SAS® GLOBAL FORUM 2016

IMAGINE. CREATE. INNOVATE.

Product Purchase Sequence Analyses By Using Horizontal Data Sorting Technique

Justin Jia **TransUnion Canada**

Amanda Lin **CIBC, Canada**

#SASGF

Product Purchase Sequence Analyses By Using Horizontal Data Sorting Technique

Justin Jia
Trans Union Canada

Amanda Lin
CIBC, Canada

ABSTRACT

Horizontal data sorting is a very useful SAS technique in advanced data analysis using SAS programming. Two years ago (Paper 376-2013), we presented and illustrated various methods and approaches to perform horizontal data sorting, and demonstrated its valuable application in strategic data reporting. However, this technique can also be employed as a creative analytic method in advanced business analytics. This paper presents and discusses its innovative and insightful applications in product purchase sequence analyses, such as product opening sequence analysis, product affinity analysis, next best offer analysis, time span analysis etc. Compared to other analytic approaches, horizontal data sorting technique has the distinct advantages of being straightforward, simple, convenient to use and producing easy to interpret analytic results. Therefore the technique can have a wide variety of applications in customer data analysis and business analytics fields.

INTRODUCTION

- Data sorting can be vertical sorting, across rows, or horizontal sorting, across columns.
- Compared to vertical sort, horizontal sort is used less frequently, and it requires the user to employ multiple sophisticated SAS skills such as Transpose, Rotate, Array, Macro, etc.
- In SAS paper 376-2013, we presented and discussed various approaches and methods to achieve horizontal data sorting through SAS programming, including Bubble Sort Method, Rotate and Transpose Method, Call Sort Routine Method etc.
- Utilization of horizontal data sorting technique can significantly enhance the format and layout of data reporting.
- Horizontal data sorting can also be used as a creative means in data analysis to leverage data and perform deep-dive business analytics such as customers' purchase tendency and preferences analyses.

CASE STUDY: FINANCIAL PRODUCT PURCHASE SEQUENCE ANALYSES

PART I. HORIZONTAL SORTING OF RAW CUSTOMER DATA

The raw financial customer data set contains 2.56 millions of records and 4 attributes (Clients_ID, Account_ID, Account_Type, Opened_Date). There are several different approaches to horizontally sort the raw data in the account opening sequence, such as Bubble Sort Method, Rotate and Transpose Method, Call Sort Method etc. We use PROC SORT and ARRAY approaches to horizontally sort the raw data due to its big flexibility.

Input Raw Financial Customer Data

Client_ID	Account_ID	Account_Type	Opened_Date
1001	601152	Chequing	2011-05-26
1001	103759	Savings	2011-05-26
1001	726789	TFSA	2011-06-01
1002	632975	Chequing	2012-06-11
1002	363108	Credit Card	2012-10-24
1002	101492	Credit Insurance	2012-10-24
1002	161079	Savings	2012-11-07
1003	987783	Credit Card	2010-11-29
1003	230104	Savings	2012-08-09
1003	599067	Credit Card	2012-09-18
1003	385030	Chequing	2012-06-05
1003	632316	Mortgage	2013-07-11
1003	281144	TFSA	2012-12-22

Horizontally Sorted Final Data Set: In An Ascending Order Of Purchase Date

Client_ID	1st Product	Purchase Date	2nd Product	Purchase Date	3rd Product	Purchase Date
1001	Chequing	2011-05-26	Savings	2011-05-26	TFSA	2011-06-01
1002	Chequing	2012-06-11	Credit Card	2012-10-24	Credit Insurance	2012-10-24
1003	Credit Card	2010-11-29	Chequing	2012-06-05	Savings	2012-08-09

Product Purchase Sequence Analyses By Using Horizontal Data Sorting Technique

Justin Jia
Trans Union Canada

Amanda Lin
CIBC, Canada

CASE STUDY: FINANCIAL PRODUCT PURCHASE SEQUENCE ANALYSES (CONTINUED)

PART II. PRODUCT PUCHASE SEQUENCE ANALYSES

1. Overall Product Purchase Sequence Analysis

A simple one-way frequency distribution by Proc Freq procedure can give us insights into the financial customers relating to their purchase sequences. Through purchase sequence analysis, we can probe and learn about the banking behavior and purchase tendency of our financial customers. Here we only show the first three purchased products and the top 5 ranking products.

Table: Banking Product Purchase Patterns of Financial Customers

Ranking	1st Purchased Product	Clients(#)	Clients(%)	2nd Purchased Product	Clients(#)	Clients(%)	3rd Purchased Product	Clients(#)	Clients(%)
1	Chequing	580,597	22.7%	Savings	318,171	12.8%	Credit Card	181,995	7.1%
2	Credit Card	304,540	13.3%	Credit Card	291,198	11.6%	Savings	173,087	6.5%
3	Savings	304,121	13.2%	Credit Insurance	280,253	11.3%	TFSA	160,478	6.3%
4	TFSA	281,670	12.2%	TFSA	249,403	10.5%	Credit Insurance	136,779	5.4%
5	Mortgage	239,208	10.7%	Chequing	158,380	8.3%	Chequing	87,606	4.1%

Chequing, Credit Card and Savings are the three most common products in their first purchase. The need of Chequing product is much higher than the Credit Card and Savings products (22.7% vs. 13.3%). Therefore, if we want to acquire new customers, Chequing, Credit Card and Savings are the top products for our promotions. In their next purchase, the top 3 preferred products are Savings, Credit Card and Credit Insurance respectively. As for the third purchased products, Credit Card is the most popular one followed by Savings and TFSA.

2. Product Affinity Analysis

Affinity analysis is a data analysis and data mining technique that discovers co-occurrence relationships among activities performed by (or recorded about) specific individuals or groups. Product Affinity Analysis can be used to design best selling/profitable product bundles for cross-selling and up-selling purposes.

CASE STUDY: FINANCIAL PRODUCT PURCHASE SEQUENCE ANALYSES (CONTINUED)

Table: Structure and Penetration of Banking Products and Product Bundles

	Structure	Ranking	Client Penetration (%)
Single Product	Credit Card	1	19.8%
	Savings	2	14.9%
	Chequing	3	13.3%
Two Products	Credit Card / Credit Insurance	1	9.3%
	Chequing / Savings	2	7.9%
	Chequing / TFSA	3	7.2%
Three Products	Chequing / Savings / Credit Card	1	5.6%
	Chequing / Credit Card / Mortgage	2	5.2%
	Chequing / Savings / TFSA	3	3.7%

Among single product customers, Credit Card, Chequing and Savings accounts are the top 3 most popular products, their client penetrations are 19.8%, 14.9% and 13.3% respectively. Among two product customers, the three most common product bundles are Credit Card / Credit Insurance, Chequing / Savings, Chequing / TFSA respectively.
A higher client penetration also suggests a higher affinity of the bundled products. If we want to design cross-sell product bundles, Chequing, Savings, Credit Card and TFSA are the high-affinity elements for bundling together according to our analysis.

3. Next Best Offer Analysis

A two-way frequency distribution based on the horizontally sorted data can tell us the probability of a customer to purchase his/her next product. Product type and event likelihood of next purchases are largely dependent on the customer's current products.

Table: Next Product Purchase Patterns Of Single Product Customers

Current Product	Next Purchased Product (Top 5)					
Chequing	Product	Savings	Credit Card	Credit Insurance	TFSA	Loan
	% of Clients	32.8%	20.7%	11.2%	10.9%	7.5%
Credit Card	Product	Chequing	Savings	TFSA	Mortgage	Line of Credit
	% of Clients	28.9%	20.1%	15.7%	13.8%	6.3%
Savings	Product	Chequing	TFSA	Mortgage	GIC	Credit Card
	% of Clients	40.7%	19.4%	13.5%	9.8%	5.2%

Product Purchase Sequence Analyses By Using Horizontal Data Sorting Technique

Justin Jia
Trans Union Canada

Amanda Lin
CIBC, Canada

CASE STUDY: FINANCIAL PRODUCT PURCHASE SEQUENCE ANALYSES (CONTINUED)

It suggests that a customer with Chequing as his/her current product is most likely to open a Savings account in his/her next purchase. The probability of this event to occur is 32.8%. Furthermore, we can see that the product type and event likelihood of next purchases are largely dependent on the current products. Chequing customers' next favorite products are Savings, Credit Card and Credit Insurance and so on, but Savings customers are most interested in Chequing, TFSA and Mortgage products in their next purchase.

Table: Next Product Purchase Patterns Of Two Product Customers

Current Products	Next Purchased Product (Top 3)			
Chequing + Savings	Product	TFSA	Credit Card	Credit Insurance
	% of Clients	29.1%	15.9%	11.3%
Chequing + Credit Card	Product	Savings	TFSA	Mortgage
	% of Clients	31.5%	28.7%	15.9%

If a customer currently has Chequing and Savings accounts, his/her next purchase is most likely to be TFSA with a probability of 29.1%. However, a customer with Chequing and Credit Card accounts may be interested in both Savings and TFSA products since the probabilities of the two next opened products are very comparable (31.5% vs. 28.7%).

4. Purchase Time Span Analysis

In customer marketing, we not only need to know what is the best product to offer our customers, we also need understand when we should provide them the offer. In short, for successful marketing, we should give our customers the right offer at the right time. Through purchase time span analysis, we can find the time span patterns and determine the correct time to contact them for marketing different products.

Table: Purchase Time Span Analysis of Financial Customers

Current Product	Next Product	Time Span To Purchase Next Products (Days)				
		N	Min	Max	Mean	Standard Deviation
Chequing	**Overall**	2,135,673	0	636	183.8	221.3
	Credit Card	491,205	2	325	93.5	81.6
	TFSA	234,924	26	629	254.6	193.3
Credit Card	**Overall**	2,236,539	0	568	257.6	221.3
	Chequing	626,231	0	512	192.1	104.9
	TFSA	351,137	68	554	235.8	163

CASE STUDY: FINANCIAL PRODUCT PURCHASE SEQUENCE ANALYSES (CONTINUED)

Financial customers with Chequing as their current product, the overall average time duration for them to purchase next products is approximately 6 months (183 days). Customers with Credit Card as their current product have longer time duration to purchase their next products, which is 257 days on average.

The average time spans differ remarkably for purchasing different next products. It usually takes Chequing customers 93 days to open a Credit Card account, while it is much longer for them to purchase a TFSA product (254 days). Through these analyses, we can find the time span patterns and determine the correct time to contact them for different marketing products.

CONCLUSIONS

As discussed in this paper, horizontal data sorting is a very useful SAS technique. Utilization of this technique can significantly enhance the format and layout of data analysis and presentation, which is very helpful in data presentation and insightful data reporting. Much more important, it is also an innovative and valuable technique in advanced data analysis and business analytics. We can use it in product sequence analyses and other business analytics fields. This data levering technique has the distinct advantages of being simple, straightforward, convenient to use, and producing easy to interpret analysis results etc. It can find important applications in a wide variety of business analytics and business intelligence fields.

REFERENCES

1. Justin Jia and Amanda Lin, "**Horizontal Data Sorting and Insightful Reporting: A Useful SAS® Technique", 376-2013, 2013 SAS Global Forum, San Francisco, CA, USA.**
2. SAS Support Website, **http://support.sas.com/documentation/cdl/en/lrdict/64316/HTML/default/viewer.htm#a000214163.htm**
3. Jiawei Han & Micheline Kamber, "**Data Mining: Concepts and Techniques", Elsevier Inc, ISBN 978-1-55860-901-3, 2006**

SAS® GLOBAL FORUM 2016

IMAGINE. CREATE. INNOVATE.

LAS VEGAS | APRIL 18-21

#SASGF