, C\} \cup \{R\})}{supp(\{L, C\})} = \frac{2/6}{3/6} = \frac{2}{3}$$

A estratégia ao minerar regras de associação é estabelecer o **suporte mínimo** para limitar o número de regras que iremos examinar e, depois disso, procuramos aquelas com a **confiança mais alta**.



PROBLEMA!

$$\text{conf}(\{L, C\} \Rightarrow \{R\}) = \frac{\text{supp}(\{L, C\} \cup \{R\})}{\text{supp}(\{L, C\})} = \frac{2/6}{3/6} = \frac{2}{3}$$

Embora a confiança acima seja $\frac{2}{3}$, temos que:

$$\text{supp}(\{R\}) = \text{Pr}(\{R\}) = \frac{4}{6} = \frac{2}{3}$$

E como

$$\text{conf}(\{L, C\} \Rightarrow \{R\}) = \boxed{\text{Pr}(\{R\}|\{L, C\}) = 2/3 = \text{Pr}(\{R\})} = \text{supp}(\{R\}),$$

Temos que $\{L, C\}$ e $\{R\}$ são eventos **independentes**.

Assim, precisamos de uma métrica adicional, que meça a **dependência** dos eventos definidos pelo antecedente e consequente da regra de associação considerada.



Insper

Lift (levante)

Medida de “surpresa” ao encontrarmos os itens X e Y juntos numa mesma sacola de compras, tomando como referência a situação em que X e Y sejam independentes.

$$lift(X \Rightarrow Y) = \frac{supp(X \cup Y)}{supp(X) * supp(Y)}$$

De modo que, se X e Y forem independentes, $lift(X \Rightarrow Y) = 1$.

Exemplo:

$$lift(\{L, C\} \Rightarrow \{R\}) = \frac{supp(\{L, C\} \cup \{R\})}{supp(\{L, C\}) * supp(\{R\})} = \frac{2/6}{\frac{3}{6} * \frac{4}{6}} = 1(\text{independência})$$

Portanto, desejamos que a medida de *lift* seja maior do que 1.



Estratégia para Mineração

- 1) Fixamos um suporte mínimo para limitarmos o número de regra de associação que iremos avaliar.
- 2) Seleccionamos as regras com confiança alta
- 3) Seleccionamos a as regras de associação com *lift* > 1

Faremos isso na mão? Não! Utilizaremos o algoritmo de data mining “apriori”



EXERCÍCIO EM GRUPO

- Varejista online (UK)
- Mix: produtos para presente
- Clientes: atacadistas/varejistas
- Base foi extraída do sistema da empresa:
 - 406,829 registros e 10 campos
 - cupom fiscal (c: cancelado)
 - id do produto (item: SKU)
 - descrição do produto
 - quantidade
 - dez/10 - jun/11
 - preço
 - id cliente
 - País
- Baixar a base: **retail.csv**

EXERCÍCIO EM GRUPO

- Quando os clientes compram (horário, dia de semana)?
- Quantos produtos os clientes costumam comprar por transação?
- Quais são os top 10 produtos vendidos?
- Criar regras de associação para a base
- Escolher um tipo de produto (pode ser top10 ou por algum outro critério de seleção) e descrever as regras de associação válidas.
- Como poderíamos aumentar as vendas deste produto/para esse tipo de cliente?

**PREPARAR UM
RMARKDOWN COM A
RESOLUÇÃO E ENVIAR
VIA BLACKBOARD ATÉ
29/11**