

IMRB: Segmenting Consumers of Bath Soap

Business Situation

The Indian Market Research Bureau (IMRB) is a leading market research agency that specializes in tracking consumer purchase behavior in consumer goods (both durable and non-durable).

IMRB tracks about 30 product categories (e.g. detergents, etc.) and within each category, about 60 – 70 brands. To track purchase behavior, IMRB has constituted about 50,000 household panels in 105 cities and towns in India, covering about 80% of the Indian urban market. (In addition to this, there are 25,000 sample households selected in rural areas; however, we are working with only urban market data.) The households are carefully selected using stratified sampling. The strata are defined on the basis of socio-economic status, and the market (a collection of cities).

IMRB has both transaction data (each row is a transaction) and household data (each row is a household), and, for the household data, maintains the following information:

- Demographics of the households (updated annually);
- Possession of durable goods (car, washing machine, etc.; updated annually) and a computed "affluence index" on this basis;
- Purchase data of product categories and brands (updated monthly).

IMRB has two categories of clients: (1) Advertising agencies who subscribe to the database services. They obtain updated data every month and use it to advise their clients on advertising and promotion strategies. (2) Consumer goods manufacturers who monitor their market share using the IMRB database.

Key Problems

IMRB has traditionally segmented markets on the basis of purchaser demographics. They would like now to segment the market based on two key sets of variables more directly related to the purchase process and to brand loyalty:

1. Purchase behavior (volume, frequency, susceptibility to discounts, and brand loyalty), and
2. Basis of purchase (price, selling proposition)

Doing so would allow IMRB to gain information about what demographic attributes are associated with different purchase behaviors and degrees of brand loyalty, and more effectively deploy promotion budgets.

The better and more effective market segmentation would enable IMRB's clients to design more cost-effective promotions targeted at appropriate segments. Thus, multiple promotions could be launched, each targeted at different market segments at different times of a year. This would result in a more cost-effective allocation of the promotion budget to different market-segments. It would also IMRB to design more effective customer reward systems and thereby increase brand loyalty.

Data File: IMRB.xls

The data profile each household – each row contains the data for one household.

Member Identification	ID		Unique identifier for each household
Demographics	SEC	1 – 4 categories	Socio Economic Class (1=high, 4=low)
	FEH	1 – 3 categories	Food eating habits (1=vegetarian, 2=veg. but eat eggs, 3=non veg., 0=not specified)
	MT		Native language (see table in worksheet)
	SEX	1: male 2: Female	Sex of homemaker
	AGE		Age of homemaker
	EDU	1 – 9 categories	Education of homemaker (1=minimum, 9 = maximum)
	HS	1 - 9	Number of members in the household
	CHILD	1 – 4 categories	Presence of children in the household
	CS	1 - 2	Television available. 1: Available 2: Not Available
	AFF		Affluence index – weighted value of durables possessed

Summarized Purchase Data

Purchase summary of the house hold over the period	brands	Number of brands purchased
	brand_runs	Number of instances of consecutive purchase of brands
	total_volume	Sum of volume
	num_trans	Number of purchase transactions; Multiple brands purchased in a month are counted as separate transactions
	value	Sum of value
	trns_per_brndrun	Avg. transactions per brand run
	vol_per_tran	Avg. volume per transaction
	ave_price	Avg. price of purchase
Purchase within Promotion	promo_none	Percent of volume purchased under no-promotion
	promo_brand	Percent of volume purchased under Promotion Code 6
	promo_other	Percent of volume purchased under other promotions
Brand wise purchase	brand_lux, etc.	Percent of volume purchased of the brand
Price category wise purchase	price_popular, etc	Per cent of volume purchased under the price category
Selling proposition wise purchase	prop_beauty, etc.	Percent of volume purchased under the product proposition category

Measuring Brand Loyalty

Several variables in this case deal measure aspects of brand loyalty. The number of different brands purchased by the customer is one measure. However, a consumer who purchases one or two brands in quick succession then settles on a third for a long streak is different from a consumer who constantly switches back and forth among three brands. So, how often customers switch from one brand to another is another measure of loyalty. Yet a third perspective on the same issue is the proportion of purchases that go to different brands – a consumer who spends 90% of his or her purchase money on one brand is more loyal than a consumer who spends more equally among several brands.

All three of these components can be measured with the data in the purchase summary worksheet.

Analysis approach (specific details are in homework assignment 9)

Use k-means clustering to identify clusters of households based on the variables that describe both purchase behavior (including brand loyalty) and basis of purchase.

How should k be chosen? Think about how the clusters would be used. It is likely that the marketing efforts would support 2-5 different promotional approaches.

How should the percentages of total purchases comprised by various brands be treated? Isn't a customer who buys all brand A just as loyal as a customer who buys all brand B? What will be the effect on any distance measure of using the brand share variables as is? Consider using a single derived variable – this has been done for you already by constructing variable “max_single_brand” as the maximum percentage purchased of a single brand.

Select what you think is the best segmentation and comment on the characteristics (demographic, brand loyalty and basis-for-purchase) of these clusters. (This information would be used to guide the development of advertising and promotional campaigns.)