Vancity

Predictive Modeling @ Vancity

October 2016

John Lo gazed thoughtfully at his computer monitor, but was not thinking about the email messages that had appeared while he was in his last meeting. He had just received the green light from his supervisor to begin developing a propensity model to support the quickly approaching 2016-2017 RRSP campaign. John had only joined VanCity earlier that month, and this project presented a great opportunity for John to demonstrate his abilities and his value to the department. At the same time, he was still adjusting to his new surroundings and was aware that this project would pose several challenges for him.

John's key concern was the time frame for the development and deployment of the project. With a deadline for delivering the mailing list within the next four weeks, John knew that he would be hard pressed to properly validate any model he would build. Fortunately, he would not be working alone. An outside consulting company had previously been retained to build a propensity model and John would be supervising their work while developing his own model. Nonetheless, to effectively manage the development of a sound model, John would first have to bring himself up to speed on the world of Credit Unions' financial services, as well as the history of Vancity's direct marketing efforts and experiences with predictive modeling.

Credit Unions

Like other financial institutions credit unions accept deposits and make loans. The main difference is that credit unions are owned and controlled by the members that use their services. Banks on the other hand, are owned by groups of stockholders whose main interest is to earn a healthy return on their investment.

Credit unions developed as the result of cooperative activities in 19th century Europe. The main premise in organizing the first credit unions was that people should be able to pool their money together and make loans to each other. Even today most credit unions are governed by three main guidelines:

- Only members can borrow from the credit union
- Loans can only be made for 'prudent and productive' purposes
- A person's desire to repay, or his/her character, is considered more important than the ability to repay.[1]

The first Canadian credit union was organized by Alphonse Desjardins in Levis, Quebec in 1900.

[1] Credit Union National Association: http://www.creditunion.coop

It was a small credit union with the first savings deposit of only 10 cents and the total collection from all members of only $26. Desjardins persevered and continued to organize other credit unions in North America. Today, total credit union membership in Canada is over 4.9 million. These credit unions are subject to provincial government regulations and as such cannot extend their operations beyond their own provincial borders.

Vancity

The Vancouver City Savings Credit Union, or Vancity, is Canada's largest credit union and the third largest in North America. Its assets in 2015 were $19.8 billion, with membership base of 519,000 and 59 branches in British Columbia.

Founded in 1945 Vancity today consists of a group of companies:

- Vancity Capital Corporation
- Vancity Community Foundation
- Vancity Enterprises
- Vancity Insurance Services
- Vancity Branch of Credential Securities Inc.
- Vancity Investment Management
- Vancity Trust Services
- Citizens Bank of Canada – the first Canadian Internet bank
- Real Assets

Vancity espouses a triple bottom line - social, environmental, and financial - and has a long history of community investing. Its vision 'calls on us to measure success in terms of how we contribute to the well-being of our members and their communities—not just financial well-being, but social and environmental well-being as well'. To achieve this, Vancity focuses on:

- ***increasing access to financial services:*** Through creation of unique services and products that enable low income and marginalized individuals to gain access to the financial services they need.
- ***advancing climate change solutions:*** Through reducing the amount of waste its employees create and financing both member and community environmental actions.
- ***strengthening other non-profits and co-operatives:*** Through offering both financial and personnel resources , help in creating strategic plans and skills for long-term financial sustainability.
- ***building successful social enterprises:*** Vancity helps build social enterprises, 'organizations that provide goods and services to earn revenue and achieve social, economic, and/or environmental goals'. Vancity directly invests in these enterprises, helps to strengthen the entire social enterprise sector, increases access of these organizations to capital and markets, and exercises its influence on the local, provincial and national environment to support development of this sector.

Competitive Environment

In 2015, there were 394 credit unions in Canada that either served specific, small geographic areas, or were exclusively used by members of an association or employees of a specific company. In BC, credit unions presented the largest financial network with more than 7,000

employees. Even though the total number of credit union branches has increased in recent years the overall number of credit unions has decreased. This is mainly due to mergers and acquisitions that resulted in an increase in the average asset size of individual credit unions.

The main competitors of credit unions are the chartered banks. In 2015, there were 48 chartered banks in Canada. Among them, they held 1/3 of all assets in the financial sector, and 92% of those assets were in the hands of the traditional 'Big Six' banks: the CIBC, the Royal Bank of Canada, the Bank of Nova Scotia, the Bank of Montreal, the Toronto Dominion Bank and the National Bank of Canada. Throughout most of Canadian bank history these six banks enjoyed a relatively undifferentiated and 'friendly competition' industry structure. Changes in federal regulations in the late 1990s put an end to this oligopoly. Fierce competition from both overseas banks as well as from small Internet startups developed. As a result, banks started moving away from being exclusively 'nuts and bolts' banking institutions and became more similar to credit unions by focusing on customer needs and creating social value based approaches to banking. To develop more customized products and create repeat purchase relationships with their customers, banks started developing and applying new market research and data mining techniques.

RRSP History

The Federal Government introduced tax-sheltered Registered Retirement Savings Plans (RRSP's) in 1957 as an incentive for Canadians to save for their retirement. The first major overhaul of the plan occurred in 1991 when the Government increased RRSP contribution limits to lesser of 18% of income or $11,500. The limit increased to $26,010 by 2015 and the unused credits can be carried forward. In 2001, the Government limited the foreign content to 30% of the total RRSP contribution, but in 2005 the foreign content limit was eliminated.

The tax advantages of RRSP's provide more incentive for high earners than for low earners to contribute, and until 2009, RRSP's were the only widely available savings vehicle with tax advantages. In that year, tax-free savings accounts (TFSA's) were introduced, which provided tax-incentives independent of income.

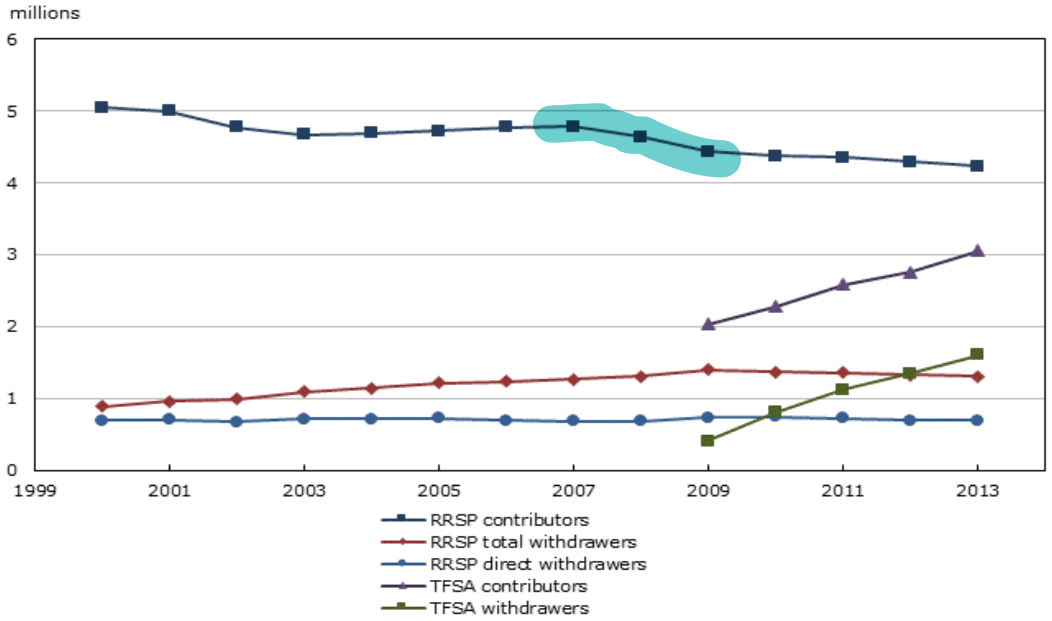
Types of RRSPs

All financial institutions offer different categories of RRSPs, such as a conservative guaranteed or term-deposit type, and self-directed investment types.

RRSP Statistics

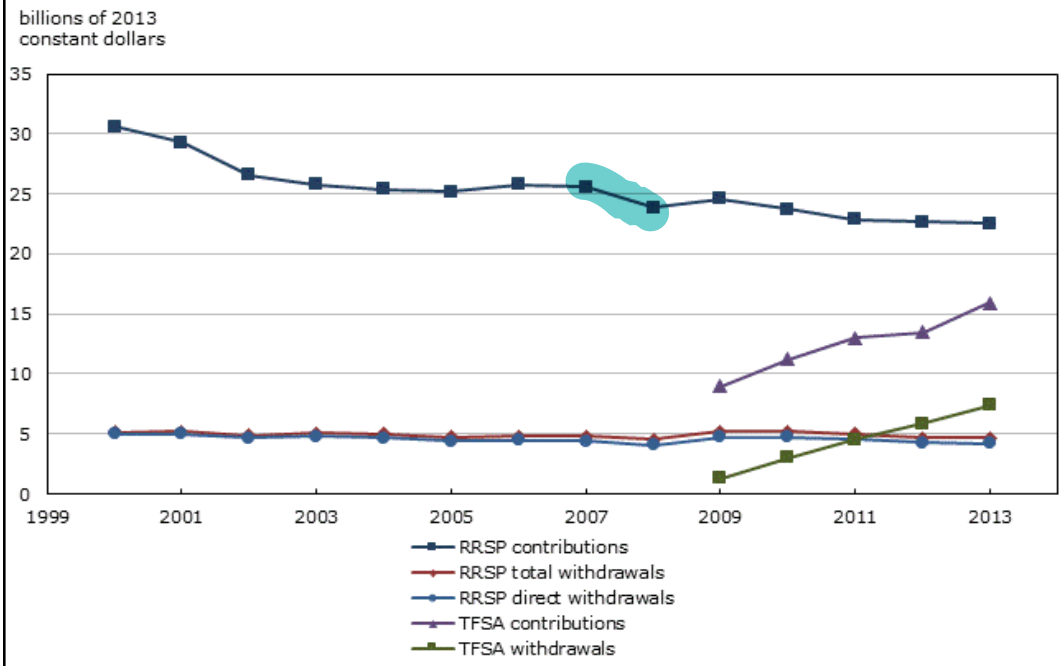
Between 2000 and 2013, the number of contributors to RRSP's declined gradually, by about 16%, from 5.0 million to 4.2 million. The largest decline occurred between 2007 and 2009, coinciding with the 2008 recession, and the introduction of TFSA's in 2009 (Chart 1-1). The total contributions also declined steadily, from $30.6 billion to $22.5 billlion (Chart 1-2).

Chart 1-1
Estimated frequencies of RRSP and TFSA contributions and withdrawals among 25- to 54-year-olds, 2000 to 2013



Notes: These data on registered retirement savings plan (RRSP) and tax-free savings account (TFSA) use are based on estimates from the Longitudinal Administrative Databank (LAD). The LAD is a 20% sample of Canadian taxfilers, so the estimates obtained from these data were multiplied by five to obtain the national estimates reported here. Direct withdrawals refers to total withdrawals minus Home Buyers' Plan defaults.
Source: Statistics Canada, Longitudinal Administrative Databank.

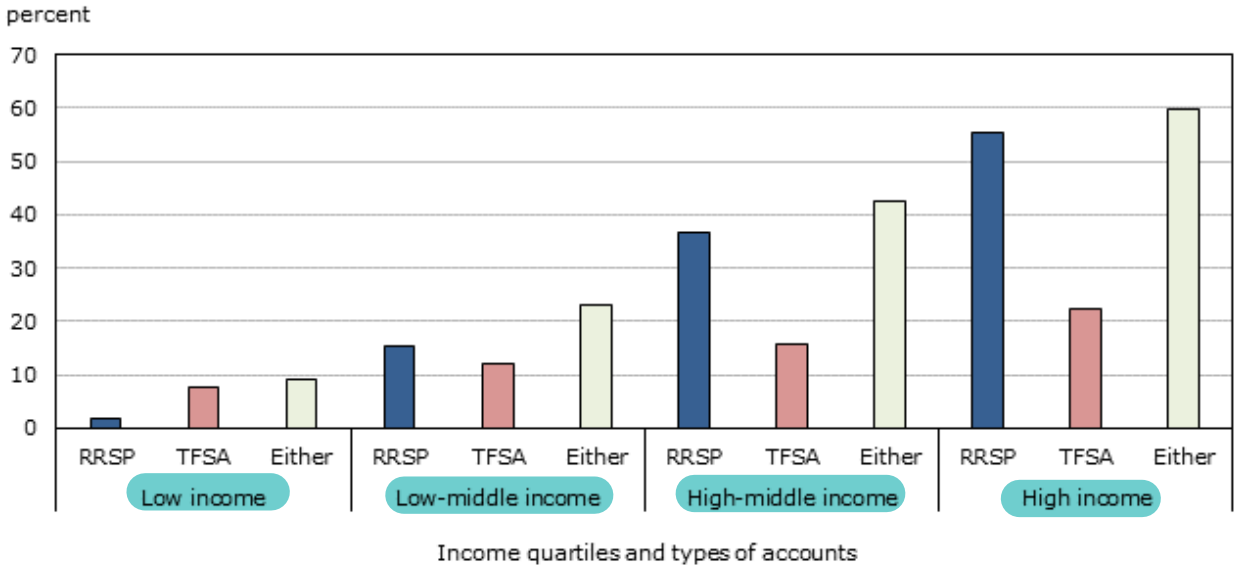
Chart 1-2
Estimated contributions to and withdrawals from RRSPs and TFSAs among 25- to 54-year-olds, 2000 to 2013



Notes: These data on registered retirement savings plan (RRSP) and tax-free savings account (TFSA) use are based on estimates from the Longitudinal Administrative Databank (LAD). The LAD is a 20% sample of Canadian taxfilers, so the estimates obtained from these data were multiplied by five to obtain the national estimates reported here. Direct withdrawals refers to total withdrawals minus Home Buyers' Plan defaults.
Source: Statistics Canada, Longitudinal Administrative Databank.

Unsurprisingly, higher income groups contribute more than lower income groups to both RRSP's and TFSA's. Chart 4 below shows the percentages of individuals in four income groups (binned by quartiles) that contribute. Fifty-five percent of those in the top quartile contribute to RRSP's, while only one percent of those in the bottom quartile contribute.

Chart 4
Incidence of contributing to an RRSP or TFSA (or either) across income groups for 25- to 54-year-olds, 2009 to 2013



Notes: Income groups are based on total income quartiles. The data on registered retirement savings plan (RRSP) and tax-free savings account (TFSA) use are restricted to the years 2009 to 2013, since TFSAs were not established before this time. For both types of accounts, an individual is deemed to be a "contributor" if positive net contributions were made in the reference year. In other words, individuals who may have contributed but also withdrew a larger amount are not counted. The RRSP withdrawals include the defaults on Home Buyers' Plan repayments.
Source: Statistics Canada, Longitudinal Administrative Databank.

Vancity's RSP Campaigns

Given the trends observed above, 2016 marked the beginning of a different way of thinking about RSP campaigns at Vancity. Looking back at the previous mass advertising and direct mail campaigns that had proven to be expensive and less effective than desired, management decided to set a new goal of finding a more efficient way of cross selling to its existing membership base. Still, penetrating the RRSP market and increasing the wallet-share wasn't the only concern facing Vancity's managers. The new marketing campaign had to agree with Vancity's overall philosophy. Besides being financially efficient, the marketing efforts needed to satisfy Vancity's goals of social and environmental responsibility, which careful targeting also satisfied. Untargeted direct mail reached many members who were not interested, consuming paper (and trees) and creating garbage unnecessarily.

Propensity Modeling

A propensity model is one of many kinds of predictive models used in database marketing and data mining. Predictive models in this context typically identify relationships between a target behaviour, often a purchase, and predictive variables, usually other customer behaviours and characteristics that are recorded in company databases. The predictive model then provides a means to estimate the likelihood of the target behaviour for each customer in the future, based on each customer's record of the relevant, or predictive, behaviours and characteristics. One natural and important target behaviour is 'propensity to purchase' (the likelihood, or probability of purchasing) and predictive models with this type of target are often called propensity models.

Vancity's RSP Cross-Sell Propensity Model

Until 2016, Vancity's RSP marketing campaigns primarily relied on mass media to get the message across to potential clients. These campaigns tended to attract new members but had little impact on existing members' RRSP purchases. Vancity had some limited experience with targeted direct marketing using two propensity models built by an external consultant a few years previously. One of the models was for the holders of RRSP's and the other one for the non-holders. The holder model worked well and generated volume by predicting which existing holders were most likely to purchase an RRSP again. However, the non-holder model was less successful. To remedy this, at the end of September 2016, Vancity hired a consulting company to help them improve 'targeted marketing campaign response rates to cross-sell RRSP products to existing Vancity clients'. Shortly after, John joined VanCity and assumed a project management role with responsibilities for supplying the consulting company with all necessary data and fielding questions that arose. John had graduated from the Simon Fraser University, where he had concentrated on marketing analytics, and he quickly suggested that it would be valuable to VanCity if he worked in parallel with the consulting company. Management agreed, and approved John's suggestion to build an in-house propensity model in parallel with the one contracted to the consulting firm.

The decision to build two propensity models simultaneously served two purposes: 1) It allowed Vancity to have an in-house model to compare with the model developed by the consulting agency and, to the extent that the models agreed, would instil confidence in their predictions; and 2) It helped the development and exploitation of the in-house capabilities of John and others, so that similar models could be built in-house in the future.

Model Objectives

The main goals of the cross-selling propensity models were to:

- Identify non-RRSP holding Vancity members who have the greatest potential for acquiring a new RRSP term.
- Develop profiles of the high-potential members that can support targeted message design.
- Increase RRSP volumes and new investment dollars, as well as product penetration and member entanglement.

- Improve targeting efficiencies and decrease campaign costs.

John restricted the database for model development to those members who had not held an RRSP with Vancity prior to the 2015-2016 campaign, since the plan was to apply the model during the following RRSP campaign period, from December 1st 2016 to March 1st 2017, to members who did not hold any RRSP with Vancity. The observed target response was the purchase of an RRSP during the 2015-2016 campaign. The calculated or predicted response would be the likelihood of purchasing an RRSP, with the calculation based only on a set of predictor variables identified by the modeling process.

Data Preparation

The analysis data set consisted of:

- 2555 Vancity members who acquired a first RRSP during the 2015-2016 RRSP campaign period
- 2555 randomly sampled Vancity members that did not own an RRSP and did not purchase an RRSP term during the same campaign

The constructed analysis database was oversampled with RRSP purchasers to an artificial 50% response rate. The actual RRSP purchase rate (for those not previously owning an RRSP) in 2015- 2016 was 2.2%. Oversampling of very low response rates ensures model estimation can pick up the factors that predict the desired response. To develop the model John used predictor variables available through the core banking system. Variables from this system included:

- ***Member personal information*** such as age, tenure, branch, gender, value segment, behavior segment
- ***Member product portfolio*** that included products held by month and balances by month
- ***Member banking transactions*** by channel and by month

Finally, this data was combined with Tax Filer Data available from StatsCan that provided average data at the neighborhood level. When interpreting the estimated model, care must be taken to remember that these geodemographic variables are not measured at the individual member level.

In the end, John tested over 70 unique variables, half of which had 12 monthly permutations. In the process of modeling, hundreds of additional variables were created through logarithmic transformations and binning.

ASSIGNMENT DIRECTIONS:

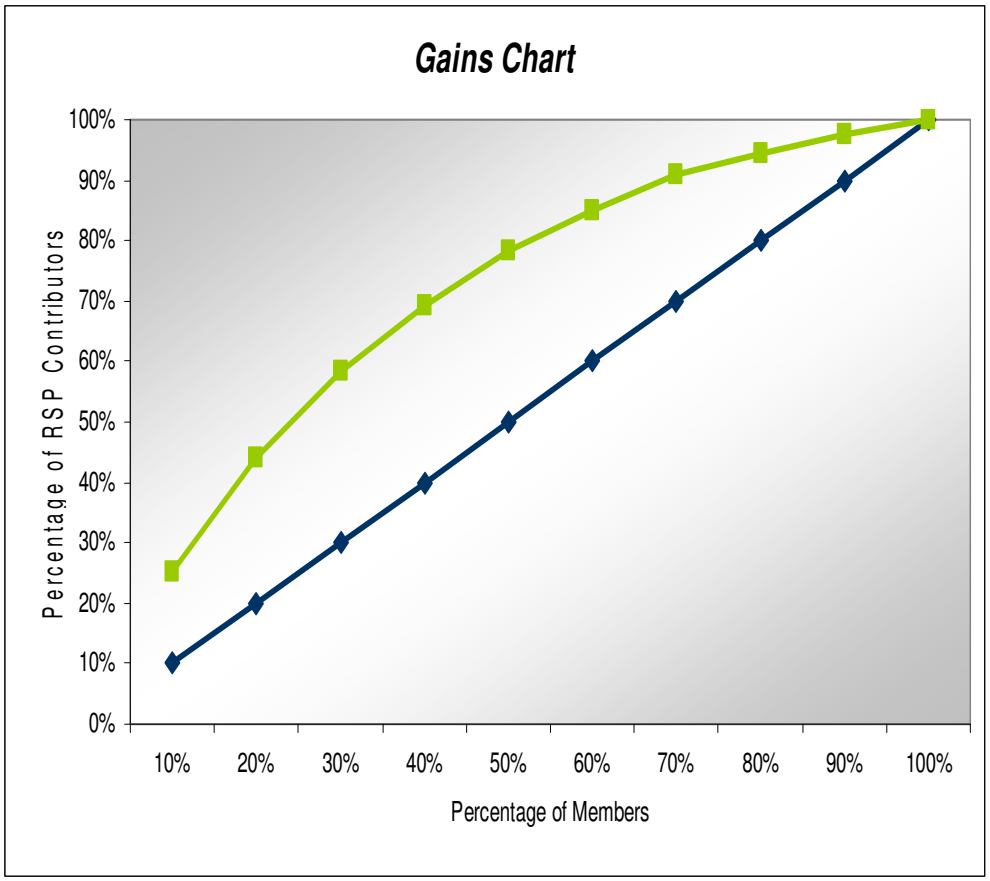
The data is provided in the csv file, ***vcRSP2017.csv***. Most of the original variables have been deleted from the data set, leaving 31 variables to work with (TABLE 1A below). This is real data and will require some thoughtful attention to missing values and other cleaning chores to get good results. Your objective is to develop—fairly quickly—a "good enough" model, and to recommend to VanCity the number of members who should be contacted to maximize profits, (contribution from RRSP sales minus contact costs). It is important to John to be able to interpret the calibrated model to check that it makes intuitive sense, and to sell it to management. That means that the final model **cannot** be a neural network model. Of course, using neural network models at intermediate stages will be a very good idea. For fair competition, everyone should use an Estimation / Validation split of **60 / 40**.

ASSIGMENT DELIVERABLES: 

1. Describe the **key** steps you went through in arriving at your final model, and **why** you took those steps. Follow the "**rapid model development**" framework. The key steps are those that advanced your modeling process, and which will allow evaluation and grading of your results. As a rough guide, 3-5 pages of text, plus figures, are expected *(60 points)*.
2. The final model's lift chart. (*10 points)*
3. Note the variables that you find to be predictive of RRSP purchase, and whether they make intuitive and managerial sense (A few sentences will be adequate). *(10 points)*
4. Using the costs and contribution in Table 2A below, recommend the percentage of members (ie., the "best X%") who should be contacted to maximize profit after contact costs, and report what this expected profit is. *(20 points)*
5. Submit your final workspace (Rdata) file to be evaluated for the competition.

The competition: Grading will be based on your reported process and logic more than your final expected profit. However, consistent with the power of real-world analytic marketing to produce dramatic results on the bottom line, there will be an incentive to maximize profit: **the student team that achieves the highest profit (using the validation set, 40% of the data) will receive a cash award of $60.** In the unlikely event of a tie, the award will be split.

The lift (called “Gains” at VanCity) chart from the final model John developed.



Can you do better? It is possible, but difficult!

Data dictionary and financial details below

TABLE 1A: Variables available in *vcRSP2017*

APURCH	=Y if RRSP purchase, N if not purchased
unique	Vancity member identification number
age	age
gendf	= 1 if female, 0 otherwise
gendm	= 1 if male, 0 otherwise
pcode	postal code
atmcrd	=1 if owns ATM card, 0 otherwise
paydep	=1 if member uses payroll deposit, 0 otherwise
BALCHQ	average monthly balance chequing account over previous 12 months
BALSAV	average monthly balance in savings account over previous 12 months
TOTDEP	average total monthly deposits over previous 12 months
BALLOAN	average monthly personal loan balance over previous 12 months
BALLOC	average monthly line of credit balance over previous 12 months
BALMRGG	average monthly mortgage balance over previous 12 months
NEWLOC	= 1 if acquired a new line of credit in previous 12 months, 0 otherwise
NEWMRGG	= 1 if acquired a new mortgage in previous 12 months, 0 otherwise
TXBRAN	number of in-branch transactions per month, previous 12 months
TXATM	number of ATM transactions per month, previous 12 months
TXPOS	number of Point of Sale transactions per month, previous 12 months
TXCHQ	number of cheque transactions per month, previous 12 months
TXWEB	number of online transactions per month, previous 12 months
TXTEL	number of telephone transactions per month, previous 12 months
TOTSERV	total number of distinct services (distinct product lines) held
CH_NM_SERV	12 month change in number of distinct services
CH_NM_PRD	12 month change in number of products (e.g, Loans, accounts, etc.)
valsegm	internal VanCity "value segment" code, using an adhoc method to rank all members; A = top 5%; B = top 6-10%; C = top 11-20%; D = top 21-70%; E = bottom 30%

GEODEMOGRAPHICS from tax file by Postal Code neighbourhood

N_IND_INC_1	number of individuals reporting income in the PC nbhd.
numrr_1	number of RRSP contributors
numcon_1	number of RRSP contributions (allows for "spousal" contributions)
avginc_1	average employment income
avginv_1	average investment income

TABLE 2A: Costs and Contribution (not the true proprietary figures!)

Contact cost (Mail and glossy brochure production cost) :	$5.30
Number of potential contacts (members without an RRSP):	120,000
Estimated average contribution from a single RRSP purchase:	$205.00