Prof. Dr. Bernd Skiera

Exercises

"Analytics in Customer Value Management"

Winter Term 2019/2020

Prof. Dr. Bernd Skiera, Chaired Professor of Electronic Commerce, Department of Marketing, Faculty of Business and Economics, University of Frankfurt, Theodor-W.-Adorno-Platz 4, 60323 Frankfurt, Germany, Phone.: +49-69-798-34649, Fax: +49-69-798-35001, E-Mail: skiera@wiwi.uni-frankfurt.de, http://www.ecommerce.wiwi.uni-frankfurt.de.

Table of contents

1. Linear Regression: Direct Marketing

Required Dataset: “direct_marketing.csv”

You are hired by a direct marketer, who sells his products only via direct mail. He sends product catalogs directly to customers, who then order directly from the catalogs. Given that this way of selling is rather costly, due to the production and dispatching costs of catalogs, the direct marketer only wants to target the most attractive customers. You receive the following data that includes n = 1000 customers and the following variables:

- Age (of customer): old/middle/young
- Gender: male/female
- OwnHome: whether customer owns home; own/rent
- Married: single/married
- Location: far/close; in terms of distance to the nearest brick and mortar store that sells similar products
- Salary: yearly salary of customer; in dollars
- Children: number of children; 0–3
- History: of previous purchase volume; low/medium/high/NA; NA means that this customer has not yet purchased (ignore this variable!)
- Catalogs: number of catalogs sent
- AmountSpent: Amount of money, that a customer spent with the direct marketer (in dollars)

a) What are the characteristics of the most attractive customers?

b) What is the expected amount of dollars spent of a customer with the following characteristics: young, female, home owner, single, far away living, $30.500 salary, 12 catalogs sent

c) Would you rather target married or single customers?

d) Which customers spent more money than you expected? could look at residuals

2. Calculation of Customer Lifetime Value and Customer Equity

2.1. Customer Lifetime Value: Catalog Company

A catalog company estimates that the costs for each email to a potential customer are 1.20€. The average return rate for email marketing campaign is 1.1%. Therefore, the costs for acquiring one customer are 109.09€ (1.20€ / 0.011). If a new customer orders for 180€ (average order) and the company has a contribution margin rate of 30%, the first investment in a customer is 55.09€ (=109.09€ – (180€ · 0.3) = 109.09€ – 54€). For the calculation of the customer lifetime value the company has to estimate the length of customer relation, the marketing spending per year per customer, the contribution margin of repeated orders and the interest rebate, which could be oriented on the minimum interest rate of investors. What lifetime value of the customer relation follows from a customer relation of three years, four mailings per year (costs 1€ per mailing; return first year: 64%, second year: 52%, third year: 44%), an average order of 200€ for repeated orders and a discount rate of 20% per year?

2.2. Customer Lifetime Value: Shop-for-less

After completing your business administration studies, you start your business "shop-for-less". Your start up's focus is on creating and distributing a shopping coupon newsletter. You publish it monthly and each newsletter includes (targeted for a female and a male customer group) 20 coupons with discounts for each of the two customer groups. For that purpose, "shop-for-less" has contacted 20 companies that agreed to offer one coupon for each newsletter and each customer group. For appearing in the newsletter, each company pays "shop-for-less" 0.10€ per newsletter for the female customer group and 0.20€ for the male customer group. You distribute the newsletter through an annual subscription of 50€ that the customers have to pay at the beginning of the year. The customers can cancel the contract at the end of the year.

The acquisition costs for "shop-for-less" are 44€ for female customers and 68€ for male customers. The costs for "shop-for-less" for preparing the coupons, marketing costs and distribution per customer and newsletter are 1.20€.

In the first year after starting the business, "shop-for-less" has acquired 5,000 female and 2,500 male customers. A consultant estimates a yearly retention rate of 70% for female customers and 80% for male customers. The yearly discount rate is 10%.

a) How high is the customer lifetime value (before and after acquisition cost) for the female and male customer group? Please describe all steps of your calculation and use yearly values.

b) According to customer equity (i.e., customer lifetime value times number of customers), which customer group is more profitable? Draft the steps and results.

c) Would the results in b) change if the customer retention rate changed and only 20% of the female customers and 10% of the male customers called their contract off in the first year? Explain your approach.

2.3. Customer Lifetime Value: Velo Mobile

Velo Mobile Telecommunications spends at a cost of $840 per new customer, on average. Each customer incurs $60 in maintenance cost and $30 record-keeping and billing costs annually. Customers can buy into three different plans: 30% of customers buy a premium package at $65 per month; 20% buy the premium package and also add a data plan for a mobile device (laptop, tablet, etc.), for an additional $35, for a total of $100 per month (the "super-premium package"); the remaining customers buy the basic service package at $40 per month. Over time, 75% of customers remain with the company from one year to the next. Assume a discount rate of 10%.

a) What is the lifetime value of a super-premium customer (premium package + data plan)?

b) What is the lifetime value of an average customer?

c) If the company currently has one million customers, what is the maximum amount of money it should be willing to spend to improve its customer retention rate from 75% to 85%?

2.4. Customer Equity Reporting

A company has a subscription-based business and sends each customer an invoice at the beginning of each month that is usually paid at the middle of the month. Customers do not start to churn before making the first payment The company calculates with a discount rate of 8% and provides you with the following information:

	Year 1	Year 2
Gross customer additions	1,600	1,800
Net customer additions	500	600
Number of customers at end of year	2,700	3,300
Average Revenue per Customer (ARPU)	$40.00	$44.00
Margin	25%	20%
Cost per Gross Addition (CPGA)	$12.00	$15.00

a) Calculate customer equity of its current customer base in both years. Outline all assumptions that you make.

b) Outline the most important shortcoming of your calculations.

2.5. Customer Lifetime Value & Customer Equity: Blue Apron

Blue Apron is an ingredient-and-recipe meal kit service that exclusively operates in the United States. The weekly boxes contain ingredients and also include suggested recipes that must be cooked by hand by the customer using the pre-ordered ingredients. Hello Fresh makes a comparable offering in Germany.

The financial report of Blue Apron discloses acquisition cost per customer of $94 as well as the following values:

Total Marketing Expenses (in million$, $M):

- 2014: $14
- 2015: $51
- 2016: $144

	Q4 ‘15	Q1 ‘16	Q2 ‘16	Q3 ‘16	Q4 ‘16	Q1 ‘17
Total Number of Active Customers	429,000	649,000	766,000	907,000	879,000	1,036,000
Number of Acquired Customers	135,638	382,979	382,979	382,979	382,979	162,234
Average Revenue per Customer	$272	$265	$264	$227	$246	$236

a) Estimate the average customer lifetime value (CLV) in Q2/2016 and Q3/2016. Make assumptions in case that some information that you require is not provided here.

b) Estimate Customer Equity of the customer base in Q2/2016 and Q3/2016.

c) Discuss one major shortcoming of your calculations.

2.6. Customer-Based Firm Valuation: Basic

A company, called Basic, with contractual customer relationships has carried over 50 existing customers from the previous year and will acquire 20 new customers in each of the following years. The acquisition of each new customer costs 30$, retaining each customer costs 20$ (also in the year of acquisition), revenue per customer and year is 200$ and the profit margin 50%. Consequently, the profit contribution per customer before retention cost is 100$. The average retention rate per customer is 80%, the company's discount rate is 10% and all revenues and costs occur at the beginning of each year. Fixed cost are 5,000$ per year and the company has debt of 10,000$. You can neglect taxes. Assume that all revenues and costs are cashed in immediately so that revenues equal cash inflows and costs equal cash outflows. The following table summarizes all values. What is the value of this company?

Variable	**Value**
Number of current customers	50
Number of new customers per period	20
Revenues per customer and period	200$
Profit margin	50%
Profit contribution per customer and period	100$
Fixed cost	5,000$
Retention rate	80%
Retention Cost	20$
Acquisition cost	30$
Discount rate	10%
Non-Operating Asset	0$
Debt	10,000$
Tax rate	0%

3. Calculation of Retention Rates, Hazard Rates and Survival Analysis

3.1. Cohort Analysis

Below you can find two alternatives to analyze cohorts (see http://christophjanz.blogspot.com/2013/10/excel-template-for-cohort-analyses-in.html). Thereby, MRR stands for "Monthly Recurring Revenue".

Alternative 1:

Conversion month	New customers	# of retained customers in month									A1
		Jan-13	Feb-13	Mar-13	Apr-13	May-13	Jun-13	Jul-13	Aug-13	Sep-13	Oct-13
Jan-13	80	78	75	72	70	69	67	66	66	65	64
Feb-13	88		88	86	82	78	77	76	73	72	70
Mar-13	105			103	103	98	94	92	90	86	82
Apr-13	110				107	106	102	99	97	92	90
May-13	115					114	112	105	98	97	96
Jun-13	128						128	122	119	115	110
Jul-13	137							136	129	122	118
Aug-13	151								149	145	135
Sep-13	161									158	154
Oct-13	168										167
		78	163	261	362	465	580	696	821	952	1,086

		% of retained customers in lifetime month									B1
		0	1	2	3	4	5	6	7	8	9
Jan-13	80	97.50%	93.75%	90.00%	87.50%	86.25%	83.75%	82.50%	82.50%	81.25%	80.00%
Feb-13	88	100.00%	97.73%	93.18%	88.64%	87.50%	86.36%	82.95%	81.82%	79.55%	
Mar-13	105	98.10%	98.10%	93.33%	89.52%	87.62%	85.71%	81.90%	78.10%		
Apr-13	110	97.27%	96.36%	92.73%	90.00%	88.18%	83.64%	81.82%			
May-13	115	99.13%	97.39%	91.30%	85.22%	84.35%	83.48%				
Jun-13	128	100.00%	95.31%	92.97%	89.84%	85.94%					
Jul-13	137	99.27%	94.16%	89.05%	86.13%						
Aug-13	151	98.68%	96.03%	89.40%							
Sep-13	161	98.14%	95.65%								
Oct-13	168	99.40%									
		98.79%	96.00%	91.36%	88.07%	86.58%	84.54%	82.25%	80.59%	80.36%	80.00%

Alternative 2:

Conversion month	New MRR
Jan-13	$ 7,851
Feb-13	$ 7,988
Mar-13	$ 9,234
Apr-13	$ 10,123
May-13	$ 11,788
Jun-13	$ 12,876
Jul-13	$ 13,212
Aug-13	$ 14,789
Sep-13	$ 15,786
Oct-13	$ 16,119

Cohort MRR in month									C1
Jan-13	Feb-13	Mar-13	Apr-13	May-13	Jun-13	Jul-13	Aug-13	Sep-13	Oct-13
$7,650	$7,536	$7,234	$6,982	$6,876	$6,780	$6,678	$6,567	$6,454	$6,342
	$7,988	$7,688	$7,232	$6,998	$6,890	$6,789	$6,679	$6,589	$6,459
		$9,145	$8,897	$8,678	$8,569	$8,432	$8,352	$8,152	$7,988
			$9,988	$9,657	$9,222	$8,901	$9,201	$9,468	$9,688
				$11,678	$11,546	$11,321	$11,678	$11,755	$11,868
					$12,876	$12,546	$12,433	$12,399	$12,976
						$12,987	$12,756	$13,345	$13,456
							$14,786	$14,201	$14,897
								$15,676	$14,988
									$15,897
$7,650	$15,524	$24,067	$33,099	$43,887	$55,883	$67,654	$82,452	$98,039	$114,559

		% of retained MRR in lifetime month								D1
		0	1	2	3	4	5	6	7	8	9
Jan-13	$7,851	97.44%	95.99%	92.14%	88.93%	87.58%	86.36%	85.06%	83.65%	82.21%	80.78%
Feb-13	$7,988	100.00%	96.24%	90.54%	87.61%	86.25%	84.99%	83.61%	82.49%	80.86%	
Mar-13	$9,234	99.04%	96.35%	93.98%	92.80%	91.31%	90.45%	88.28%	86.51%		
Apr-13	$10,123	98.67%	95.40%	91.10%	87.93%	90.89%	93.53%	95.70%			
May-13	$11,788	99.07%	97.95%	96.04%	99.07%	99.72%	100.68%				
Jun-13	$12,876	100.00%	97.44%	96.56%	96.30%	100.78%					
Jul-13	$13,212	98.30%	96.55%	101.01%	101.85%						
Aug-13	$14,789	99.98%	96.02%	100.73%							
Sep-13	$15,786	99.30%	94.94%								
Oct-13	$16,119	98.62%									
		99.09%	96.30%	96.02%	94.40%	93.77%	92.07%	88.64%	84.33%	81.53%	80.78%

a) Briefly outline how to interpret the numbers in each of the two alternatives.

b) Outline which of the two alternatives you prefer and why you do so.

c) For your preferred alternative, outline briefly what can be learned by looking at the values in each row and each column.

d) Outline briefly how a hazard model could help you in getting insights. Outline the dependent as well as the independent variables of your hazard model.

e) What is the major difference between a parametric distribution according to an Exponential and a Weibull distribution? Which of those two distributions would you prefer to use?

3.2. Logistic Regression: Pilgrim Bank

Related dataset: “pilgrim_bank_b.csv”

This question is related to the case study “Pilgrim Bank” that was used in class. The data set to this case study still consists of 31,634 observations and the following variables that you know from the class:

- 9Online: 1 if online banking is used, 0 otherwise
- 9District: three districts which are coded as “1100”, “1200”, and “1300”
- 9Age: an integer value from 1 to 7 that represents different age groups
- 9Income: an integer value from1 to 9 that represents different income groups
- 9Tenure: the number of years that customer is with the bank
- 9Profit: the profit that the bank made with the customer in 1999

In addition, the case now contains a variable “0profit”, which is the profit that the bank made with the customer in 2000. It contains a missing value if the customer left the bank.

a) Develop a model to predict the likelihood to churn for each customer. Briefly explain your motivation behind the use of your model.

b) Use your model to predict the likelihood to churn for one customer and use “manual calculations” (e.g., via calculations done in excel or your pocket calculator) to outline that the model indeed predicted the correct value.

c) Use the classification matrix and evaluate the prediction quality of your model.

3.3. Survival Analysis: HappySound

Related dataset: “happy_sound.csv”

The online music streaming service “HappySound” wants to decrease customer churn. Therefore “HappySound” upgraded a random group of 12 users, by providing these users with a new feature in their streaming service for four weeks. “HapySound” provided some of the upgrades to new users, and some to existing users. The users were able to churn any moment in time. Your task is now to perform a survival analysis of “HappySound” users. You therefore receive a small data set with in total 26 users and the following variables (see “churn_data.dta”):

- futime: Time variable, indicating the survived time
- fustat: Dummy variable, indicating if the customer churned (=1) or not (=0)
- age: Age of the customer

- won_back: Dummy variable, indicating if the customer previously wanted to churn and was won back, with another marketing action (=1) or not (=0)
- upgrade: Dummy variable, indicating if the customer received the feature upgrade (=1) or (=0)
- gender: Dummy variable, indicating of a customer is female (=1) or male (=0)

Hint: Use the "survminer" R-package to answer the question. The "survminer" Cheat sheet is available online.

a) What challenge regarding the data structure do you have in this case?

b) Create and interpret a Kaplan-Meier-Curve for won back customers and for customers that received the feature upgrade. Hint: Use the ggsurvplot()-command in R.

c) Estimate a Cox-Proportional-Hazard Model. Is the feature upgrade useful to decrease the probability of a customer to churn? Hint: Use the coxph()- command in R.

d) How strong does age (i.e., one additional year of age) affect the hazard of a customer to churn?

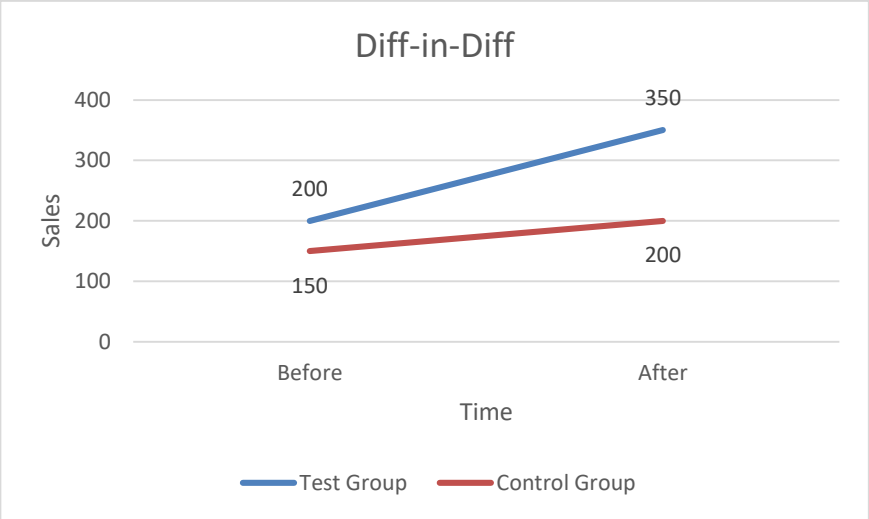
4. Difference-in-Difference and Before-and-After Analysis

4.1. Sun Cream

The sun cream producer "SunProtect" recently launched a redesigned product packaging. SunProtect now wants to estimate the effect of the new design on sun cream sales.

a) Would it be enough to compare sales before and after the launch of the new product packaging? Justify your answer.

b) The "SunProtect" customer analytics unit runs a field experiment, with a control group (supermarkets, that only sell the old product packaging) and an observation before and after the treatment. As result, you can observe the following numbers:



Note: Sales is the average monthly number of sold sun cream units per supermarket.

How big is the observed effect of the packaging change on the average monthly sales per supermarket?

4.2. Mississippi Retailer

Related dataset: "mississipi_data.csv"

Mississippi is an online retailer selling products from various categories, from books to electronics and from clothing to jewelry. Mississippi runs many online advertising campaigns, all aimed at the specific product categories which they sell. Especially for the fashion category Mississippi is investing heavily in online advertising, given this category has a high gross profit margin of 30% and a CPM of €2.50 on average. A problem is that Mississippi does not know how profitable online advertising actually is, and what ROI their advertising has.

Recently Mississippi faced a budget constraint problem: they temporary run out of budget for the advertising campaign which focuses on their fashion category. As a result, there was not anything spend on advertising in this category for a total of 20 days. For the other categories the advertising campaigns did however continue as normal.

Mississippi has asked you to have a look at this data. They have provided you with the daily gross sales revenues (measured in euros) for two categories, namely fashion and electronics. The data consists out of 25 days before the period of the budget constraint, 20 days during the budget constraint, and 15 days after the budget constraint. On the days where advertising was active, they spend on average €5,000 per day on fashion related advertising and €0 during the

days of the constraint. Mississippi wants to know to what extend advertising in the fashion category contributes to sales and what the ROI is.

a) Perform a before-during-after analysis on the provided data. Apply log-transformation of variables, if needed.

b) Motivate your decision on the setup of the model!

c) Interpret all parameters of the model. Are the parameters in line with what you would expect? Why (not)?

d) Use the model to calculate the ROI of advertising.

e) How certain are you about the causality of the findings? Explain why!

4.3. Amazon

Related dataset: "amazon_data.csv"

Amazon Prime has been launched in different years for the different regions where Amazon is active; in January 2005 it was launched in the US, while in January 2013 it was launched in Canada. Given that these two markets are very similar to each other in terms of growth rate before the launch of Amazon Prime, Amazon's CMO wants to investigate the impact of Amazon Prime on revenue with a before-and-after analyses and a difference-in-differences analysis. The result should show the true value of Amazon Prime. You get a dataset with the revenue (in millions of dollars) of Amazon from the US and from Canada from January 2003 until December 2006 (see "amazon_data.csv").

a) Use the dataset to perform a before-and-after analysis to examine the impact of introducing Amazon Prime on revenue of Amazon in the US. Interpret the model (i.e. all parameters, and indicate if they make sense in terms of direction and (in)significance and what each parameter means) and explain the reasons for the choices that you have made in terms of the independent variable(s) and transformation of variables.

b) Perform on the dataset a difference-in-differences analysis to find out what the impact of introducing Amazon Prime is for the revenue of Amazon in the US; also interpret the model (i.e. all parameters, and indicate if they make sense in terms of direction and

(in)significance and what each parameter means). Tip: to perform a difference-in-differences analysis on the dataset, you do need to restructure the dataset.