Module 3.2

Customer Segmentation and Profiling
Customer Relationship Management

Eric Tham
Institute of Systems Science
National University of Singapore

isstyc@nus.edu.sg

Content

- Customer segmentation & profiling
- Cluster Analysis
  - K-means
- RFM segmentation method
- Summary

Customer Segmentation & Profiling

Some Key Customer Analytics Techniques

Customer Analytics **is a process by which customer data is used to help make business decisions through Segmentation, Customer Profiling and Predictive Analytics**

Segmentation & Profiling

Customer Segmentation

One size does not fit all

Customer segmentation is the practice of dividing a customer base into groups of individuals that are similar in specific ways relevant to marketing, such as age, gender, interests, spending habits and so on

Types of Customer Segmentation

Demographic

Age, sex, education, income and relationship status.

Attitudinal

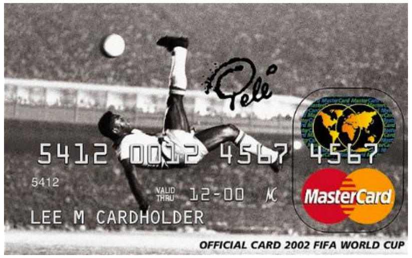
Values, interests, attitudes, lifestyle and preference.

Behavioral

Buying patterns, purchase history, and contexts of use.

Demographic Segmentation

Lifestyle Segmentation



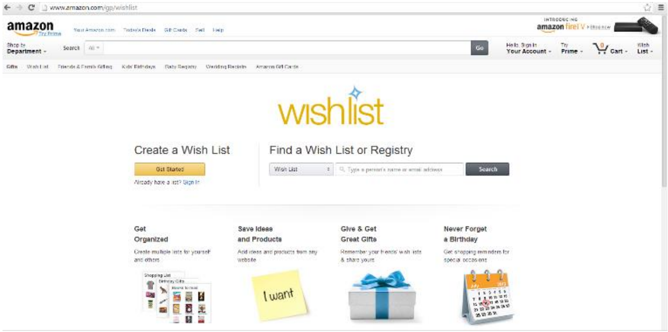
Attitudinal Segmentation

Common Interest:

- User group is a group of individuals with common interests in a particular technology, usually computer-related

Customer Preference:

- Amazon uses Wish Lists that capture customer desires and preferences, and can be used to segment customers more precisely



Behavioral Segmentation (1)

- Relationship Status
  - Major life events like getting engaged or married.

- Buying Patterns
  1. Classic upmarket
  2. Functional above all
  3. Fashionable originality
  4. Get money's worth

1

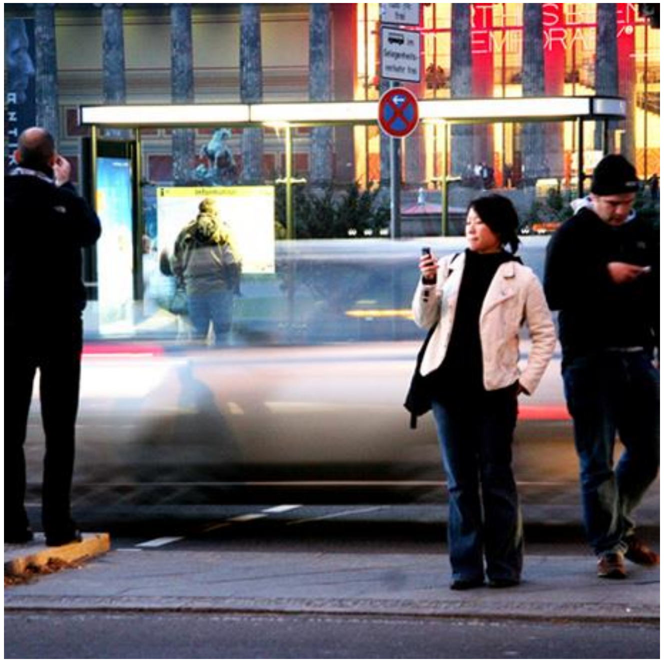
3

2

4

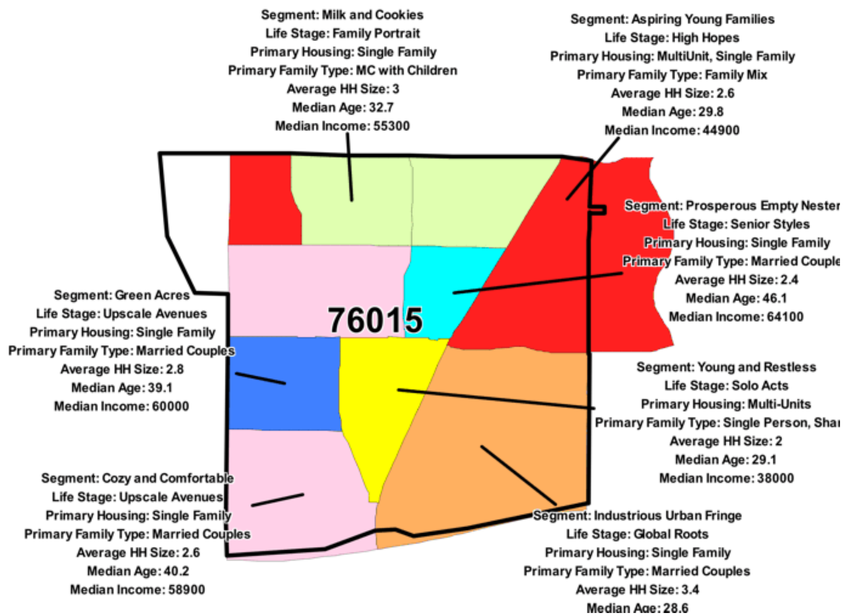
Behavioral Segmentation (2)

- Context of Use
  - The Context of Use is the actual conditions under which a given product is used, or will be used.



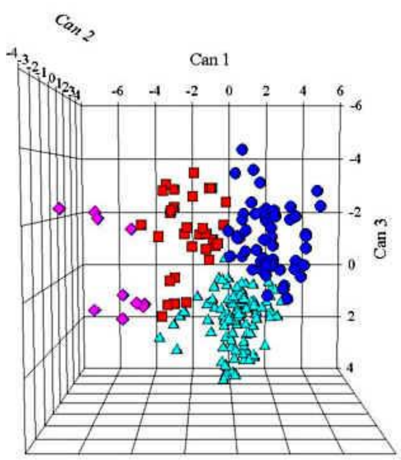
Customer Profiling (1)

Customer profile refers to a customer or set of customers including demographic, geographic, and psychographic characteristics, as well as buying patterns, creditworthiness, and purchase history - Helps To Describe Segments Meaningfully for Strategy Formulation

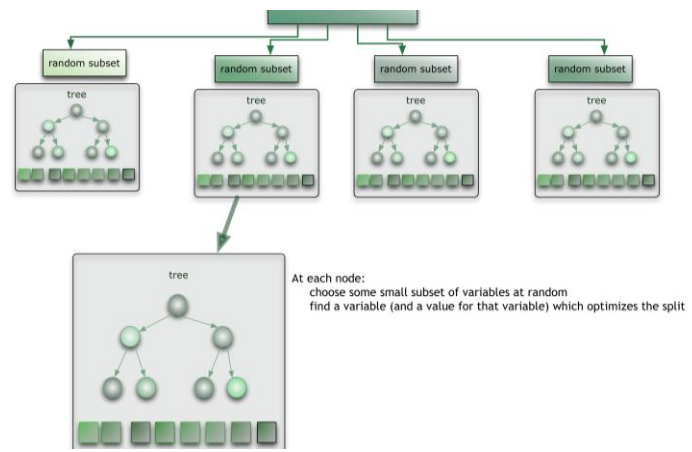


Customer Segmentation Techniques

- Common statistical techniques for segmentation analysis include:
  - Clustering algorithms such as K-means or other Cluster analysis
  - Statistical mixture* models such as Latent Class Analysis
    - It identifies unobservable subgroups within a population.
  - Ensemble* approaches such as Random Forests
    - It constructs a multitude of decision trees.



Latent Class Model



The random forest

Cluster Analysis

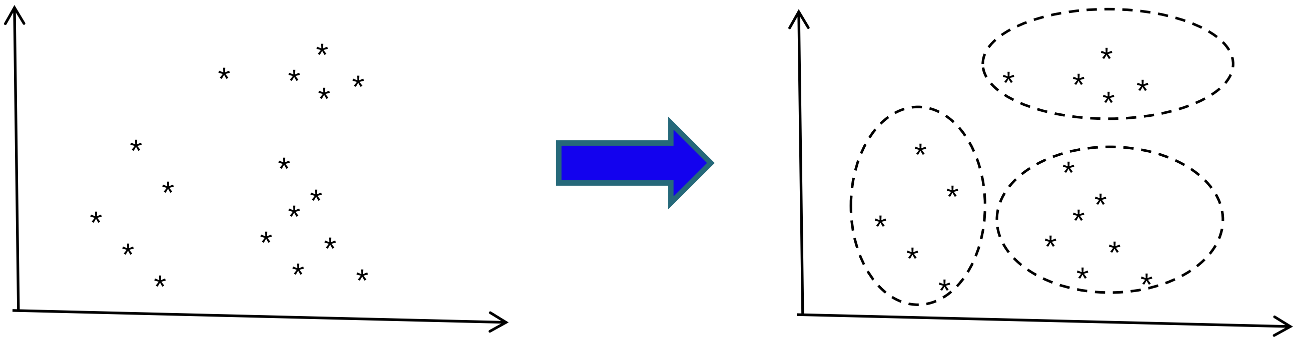
Types of Cluster Analysis

1. K-Means Method
2. RFM Analysis
3. Hierarchical Clustering
    - ✓ Agglomerate Method
    - ✓ Divisive Method
4. Kohonen's Clustering Method (Self Organising Maps SOM)

What is Cluster Analysis?

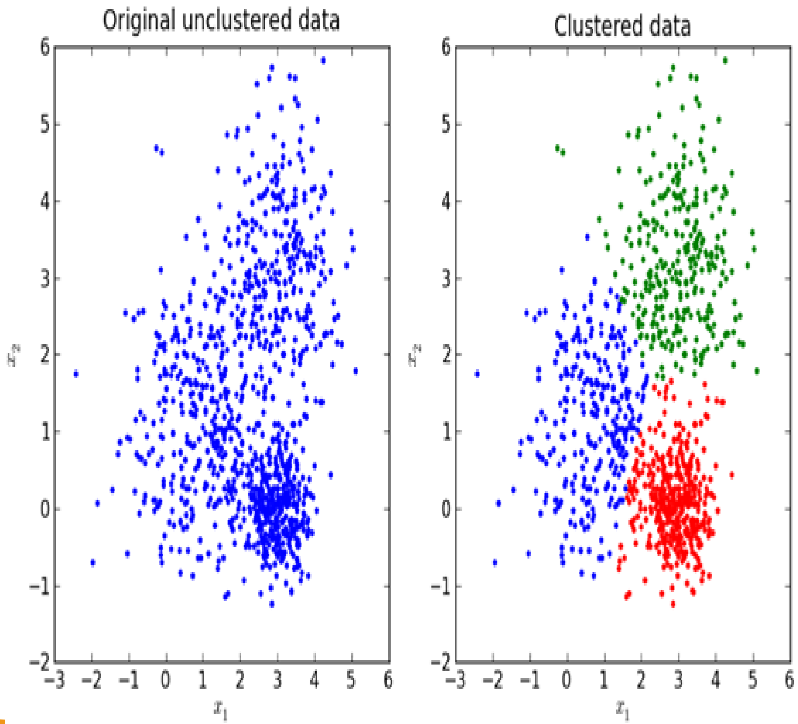
Goal of Clustering

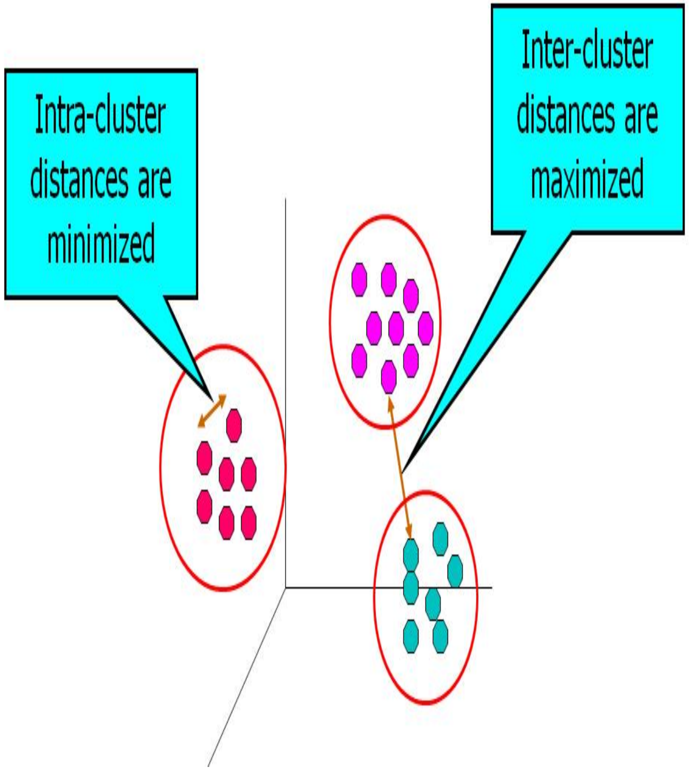
- To segment the data into a set of ***homogeneous*** clusters (i.e.: members within same clusters are similar enough) of observations for the purpose of generating insight to ***formulate customer strategy***
- e.g. customers should have more similarity within the same group such as lifestyle, purchase preferences etc. **but**
- Less similarity between groups



Characteristics of Cluster Analysis

Cluster Analysis **is an unsupervised learning - there are no pre-defined classes**





Technicalities of Cluster Analysis

- **Need to begin by determining a distance measure** which will represent how far apart two data points are

- If each data point is represented by straightforward **numerical variables** then it is relatively easy;

- Measures of Distance
  - ***Euclidean distance.***
  - Squared (or absolute) Euclidean distance.
  - City-block (Manhattan) distance.
  - Mahalanobis distance (D2)
  - …………………

- In most cases we use Euclidean distance

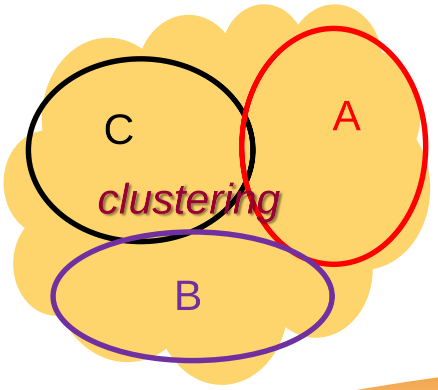
- However it is important that the measures of distance are ***scaled or normalized*** so that one component of distance does not dominate
  - One measure of scale is the Z score
  - Transform the data so that the distances have mean 0, and variance of 1

Utility of Cluster Analysis

Decide action

- Specific action on particular group, e.g.:
  - Targeting on group B (customers with potential interest in some new service)
  - Increasing influence to C (customers with possibility in developing some interest in new service)

Description of each group is achieved through **profiling**



Cluster Analysis : Be Mindful About Data Type

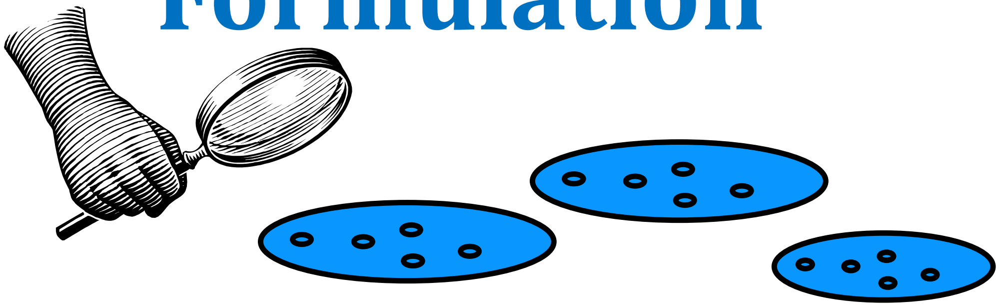
With categorical variables it is more difficult to define distances

- If the categories can be related to levels of some measurement then we can define distance/association measures based on this
- Example: Your measurement scale may be
  - Strongly like
  - Like
  - Neutral
  - Dislike
  - Strongly dislike

- If one object scores “Like” and another scores “Neutral” then the distance could be 1
- If one object scores “Like” and another scores “Dislike” then the distance could be 2

Cluster Interpretation For Strategy Formulation

Interpreting Clusters



Can the clustering be explained in practical terms?

Examine the value of each variable of the cluster centroid, (the cluster profile) of each cluster.

Look at the distinguishing characteristics of each cluster's profile and identifying substantial differences between clusters

Cluster solutions failing to show substantial variation between clusters indicate that this may be a spurious clustering

Other cluster solutions should be examined.

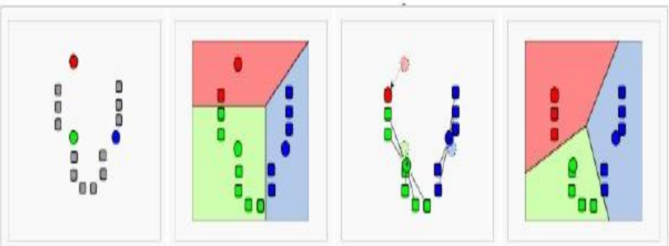
K–Means Clustering

This is an efficient & perhaps the fastest clustering algorithm that can handle both long (many records) and wide datasets (many input fields)

It is distance based and unlike the hierarchical algorithms, it does not need to calculate the distance between all pairs of records

The number of clusters to be formed is predetermined and specified in advance

Non-Hierarchical Divisive

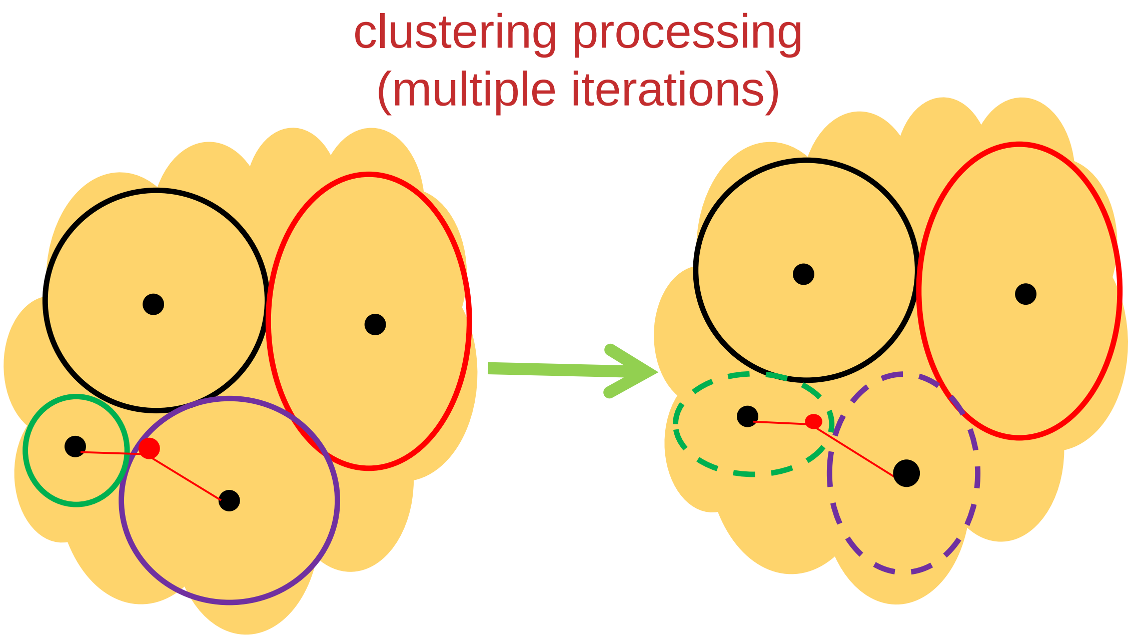


- The **simple K-means** Non-Hierarchical Divisive Process
  - Pick K*, which is our initial "seed" number which we hope is the true cluster#
  - Carry out RANDOM SAMPLING of the data set to establish "seed" Centroids (average location of cluster members)
  - Go through list of items and reassign those that are closer to a "competing" cluster's Centroid
  - Calculate an updated Centroid value and repeat the process until no reassignments of items take place.
- GOOD: Membership of items in clusters is fluid and can change at ANY time.
- BAD: Since K-means relies on RANDOM SAMPLING you may not pick up on RARE groups and so they may not appear as distinct clusters: so → K* < K
- BAD: In simple K-means, K is fixed in advance, however newer methods to iteratively adapt K seem to be available (P696, Dean & Wichern).
- **For examples see P696-701 Dean & Wichern**

K – Means Clustering

- Usually a number of different solutions should be tried & evaluated before approving the most appropriate.

- K-Means Clustering is best for handling **continuous** clustering fields



K – Means Clustering

Step 1 : Start with *k* initial clusters (user choose *k*)

Step 2 : At every step, each data record is reassigned to the cluster with "closest" centroid

Step 3 : Re-compute the centroids of clusters that lost or gained a record, and

- repeat the previous step as the next iteration

Step 4 : Stop when moving any more records between clusters increases cluster dispersion

K – Means Clustering

The choice of the number of clusters can

- Either be driven by external consideration (e.g.: previous knowledge, practical constraints, etc)
- Or we can try a few different values for *k* and compare the resulting clusters

Initial partition into *k* clusters

- Any available external information suggesting a certain partitioning, or the centroids of the *k* clusters, should be used
- In case no information is available for the initial partition, the algorithm can rerun with different randomly generated starting partitions to reduce the chance of resulting poor solution

NUS National University of Singapore | ISS INSTITUTE OF SYSTEMS SCIENCE

Critique

Advantages

- Clustering of customers according to their attribute preferences
- Cluster customer with similar behaviours/characteristics
- Clusters of similar brands/products can help identifying competitors / market opportunities

Disadvantages

- Choice of cluster-forming variables can be random
- Determine the right number of cluster often time-consuming
- Highly dependent on the analyst’s interpretation

National University of Singapore
INSTITUTE OF SYSTEMS SCIENCE

Recency, Frequency and Monetary

Recency, Frequency and Monetary

- RFM is a transaction data analysis method widely applied for customer segmentation.
- **Recency (R)**
  - The time interval between the last purchase and a present time reference. The freshness of customer activity.
    - E.g. When was the customer's last purchase?
- **Frequency (F)**
  - The frequency of customer's transactions
  - Eg. The total no of recorded transactions
- **Monetary (M)**
  - The total amount of money spent by the customer over a particular time period. Shows the willingness to spend
    - E.g. How much did the customer spend?

What is RFM Analysis?

- It is the most basic and common form of customer database segmentation.
- It can be used to activity-related data with measurable value and is repeatable.
- It is used in areas like e-commerce, faming (in-app purchases, levels played), lead management.
- Need not be confined; can be used not only for purchase history, website but also visits, social engagement

What can RFM Analysis do?

- RFM Analysis helps organizations *target* which customers to focus on
  - E.g. give select offers and direct promotions at them

- Companies can use it to reduce *churn* “lost” customers and incentivize them to purchase items

- RFM Analysis can help companies keep track of their customers and build *customer relationship* that can increase sales and productivity.

- It also identifies not-so-important customers for *prioritisation*
  - Who spend low dollar amounts in small quantities at irregular intervals

How to do RFM segmentation?

1. First create the RFM metrics table from the customers characteristics.

2. Compute the quartiles (quintiles) for the customers.

3. Segment the customers according to the distributions for each metric.

4. Allocate the customers into the appropriate ‘quartile’ for each of the ‘R’, ‘F’ and ‘M’.

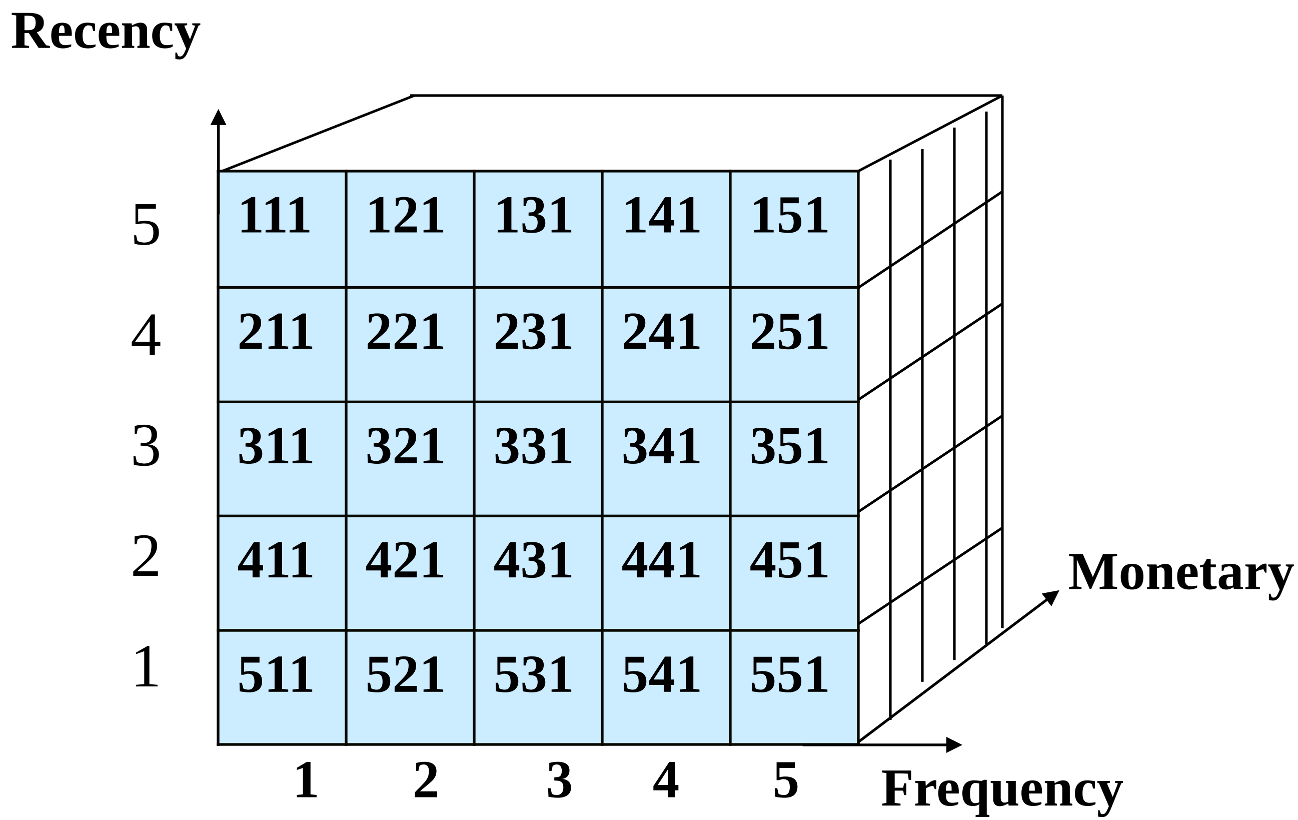
5. The customers are then labelled with ‘RFM’ categories.

Calculate RFM Score

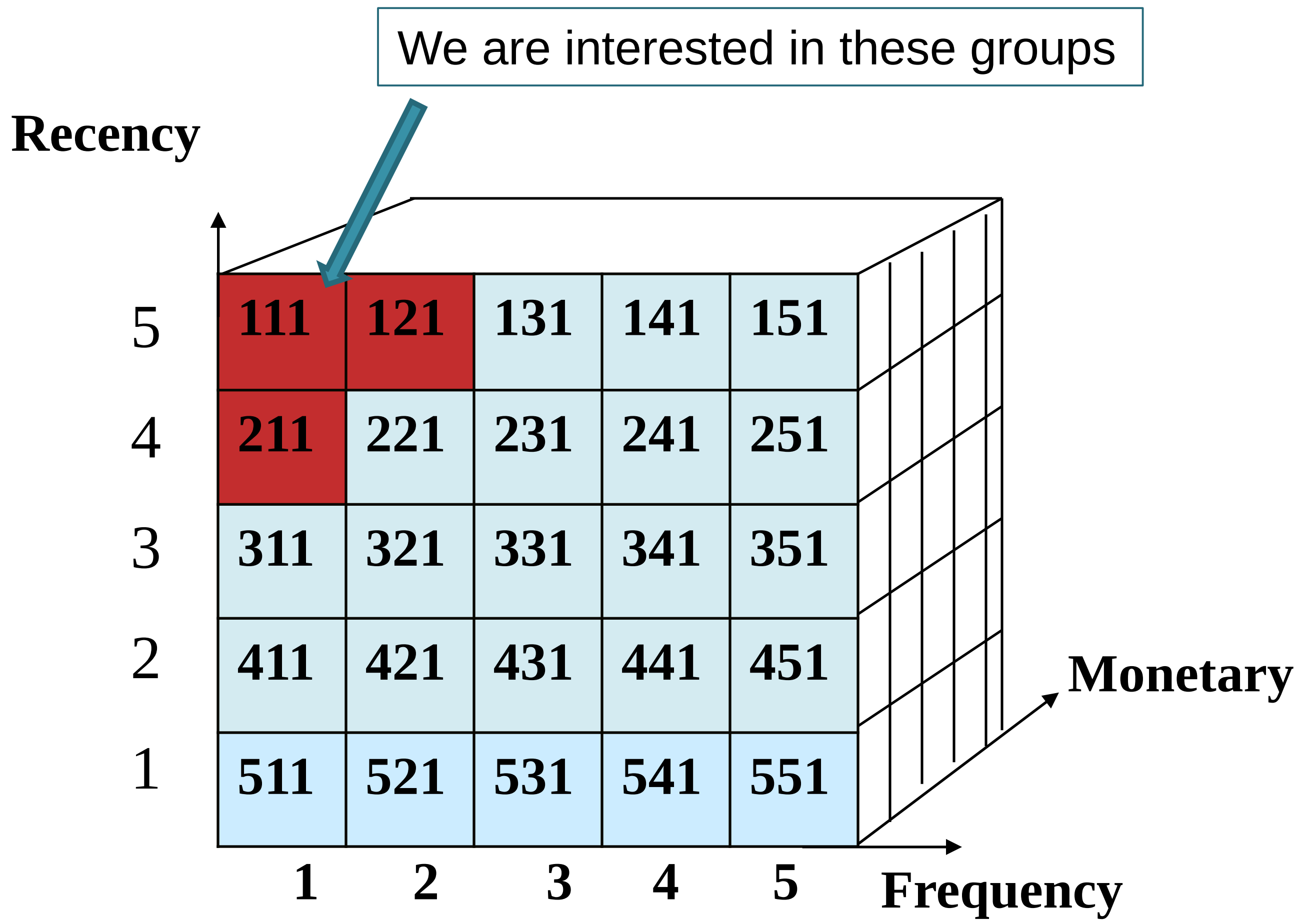
- Through a Recency, Frequency and Monetary table

CustomerID	Recency (Day)	Frequency (Number)	Monetary (TL)
1	3	6	540
2	6	10	940
3	45	1	30
4	21	2	64
5	14	4	169
6	32	2	55
7	5	3	130
8	50	1	950
9	33	15	2430
10	10	5	190
11	5	8	840
12	1	9	1410
13	24	3	54
14	17	2	44
15	4	1	32

A 3-D RFM table



A 3-D RFM table



Hurdles and dealing with Outliers

- Outliers refer to those customers with extremely low values.
- It could be argued that these ‘low visit’ people are just “noise”, they are not valuable customers.
- We can therefore create a “**hurdle**” where we remove those customers who have unusually low values.
  - For example, if you have 1,000 customers but the bottom 10% customers visit you only once or twice
    - Then these customers probably don’t contribute much to your business.
- Hurdle Rate is a much more “fair” representation of average customer activity if extreme customers who
  - really aren’t active
  - don’t contribute significantly to profitability

  are eliminated from the data set

Hurdle Rate

- Formal Definition
  - A Hurdle Rate is the percentage of your customers who have at least a certain activity level for Recency, Frequency, and Monetary value
  - i.e. the percentage of customers who have
    - Engaged in a behaviour since a certain date (Recency),
    - Engaged in a behaviour a certain number of times (Frequency),
    - Purchased a certain amount (Monetary value).
- Set up Hurdle Rates for these activities & look at their performance
- It is a simple way to track the overall health of your business or your promotion activities over the time.

Frequency Hurdle Rate

Using the database example:

- Filter out customers that have not visited for a long time
- The number of customers over the hurdle

e.g. Recency Hurdle rate= 44/100 = 44%

How do you set a hurdle on the F and M at the same time? Why do you want to do this?

FM Table

ID	R	F	M
18	90	3	823
19	60	16	156
20	61	2	250
21	74	14	380
22	95	2	658
24	98	20	883
25	84	2	733
26	86	17	645
28	69	7	126
29	88	5	1000
31	81	20	201
35	63	10	556
36	90	11	778
37	76	7	149
38	79	18	306
40	84	13	753

Exercise

1. Who are the most responsive customers?
2. What is the frequency hurdle rates if you want to consider only the top 60% frequent customers?
3. Who are the most valuable customers?

ID	Recency (Day)	Frequency (Number)	Monetary ($)
1	3	6	540
2	6	10	940
3	45	1	30
4	21	2	64
5	12	4	169
6	32	2	55
7	5	3	130
8	50	1	950
9	33	15	2430
10	10	5	190
11	5	8	840
12	1	12	1410
13	24	3	54
14	17	2	44
15	4	1	32

NUS National University of Singapore | ISS Institute of Systems Science

Uses of RFM score I

- The RFM process is both **descriptive** and **prescriptive**.
  - To connect the 'RFM' to future customer behaviour.
- You can analyse the characteristics and purchasing behaviour of the best customer group and try to understand what distinguishes them from typical customers, e.g.
  - Do they tend to buy a subset of your products or services?
  - Do they live in demographically similar neighbourhoods?
  - What are the most important metrics – R, F or M?

It is also a characteristic of an industry (to describe later)

Industry-Based RFM

On-line Fashion retail

Luxury Goods

Banking Services

Low R, Low F, High M

High R, High F, Low M

High R, Low F, Low or High M

Go to menti.com code 194732

Uses of RFM score II

- For marketers, RFM model is identified for discovering more responsive customers who are most likely to respond to your offer.
  - This enables you to target your marketing campaigns at the relevant customer segment
- In particular RFM analysis was used by Charities to target mailings to customers most likely to make donations.
  - The reasoning behind RFM was simple
    - People who donated once were more likely to donate again.

Analyzing Frequency

- Suppose you have a set of response rates ranked by Frequency of Purchase

Analysing Frequency			
Total Number of Purchases	Number Mailed	Number Responding	Response Rate
5+	11101	1935	17%
4	15204	1829	12%
3	18209	1193	7%
2	44220	1427	3%
1	116103	24060	21%
Total	204837	9790	5%

Analysis / Inferences:

- It seems that the mailers do a good effect with first time mailers getting a response rate .
- The marketing is done well, but the product may need improvement as the customers do not seem to frequent again.

This is a one-dimensional analysis based on frequency purely.

Analyzing Recency

- Suppose you have a set of response rates showing when customers bought from you last.

Analysing Recency			
Most Recent	Number Mailed	Number Responding	Response Rate
0-3 Months	11101	1935	17%
4-6 Months	14004	1829	13%
7-9 Months	14201	1193	8%
10-12 Months	16023	1427	9%
13+ Months	112431	5000	4%
Total	167760	9790	6%

- The index of response is computed by dividing the average response rate into the actual response rate (17%, 13%, etc.) for each group and multiplying by 100.
- Analyse the results shows that the most recent purchasers are almost 3x as likely as the average to buy again.

  ⟹ Focus your attention on your most recent customers.

  ⟹ Provide them with superior service and contact them frequently

  ⟹ What is the product 'popularity' last?

Monetary

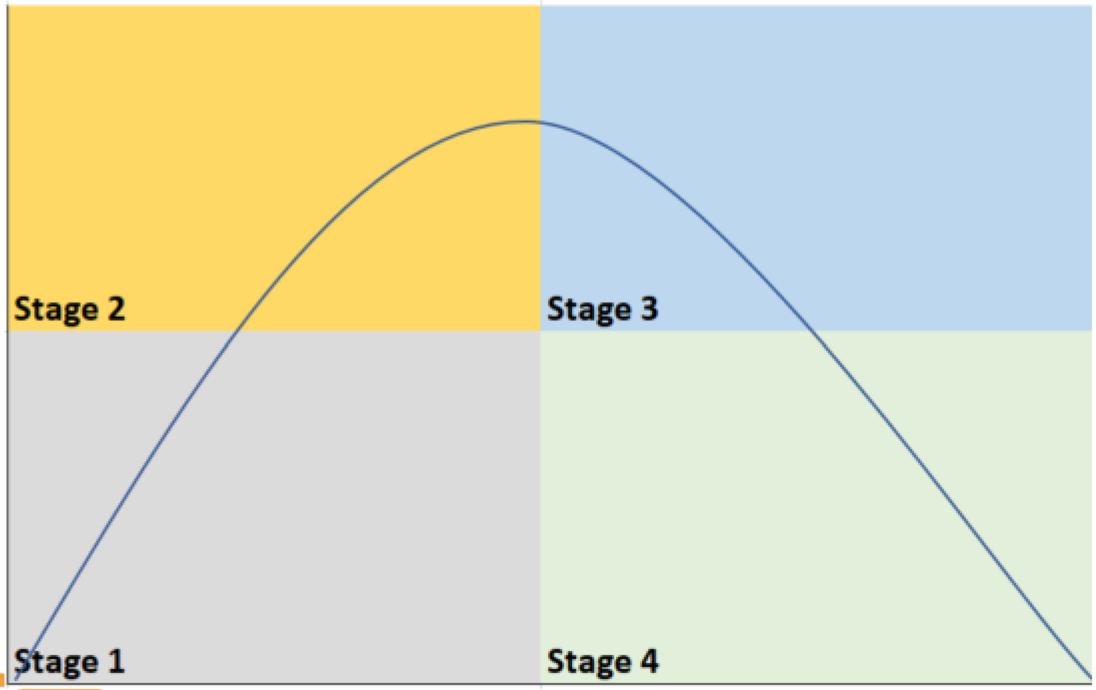
- If we rank response from customers by monetary sales, (i.e. the total dollar value of their purchases since they first started buying from your company)

Lifetime Purchases	Number Mailed	Number Responding	Response Rate
$500+	4151	788	19%
$300 - $499	12254	1829	15%
$200 - $299	1277	1493	117%
$100 - $199	24220	1735	7%
$1 - $99	138935	3945	3%
Total	204837	9790	5%

- It seems that the more customers spend, the better they respond.

Life Stages for RFM Model

Stage 1: Advertising campaign started. High potential with low F, low R and low M.
Stage 2: Advertising campaign starts to bear fruit. High R, High F and M increasing
Stage 3: Interest starts to wane with High R, Lower F still high M
Stage 4: On a trough. RFM all low.



Using RFM score to analyze customer retention

- The RS (Recency, Sales) matrix calculates for each customer, recency and the average inter-purchase time
- The greater the recency is than the inter-purchase time, the more likely is that customer has been lost
- So this will identify those “lost” customers, <u>or</u> those customers whom you may be about to loose
- Advantages of this approach:
  - Cheap, easy to implement and easy to understand
- Limitations of this approach:
  - Overly simplistic and in most cases highly inaccurate.
    - Does not consider fad, trends or advertising campaigns but useful nevertheless!

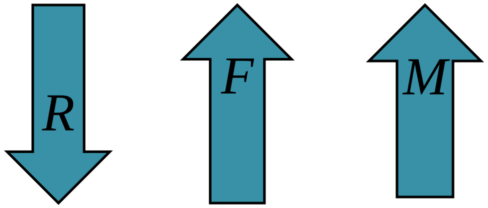
Another use of RFM – Evaluating Sales staff performance

- A RFM Analysis of customers can be done for Salesmen or account managers to give a relatively objective picture of how they are performing
- You can analyze the amount of revenue generated per person and compare different salespeople.

  This provides for a more objective measurement scale; then just a monetary value.

Use RFM for 'Behavioral Clustering'

- RFM analysis assigns value-scores to each customer on basis of his/her purchase behaviour.
- The value-score indicates the customer's contribution.
- The importance (weight) of each RFM variable depends on the characteristics of the industry.
- The weight of the three variables can be equable or different
- Segment the entire customer list
    - Either based on the overall RFM value.
    - Or the combination of individual recency, frequency and monetary scores



- Indeed RFM can be used as the basis clustering (described earlier)

Strengths and Weaknesses of RFM Analysis

Strengths	Weaknesses
• Makes use of existing data in companies • Allows marketing and advertising focus • Specific targeting can increase profit and reduce costs; companies gain by not focusing on customers who will not be greatly profitable • You can offer incentives to middle scoring customers to increase their purchases • Analysis is relatively quick and easy to interpret • RFM is best suited for companies who offer a rewards program. They are able to track spending and can offer their high profile clients incentives to spend more	• It only looks at three variables and there may be others that are more important • Customers with low RFM scores may be ignored, even though they may have legitimate reasons for spending more with other vendors. • Missed opportunities for business relationships with potential customers leading to loss of future sales • Not suited to companies who provide products that are unique and will not be purchased in large quantities

An extension of RFM - RFMPD Analysis

- RFMPD includes 2 additional variables.
  - P stands for Payment. This measures when the company receives payment.
  - D stands for Date. This is the date of the customers last payment.
- The final score is based on a value for R, F, M, & P
- RFMPD also takes into account the customer's payment history
  - If a customer pays on time, you know that there are no cash flow issues
  - Slow payers may be having financial problems which may increase in the future
  - This will matter when targeting desirable /less desirable customers

This is useful for assessing the 'credit health' of customers and also beneficial to the cashflow for the company.

Summary

- It is much more expensive to attract new customers than to retain existing customers.
- Customer segmentation is an inexpensive method to retain customers by understanding their behaviours and characteristics.
- It allows for targeted marketing, and also helps to understand their reaction towards the products.