Churn and Customer Lifetime Value Modeling for an Online Publisher[1]

Ross Hochwert

Introduction:

MagazineCo, LLC creates women's lifestyle content in the cooking and crafting niches across multiple platforms including digital websites and e-mail newsletters. With over 8 million subscribers to their many opt-in branded newsletters, MagazineCo sends out millions of emails every day with advertisements in their newsletters serving as one of their main revenue drivers. Yet, with their daily deluge of data concerning click-through-rates, open rates, and unsubscribe rates, MagazineCo had not yet developed a customer lifetime value model to estimate each subscriber's worth. My research sought to model the churn and revenue patterns of customers and utilize these models to develop a more sophisticated customer lifetime value to better direct marketing dollars for customer acquisition.

Client Background:

As a small business, MagazineCo's main concern regarding customer acquisition had been to recoup acquisition costs as quick as possible. Assuming that each subscriber would stay for an average of 6 months and given a fixed average revenue-per-thousand (RPM) for emails to these subscribers, MagazineCo was willing to spend up to $2.00 to acquire a subscriber through paid platforms such as Google AdWords and Facebook Ads. They knew that the actual customer lifetime value (CLV) was higher than this back-of-the-envelope calculation, but in the absence of more sophisticated modeling techniques and with constrained capital, they focused on

---

[1]

maintaining cash flows initially over precise estimates for the long-term. As the business has expanded and grown over the last 6 years, they have seen more flexibility with capital and would like to know how and where to effectively spend their marketing dollars for acquiring customers.

Additionally, MagazineCo was interested in analyzing CLVs among customers acquired from different sources. MagazineCo acquires subscribers mainly through three channels: paid search, organic search, and referral. Paid search refers to bidding on different keywords and advertising under those keywords on search engines, while organic subscriptions occur when one of MagazineCo's websites organically appears under search results. Referral traffic comes mainly from social media sites like Facebook and Pinterest and from a network of bloggers affiliated with MagazineCo. By segmenting the data into these three different cohorts, I could analyze any differing behaviors and more accurately predict CLV. Finally, a model that accurately forecasted how many emails new and existing customers would receive over their life could help MagazineCo estimate the total value of its subscribers if they were looking to be acquired. While the industry typically utilizes multipliers for acquisitions, the expected future revenues of customers could represent a powerful tool for increasing the value of MagazineCo.

Data:

MagazineCo oversees 27 different online brands within the cooking and crafting niches, each with its own website and newsletter. I choose to focus on one of the more popular brands, AllFreeCrochet.com, which specializes in free crochet patterns. I received a data file with every subscriber, the source of their subscription (referral, organic or paid) and the date they unsubscribed (or NA if still alive). I then set up cohorts for all subscribers who subscribed in

January 2014 and July 2014, and modeled their churn rates through March 2016 (See Figure 1). This graph represented a wealth of knowledge for the client as, heretofore, they hadn't had access to this type of report or chart. Just by eyeballing the data, one could see that about 80% of subscribers make it to 6 months and about 50% to one year. I additionally created cohorts within the January 2014 data based on their source of acquisition: referral, organic or paid. Both January 2014 and July 2014 cohorts and paid and organic cohorts showed similar churn patterns, showing a lack of seasonal differences. That being said, the shape of the curves is perplexing. Between month 0 and month 1, there is a steep drop-off, with another steep drop-off occurring at month 6. Between these drop-offs, the curve is smooth and shallow, rather than steep or bumpy. My task then became to model these differences and provide managerial conclusions for the shape of the graph.

Analysis:

Newsletter subscribers operate in a contractual discrete fashion, meaning we know when the subscriber is no longer a subscriber (they hit unsubscribe) and we assume they receive daily emails and therefore daily opportunities to unsubscribe in a discrete fashion (every time they receive an email). Of course, they can unsubscribe anytime they like, but to simplify the model, I assumed discrete churn behavior and aggregated churns at the monthly level. Given the nature of subscribers, I initially began modeling the data with a shifted Beta Geometric curve across cohorts. All four cohorts (January 2014, July 2014, January 2014 paid-only and January 2014 organic-only) yielded similar alpha and beta coefficients, around 1 and 24 respectively. The predicted curve fit actual churn well, but did not account for spikes in month 1 and month 6.

Given the similar estimated coefficients and the strikingly similar actual churn patterns between months and sources, I directed my focus toward modeling one cohort, the January 2014 cohort, rather than explain differences between cohorts. Within the scope of this research, I assumed little to no differences in subscriber behavior and treated all subscribers regardless of time or source of acquisition the same.

Upon discussion with the client, the spike at month 6 was driven by MagazineCo themselves: they have a strict policy to remove any subscriber from their list that has not opened an email from MagazineCo within 6 months in order to keep the lists of the highest quality. Given this external driver, I deemed it appropriate to insert a spike into my model at month 6 to improve the fit. The next step after a spike was to shift into Beta discrete Weibull (BdW) models to view time effects of churn propensity. Specifically, I wanted to answer the question of whether or not customers experience time dependence effects wherein their churn propensity decreases or increases over time. The initial model yielded alpha and beta values well above 6,000, indicating spiky behavior and no need for a beta distribution to simulate heterogeneity, therefore I re-ran the model as a discrete Weibull.

Finally, I integrated two latent classes to the discrete Weibull, each with a varying churn rate (theta) and time dependency (c). The latent class model fit the data incredibly well, revealing two types of subscribers: one class of subscribers who either immediately unsubscribed or exhibited low churn rates after the first month, and then another class of subscribers who followed a more typical churn rate and time dependence. Put another way, a subscriber will either unsubscribe immediately, stay very loyal for life, or unsubscribe at an average decreasing rate over time (See Figure 2).

While this model fits the two years of data the best of all models trialed, it predicts almost no churn after 36 months on the list, whereas all other models predict a reasonable amount of churn into the future. To test out-of-sample-fit, I generated a churn graph from the actual January 2012 cohort and extended the January 2014 parameters to 52 periods, in order to model 4 years of churn, which I then overlaid on the actual 2012 data. (See Figure 3). Impressively, both the simple discrete Weibull with spike and the discrete Weibull latent class with spike modeled the January 2012 data well for 2 years, even though they were based on January 2014 data, implying similarity in customers across time. However, at the two-year point, the discrete Weibull latent class failed to predict churn, while the simple discrete Weibull followed churn very closely into the future. Ultimately, the latent class discrete Weibull yielded the lowest BIC, yet the out-of-sample fit and complicated story lead me to choose the simple discrete Weibull as the best model for MagazineCo. (See Table 1 for comparison of models). The out-of-sample fit, parsimony and understandable business story resonate well with the discrete Weibull and therefore my model of choice for CLV.

Discussion:

The final model points to important behavioral differences within the January 2014 cohort and their churn propensities. In any given cohort, around 8% of subscribers will immediately unsubscribe within the first month. Managerially, these subscribers probably did not expect this type of content, did not know what they were signing up for in the first place, or were generally unsatisfied with the newsletter and wanted to be removed as soon as possible. Over the next five months, there will be a steady churn rate of around 3%, until month 6 where an additional 10% of subscribers will then automatically be removed from the list due to inactivity. These inactive users reduce the quality of the email list as they will not open emails or click on ads, making AllFreeCrochet seem less attractive to advertisers. Together, around 20% of any cohort, or about 1 in every 5 subscribers, will either immediately unsubscribe or be removed given inactivity. The remaining 70% of subscribers will churn at a rate of 2% per month, decreasing every month due to time dependency, eventually churning at about 1% by the end of year 2 and and 0.01% at year 3.

With the churn rates and revenue per customer per month completely modeled, I calculated a discounted expected lifetime value of a customer to be around $8.00. This CLV indicates a large gap between acquisition costs and actual value. Given average acquisition costs of only $2.00, MagazineCo makes almost 4x back over lifetime of each customer. With such a large gap, it may be in MagazineCo's best interest to increase spending on acquisition costs in order to acquire more subscribers. At the same time, however, increased acquisition must be closely examined in order to determine the quality of subscribers. If increase acquisition results in lower quality subscribers with differing CLVs, the model changes and must be re-examined.

Limitations and Further Research:

The final discrete Weibull model with a spike predicted churn quite well alongside a parsimonious story and gives MagazineCo a more refined view of CLV. However, the underlying assumptions I made to simplify the model may be fleshed out in additional research to further improve the model's accuracy. One such assumption is a constant revenue per customer, based on the assumption that MagazineCo is only paid per thousand emails sent to customers. In reality, an additional revenue source for MagazineCo is when subscribers open the emails and then click on the links back to AllFreeCrochet.com, which generates even more impression and more revenue. Instead of modeling these open rates and click rates, I utilized averages and added this additional stream of revenue to the total revenue per month. A more sophisticated model may incorporate these effects in one of two ways: a beta-binomial beta-geometric (BG/BB) or a hidden Markov model (HMM). A BG/BB model would allow the churn rate (beta-geometric) to be estimated alongside an open rate (beta-binomial), with another extension for click rate. An HMM would allow the modeling of hidden states of users such as low activity, medium activity and high activity, correlating to low, medium and high open rates and click rates. Both methods would more precisely predict revenue derived from clicks to the website.

As alluded to earlier, I concluded no difference in churn propensity and therefore no difference in customer lifetime value amongst subscribers acquired in different months or from different sources. This conclusion acted more as an assumption given the limited scope of this project; while on the surface no differences appeared when looking at the high level data, perhaps further examination is needed to verify the homogeneity claim of subscribers. No research was done in the realm of referral traffic, which could yield interesting differences in CLV. Additionally, within paid search, there remain many different campaigns on different keywords that could also yield interesting differences in CLV.

Finally, analysis was completed on a monthly basis for two months, January and July, rather than on a continuous weekly basis for each week of the year. Given the seasonality of the craft and food industries, subscribers occurring right before holiday season may in fact act differently than subscribers acquired during non-peak times. Further research on this topic could explain differences in spikes at time 1 (people who unsubscribe immediately), and at time 6 (people who are removed from the list) and help direct marketing dollars where these type of subscribers are avoided in acquisition. Along this pattern of thinking, further exploration into residual value should be analyzed, specifically looking at subscribers who pass the 1-month spike and 6-month spike periods. The 20% of customers who unsubscribe during these times severely dampen the customer average lifetime value, given these customers represent such a large proportion unsubscribing so soon.

Lastly, the spike at month works well for the January 2014 cohort, but MagazineCo policy regarding when to remove users has shifted and become more specific over the last two years. They now have differing rules for when to remove users based on ISPs. The simple spike at month 6 will no longer suffice as some subscribers will be removed at 90 days of inactivity, while others may be 120 days or 150 days. Further analysis at the ISP level should be conducted to evaluate CLV differences amongst a gmail.com email and a hotmail.com and so on.

Conclusion:

To conclude, my research set out to accurately forecast customer churn for MagazineCo and attach reasonable revenue figures for each customer. A discrete Weibull model with a spike fit the data well both in-sample and out-of-sample while still providing a succinct story about customers. With this model, I calculated the average lifetime of a customer and multiplied this lifetime by a fixed discounted revenue per month in order to yield a customer value of $8.00. Armed with this knowledge, MagazineCo is better positioned to spend marketing dollars and forecast subscriber behavior into the future.

Appendix

Figure 1: Survival Rate for January 2014 Subscribers

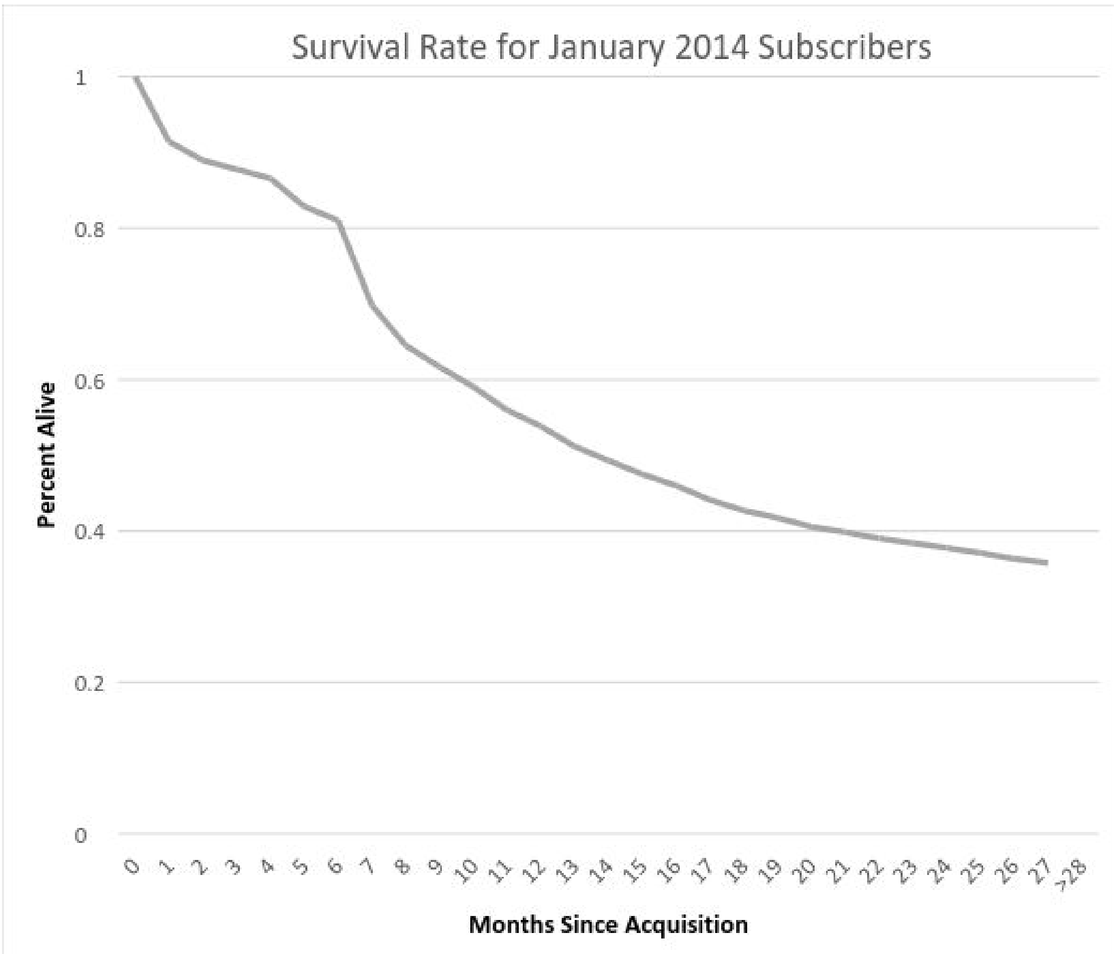

Figure 2: 2-Segment dW Churn Model

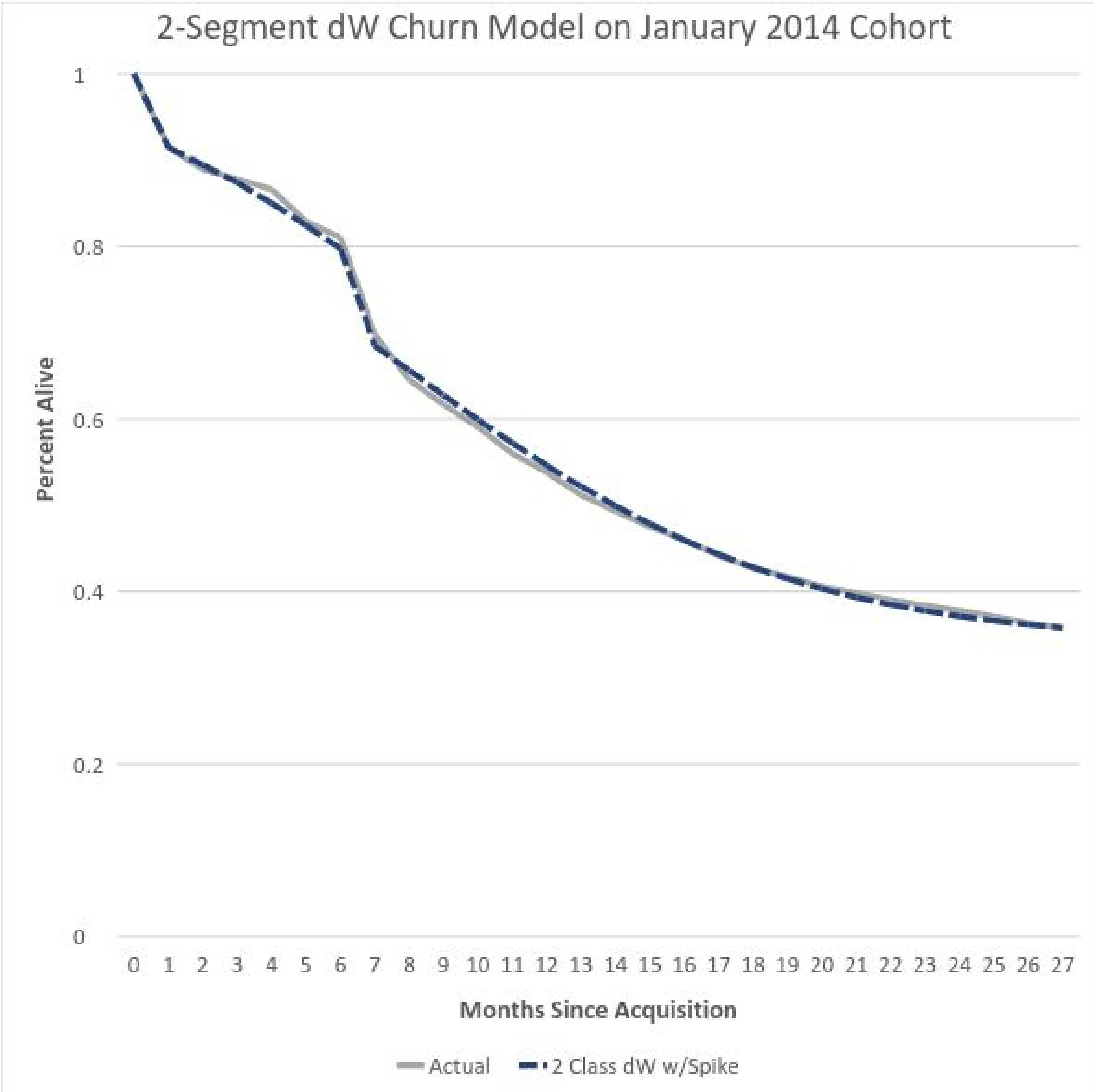

Figure 3: Survival Rates of January 2012 Data vs Predicted Rates of Survival for January 2014

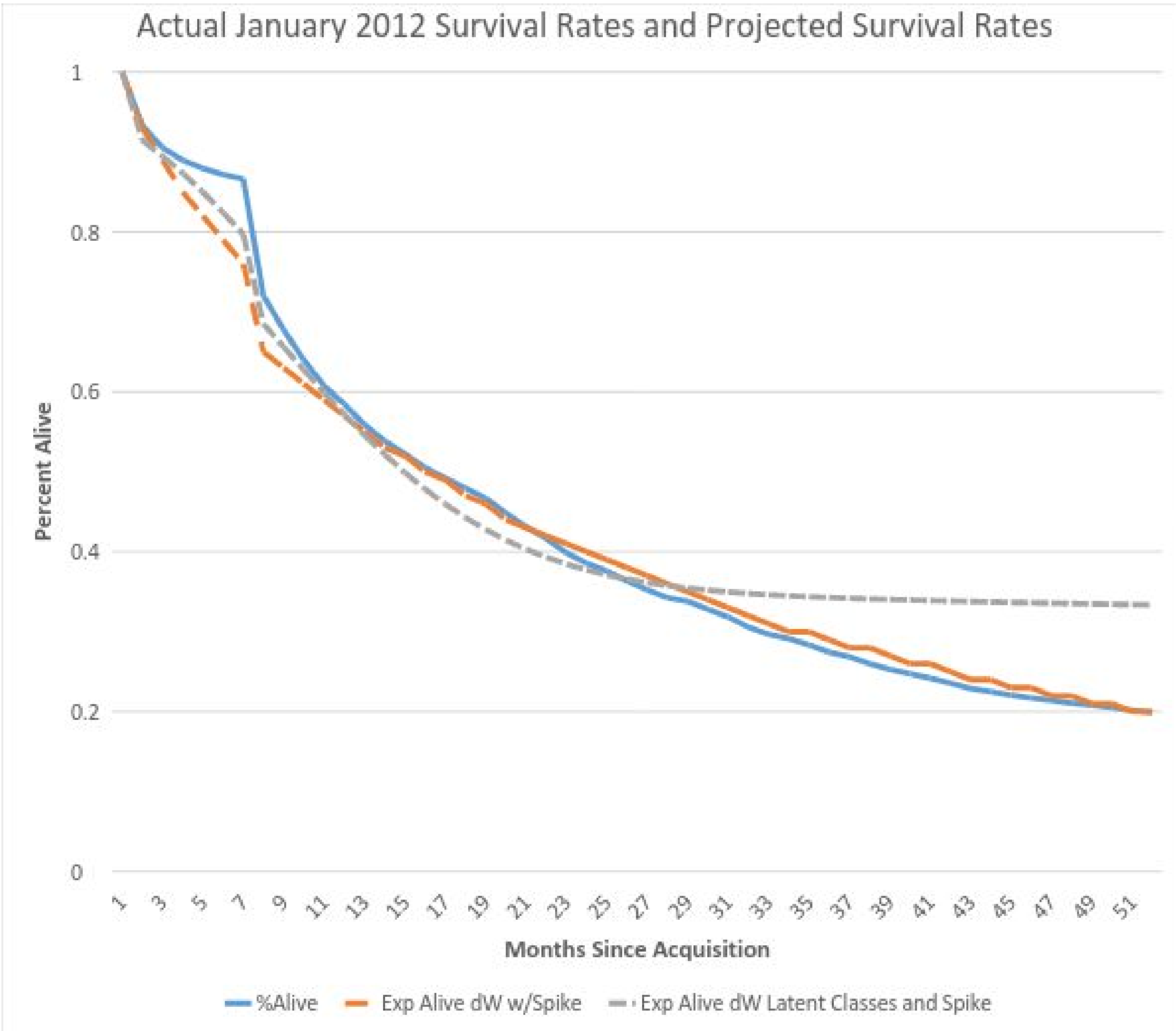

Table 1: Comparison of Models on January 2014 Cohort

Model Name	Parameters	LL	BIC
sBG	2	-127728	255489
sBG w/Spike at 6	3	-124605	249244
dW	3	-127846	255725
dW w/Spike at 6	3	-124315	248663
dW w/Spike at 1 and 6	4	-123909	247862
2 Segment Latent Class dW w/Spike at 6	6	-122612	245290