Attrition Campaigns: Who Should We Target? [white paper]

Annegret Henninger

December 17, 2020

Introduction

In the mid-twentieth century, mass production and mass marketing increased product availability which, in turn, decreased the time spent getting to know customers in order to find products that meet their needs (Chen & Popovich, 2003). Since then, the development of the data-driven economy and the advancement of analytical tools has re-introduced the value of information on customers for Customer Relationship Management (CRM) (Chen & Popovich, 2003). This is especially valuable in competitive markets that invest heavily in customer retention and acquisition campaigns. For these reasons, CRM became a dominant research area in the 1990s with relationship-based competition research continuing as a dominant area into the 2000s (Merali, Papadopoulos, & Nadkarni, 2012). One of the main goals of CRM is to understand and reduce customer churn as “the central purpose of managing customer relationships is for the enterprise to focus on increasing the overall value of its customer base—and customer retention is critical to its success” (Peppers & Rogers, 2004, p.15).

Knowledge management is one of the most critical challenges and success factors of CRM and includes information processing, storing, and communicating knowledge (Gebert, Geib, Kolbe & Brenner, 2003). Furthermore, Davenport, Harris, and Kolhi, (2001) point out that a key challenge is capturing useful knowledge and not wasting resources gathering and storing useless knowledge. Valuable knowledge about customers includes customer connections, histories, requirements, expectations, and purchasing activity (Gebert et al., 2003; Davenport, et al., 2001; Day, 2000).

While churn rates are an important and popular research area, most of the literature focuses on the attrition problem from a binary perspective: churn/not churn. In this paper, I will investigate which telcom customers are the most valuable before determining their level churn risk, to equip managers with more sophisticated information enabling them to develop more strategic attrition campaigns. I will especially focus on the customer's age, and life expectancy, and the services they are likely to use based on their demographic. As such, the research question is: Who *should* be targeted with preventative attrition campaigns based on a combination of the customer’s value and churn risk?

I will be analyzing the IBM dataset for telcom customer churn. Knowledge-based churn prediction is particularly valuable for telcom companies as they operate in highly competitive markets (Jafari-Marandi, Denton, Idris, Smith & Keramati, 2020).

Literature Review

Globalization, digitization, and digitalization have supported entrepreneurship by lowering uncertainty and creating more fluidity in the entrepreneurship process (Nambisan, 2017). The distinction between the related concepts of digitization and digitalization should be drawn. Tilson et al., (2010) define digitization as a technical process that renders technology that is analog or physical into a digital, machine-readable format. Meanwhile, they define digitalization as a socio-technical process of integrating digital processes, applications, and techniques to the broader social, economic, and organizational contexts rendering the infrastructure digital. With waves of digitalization as identified by Legner et al. (2017), there exists ubiquitous underlying infrastructure to support expansive, scalable entrepreneurial activities. Digitalization weakens economic barriers to entry and enables new, affordable avenues to reach consumers, leading to market saturation in certain industries (Hervé, Schmitt & Baldegger, 2020; Nambisan, 2017; Manyika, Lund, Bughin, Woetzel, Stamenov & Dhingra, 2016).

While market saturation can increase competition which can have positive impacts such as driving innovation, the impact is also strongly felt on customer retention and churn rates (Colgate, Stewart, & Kinsella, 1996). De Caigny, Coussement, and De Bock (2018) found that there are two streams of work in addressing customer churn: (1) improving existing churn prediction models, and (2) understanding customer churn and identifying main drivers. I cover both in the background literature but approach the problem from the managerial perspective by using existing data analytical tools to explore the variables in my dataset in order to identify patterns and main predictors of customers abandoning a telcom brand.

Model performance approach

Table 1 below shows a list of heavily cited academic papers that evaluated churn rates, and the analysis techniques they used.

Table 1: Analytics used in churn evaluations in the literature

Authors	Dataset	Logistic regression	Decision trees	Random forests	SVM	Classification trees	Logistic model trees	Bagging	Random subspace method	Generalized additive models	Naive Bayes	Bayesian networks	ANN	KNN	Nearest neighbors	Logit leaf model
Verbeke, Dejaeger, Martens, Hur, & Baesens (2012)	Telcom	*	*	*	*		*	*			*	*	*		*	
Chen, Fan, & Sun (2012)	Food, telcom, adventures	*	*	*	*								*			
Coussement, Lessman, & Verstaeten (2017)	Telcom	*	*	*	*			*			*	*	*			
Moeyersoms & Martens (2015)	Energy	*	*		*											
Sundararajan & Gursoy (2020)	Telcom	*	*	*	*											
De Bock & Van den Poel (2012)	Bank, DIY, telcom, mail order	*		*				*	*	*						

	clothes															
Ballings & Van den Poel (2012)	Newspaper	*				*		*								
De Caigny, Coussement, & De Bock (2018)	14 different industries	*	*													*
Jafari-Marandi, Denton, Idris, Smith, & Keramati (2020)	Telcom												*			
Keramati, Jafari-Marandi, Aliannejadi, Ahmadian, Mozaffari, Abbasi, (2014)	Telcom		*		*								*	*		
Coussement & Debock (2013)	Online gambling		*	*						*						

The table shows that logistics regression, decision trees, random forest, and SVM are the most popular tools across the papers reviewed. Though De Caigny et al. (2018) argue that comprehensibility and predictive performance are the two most important factors in models used for customer churn prediction, predictive performance is often the main or even sole performance measure.

Logistic regression and decision trees are strong in both predictive performance and comprehensibility models and, as such are widely used algorithms (Verbeke et al., 2012). However, while decision trees are effective at handling interaction effects between variables, it struggles with the linear relationships between them (De Caigny et al., 2018). Logistic regression, which is regarded as the gold standard in churn prediction because of its strong performance, does the reverse -it is suited for linear relationships but fails to understand the relationships between variables (Coussement, Lessman, & Verstraeten, 2017; Neslin et al., 2006). De Gaigny et al., (2018) proposed a hybrid model of the two models named the Logit leaf model (LLM). Their model splits the data into subsets to which logistic regression is then applied.

Random forest also returns high predictive performance and is often used in customer churn prediction models (Coussement & Van den Poel, 2008). Sundararajan & Gursoy (2020) argued that logistic regression, random forest, and SVM all performed better than decision trees for churn prediction. While logistic model trees are less used, Verbeke et al., (2012) did a comparative study of algorithms and found it to also have strong predictive performance in churn analysis.

Comprehensibility is equally important, especially for the portion of the research that focuses on identifying churn drivers for managing attrition rates. The comprehensiveness of a model is a result of the size and classification output (Martens et al., 2011). Output can be in the form of rule-based models, tree-based models, linear models, and non-linear models. Meanwhile, the size, which Zaheer, Ahmed & Smola (2017) argue can render a model uninterpretable, correlates to the output type. The number of terms determines the size of linear and non-linear models, while the number of leaves and rules determine the size of tree-based and rule-based models (De Caigny et al., 2018).

It is generally accepted that smaller models are more comprehensive which stems from seminal works in psychology on working memory and short-term memory (Zaheer et al., 2017; Baddeley, 1996). Working memory which evolved from the earlier concept of short-term memory is the short-term processing, and storage of information enabling us to perform cognitive skills such as reasoning, learning, and comprehending (Baddeley, 1996). Cowen's (2001; 2016) works evaluate memory capacity and find that people can retain three to four unlinked, chunked, or separate items of information. Chunking is the grouping of items into larger groups like categories of variables, or chunks of digits in a phone number (Baddeley, 1996). Miller's (1956) "magical number 7" is often cited. However, Miller clarifies in his 1989 autobiography that it was meant in a tongue-in-cheek fashion, which readers have misinterpreted at face value.

The comprehensivity of the output type is not widely agreed upon and can depend on the subject matter needed to be communicated. However, non-linear models seem to be regarded as the least comprehensive (De Caigny et al., 2018). For these reasons, the effective analysis of churn data requires data-driven decisions throughout the analytical process to determine the most appropriate models with regards to their performance, and comprehensibility (Lima, Mues & Baesens, 2011; Lessman & Voß, 2009).

Managerial-based churn research

Analyzing customer information has proven to be one of the most valuable tools in understanding and preventing customer churn (Ganesh, Arnold, & Reynolds, 2000; Verhoef, 2003). It can help identify those customers who have less brand loyalty, and who are more likely to leave (Shaffer & Zhang, 2002). This enables marketers to strategically build targeted customer retention campaigns based on a better understanding of those customers' inclinations (Blattberg, Kim & Neslin, 2010). Many authors point out that this additional information can improve customer segmentation (Hansen, Samuelsen & Sallis, 2013; Seret, Verbraken, Versailles, & Baesens, 2012).

The performance and comprehensibility of models discussed in the model predictive performance approach is also a crucial component of the managerial approach. Jafari-Marandi et al., (2020) used artificial neural networks for telecom customer churn prediction. They argue that the action capability and comprehensibility needed to communicate the information of predictive models depend on the successful identification of patterns in the data.

Determining the root causes of churn rates is an important research discipline as customer acquisition is several times more costly than customer retention. (Torkzadeh, Chang, & Hansen, 2006). Van den Poel and Larivière (2004) demonstrated that increasing customer retention can have an impact on growth that is up to four times higher than customer acquisition. This is true for several reasons:

- Nurturing existing customer relationships enables organizations to focus on existing customer needs (Gebert et al., 2003; De Caigny et al., 2018).
- Searching for new customers is expensive (Dawes & Swaukesm 1999);
- New customers are often associated with higher attrition rates (De Caigny et al., 2018);
- If existing customers leave, they may influence others within their social circles to follow suit (Nitzan & Libai, 2011).
- Long term customers are likely to spend more (Ganesh et al., 2020);
- Existing customers may recruit more customers from word of mouth (WOM) (Ganesh et al., 2000);
- Organizations have more data on long-term customers which can be used to decrease service costs (Ganesh et al., 2020);
- If organizations can keep existing customers to become long-term customers, they are less susceptible to competitive marketing (Colgate et al., 1996);
- A loss of existing customers increases the need to spend on customer acquisition to maintain revenue sources

However, Ascarza et al. (2018) find that targeting customers most likely to abandon the company may be ineffective and even harmful. They find that a more holistic approach should be considered to determine who is at risk, why they are at risk, who should be targeted, when should they be targeted, and with what incentives should they be targeted. I take a holistic approach to determining which customers should be targeted based on their risk and value as customers. Also, I agree that the timing (preventative or reactionary campaigns), and incentives used should be determined before taking action.

The difficulty in determining causal inference for customer churn to take action stems from the counterfactual problem; we must try and predict what the customers would do in an alternative hypothetical situation. For example, if we predict variable *x* (e.g., payment method) to be a main influencing factor in churn, we do not know with certainty if the customer would churn or not if we offered the customer an alternative payment method (Li & Pearl, 2019). The problem is thus predicting the accurate counterfactual behaviour of the customers with the given characteristics available in the dataset.

Methods and Data

I use the IBM Telco Customer Churn dataset (https://www.kaggle.com/blastchar/telco-customer-churn). The dataset contains 7043 customers (rows), 21 features (columns), and the target variable is ‘churn’. It includes information on customers who left in the last month (churn), services that customers were signed up for (phone, internet, multiple lines, online security, device protection, tech support, online backup, and streaming TV and movies), demographic information (senior/not senior, gender), their social connections (dependent(s)/no dependent(s), partner/no partner), and the customer account information comprising of the tenure (length of customer-company relationship), contract (month-to-month, 1-year, 2-year), billing method (paper or paperless), monthly charges, payment method (paper, automatic) and total charges). Table 2, below, shows the variables of the dataset and related background literature.

Table 2: Dataset variables

Variable	**The customer**	**Background**
Gender	Is male or female	Sundararajan and Gursoy (2020) found there to be no correlation between gender and churn rates.
SeniorCitizen	Is a senior citizen or not	Younger customers have the potential to be longer-term customers (Jafari-Marandi et al., 2020). The senior age (65) is slightly lower than the average life expectancy of many countries (Jafari-Marandi et al., 2020).
Partner	Has a partner or not	Partner and dependents: A customer is more likely to churn if their connections churn (Verbeke, Martens, Baesen, 2014).
Dependents	Has dependent(s) or not	Partner and dependents: In telcom customer churn research, “social embeddedness” has demonstrated to be negatively correlated to attrition (Benedek, Lublóy & Vastag, 2014).
Tenure	Number of months the customer has been with the company	Long-term customers are more likely to spend more (Ganesh et al., 2020). New customers are associated with higher attrition rates (De Caigny et al., 2018) Organizations have more data on long-term customers which can be used to decrease service costs (Ganesh et al., 2020). Long-term customers are less susceptible to competitive marketing (Colgate et al., 1996). Tenure plays a large role in customer churn (Sundararajan & Gursoy, 2020).
PhoneService	Has a phone service or not	
MultipleLines	Has multiple phone lines or not	
InternetService	Type of service: DSL or fiber optic	Internet service plays a large role in customer churn (Sundararajan & Gursoy, 2020).
OnlineSecurity	Has online security (Yes), or (No), or (No internet service)	
OnlineBackup	Has online backup or not	
DeviceProtection	Has device protection or not	
TechSupport	Has tech support or not	

StreamingTV	Has streaming TV or not	
StreamingMovies	Has streaming movies or not	
Contract	The contract term of the customer: month-to-month, year, or two year	Sundararahan and Gursoy (2020) found that the contract term, together with the tenure and billing method affected churn.
PaperlessBilling	Has paperless billing or not	Automatic billing can play a large role in reducing customer churn (Sundararajan & Gursoy, 2020).
PaymentMethod	Customer's payment method	
MonthlyCharges	Monthly charges	Monthly charges appear to play a large role in customer churn (Sundararajan & Gursoy, 2020).
TotalCharges	Total charges	
Churn	Has churned in the last month	

A customer does not usually churn due to one element. Sundararajan and Gursoy (2020) were able to determine the likelihood of churning through a combination of the customer's tenure, length of time customers were customers and their payment method customer. However, we will explore these individual variables separately. The table shows that long-term customers are less susceptible to attrition campaigns and are likely to spend more, leading to hypotheses 1 and 2.

H1: Customers with longer tenure (> 1 year) are less likely to churn than newer customers (<1 year). Though it still needs to and will be explored the threshold at which the length of tenure becomes sufficient to significantly impact churn prediction.

H2: Customers with longer tenure (>1 year) spend more than newer customers (< 1 year). Again, if true, the difference still needs to be quantified.

Baby-boomers who are currently seniors are often thought of as the age range with the most disposable income, and thus a desirable target demographic. However, they may not be the most lucrative target for telcom companies. With digitization and digitalization ramping up in waves since the IT industrialization era of the 1990s, I predict the social aspects of seniors' lives are less digitalized, and thus have less use for the additional, internet-based services (Legner, et al., 2017). This leads to hypothesis 3.

H3: Seniors have lower total charges. With senior citizens, the customer's value should be based on their average annual income, multiplied by their expected remaining life expectancy. This is needed to determine the economical significance of this demographic. The IBM dataset does not give the specific age of each customer, but we can assume that customers who are not seniors have a longer life expectancy.

The research also shows that long-term customers are less likely to churn and are less susceptible to competitive acquisition campaigns, leading to hypothesis 4.

H4: seniors are less likely to churn than non-seniors. If this is true, and coupled with the diminishing life expectancy of seniors, seniors may be more suited to reactionary campaigns as opposed to preventative campaigns. If H3 and H4 are true, I will explore the marginal benefits of attrition campaigns, once customers pass the 'senior' age threshold. This is needed to determine the economical significance of this demographic.

Table 2 also indicates that social networks (embeddedness and connections) within a telcom brand has an impact on customer churn. However, the IBM dataset does not include information on the partner's or dependent's telcom company, limiting any analysis. Nonetheless, these variables will still be explored to see if they have any impact on churn.

In the data exploration and analysis, I will evaluate the observations in each of the variables, and observe the distribution of the respondents' characteristics. This may indicate if a group of respondents need to be weighted. I will also explore the variables to determine which customers are high-value customers, meaning customers with a large total spend, and with a wide range of service subscriptions indicating high upsell potential. Then, I will determine the churn rates of both groups to determine who should be targeted with attrition campaigns based on a combination of the customer's value and churn risk.

Based on the performance evaluation of analytical tools in previous works on churn analysis, I will use logistic regression with marginal analysis to address the counterfactual problem. I may also use decision trees, and SVM if logistic regression does not adequately show the relationship between variables when evaluating the interactions among variables in a linear model. Because smaller models are more comprehensive I will be focusing on building a model based on significant variables. Again, I may also explore using linear, rule-based, or tree-based models which will be determined on the consideration of a combination of the comprehensivity and accuracy of the model. Martens et al. (2011) suggested using a combination of different output types to be able to represent robust, complex models that can be flexible and generalizable, together with more comprehensive simpler models for communication and visualization. As highlighted in the literature review, the best models on a measure of performance and comprehensiveness will need to be determined during the analytical process (Lessman & Voß, 2009; Lima, Mues & Baesens, 2011).

If I notice significantly higher predictive performance and accuracy amongst complex models, I will consider chunking variables into related categories to increase the comprehensibility of models while maintaining performance to meet both managerial value and technical prediction value.

Results

In the dataset, 5174 customers churned in the past month, and 1869 did not churn. Figure 1 shows that there is a large number of individuals who churned who had low total charges, and customers who did not churn are distributed more heavily amongst the high total charges. The total charges are a combination of monthly charges and tenure. The distribution of monthly charges in Figure 2 below shows that most monthly charges are between $10 and $30. After that, most monthly charges are between $70 and $100. This study will focus on the latter group which is cumulatively larger. However, Figure 3, below, shows that of those who churned a high density of them had higher monthly charges. The goal is to determine which customers have high charges and do or do not churn to determine who to focus attrition campaigns on. The results are further broken down by hypothesis.

Figure 1: Total Charges of Churn vs. Not Churn

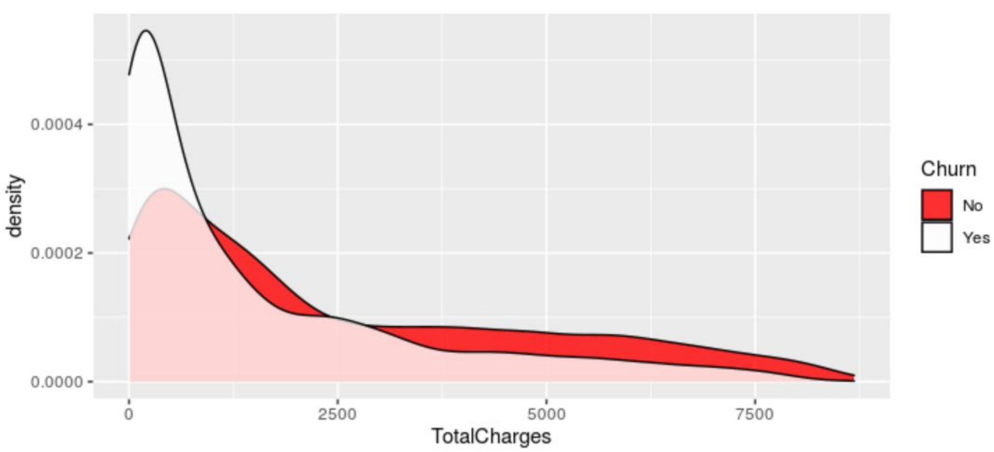


Figure2: Distribution of Monthly charges

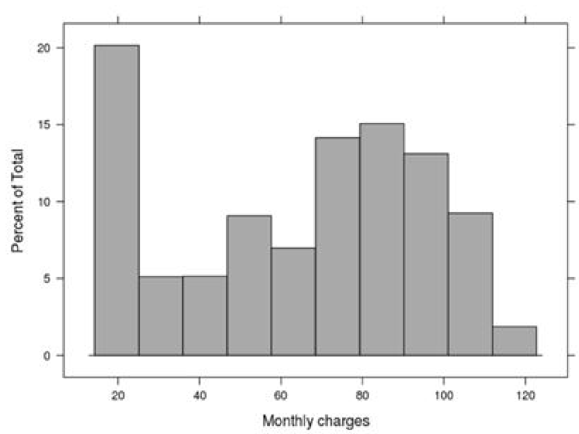
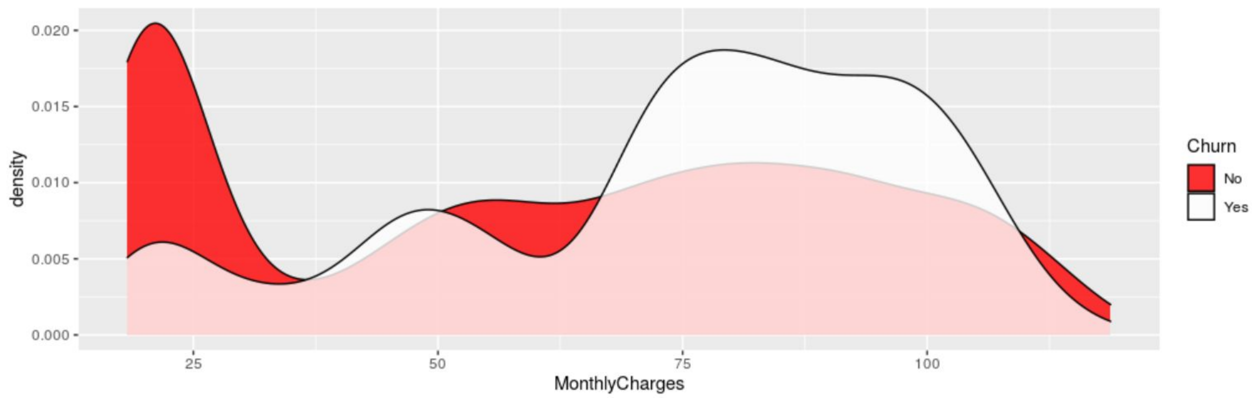


Figure 3: Density of Monthly Charges



H1: Customers with longer tenure (> 1 year) are less likely to churn than newer customers (<1 year).

The mean tenure (in months) is 32 with a standard deviation of 25, a minimum tenure of 0, and a maximum tenure of 72. The first quartile is 9 and the third quartile is 55. By using a linear function, it was found that the mean tenure of customers who churned was 17.98 months while for customers who did not churn it was 37.57. Both values were statistically significant with a p value <0.01. Table 3, below, shows the number and percentages of people who churned and did not churn.

Table 3:Churn by Tenure

	0-1 Year	**1-2 Years**	**2-3 Years**	**3-4 Years**	**4-5 Years**	**5 + Years**
Not churned	1149	730	652	617	712	1314
Churned	1037	294	180	145	120	93
Of those who did not churn	22.20%	14.10%	12.60%	11.90%	13.80%	25.40%
Of those who churned	55.50%	15.70%	9.60%	7.80%	6.40%	5%
Of the given tenure range, customers did not churn	52.60%	7.13%	7.84%	8.10%	85.60%	93.40%

Of the given tenure range, customers churned	47.40%	28.70%	21.60%	19%	14.40%	6.60%

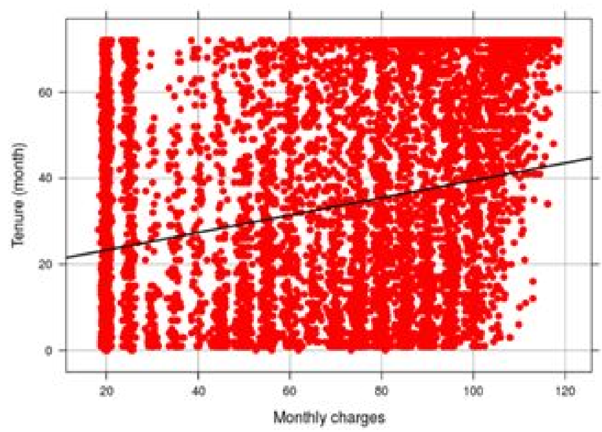
H2: Customers with longer tenure (>1 year) spend more than newer customers (< 1 year).

A linear model was run against a categorical dummy variable of tenure where each category equals 1 year. The results are in Table 4, below. Figure 4 below illustrates the positive correlation between monthly charges and tenure.

Table 4: Tenure Linear Regression

	Dependent variable: Monthly charges
1-2 year,	5.259***
	(1.105)
2-3 years	9.478***
	(1.189)
3-4 years	10.220***
	(1.228)
4-5 years	14.453***
	(1.189)
5+ years	19.855***
	(0.997)
Constant	56.098***
	(0.624)
Observations	7,043
R2	0.060
Adjusted R2	0.059
Residual Std. Error	29.184 (df = 7037)
F Statistic	89.808*** (df = 5; 7037)
Note:	*p<0.1; **p<0.05; ***p<0.01

Figure 4: Monthly Charges vs. Tenure



To determine if tenure contributes to monthly charges all else being equal, the other variables were first explored to determine which models should be added to the linear function. Each variable was explored using the tapply function to determine the variables that contributed to an increase in higher mean monthly charges. Table 5, below, shows the mean of monthly charges for each variable as well as for seniors vs. non-seniors to answer H3, following H2.

Table 5: The Mean Monthly Charges Related to the Variables

	Mean Monthly Charges		
	Sample	**Seniors**	**Non-seniors**
Dependents			
Dependents	59.52	82.29	58.5
No dependents	67.00	79.61	63.59
Paperless billing			
Paperless billing	73.55	83.16	70.99
No paperless billing	51.99	68.82	50.28
Partner			
Partner	67.77	64.73	82.78
No partner	61.94	76.84	59.19
Contract			

Month-to-month	66.44	78.88	63.11
1 year	65.04	62.32	83.46
2 year	60.77	58.94	80.29
Phone service			
Phone service	42.03	83.69	63.98
No phone service	67.02	42.18	42.18
Internet			
DSL	58.1	58.48	54.98
Fiber Optic	91.5	91.18	91.62
No internet	21.08	21.9	21.05
Online backup			
No online backup	71.94	78.41	70.33
No internet	21.08	21.9	21.05
Online backup	83.08	87.97	81.89
Tech support			
No tech support	74.59	80.06	72.87
No internet	21.08	21.9	21.05
Tech support	80.68	90.65	79.23
Streaming movies			
Not streaming	65.43	70.8	64.27
No internet	21.08	21.9	21.05
Streaming movies	88.48	92.38	87.39
Multiple lines			
No multiple lines	54.2	72.43	51.94
No phone service	42.03	41.18	42.18
Multiple lines	82.04	90.01	79.74
Online security			
No online security	75.69	81.01	74.09
No internet service	21.08	21.9	21.05
Online security	78.84	87.08	77.5
Device protection			
No protection	70.6	77.03	68.97
No internet	21.08	21.9	21.05
Device protection	84.82	90.02	83.59

Streaming TV			
Not streaming	65.39	71.15	64.08
No internet	21.08	21.05	21.9
Streaming TV	88.74	92.93	87.61
Payment method			
Bank transfer	67.19	81.13	64.71
Credit card	66.51	79.6	64.29
Electronic check	76.26	82.13	74.28
Mailed check	43.91	62.47	42.77

Variables which Table 5 indicated to have the most impact on monthly charges were added to the linear model and include: paperless billing, phone service, internet, streaming movies, and streaming TV. Each variable was first run in a linear model with monthly charges alone to ensure only significant variables were added to the aggregate linear function for all else being equal. All four variables were significant, as seen in Table 6, below. Since the results of streaming TV and streaming movies were highly similar, they were combined into one model to improve the comprehensibility of the output in Table 6 below.

Table 6: Linear regression of Monthly Charges vs. Tenure

	Dependent variable:			
	Monthly charges			
	(1)	(2)	(3)	(4)
1-2 year,	5.300***	5.100***	2.200***	2.500***
	(1.100)	(1.000)	(0.570)	(0.300)
2-3 years	9.500***	9.500***	3.300***	3.800***
	(1.200)	(1.100)	(0.610)	(0.320)
3-4 years	10.000***	10.000***	3.400***	4.900***
	(1.200)	(1.100)	(0.640)	(0.330)
4-5 years	14.000***	14.000***	4.500***	5.900***
	(1.200)	(1.100)	(0.620)	(0.330)
5+ years	20.000***	20.000***	7.600***	10.000***
	(1.000)	(0.930)	(0.540)	(0.280)
Paperless billing		22.000***		-0.660***
		(0.660)		(0.210)
Fiber Optic				30.000***
				(0.220)
No internet				-28.000***
				(0.280)

No internet repeat			-41.000***	
			(0.500)	
Streaming TV			15.000***	11.000***
			(0.460)	(0.240)
No internet repeat				
Streaming movies			15.000***	11.000***
			(0.460)	(0.240)
Constant	56.000***	43.000***	59.000***	45.000***
	(0.620)	(0.700)	(0.380)	(0.250)
Observations	7,043	7,043	7,043	7,043
R2	0.060	0.180	0.750	0.930
Adjusted R2	0.059	0.180	0.750	0.930
Residual Std. Error	29.000 (df = 7037)	27.000 (df = 7036)	15.000 (df = 7034)	7.800 (df = 7032)
F Statistic	90.000*** (df = 5; 7037)	263.000*** (df = 6; 7036)	2,672.000*** (df = 8; 7034)	9,697.000*** (df = 10; 7032)
Note:				*p<0.1; **p<0.05; ***p<0.01

H3: Seniors have lower total charges.

First, the monthly charges which are total charges divided by tenure was evaluated to see the impact of each variable, separate to tenure. The mean monthly charges of non-seniors is $61.84 with a standard deviation of 30.32 and $79.82 for seniors with a standard deviation of 20.316. The results are statistically significant with a p value of 0f <0.01. For seniors who were female, the mean was $81.11, and for seniors who were male, it was $78.54 showing little difference. Therefore, H3 could be incorrect and thus was explored further by determining all the variables that may additionally contribute to increased or decreased monthly charges for seniors and non-seniors (Table 6, above).

A linear regression function was created to determine the significance of the senior citizen variable on monthly charges and on total charges. The most significant variables that affect seniors, were included. First, the variables were run independently against monthly charges to test their validity and observe the impact. For comprehensibility, the linear models have been condensed into four in both Table 7 and 8 below.

Table 7: Correlation Between Seniors and Monthly Charges, All Else Equal

	Dependent variable:			
	Monthly Charges			
	(1)	(2)	(3)	(4)
Senior	18.000***	-1.200***	-1.300***	-0.070
	(0.950)	(0.430)	(0.370)	(0.280)
1-2 year			4.700***	2.800***
			(0.420)	(0.320)
2-3 years			7.200***	4.200***
			(0.450)	(0.350)
3-4 years			9.400***	5.000***
			(0.460)	(0.380)
4-5 years			12.000***	6.400***
			(0.450)	(0.380)
5+ years			18.000***	9.200***
			(0.380)	(0.390)
1-year contract				3.100***
				(0.300)
2-year contract				4.600***
				(0.350)
paperless billing				1.500***
				(0.220)
Fiber optic		34.000***	34.000***	28.000***
		(0.350)	(0.300)	(0.260)
no internet		-37.000***	-37.000***	-39.000***
		(0.420)	(0.360)	(0.320)
Phone service				22.000***
				(0.370)
NA				
Device protection				10.000***
				(0.240)
Constant	62.000***	58.000***	51.000***	31.000***
	(0.380)	(0.260)	(0.300)	(0.380)
Observations	7,043	7,043	7,043	7,043
R2	0.048	0.820	0.870	0.930
Adjusted R2	0.048	0.820	0.870	0.930
Residual Std. Error	29.000 (df = 7041)	13.000 (df = 7039)	11.000 (df = 7034)	8.200 (df = 7029)
F Statistic	359.000*** (df = 1; 7041)	10,753.000*** (df = 3; 7039)	5,718.000*** (df = 8; 7034)	6,762.000*** (df = 13; 7029)
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 8: Correlation Between Seniors and Total Charges, All Else Equal

	Dependent variable:			
	TotalCharges			
	(1)	(2)	(3)	(4)
Senior	633.000***	-41.000	-80.000***	-25.000
	(73.000)	(69.000)	(31.000)	(29.000)
1-2 year			833.000***	708.000***
			(35.000)	(33.000)
2-3 years			1,640.000***	1,432.000***
			(38.000)	(36.000)
3-4 years			2,523.000***	2,233.000***
			(39.000)	(39.000)
4-5 years			3,476.000***	3,152.000***
			(38.000)	(40.000)
5+ years			4,830.000***	4,374.000***
			(32.000)	(41.000)

1-year contract				290.000***
				(31.000)
2-year contract				142.000***
				(37.000)
paperless billing				46.000**
				(22.000)
Fiber optic		1,096.000***	1,095.000***	903.000***
		(57.000)	(26.000)	(27.000)
no internet		-1,456.000***	-1,292.000***	-1,227.000***
		(67.000)	(30.000)	(33.000)
Phone service				745.000***
				(38.000)
NA				
Device protection				676.000***
				(25.000)
Constant	2,177.000***	2,120.000***	128.000***	-634.000***
	(29.000)	(42.000)	(25.000)	(39.000)
Observations	7,043	7,043	7,043	7,043
R2	0.011	0.180	0.830	0.860
Adjusted R2	0.010	0.180	0.830	0.860
Residual Std. Error	2,255.000 (df = 7041)	2,046.000 (df = 7039)	922.000 (df = 7034)	849.000 (df = 7029)
F Statistic	76.000*** (df = 1; 7041)	534.000*** (df = 3; 7039)	4,441.000*** (df = 8; 7034)	3,319.000*** (df = 13; 7029)
Note:				*p<0.1; **p<0.05; ***p<0.01

H4: Seniors are less likely to churn than non-seniors.

Using the tapply function, the variables were initially assessed for their association to churn. The results are below in Table 9.

Table 9: Which Customer Attributes are Most Associated with Churn?

	Percentage of Churn/Not Churn		**Percentage of Variable**	
	Churn	**Not Churn**	**Churn**	**Not Churn**
Contract				
Month-to-month	88.60%	42.90%	42.70%	57.30%
1 Year	8.90%	25.30%	1.30%	98.70%
2 Year	3.60%	31.80%	2.80%	97.20%
Partner				
Partner	36%	53%	20%	80%
No partner	64%	47%	33%	67%
Phone Service				
Phone service	90.90%	9.10%	73%	27%
No phone service	9.10%	90.90%	72%	25%
Online Backup				
No online backup	66%	36%	39.90%	60.10%
Online backup	28%	37%	25.50%	78.50%
No internet service	6%	27%	7.40%	92.60%
Online security				
No online security	78%	39%	41%	58.20%

No internet service	6%	27%	7.40%	92.60%
Online security	16%	33%	14.60%	85.40%
Tech support				
No tech support	77%	39%	41.60%	58.40%
No internet service	6%	27%	7.40%	92.50%
Tech support	17%	34%	15.20%	82.80%
Gender				
Female	50%	49%	27%	73%
Male	50%	51%	26%	74%
Dependents				
Dependents	17%	34%	15%	85%
No dependents	83%	66%	31%	69%
Multiple lines				
No multiple lines	45%	49.10%	25%	75%
No phone service	9.10%	9.90%	25%	75%
Multiple lines	45.50%	41.00%	29%	71%
Streaming movies				
Not streaming movies	50%	46%	33.70%	66.30%
No internet service	6%	27%	7.50%	92.50%
Streaming movies	44%	35%	29.90%	70.10%
Streaming TV				
Not streaming TV	50%	46%	33.50%	66.50%
No internet service	6%	27%	7.40%	92.60%
Streaming TV	44%	35%	30.10%	69.90%
Paperless billing				
Paperless billing	75%	54%	34%	66%
No paperless billing	25%	46%	16%	84%
Senior citizen				
Senior citizen	25%	13%	42%	58%
non-senior citizen	75%	87%	24%	76%
Internet service				
DSL	35%	38%	19%	81%
Fiber optic	69%	35%	41.90%	58.10%
No internet service	6%	27%	7.40%	92.60%
Device protection				

No device protection	65%	36%	39.10%	60.90%
No internet service	6%	27%	7.40%	92.60%
Device protection	29%	36%	22.50%	77.50%
Payment method				
Automatic bank transfers	14%	25%	17%	83%
Credit cards	12%	25%	15%	85%
Electronic checks	57%	25%	45%	55%
Mailed checks	16%	25%	19%	81%

Initially, H4 appears to be incorrect. To evaluate further, a regression model was created with other variables that have a strong association with churn to determine if the seniors variable has a significant impact. This functions by using the binary variable where there is a threshold between the two values. Because of the threshold, a logistic, not linear, regression must be used.

Online backup, online security, tech support, and device protection were represented by tech support as they showed highly similar results in the initial analysis, suggesting they are part of a bundle. The first model showed that the variable, dependent, was not significant so the model was rerun without it.

Table 10: Logistic Model for Churn

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.5051  -0.7347  -0.3095   0.8819   3.0235

Coefficients: (1 not defined because of singularities)
                                          Estimate Std. Error z value Pr(>|z|)
(Intercept)                               -0.84371    0.09969  -8.464  < 2e-16 ***
factor(TechSupport)No internet service    -0.90560    0.12244  -7.396 1.40e-13 ***
factor(TechSupport)Yes                    -0.44703    0.08112  -5.510 3.58e-08 ***
factor(SeniorCitizen)1                     0.19244    0.07882   2.442   0.0146 *
factor(Contract)One year                  -1.34666    0.09430 -14.280  < 2e-16 ***
factor(Contract)Two year                  -2.43335    0.15810 -15.392  < 2e-16 ***
factor(Partner)Yes                        -0.31181    0.06581  -4.738 2.16e-06 ***
factor(InternetService)Fiber optic         0.81766    0.07249  11.279  < 2e-16 ***
factor(InternetService)No                       NA         NA      NA       NA
factor(PaymentMethod)Credit card (automatic) -0.06585 0.11043  -0.596   0.5510
factor(PaymentMethod)Electronic check      0.57753    0.09012   6.408 1.47e-10 ***
factor(PaymentMethod)Mailed check          0.22935    0.10877   2.108   0.0350 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 8150.1  on 7042  degrees of freedom
Residual deviance: 6202.4  on 7032  degrees of freedom
AIC: 6224.4
```

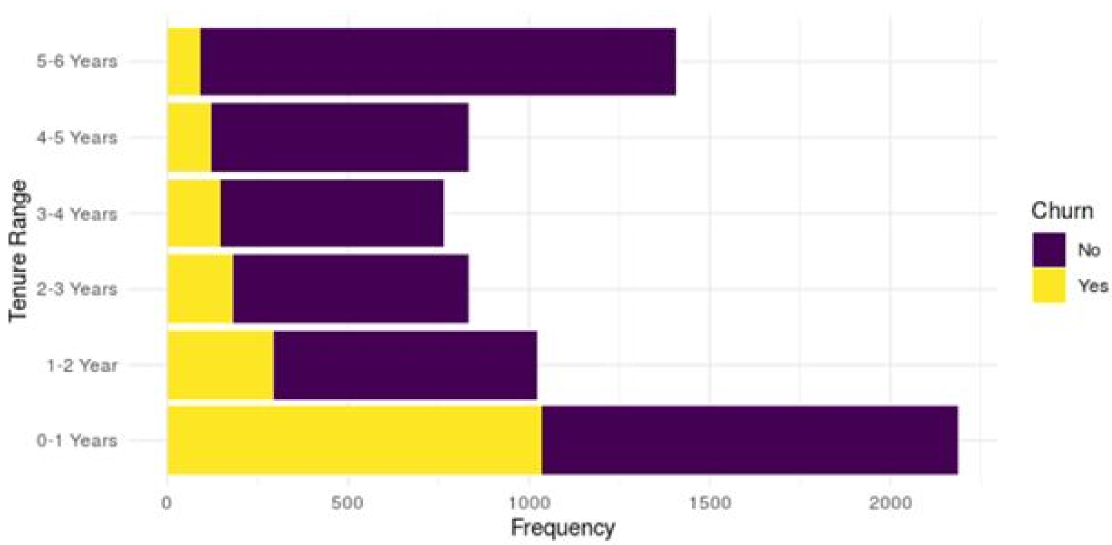
Discussion

The discussion follows the same hypothesis headings as the results section.

Hypothesis 1

A clear pattern was seen in the relationship between tenure and churn. Figure 5, below shows a bar graph that indicates that customers whose tenure is below a year have the highest churn rates.

Figure 5: Frequency Distribution of Tenure.



A linear function found that the mean tenure of customers who did not churn (17.98) was over double the tenure of customers who churned (37.57). Table 3: Churn by Tenure, shows that 55.5% of customers who churned had a tenure of under 1 year. At the same time, when looking at it from the perspective of the year, out of customers who were tenures 0-1 year, 47.4% churned. Therefore H1 is true, *Customers with longer tenure (> 1 year) are less likely to churn than newer customers (<1 year).*

Hypothesis 2

When dividing tenure into categorical values divided by year, a linear regression model showed that the mean monthly charges for telcom customers were $56 for customers with a tenure of less than 1 year, $59.30 for customers tenured 1-2 years, $65 for customers tenured 2-3 years, $66 for customers tenured 3-4 years, $70 for customers tenured 4-5 years, and $76 for customers tenured 5 or more years. All of the means were statistically significant with a p value of <0.01. Therefore, H2 was initially suggested to be correct. By dividing the tenure into categories, and running a linear regression with monthly charges, it was clear that the mean monthly charges increased for the category of 12 months.

However, this does not show the impact of tenure, all else being equal. The other variables were explored to create a linear model with all of the variables which had a significant influence on mean monthly charges. The results of the aggregate linear regression (Table 6 above), show that all else being equal, the tenure still impacts the monthly charges and is statistically significant. The monthly charges are $45 for customers who are tenured for or less than one year. This increases by $2.50 when they are tenured for 1-2 years, by $3.80 when they are tenured for 2-3 years, by $4.9 when they are tenured 3-4 years, by $5.9 when they are tenured 4-5 years, and by $10 when they are tenured over 5 years. Therefore, we can say with certainty that H2 is correct: all else being equal, *customers with longer tenure (>1 year) spend more than newer customers (< 1 year).*

Hypothesis 3

In the initial observations, the mean monthly charge of senior citizens ($80) is higher than the mean spend for non-senior citizens ($60) - thereby contradicting H3. When gender was added there was little change, though female seniors had slightly higher monthly charges ($81 vs.$79). The difference between seniors and non seniors was further explored to see if there are certain factors that seniors have in common which contribute to their higher mean monthly spend. Table 11, below, shows the observation of each variable.

Table 11: Monthly Charge of Seniors vs. Non-seniors

Variable	**Observations**
Dependents	Overall and amongst those who are non-seniors, having no dependents resulted in higher mean monthly spend. The reverse was seen amongst seniors.
Paperless billing	Customers who use paperless billing had higher mean monthly charges, especially among non-senior customers.
Partner	There were higher mean monthly charges both overall and for non-seniors. Among seniors, the reverse was true.
Contract	Overall, and amongst seniors, the highest mean monthly charges were amongst those who had month-to-month contracts. Amongst non-seniors, customers with 1-year contracts had the highest mean monthly charges.
Phone service	Overall and amongst non-seniors, there was a higher mean of monthly charges for those with phone service. The reverse was true for seniors.
Internet service	Overall and amongst both seniors and non-seniors, the mean monthly charges of those who had fiber optic internet was highest, followed by DSL.
Online backup	Overall, and amongst seniors and non-seniors when looked at individually, those who had online backup had the highest mean monthly charges.
Tech support	Overall, and amongst seniors and non-seniors when looked at individually, those who had tech support had the highest mean monthly charges. However, those with no tech support still had significantly higher mean monthly charges than those with no internet.
Streaming movies	Overall, and amongst seniors and non-seniors when looked at individually, those who were streaming movies had the highest mean monthly charges. Those who were not streaming movies still had significantly higher mean monthly charges than those with no internet. The difference is mostly seen among seniors.
Multiples lines	Overall, and amongst seniors and non-seniors when looked at individually, those who had multiple lines had the highest mean monthly charges.
Online security	Overall, and amongst seniors and non-seniors when looked at individually, those who had online security had the highest mean monthly charges.
Device protection	Overall, and amongst seniors and non-seniors when looked at individually, those who had device protection had the highest mean monthly charges.
Streaming TV	Overall, and amongst seniors and non-seniors when looked at individually, those who were streaming TV had the highest mean monthly charges. This was especially seen amongst seniors.

Payment method	The mean monthly charges were lowest among mailed checks and highest among electronic checks. Bank transfers and credit cards were more or less equal.

Both seniors and non-seniors had a higher mean monthly spend when they had online backup, tech support, streaming movies and TV, multiple lines, online security, device protection, and who used electronic checks, paperless billing, and internet service- especially when it was fiber optic. Seniors, uniquely, had higher mean monthly charges when they had a partner, month-to-month contract, and no phone service. Non-seniors uniquely had higher mean monthly charges when they had a partner, a 1-year contract, and phone service. Linear models were then run both on seniors vs. monthly charges, with additional significant variables, and seniors vs. total charges, with additional significant variables.

The results in Table 8 show that the monthly charges of seniors are, on average, $18 higher than non seniors and it is statistically significant with a p value of less than 0.01. However, when the fiber optic internet is taken into account, seniors actually spend less. Both values are statistically significant, but the fiber optic internet has a larger impact on monthly charges. This suggests that one of the factors making seniors higher spenders by month is their propensity to purchase fiber optic internet.

Once tenure is taken into account, seniors do spend more, all else being equal for customers tenured more than 1 year. Customers under 1 year of tenure spend more if they are non-seniors, which is consistent with the spread of charges by customers tenures less than one year and the heavy density of non-seniors being in that group. However, after one year, the charges do increase per tenure for seniors who spend more than non-seniors: $2.8 more non 1-2 year tenure, $4.2 more for those tenured 2-3 years, $5.0 more for those tenured 23-4 years, 6.4 more for those tenured 4-5 years, and $9.2 more for those tenured over 5 years.

When looking at total charges (Table 9, above) there are similar results. Alone, senior citizens spend more than non-senior citizens. However, when fiber optic internet is taken into account, they spend less. This reverses again when tenure is considered and when tenure is above one year, then seniors spend more than non-seniors. The results are statistically significant with a p value of less than 0.01. Therefore, H3 is correct, Seniors spend more than non-seniors- when the tenure of the customers is more than 12 months.

Hypothesis 4

The factors that related to the highest percentages of churn in their variable were individuals with a month-to-month contract, no partner, no online backup, no online security, no tech support, those with dependents, paperless billing, fiber optic internet service, no device protection, electronic checks, higher month-to-month charges, low tenure, and who are seniors. The results were very similar for tech support, online security, device protection, suggesting these may be part of a bundle.

Narrowing in on seniors, of the people who churned, 25% were seniors, and 75% were non-seniors. Of the people who did not churn, 13% were seniors, and 87% were non-seniors. Of the people who were seniors, 42% churned, and 58% did not churn. Of the people who were not seniors, 24% churned and 76% did not churn.

The numbers for tech support were very similar to online security, device protection, and online backup suggesting many people either are similar shoppers or more likely, these features are available as part of a bundled add-on. There was a stronger correlation to not churning when customers had these extra features. and less likely to churn.

The results of the logistic regression model in Table 10, above, shows that Senior citizens have a positive correlation to churn, which rejects H4. However, the significance is minimal with a p value of 0.0146. This is still less than 0.05, and thus acceptable, and thus H4 is rejected: seniors are not less likely to churn than Non-seniors. Perhaps this is because seniors have more time to do things like evaluating telecom options continually. Stronger statistical significance was seen for Fiber Optic internet service and Electronic check which are contributors to customer churn.

Concluding Remarks

In this paper, I reviewed existing research on churn rates which can be categorized into two distinct research disciplines improving model predictive performance and understanding the drivers of churn to determine actionable strategies. The most popular algorithms used in churn prediction are logistic regression, decision trees, SVM, and random forest as a result of their strong predictive performance and comprehensibility. However, background research shows that the researcher must make data-driven decisions throughout the research process to ensure their models perform well and can effectively communicate the needed information. The background research also shows that most work focuses on customer churn as a binary problem (churn/not churn). I find this to be a weakness as it does not support sophisticated or strategic action. I will systematically evaluate the impact of attrition while taking into account the value of the different types of customers.

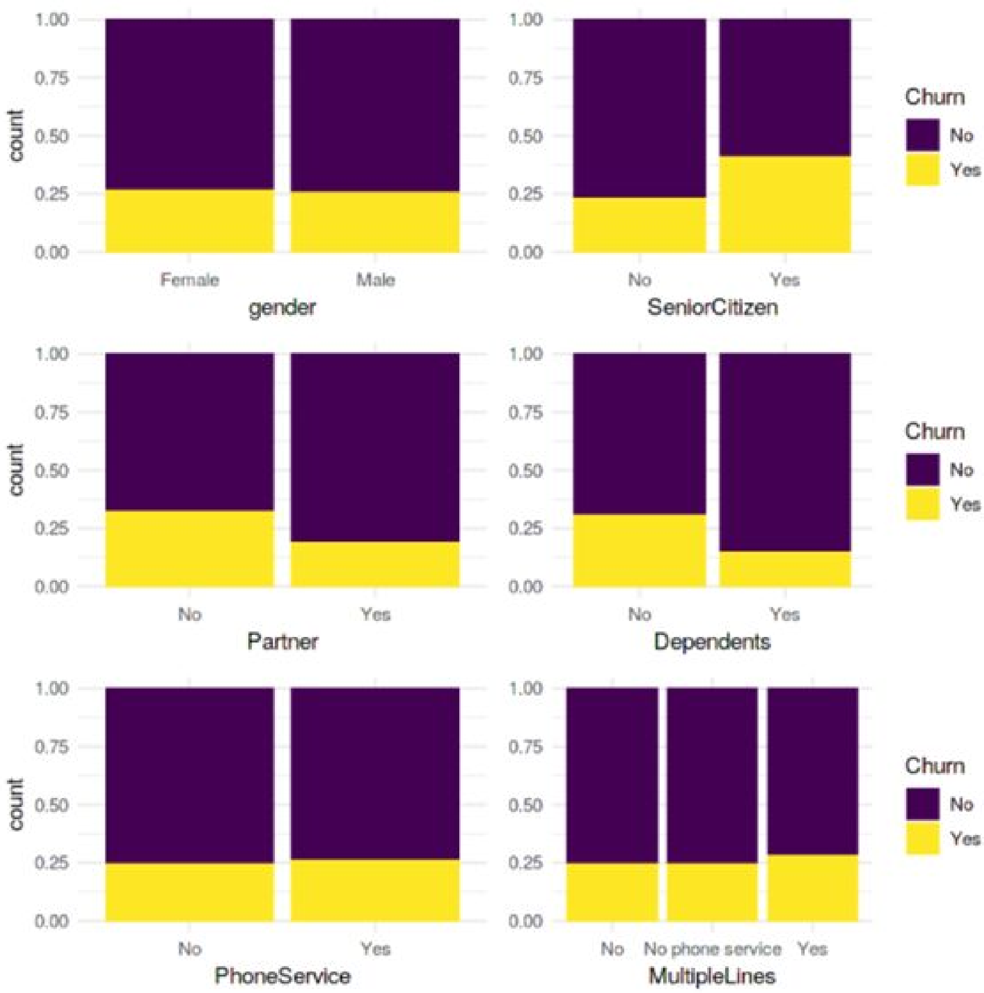
This study of churn found that most hypotheses were correct. Customers with longer tenure are less likely to churn, and spend more than customers with shorter tenure. While the hypothesis differentiated short and long tenure by the threshold of 1 year for testing, the increase in retention and in consumer spend is seen to increase as tenure increases. Seniors were also seen to be less likely to churn. However, H3 was incorrect, seniors did not have lower total charges, overall. When the tenure was less than one year, non-seniors had higher charges, but after one year, seniors had higher total and monthly charges.

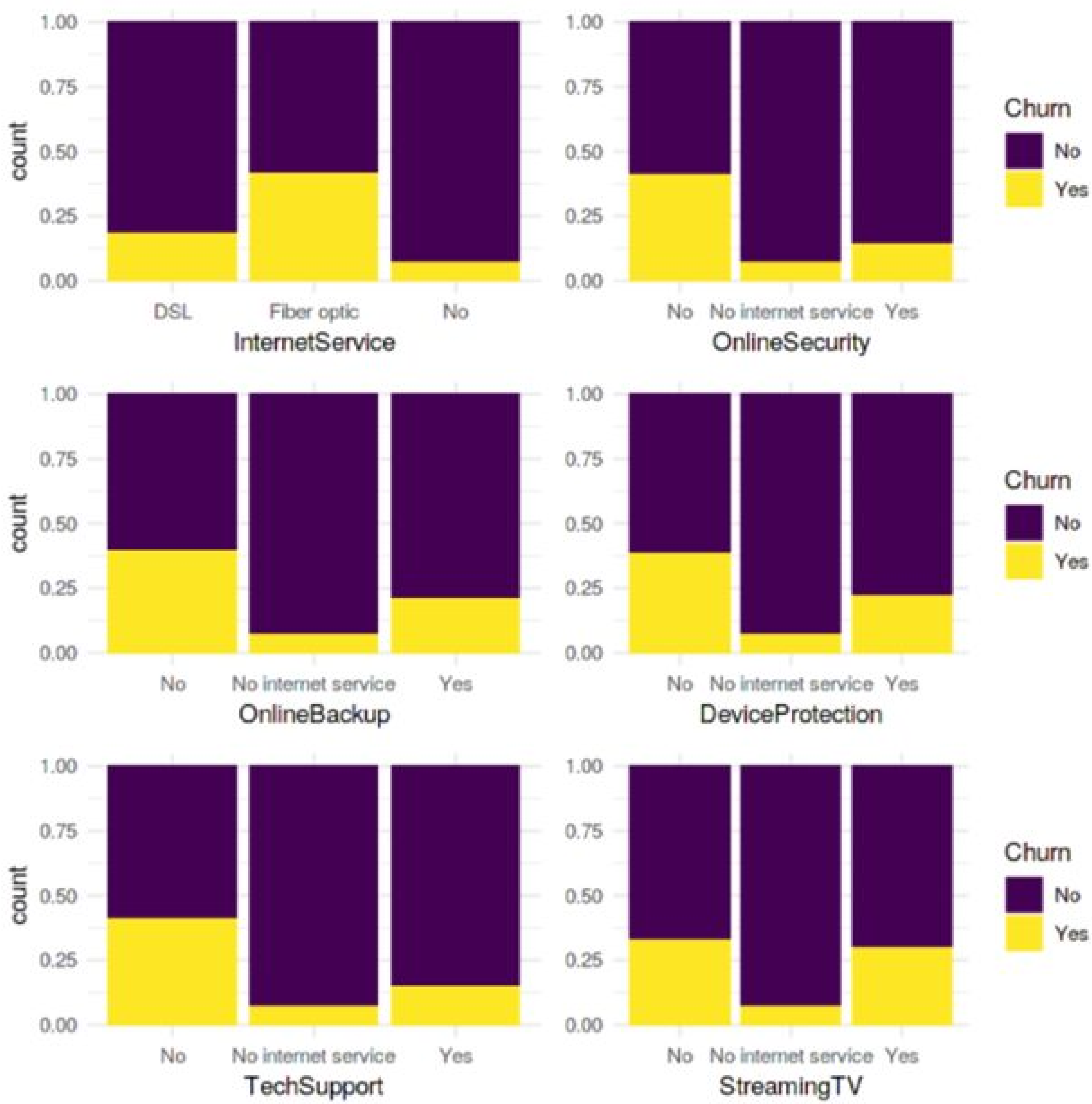
Seniors have proven to be the high-value customer. However, their attrition rates are also low. Therefore, it does not make sense to spend attrition campaigns on this demographic. One option is attrition campaigns aimed at seniors, who also share some of the attributes of customers likely to churn. These variables include the use of electronic checks, tenure of less than one year, on month-to-month contracts, who had no online backup, online security, tech support, and paperless billing. These variables should be explored further with predictive modeling to determine which variables of a senior data subset are most likely to contribute to churn.

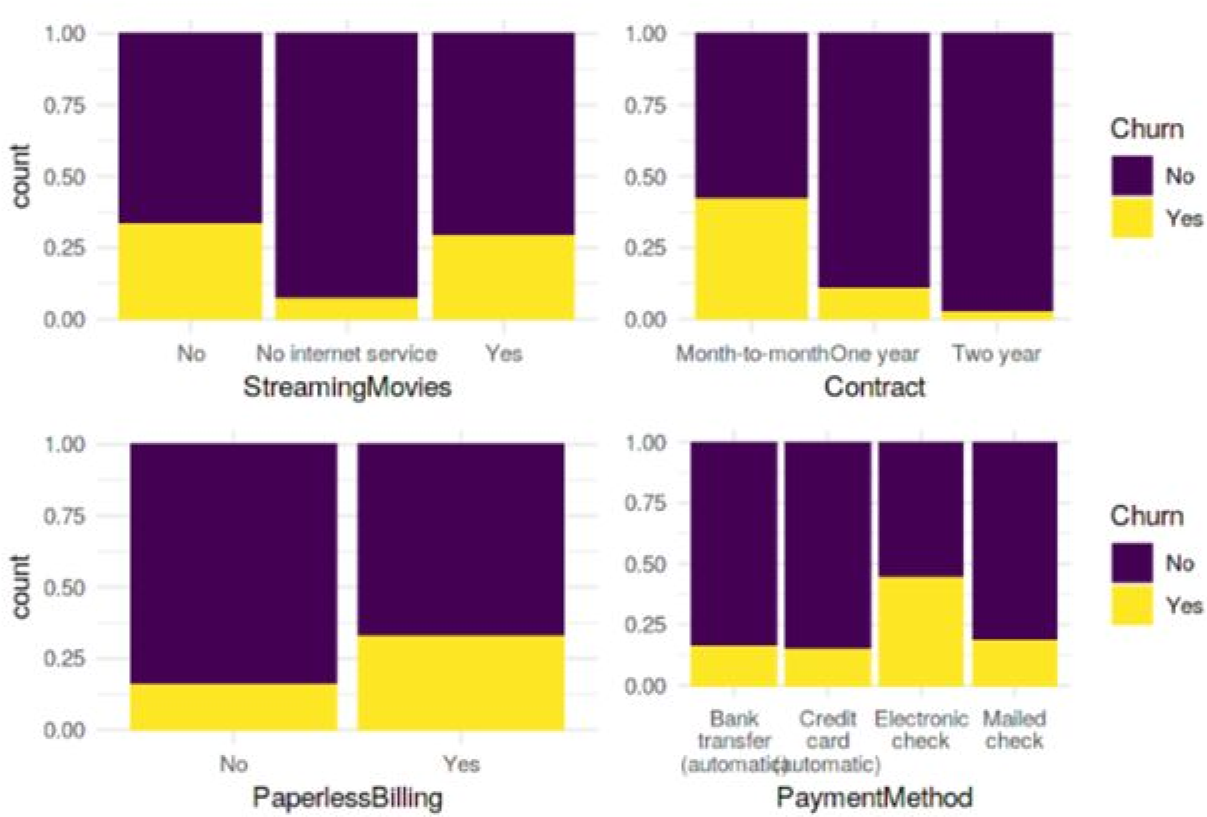
Other variables that statistically significantly contributed to monthly spend but were not related to strong retention were the subscription to fiber optic internet, and TV and Movie. Further research could use predictive modeling to determine the value of attrition campaigns for customers with these variables.

Appendix

Churn by Variable




1.00
0.75
0.50
0.25
0.00
count
DSL
Fiber optic
No
InternetService
No
No internet service
Yes
OnlineSecurity
Churn
No
Yes
OnlineBackup
DeviceProtection
TechSupport
StreamingTV


1.00
0.75
0.50
0.25
0.00
count
No
No internet service
Yes
StreamingMovies
Month-to-month
One year
Two year
Contract
Churn
No
Yes
PaperlessBilling
Bank transfer (automatic)
Credit card (automatic)
Electronic check
Mailed check
PaymentMethod

References

Ascarza, E., Neslin, S. A., Netzer, O., Anderson, Z., Fader, P. S., Gupta, S., ... & Provost, F. (2018). In pursuit of enhanced customer retention management: Review, key issues, and future directions. *Customer Needs and Solutions*, *5*(1-2), 65-81.

Baddeley, A. D. (1996). The concept of working memory. Models of short-term memory, 1-27.

Benedek G, Lublóy Á, Vastag Y (2014) The importance of social embeddedness: churn models at mobile providers. Decis Sci 45(1):175–201Return to ref 10 in article

Caigny, A., Coussement, K., & De Bock, K. W. (2018). A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees. European Journal of Operational Research, 269(2), 760-772.

Chen, I. J., & Popovich, K. (2003). Understanding customer relationship management (CRM). Business process management journal.

Colgate, M., Stewart, K., & Kinsella, R. (1996). Customer defection: a study of the student market in Ireland. International Journal of Bank Marketing.

Coussement, K., Lessmann, S., & Verstraeten, G. (2017). A comparative analysis of data preparation algorithms for customer churn prediction: A case study in the telcommunication industry. Decision Support Systems, 95, 27-36.

Coussement, K., & Van den Poel, D. (2008). Churn prediction in subscription services: An application of support vector machines while comparing two parameter-selection techniques. Expert systems with applications, 34(1), 313-327.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and brain sciences, 24(1), 87-114.

Cowan, N. (2016). Working memory capacity: Classic edition. Psychology press.

Davenport, T. H., Harris, J. G., & Kohli, A. K. (2001). How do they know their customers so well?. MIT Sloan Management Review, 42(2), 63.

Dawes, J., & Swailes, S. (1999). Retention sans frontieres: issues for financial service retailers. International Journal of Bank Marketing.

Day, G. S. (2000). Capabilities for forging customer relationships (No. 118). Cambridge, MA: Marketing Science Institute.

Ganesh, J., Arnold, M. J., & Reynolds, K. E. (2000). Understanding the customer base of service providers: an examination of the differences between switchers and stayers. Journal of marketing, 64(3), 65-87.

Gebert, H., Geib, M., Kolbe, L., & Brenner, W. (2003). Knowledge‐enabled customer relationship management: integrating customer relationship management and knowledge management concepts [1]. Journal of knowledge management.

Hansen, H., Samuelsen, B. M., & Sallis, J. E. (2013). The moderating effects of need for cognition on drivers of customer loyalty. European Journal of Marketing.

Hervé, A., Schmitt, C., & Baldegger, R. (2020). Digitalization, Entrepreneurial Orientation and Internationalization of Micro-, Small-and Medium-Sized Enterprises. Technology Innovation Management Review, 10(4).

Jafari-Marandi, R., Denton, J., Idris, A., Smith, B. K., & Keramati, A. (2020). Optimum profit-driven churn decision making: innovative artificial neural networks in telecom industry. Neural Computing and Applications, 32(18), 14929-14962.

Keramati, A., Jafari-Marandi, R., Aliannejadi, M., Ahmadian, I., Mozaffari, M., & Abbasi, U. (2014). Improved churn prediction in telecommunication industry using data mining techniques. Applied Soft Computing, 24, 994-1012.

Keramati, A., & Ardabili, S. M. (2011). Churn analysis for an Iranian mobile operator. Telecommunications Policy, 35(4), 344-356.

Legner, C., Eymann, T., Hess, T., Matt, C., Böhmann, T., Drews, P., ... & Ahlemann, F. (2017). Digitalization: opportunity and challenge for the business and information systems engineering community. Business & information systems engineering, 59(4), 301-308.

Lessmann, S., & Voß, S. (2009). A reference model for customer-centric data mining with support vector machines. European Journal of Operational Research, 199(2), 520-530.

Li, A., & Pearl, J. (2019, July). Unit Selection Based on Counterfactual Logic. In IJCAI (pp. 1793-1799).

Lima, E., Mues, C., & Baesens, B. (2011). Monitoring and backtesting churn models. Expert Systems with Applications, 38(1), 975-982.

Martens, D., Vanthienen, J., Verbeke, W., & Baesens, B. (2011). Performance of classification models from a user perspective. Decision Support Systems, 51(4), 782-793.

Manyika, J., Lund, S., Bughin, J., Woetzel, J. R., Stamenov, K., & Dhingra, D. (2016). Digital globalization: The new era of global flows (Vol. 4). San Francisco, CA: McKinsey Global Institute.

Merali, Y., Papadopoulos, T., & Nadkarni, T. (2012). Information systems strategy: Past, present, future?. The Journal of Strategic Information Systems, 21(2), 125-153.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological review, 63(2), 81.

Miller, G. A. (1989). George A. Miller. In G. Lindzey (Ed.), A history of psychology in autobiography Vol. 8 (p. 390–418). Stanford University Press. https://doi.org/10.1037/11347-011

Nambisan, S. (2017). Digital entrepreneurship: Toward a digital technology perspective of entrepreneurship. Entrepreneurship Theory and Practice, 41(6), 1029-1055.

Neslin, S. A., Gupta, S., Kamakura, W., Lu, J., & Mason, C. H. (2006). Defection detection: Measuring and understanding the predictive accuracy of customer churn models. Journal of marketing research, 43(2), 204-211.

Nitzan, I., & Libai, B. (2011). Social effects on customer retention. Journal of Marketing, 75(6), 24-38.

Peppers, D., and Rogers, M. (2004). Managing customer relationships: a strategic framework. John Wiley & Sons, Hoboken

Tilson, D., Lyytinen, K., & Sørensen, C. (2010). Research commentary—Digital infrastructures: The missing IS research agenda. Information systems research, 21(4), 748-759.

Torkzadeh, G., Chang, J. C. J., & Hansen, G. W. (2006). Identifying issues in customer relationship management at Merck-Medco. Decision Support Systems, 42(2), 1116-1130.

Blattberg, R. C., Kim, B. D., & Neslin, S. A. (2008). Why database marketing?. In Database marketing (pp. 13-46). Springer, New York, NY.

Reinartz, W. J., & Kumar, V. (2003). The impact of customer relationship characteristics on profitable lifetime duration. Journal of marketing, 67(1), 77-99.

Sundararajan, A., & Gursoy, K. (2020). Telecom customer churn prediction.

Seret, A., Verbraken, T., Versailles, S., & Baesens, B. (2012). A new SOM-based method for profile generation: Theory and an application in direct marketing. European Journal of Operational Research, 220(1), 199-209.

Shaffer, G., & Zhang, Z. J. (2002). Competitive one-to-one promotions. Management Science, 48(9), 1143-1160.

Van den Poel, D. and Larivière, B. (2004). Customer attrition analysis for financial services using proportional hazard models. European Journal of Operational Research, 157(1),pp.196-217.

Verbeke, W., Dejaeger, K., Martens, D., Hur, J., & Baesens, B. (2012). New insights into churn prediction in the telecommunication sector: A profit driven data mining approach. European journal of operational research, 218(1), 211-229.

Verbeke W, Martens D, Baesens B (2014) Social network analysis for customer churn prediction. Appl Soft Comput 14:431–446

Zaheer, M., Ahmed, A., & Smola, A. J. (2017, July). Latent LSTM allocation: Joint clustering and non-linear dynamic modeling of sequence data. In International Conference on Machine Learning (pp. 3967-3976).