Retail Customer Segmentation using SAS

April 2014

Calgary SAS Users Group meeting

Jenny Chen

Data Science, LoyaltyOne

a LoyaltyOne program

Agenda

- Overview
- Applications
- Objectives
- Types of Segmentation
- A Real Example with SAS Code
- Further Reading

a LoyaltyOne program

Overview

- **Customer Segmentation is the practice of classifying your customers into distinct groups based on the similarities they share with respect to any characteristics you deem relevant to your business**
- **Key components in developing proper, actionable segmentation**
  - Understand business needs and objectives
    - Customer satisfaction
    - Share of wallet
    - Market share
  - Available customer information
    - Demographic
    - Geographic
    - Behavioral
    - Attitudinal

AIR MILES a LoyaltyOne program

Applications

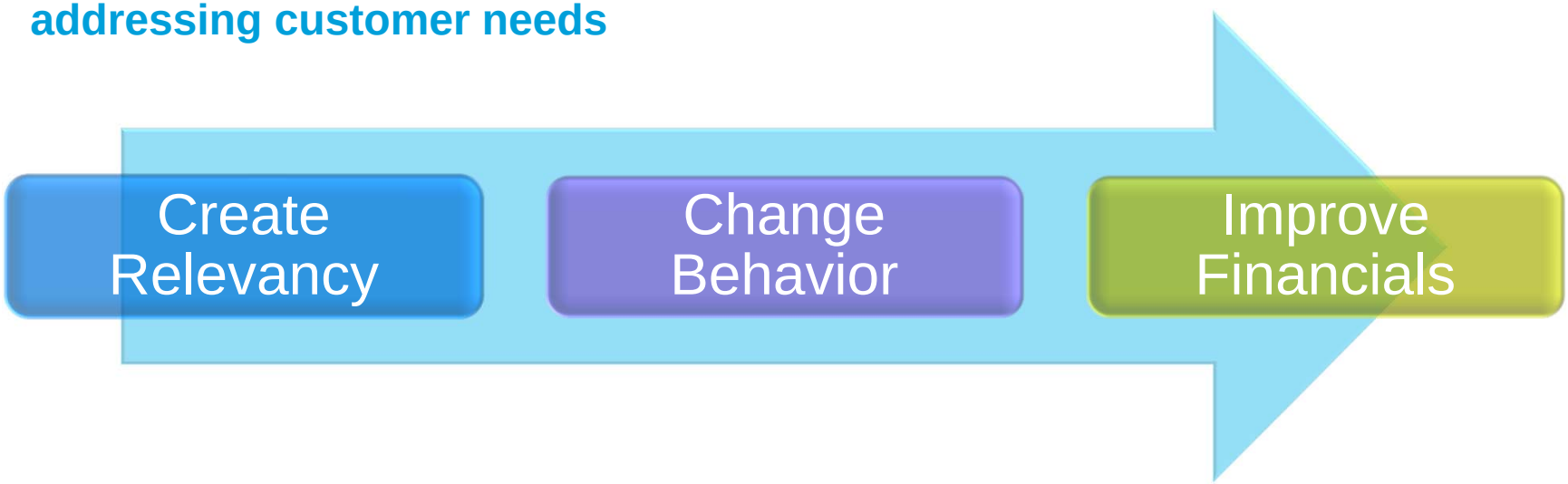
Applications	Examples
Customer Investment Allocation	• Determine segment's current and potential profitability • Define differentiated value proposition by segment
Customer Relationship Management	• Better meet customer needs, improve customer satisfaction • Develop customer loyalty to your brand • Improve retention strategy
Tailor Marketing Program	• Build relevant campaign, communication • Optimize marketing channel mix
Guide New Product Development	• Product preference • Potential targeting audience
Guide Product, Service Pricing	• Price sensitivity • Stimulate customer demand

…and many more

a LoyaltyOne program

Objectives

Segmentation enables business to increase profitability while positively addressing customer needs



- Better defined target universe
- Right offer, communication, channel, time for right customers
- Personalized experience between clients and their customers

- Higher response rates
- Combat attrition of high profit customers
- Grooming lower current value high potential customers
- Identify cross-sell & up-sell opportunities by customer

- Measureable improvement in sales, traffic, revenue
- More cost effective campaign deployment and ROI
- Opportunity identification by category, department, store, etc.

AIR MILES a LoyaltyOne program

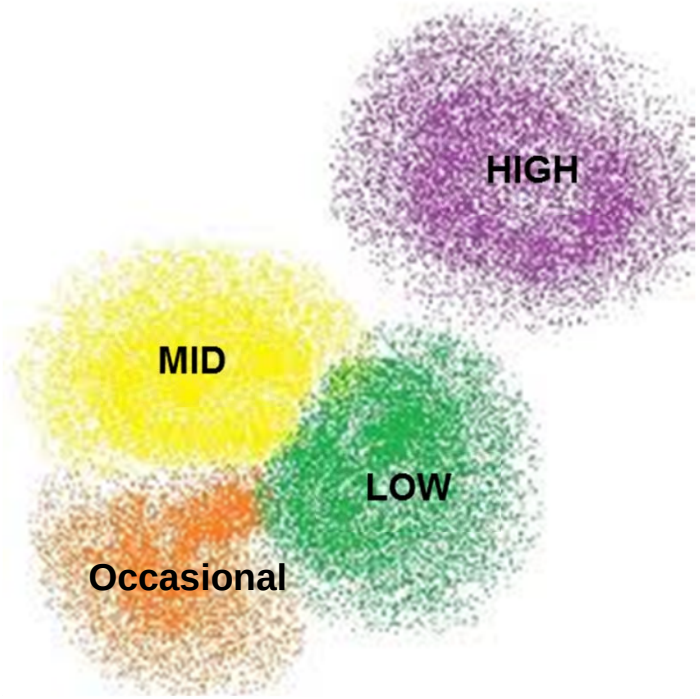
Segmentation Type

Data source

- Demographic, geographic
- Behavior, attitudinal
- Life stage, life style

Approach

- Business rules
  - Profit ranking
  - RFM (Recency, Frequency, Monetary)
- Supervised clustering
  - Decision tree
- Unsupervised clustering
  - K-means clustering



a LoyaltyOne program

A Behavior Based Segmentation Example

Using unsupervised clustering segmentation for a grocery chain which would like better product assortment for its high profitable customers

Potential Inputs

Value	• Basket Size • Visit Frequency
Basket	• Spend by category • Type of category • Brand spend (i.e. private label)
Promotions	• % bought on targeted promotion • % bought from flyer
Time	• Time of day • Day of week
Location	• Store format • Area population density

Clustering approach

Deal Seeking Mom

Key Differentiators

- Full store shop
- High avg. basket size / # trips

- High % purchased on promotion
- Rewards seeker

- High spend categories
  - Fresh produce
  - Organic food
  - Multipack juice, snack

AIR MILES a LoyaltyOne program

Segmentation Building Process I

Define business objective

- Identify high profitable customer groups
- Improve product assortment for them
- Design relevant, effective marketing campaigns for them
- Quantify campaign measurements

Gather information

- Customer transactional data
- Promotional data
- Client store, department, category merchandising hierarchy

Align with client on project scope

- Business Objectives
- Data Sources
- Methodology
- Measurement plan

Segmentation Building Process II

Data preparation

- Data merge purge, data cleaning, dealing with missing
- Convert all attributes to numeric metrics, cap outliers
- Standardize all metrics/variables
- Standardization enables each metric have similar contribution to the future cluster building

```
PROC STDIZE DATA=SAS_DATA.ALL_INFO
OUT=SAS_DATA.STANDARD_STD
OUTSTAT=SAS_DATA.SUMMARY
(WHERE=(_TYPE_ IN ('LOCATION', 'SCALE')));
VAR &ALL_VAR;
RUN;
```

a LoyaltyOne program

Segmentation Building Process III

Data exploration, variable reduction

- Check descriptive statistics, correlation among all potential variables
- Use principle component, factor analysis (**`PROC PRINCOMP, PROC FACTOR`**)
- Use the Variable Clustering node in SAS Enterprise Miner to create variable cluster constellation plot and variable cluster tree diagram

Measure similarity among customers

- Using Euclidean distance - this measures the distance between 2 points in multi-dimensional space
- A very common method of unsupervised clustering with large data using Euclidean distance is K-means clustering

Segmentation Building Process IV

K-means clustering

- Using the distances to group customers into K clusters where each customer is with the nearest centroid
- The centroid is calculated as the multi-dimensional set of the means of the variables used for the particular cluster
- Pre-determine a range of number of clusters, use bottom-up approach
- Test the FASTCLUS procedure repeatedly using different starting points and different number of clusters until we obtain stabilized centroids and desired distance between clusters

```
PROC FASTCLUS DATA=SAS_DATA.STANDARD_STD MAXCLUSTERS=29
OUT=SAS_DATA.RESULT OUTSEED=SAS_DATA.CENTRES
OUTSTAT=SAS_DATA.CLUSTER_STAT REPLACE=RANDOM
RANDOM=187413849 CONVERGE=0.001 MAXITER=350;
VAR &VAR_LIST;
RUN;
```

Output of K-means Clustering

- Centroids info

Cluster Mean									
Cluster	Var1	Var2	Var3	Var4	Var5	Var6	Var7	Var8	Var9
1	(0.68)	(0.28)	(0.37)	(0.20)	(0.77)	0.46	(0.29)	1.11	1.70
2	0.81	0.91	2.44	(0.19)	0.93	0.95	0.41	(0.62)	(0.56)
3	(0.68)	(0.62)	(0.73)	(0.44)	(0.77)	(0.67)	(0.53)	1.11	1.66
4	(0.49)	(0.54)	(0.45)	(0.45)	(0.51)	(0.33)	(0.48)	(0.46)	(0.56)
5	(0.17)	0.04	0.10	(0.02)	(0.13)	0.73	0.70	(0.43)	(0.55)
...									
29	(0.68)	(0.68)	(0.84)	(0.47)	(0.77)	(1.64)	(0.58)	2.27	2.40

- Statistics for variable comparison and further variable reduction
- RSQ/(1-RSQ) is the ratio of between cluster variance to within cluster variance
- Get client's buy-in for the final list of variables for cluster building

Statistics for Variables	
Variable	RSQ/(1-RSQ)
Var1	13.82
Var2	16.51
Var3	6.69
...	
Var 9	56.76

Segmentation Building Process V

Bottom-up approach

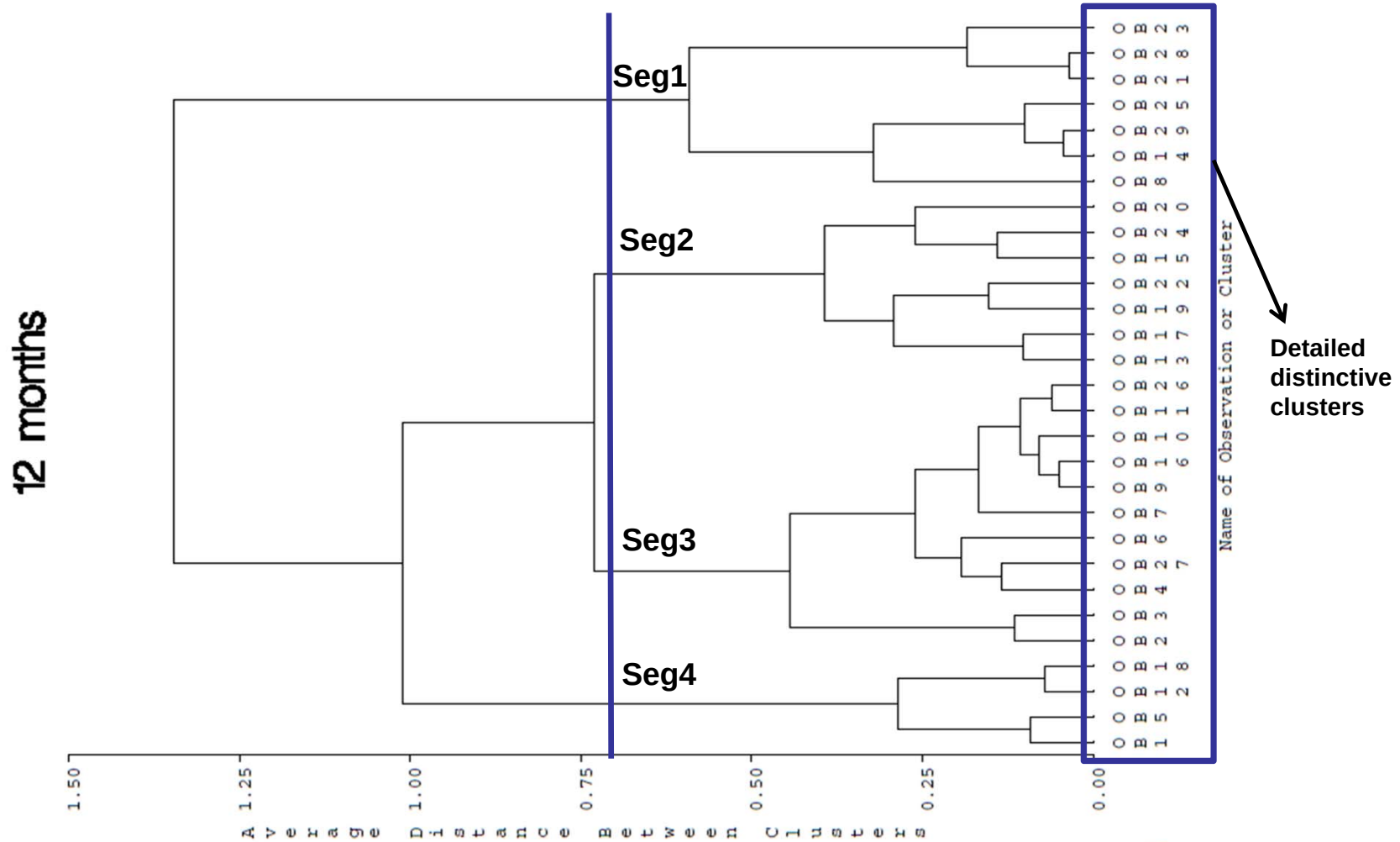
- Flexible with how many clusters to choose initially
- Aggregate clusters up afterwards using hierarchy clustering **PROC CLUSTER**
- Create a tree diagram for clusters using **PROC TREE**
- Choose proper distance to fit into client's request of final number of segments and distribution of segments

```
PROC CLUSTER DATA=SAS_DATA.CENTRES METHOD=AVE OUTTREE=TREE;
VAR &VAR_LIST;
RUN;

PROC TREE DATA=TREE;
RUN;
```

Bottom-up Approach

Tree diagram to illustrate the arrangement of the clusters



a LoyaltyOne program

Segmentation Building Process VI

Segments profiling

- Using cluster building variables to profile the segments
- Using additional data source to profile them as well
- Paint a clear picture of your segments

Validate segmentation effectiveness

- Test campaigns with segments strategy
- Measure campaign results by segments
- Validation segments with market research survey

Review, iterate, deploy again

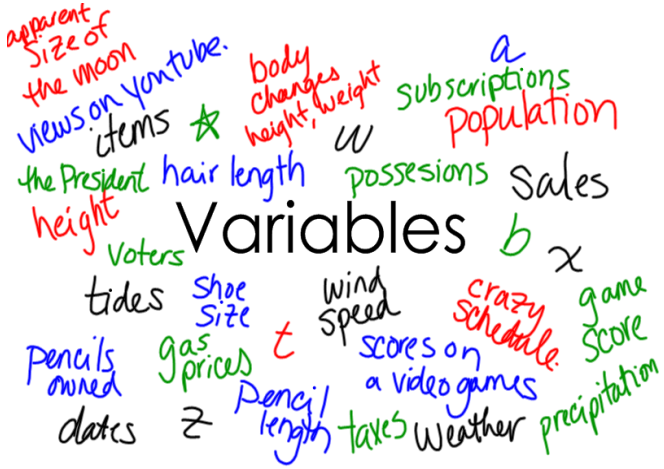
- Update segments regularly
- Monitor the migration of segments
- Gather feedback, campaign response results
- Maintain and improve the process

a LoyaltyOne program

Pros and Cons of K-means Clustering

Pros

- Flexible with fewer assumptions about your customer population
- Can incorporate many relevant metrics
- Provide more comprehensive, diverse segments
- Simple algorithm fast to run with large dataset
- Prove to work well for behavioral targeting



Cons

- Purely data driven
- Pre-specify number of clusters or the minimum distance between clusters
- Time consuming to maintain: for example, certain metrics used to build the clusters were changing overtime dramatically
- May have difficulty to explain the concept to your marketing clients

a LoyaltyOne program

Question or Comment

a LoyaltyOne program

Reference

SAS website

http://support.sas.com/documentation/

http://support.sas.com/resources/papers/

http://support.sas.com/events/sasglobalforum/previous/online.html

- http://support.sas.com/resources/papers/proceedings12/103-2012.pdf
- http://support.sas.com/resources/papers/proceedings12/200-2012.pdf
- http://support.sas.com/resources/papers/proceedings12/329-2012.pdf

A video about LoyaltyOne: Enriching Relationships

http://www.youtube.com/watch?v=NSed-m3Z8Ik

a LoyaltyOne program