How to maximize customer retention and value

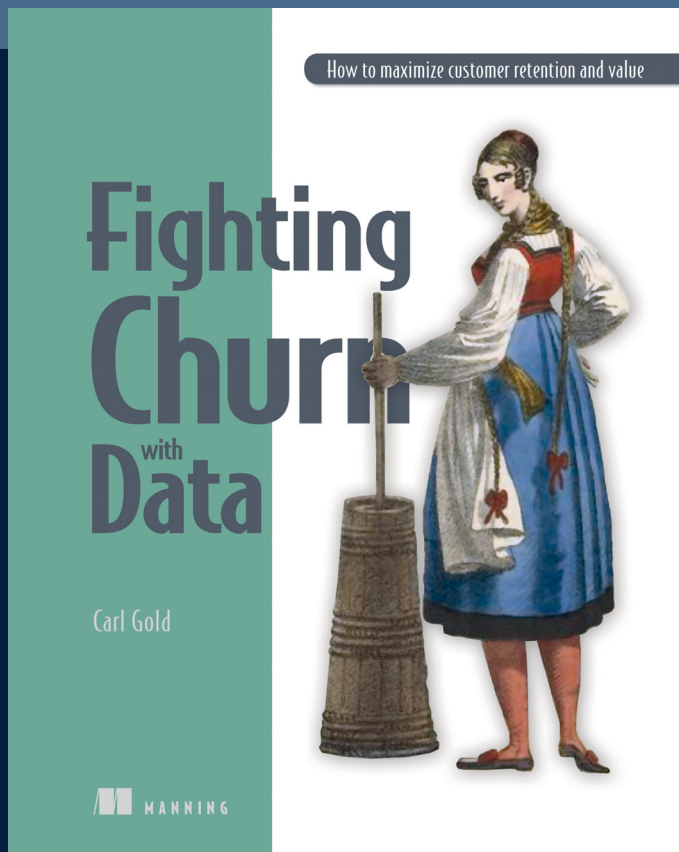# Fighting Churn with Data

Carl Gold

MANNING

How to maximize customer retention and value

Fighting Churn with Data

Carl Gold

MANNING

Fighting Churn with Data

by Carl Gold

ISBN 9781617296529
350 pages (estimated)
$39.99
available in summer 2020

Fighting Churn with Data

Carl Gold

Chapter 1

contents

The world of churn 1

What is churn? Why do we fight it? And how can data help? In short, why are you reading this book? If you're reading this book you're probably:

- A data analyst, data scientist, or machine learning engineer
- Working for an organization that offers a product or service with repeat customers or users

Or maybe you're studying to get one of those jobs or filling such a role even though it's not your job.

Such services are often sold by subscription, but your organization doesn't need to sell subscriptions to take advantage of this book: All you need is a product with repeat customers or users and a desire to keep them coming back. This book teaches many techniques around subscriptions, but in every case I'll show how the same concepts apply to retail and other non-subscription scenarios.

To get the most out of the book, you should already have a background in data analysis and programming. If that's you, then get ready for a game-changing breakthrough in the way you think about customers and data. This isn't your usual book about data analysis and data science because, as you'll learn, the usual approach doesn't work for churn. But you don't need to have a degree in data science to take advantage of this book: I'll review enough of the basics so that anyone with a little programming experience can get great results. With that in mind, I'll refer to you, the reader, as a "data person" since this book is written from the point of view of the person who works with the data. That said, this book is packed with business insights from real-world case studies, so even if you don't program you can still learn from reading the book! Give the book to your developer when it comes time to put the theory into practice, because this is a hands-on approach to the subject of churn and data.

If you already work with an organization that offers a live service, you probably know all about churn and want to get on with the fight. But I need to provide context for those who are starting out, and even if you already know about churn, I need to dispel a few common misconceptions before we begin. The first part of this chapter (sections 1-3) will provide you with the context for the rest of the book: What churn is, how you fight it, why fighting churn is hard and why I have selected the topics for the book. The middle of this chapter (sections 4-6) will start to make the theory concrete by describing the business contexts where these strategies apply and what data different types of companies have to work with. The end of the book (sections 7 and 8) will bring the theory to life by taking a first look at company case studies that will be featured throughout the book. By the end of the book you'll create results like those for your own product, or at least on a simulated case study.

1.1   *Why you're reading this book*

A primary goal for any product or services is to grow. Growth occurs when new customers are added through marketing and sales, but when customers leave it counteracts the growth and can even lead to contraction. For those in the business of providing such services, customers quitting a service has become known as "churn".

> **DEFINITION**   *Churn* is when a customer quits using a service or cancels their subscription.

Most service providers focus on acquisitions, but to be successful, a service must also work to minimize churn. If churn isn't addressed in an ongoing and proactive way the product or service won't reach its full potential.

The term churn actually originated with the "churn rate": The churn rate refers to the proportion of customers departing in a given period, as will be discussed in more detail later. This leads to the customer or user population changing over time, which is why the term "churn" made sense: because the word churn originally meant "to move about vigorously", as in churning butter. But in the business context the term churn is now also used as verb: "the customer is churning"; "the customer churned"; and as a noun: "the customer is a churn" or "make a report on last quarter's churns".

Customers not churning from a service can also be framed in a positive sense, if you prefer to see a glass as half full. In that case, people talk about "customer retention".

> **DEFINITION**   *Customer retention* refers to keeping customers using a service, and renewing their subscriptions, if there are subscriptions. Customer retention is the opposite of churn.

Reducing churn is equivalent to increasing customer retention, and the terms are interchangeable to a large degree. When the goal is stated as retaining more customers for longer, it's clear that in addition to "saving" customers who are at risk of churning, there should also be a focus on keeping customers more engaged generally.

There is even the possibility "upselling" the most engaged customers to more advanced versions of the service, typically for more money. Saving churns, increasing engagement, and even upsells are all important goals for services with repeated customer interactions. The difference between these is a matter of the area of focus and not the strategic intent. This book will give you the ability to address engagement and upsells with data as well. It may come as a surprise but note the following.

> **TAKEAWAY** Despite the wide variety of products and services with repeat customers, a single set of techniques exists for using data to fight churn and increase engagement, retention, and upsell.

This book will give you the skills to fight churn effectively using data in any kind of recurring user interaction scenario.

1.2   *The typical churn scenario*

If you already work in an organization that creates a subscription product, your situation probably looks something like the one shown on the left side of the diagram figure 1.1. The key ingredients are:

1 A *Product* or service that's offered and used on a recurring basis.
2 *Customers* who are interacting with the product.
3 The customers may have entered into *Subscriptions* to receive the product or service. Subscriptions often (but not always) have a monetary cost associated with them.
4 Subscriptions can be ended or canceled, known as churn. If no subscriptions exist, a customer churns when they stop using the product.
5 The timing, prices, and payments for the customers and subscriptions (if any) are captured in a database; typically, a transactional database.
6 When customers use or interact with the product or service, these events are often tracked and stored in a data warehouse.

We'll look at a wide variety of products that fit this description later in this chapter (1.4). If your scenario isn't quite like this but has several of the elements that's fine too—there are related scenarios to which the techniques in the book also apply described in section 1.5 (Non-Subscription Churn Scenarios). What's described until then are the most common techniques.

Throughout the book I may interchange the terms "subscriber", "customer", and "user". Each of these has slightly different connotations, but in general the same ideas will always apply (a subscriber has a subscription, a customer pays and a user may do neither—but you still want them coming back!) The techniques in this book apply regardless of which relationship you have with your customer, so if I make an example out of a persona that isn't relevant to you, then you should mentally substitute the persona to the one that is most appropriate for your product.
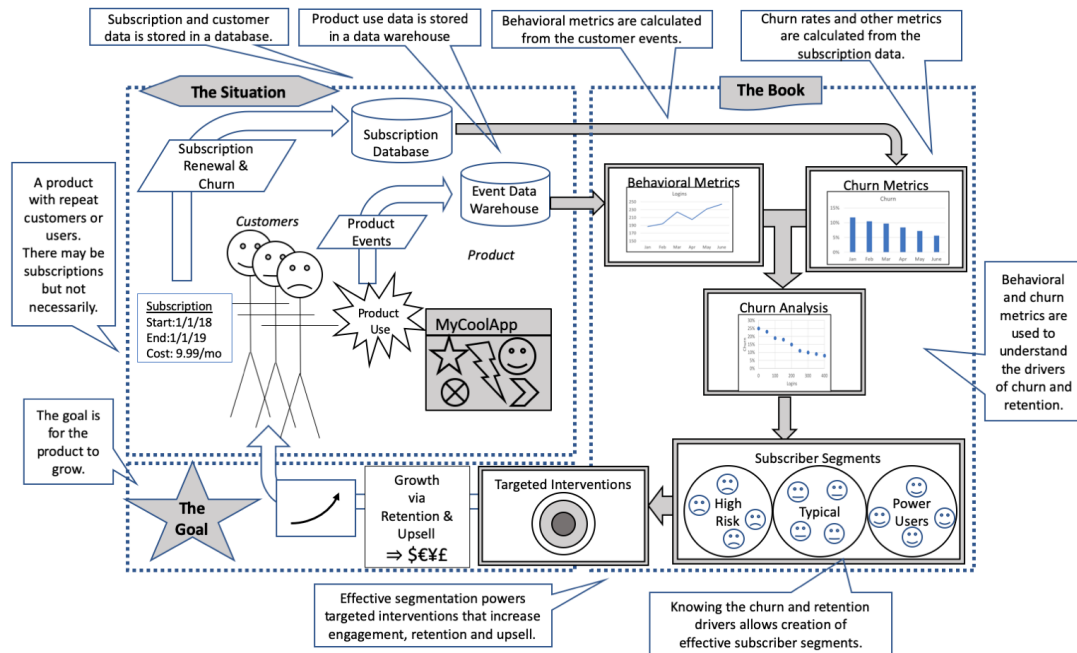
Figure 1.1   What this book is about.

1.3   *What this book is about*

The right side of figure 1.1 shows how the techniques in this book work together. The following describes each step in the process:

1   *Churn Measurement*: Subscription data is used to identify churns and create churn metrics. The churn rate is an example of a churn metric. The subscription database also allows identification of customers who churned and who renewed and exactly when they did, which are needed for further analysis.

2   *Behavioral Measurement*: The event data warehouse is used to create behavioral metrics that summarize the events pertaining to each subscriber. Creating behavioral metrics is a crucial step that allow the events in the data warehouse to be interpreted.

3   *Churn Analysis*: The behavioral metrics for identified churns and renewals are used together in a churn analysis. The churn analysis identifies which behaviors of subscribers are predictive of renewal and which are predictive of churn and can create a churn risk prediction for every subscriber.

Not shown in figure 1.1: At this stage, additional sources of information apart from the subscriber database and event data warehouse can also be brought into the analysis. This includes either *demographic* information about customers or users who are individual consumers (age, education, etc.) or *firmographic* information about subscribers who are businesses (industry, number of employees, etc.)

4 *Segmentation*: Based on their characteristics and risks, subscribers are divided into groups or segments that combine aspects of their risk level, their behaviors and any other characteristics found to be significant. The purpose of these segments is to target them for interventions designed to maximize the subscriber lifetime and engagement with the service.

5 *Intervention*: Using the insights and subscriber segmentation rules derived from churn analysis, churn reducing interventions are planned and executed. Interventions can include email marketing, call campaigns, and training. Another type of long-term intervention is making changes to the product or service itself. The information from the churn analysis is useful for this too. This is the crucial step that drives the desired outcome (growth!). More information about different types of interventions begins in the next section and is provided throughout the book, but you'll find that interventions are only covered at a general level, which is why interventions are shown as partly outside the scope of the book in figure 1.1.

I'll refer back to figure 1.1 in each chapter to make it clear what part of the process the chapter covers.

1.4 *Fighting churn*

This book is motivated by the difficult challenge facing people trying to reduce churn. That said, my motto is to "underpromise and over-deliver". I'll begin with warnings about how hard it will be. But eventually I'll show that the imperfect options available can still lead to a material impact on your churn and user engagement.

1.4.1 *Interventions that reduce churn*

Companies use five main strategies to reduce churn. I'll summarize them here and explain more on these topics throughout the book.

1 *Product improvement*: Product managers and engineers (for software) and producers, talent and other content creators (for media) reduce churn by making changes to the product features or content that improve the value or enjoyment the customers receive. That can include both brand new features and content or repackaging to make sure users *find* the best parts. This is the most primary or direct method of reducing churn. Another method to reduce churn in software is to try to increase "stickiness", which roughly means modifying the product to increase the switching cost for customer to change to an alternative. Switching cost is increased by providing valuable features that are hard to reproduce or are difficult to transfer from one system to another.

2 *Engagement campaigns*: Marketers reduce churn with mass communication that directs subscribers to the most popular content and features. This is more of an educational function for marketing than a traditional approach of marketing to increase sales: Remember, the subscribers already have access and know what

the service is like, so promises won't help. Still, this function is often undertaken by marketing because these are the people with skill in crafting effective mass communication.

3   *One-on-one customer interactions*: Customer success and support representatives prevent churn by making sure customers adopt the product and help them if they can't. Customer Support is the department that traditionally has helped customers, but Customer Success is a new and separate function in many organizations that's explicitly designed to be more proactive: Customer Support helps customers when the customer's ask for help, and Customer Success tries to detect customers that need help and reach out to them *before* they ask for it. Customer Success is also responsible for "onboarding" a customer or making sure they do everything necessary to start taking advantage of the product.

4   *Right-sizing pricing*: Account managers in the sales department (if one exists) may be the last resort in stopping churn, assuming the service isn't free. Account managers are the people who can actually reduce the price or change the subscription terms, managing the process through which a customer may downsell to a less expensive version. In consumer products without a sales department this role is usually taken on by senior Customer Support representatives who have similar authority. While we describe the role of sales as a "last resort" in preventing churn, a more proactive approach in many organizations is to "right size" the sales in the first place: do a better job of selling the product version that's optimal for the customer rather than selling the most expensive version possible. This may hurt the short-term gain to the service from each sale, but if done correctly it will reduce churn and ultimately improve the lifetime value of the customers.

5   *Targeting acquisitions*: You may find that different channels that you acquire customers through produce different qualities of customers in terms of retention and churn. If that's the case, it makes sense to focus your acquisition on the best channels. This approach is different because rather than trying to keep the customers you have longer, it tries to find better customers to replace them with. It's the least direct method to reduce churn and limited because most products cannot get unlimited customers from their preferred channels. Still, it's an important tool and you should take advantage of it if you can.

All of these methods are most effective when they're *data driven*, meaning that your organization picks the targets and tailors the tactics based on correct reading of the available data. Being data driven doesn't require you to have a certain amount or type of data or to use any particular technology. The emphasis of the book is on using the available data correctly, regardless of what type of product you work on or what type of intervention you ultimately employ to reduce churn.

> **TAKEAWAY**   Being *data driven* when fighting churn means designing product changes, customer interventions, and acquisition strategies based on a sound reading of available data.

One thing to note: Interventions and service modifications are the final crucial step to achieving the goal of lower churn and longer retention. But how to execute interventions is beyond the scope of this book. Note that unlike the data analysis techniques, *interventions to influence subscriber behavior are generally specific to different types of subscription services.* There's no one size fits all intervention. Also, in general other people (other than the data person) have the job of making those interventions (that is, product designers, marketers, and so on).

> **TAKEAWAY** General principles do exist for churn-reducing interventions, but these require customization to each product's circumstances.

The circumstances that shape interventions include not only the particular features of the product or content but also the technology and resources available for making the interventions. To give adequate coverage to interventions would be another book (or even a separate book per industry), and it would be a book aimed at business managers, not a technical book like this one. Interested readers should look for titles in "Customer Success" in the business section, or more specifically under product design, marketing, customer support, and so on. The tools and techniques in *this* book will revolutionize your products' performance in every one of those areas, but don't expect the data person to do it all on their own!

1.5 Why churn is hard to fight

Now that you know the goal and strategies available, I'll introduce you to the difficulties you'll face. The difficulties are what motivates my recommendations (in the next section) for how you use your data to fight churn and the topics you'll learn in this book.

1.5.1 Churn is hard to prevent

The bad news here is that people are (mostly) rational and self-interested, and your customers already know your product. To reduce churn in a long-term and reliable way you have to either improve the value delivered by your product or reduce the cost. Think of what it would have taken to stop *you* from churning off the last product that you churned from: Would better content and features have stopped you from churning? Maybe. Would a lower price have stopped you from churning? Maybe. How about an improved user interface? Would that stop you from churning? Probably not, unless the user interface was terrible to begin with. And would more frequent emails and notifications from the product stop you from churning? Again, probably not, unless those emails contained information that you found valuable. (There's that value word again!)

To reduce churn you need to increase value the users receive, but this is going to be harder than getting people to sign up for a service in the first place. Because these are people who already know exactly what the service is like, promises made by marketing or sales representatives aren't going to get much traction. As the data person, you may be asked for "silver bullets" to reduce churn, and is the following gives you the bad news in short.

TAKEAWAY   You have no silver bullets to reduce churn.

If by a silver bullet we mean a low-cost and reliable method. In the words of the famous startup CEO and venture capitalist Ben Horowitz, "There are no silver bullets for this, only lead bullets." He was talking about delivering competitive software features in his startup memoir, "The Hard Thing About Hard Things", but I think this applies equally to fighting churn: What it means is that no quick "one and done" fixes exist. You have to do the hard work of increasing the value you provide to subscribers continuously. It's not to say that simple fixes for problems in subscription services never exist, but these types of issues are often addressed by the normal job of people such as product managers and content producers. When the service turns to a data person for help in the effort to reduce churn, these kinds of low-hanging fruit have usually already been picked, so to speak. If a data person *does* discover easy fixes, it's a sign that those who created the service in the first place have not been doing their jobs well (so it's possible you'll find easy fixes, but you *shouldn't* find them...).

The alternative is to reduce the cost of the service. But reducing the monetary cost is the nuclear option for a paid service—revenue churn or downsells may be better than complete and total churn, but it's still churn.

WARNING   Price reduction is a "diamond bullet" against churn: it always works, but you can't afford it.

As you'll see in the next chapter, most services consider downsells another form of churn.

1.5.2   *Predicting churn doesn't work (very well)*

Now let's talk about the usual tool in the data scientist's toolkit: prediction with a machine learning system. Predicting churn doesn't work well here for two reasons. First and most important, predicting churn risk doesn't help with most churn-reducing interventions. Because there's no such thing as a one-size-fits-all churn intervention, interventions need to be targeted by factors other than the likelihood of churn. This is different from other areas where yes/no predictions tell you enough to choose an action, such as spam email or fraud detection. If you classify an email as spam you put it in the spam folder: done! But if you predict a customer is at risk for churn, then what? To reduce churn, you may run an email campaign to promote use of a product feature. But a campaign like that should be targeted at users who don't use the feature, not spammed to *all* users who are churn risks for any reason. Clogging users' inboxes with inappropriate content is going to drive them away, not save them! Churn risk prediction may be a useful variable in choosing customers for one-on-one interventions by customer success teams, but even then, it's only one variable defining the targets.

This may disappoint certain readers: To reduce churn it will not be sufficient to deploy an "AI" system that will win a data science competition. If you try to deliver an analysis that predicts churn without providing more actionable information, the business can't use it easily and they probably won't use it at all. Believe me when I tell you

that predicting churn isn't the focus of fighting churn with data—this is one of the most important lessons I had to learn when I started working in this area.

> **TAKEAWAY** You can't often find one-size-fits-all churn interventions, so predicting customers at risk of churn is only a little bit helpful when it comes to reducing churn.

The second reason predicting churn doesn't work well is that churn is hard to predict with high accuracy, even with the best machine learning. It's easy to see why predicting churn is going to be difficult if you recall your own behavior the last time you churned from something. You probably weren't taking full advantage of the product for a long time, but it took you so long to cancel because you were too busy, you spent time researching your alternatives, you couldn't make up your mind, or you forgot. If a predictive system was observing your behavior during that time, it would have flagged you as a risk and been wrong for the time it took you to make up your mind and cancel. The moment of churn was shaped by too many extraneous factors to be predicted.

  Apart from extraneous factors influencing timing, what makes churn hard to predict is that utility or enjoyment is a fundamentally subjective experience. The likelihood of churn varies from individual to individual even under the exact same circumstances. This is especially important for consumer services, where churn is usually hardest to predict. For business products, the customers tend to be rational. But neither the customer nor you have enough information to do a precise cost benefit analysis on their use of product. Finally, churn is normally rare in comparison to retention; actually, it has to be for any paid subscription that remains in business. Because churn is a rare outcome, false positive predictions will be common no matter how you make predictions. Given all these things going against you, it's inevitable that predictions of which customers will churn will be relatively crude. If you worked on a project where you predicted churn in the past and found it easy to predict with high accuracy, you might have been predicting churn too late, when it isn't actionable (see chapter 4). I'll provide actual data on churn prediction accuracy and what constitutes accurate versus inaccurate churn prediction in chapter 9. For now, I hope I gave you enough anecdotal arguments for why highly accurate prediction isn't usually possible.

> **TAKEAWAY** Extraneous factors, subjectivity, incomplete information, and rarity make it hard to predict churn accurately.

1.5.3 *Reducing churn is a team effort*

One of the hardest things about preventing churn is that it's no one's job, in the sense that no one person or job function can do it alone. Consider the different strategies for churn reduction described in the last section: product improvement, engagement campaigns, customer success and support, sales, and pricing. Those functions span more than half the departments in a typical organization! Churn reduction is going to suffer from problems of communication and coordination. If left unchecked, there will be a tendency for different teams to come up with uncoordinated approaches to

reduce churn. It would be counter-productive, for example, for the product and marketing teams to decide to focus on driving use of *different* features or content! And those approaches may be based on limited or flawed information: because they aren't the data experts (that's you, remember?), there's no guarantee choices made in independent teams are going to be properly data driven.

> **TAKEAWAY**   Churn reduction efforts are at risk of miscommunication and lack of coordination between the multiple teams involved.

Also, in a typical situation the data person can't do *anything* to reduce churn on their own. Because reducing churn depends on actions taken by the specialists in different parts of business, not by a data person who's wrangling the data. These coworkers you have to work with to reduce churn are diverse, and I'll refer to all of those people as "the business", for lack of a better term. It isn't to imply that the data person isn't a part of the business, but data people usually have no *direct* responsibility for concrete business outcomes (such as revenue) while the people in those other roles often do. From the point of view of the data person, "the business" is the end user of the data analysis results.

> **TAKEAWAY**   For the data person, the goal is to make the business people more effective at churn reducing interventions.

1.5.4   *Great customer metrics: weapons in the fight against churn*

Churn is hard to fight because different parts of the business are responsible for reducing churn in different ways. All of these teams have different tools and methods, and these different teams may not align on the situation and strategies. Also, all of the methods to reduce churn require the business people to target the interventions to customers most likely to respond. As a result of all these factors, what the business people need to help them fight churn are a shared set of facts or rules for understanding customers and their engagement with the product. The best way to make data into a weapon in the fight against churn is to use the data to produce effective measurements of the customers and get those measurements in the hands of the business' churn fighters. As will be explored fully in chapter 3, a measurement of customers is called a *metric* for short.

> **DEFINITION**   A *customer metric* is any measurement you make on all customers individually.

To make a simple example, a metric can be something such as how many times per month every customer uses a software feature or watches episodes from a certain series. But not any metric is great for fighting churn.

To continue the simple example, you may discover business rules such as "customers who use (view) the product feature more than five times a month churn at half the rate of customers who use it only once a month or less." Something such as using or viewing more of a particular feature isn't complicated, and the finding about churn

> **Metrics for fighting churn**
>
> For a customer metric to be really great for fighting churn the metric should be:
>
> - Easily understood by the business
> - Clearly associated with churn and retention, so it's obvious what a healthy customer looks like
> - Segments customers in a way useful for targeted interventions that increase engagement
> - Useful to multiple different functions of the business (for example, product, marketing, support, and so on)

makes it obvious what is a healthy customer. Each part of the business can use a fact like that differently. Product creators will know the feature is providing value and either replicate it or make it easier to find. The marketing department can design a campaign to drive users to the feature. And when Customer Success/Support people are talking to a customer, they can ask if the customer is using that feature and encourage the customer to try it if not.

That may sound easy but coming up with actionable findings that *sound* simple is harder than it sounds. Certain findings can be misleading (see section 1.8.2 for an example), and the more common problem is there can be too many potential metrics and rules, so the challenge is to find a *concise* set of metrics for the business to follow. Looking for easy-to-understand facts and rules doesn't mean your job will be easy!

I arrived at this focus on delivering great customer metrics by experiencing situations where the metrics weren't great. When I was starting out, the situation often went like this: customer metrics were already decided by a company before we started analyzing churn, and those would be the inputs for creating a predictive model. The predictive model wouldn't be used (as described in the previous section), and along the way I often found that the customer metrics had been poorly designed and weren't good for predicting and understanding churn. But at the end of the data science project, everyone would ignore the predictive model and continue using the mediocre metrics anyway, because they needed something to do their segmenting. That's when the lightbulb went off for me: the data analysis should focus on making sure the metrics are great for churn, because that's what people are going to use to do their jobs. And I knew that as a data expert I could do a better job at creating customers metrics than the business people—and so can you.

The approach that I'm going to teach you is more like a traditional statistical or scientific analysis—data people who trained as statisticians will likely find this approach more natural than computer scientists. The process is to iteratively test different customer metrics, analyze their relationship to churn and to each other, as well as evaluate them for interpretability and how useful they are for segmenting and making interventions. When you find the best set of metrics, that's the main deliverable to the business. And you'll also be in a good position to run predictive models for additional use cases, which will be described in later chapters.

> **TAKEAWAY**   The main deliverable to the business from the data analysis project will be a set of great customer metrics.

1.6   Why this book is different

You may have realized by now that this isn't your usual book about data science or data analysis. Table 1.1 summarizes how this book compares to a typical book in this field.

Table 1.1   How This Book Compares to Other Books About Data Analysis

This book	Other books
One scenario, from start to finish, including applications	Helicopters into many different scenarios without details
Focus on understanding the data and designing metrics (for example, feature engineering)	Focus on algorithms
Datasets are created from raw data, in an iterative process	Datasets are fixed benchmarks
Emphasis on interpretability, parsimony, and agility	Emphasis on maximizing accuracy or other technical metrics

First, this entire book is focused on one scenario and process: using your data to fight churn. In contrast, most books on data analysis cover a wide variety of use cases and emphasize teaching statistical and computer science algorithms. That said, section 1.5 will explain that a whole family of use cases exists that are similar to churn to which the same techniques apply. But this *book* will stay focused on churn to make the learning easy—once you're an expert it won't be hard to modify the techniques to a related scenario.

In typical data analysis books, the focus is on teaching algorithms and the data sets used to demonstrate the techniques are given, also known as a "benchmark" data set. The method in this book starts from raw data and turns it into an analytic dataset. That is a large portion of the work: the data about customers is stored in a database of events and this raw data must be converted into summary metrics in to be used in an analysis. Several statistical/machine learning algorithms will be explained at the level required to use them, but this book won't teach too much theory. Instead, the focus is on teaching the entire analytic process, including application of the results, in a real-world scenario. This book is an applied approach, not theoretical.

An important way in which real world data problems differ from training is that in the real world the job never ends: As soon as one analysis of churn is complete, new product features or content will be created that will require re-analysis. Or an entirely new type of data may become available to augment the original. Also, constant changes occur in the business environment, such as competition or changing economic conditions. Such changes can require re-analysis even if there are no changes in the product. To succeed in this environment, the process to use the data must be *parsimonious* and

agile. Parsimony means to the use of the minimum amount of data and analytic steps required to get the job done. *Agility* means responding to change quickly and efficiently. Achieving parsimony and agility has important consequences. I'll return to these themes throughout the book and for now only point out two key takeaways.

> **TAKEAWAY** The goal is to deliver actionable knowledge to the business people. Listen to them and try to answer their questions first. Don't exhaustively test every hypothesis or evaluation metric.

> **TAKEAWAY** Write code to automate the process. This makes it much easier to accommodate the inevitable corrections or requests for changes.

Because of the need to communicate with business people to fight churn, the analysis performed by a data person only achieves maximum impact when it provides actionable knowledge to the business. For this reason, a theme throughout the book is how to communicate the analysis results in a way that non-technical people can understand. I'll use relatively simple visualizations and avoiding technical jargon in favor of common language. I even recommend making big simplifications in explaining the analysis to business people (but no corners are cut in actually doing this analysis!). Specific examples of explanation strategies I have found to work are explained throughout the book.

1.6.1 *The importance of data and metric design (feature engineering)*

Many people who take academic classes on data science or data analysis get a surprise when they start working at a company or on a real research project: the data they need isn't sitting waiting for them in a CSV file or database table, all ready to run through an algorithm. In fact, most real-world projects involve locating and merging data from multiple databases or systems and this process is a large part (usually more than half) of the work of the analysis. In academic data science this is referred to as feature engineering, but I'll stick to the term metric design because of the need to communicate to business people (more about this in chapter 3).

Another common misconception picked up in a typical data science education is that the choice of algorithm or analytic method is the most important thing determining the model power or accuracy. In fact, the design of the summary metrics (known as data features to academics) that allow the data to be analyzed is the most important part of the process, even when accuracy is concerned.

Data people who are academically inclined may see it's inefficient or even unworthy of them to prepare their own data, seeing it as "janitor work". But many small decisions must be made when preparing the data, and some of these may have tremendous consequences for the analysis if those decisions aren't made correctly or not in the way the data analyzer expects. It's okay to have another person prepare your data, but only if they're going to be okay with you telling them to change and correct the way they do things repeatedly, as typically happens over the course of a project.

> **WARNING** It's extremely risky to delegate crucial tasks of data preparation to another team or to any individual who doesn't report directly to the person responsible for the final analysis.

Finally, I want to add that in my own experience understanding the data and designing the metrics is the most fun and creative part of the entire process! In my opinion, this is the true "science" in data science: learning from your data through experimentation.

1.7 *Products with recurring user interactions*

For those not already familiar with this area, I'm going to summarize the current state of the world of products that have recurring user interactions or subscriptions. Subscription and recurring payment business models are definitely not new: subscription news services have existed since at least the 16th century, and the recurring payment of insurance premiums was established in the 17th century. The 20th century saw the rise of ubiquitous recurring payment services for a wide range of utilities made available by the second industrial revolution: first water, gas, electricity, and telephone service; and in the late twentieth century cable television, mobile phone service, and, of course, internet service. These services are all based around recurring payment relationships between the consumer of a service and the provider. All such relationships can be referred to as "subscriptions".

Although when we think of subscriptions, we usually think of a fixed-fee paid periodically, subscription services may actually collect three types of payments between the consumer and provider:

1. Recurring payments: Fixed payments of the same amount for each period of service
2. Usage based payments: Pay for the amount of the service used, based on a unit of measure
3. One-time payments: Usually fees for setup, but also for temporary (non-recurring) upgrades to service or one-time ("in app") purchases

The services are sometimes paid in advance at the start of each service period and sometimes in arrears after the service is provided. But what's common in all of these services is the continuing relationship between consumer and service provider.

The 21st century has seen a new explosion in subscription services of an even greater variety, largely delivered over the internet and created (or at least managed) using cloud computing platforms. One important characteristic that sets these 21st century subscriptions apart from most of their predecessors is the fact that the new products are often of a more discretionary nature. While for 20th century recurring payment services there were usually few or no options (many utilities are still regulated monopolies), we have a variety of options when it comes to 21st century subscriptions. We often have alternative services to choose from, such as switching from one streaming music service to another. Also, we have alternative means of achieving the same ends, such as a business writing its own software rather than buying it on a

subscription. And finally, the fact is that many of these modern subscriptions are things we can live without: do you even need that food delivery subscription when you could shop in a store?

In the following sections I'll describe several of the wide variety of subscriptions that exist today, because we'll consider these business models in the rest of the book.

1.7.1 *Paid consumer products*

Most people these days are familiar with consumer subscription services. These products typically cost a modest amount each month (less than a fancy meal), and the price often ends in "99" (9.99, 49.99, 99.99, and so on) This is how most consumers get much of their entertainment these days, and a wide variety of additional products have become available:

- Desktop software, which was formerly sold through perpetual licenses, for example, word processors and spreadsheets, graphics creation tools, anti-virus, and so on.
- New types of software as a service, such as identity theft protection, cloud storage, home security camera video monitoring, and so on.
- "Boxes" of physical products: shaving and personal grooming, meals, crafts, gifts, and so on.
- Personal apparel items, including clothing and watches.

These products are often referred to as Business to Consumer services, or B2C; another related term for is Direct to Consumer (D2C) although it emphasizes a market in which not all producers are D2C (at the time of this writing many television channels are available only through cable or satellite subscription, but a streaming service is D2C.)

1.7.2 *Business-to-business services*

In terms of the value of the market, subscription services for businesses an enormous market segment. These are commonly known as Business-to-Business products, or B2B. Starting with Salesforce, which created the first cloud-based Customer Relationship Management system in the 2000's, this market has exploded. Now nearly all new software products for businesses are offered as a service or SaaS: "Software as a Service". And at the same time, existing on-premises software products have started to shift to the new model due to the efficiencies of cloud deployment and upgrades. These products exhibit a dizzying variety of payment terms, because business products don't shy away from complex contracts (indeed they often seem to favor them. . .) The most well-known categories of B2B subscriptions include:

- *Customer Relationship Management (CRM)*: SaaS to coordinate sales team and marketing interactions
- *Enterprise Resource Planning (ERP)*: SaaS for accounting, logistics and production

- *Subscription Business Management Systems (SBM)*: SaaS to manage subscriptions are also sold on a subscription basis.
- *Human Resource Management (HR)*: SaaS for managing employees, including hiring
- *Support Issue Tracking Systems (ITS)*: SaaS for tracking and managing customer support interactions, or "tickets"
- *Desktop software*: Spreadsheets, word processors, email, and illustrations programs sold through multi-user subscriptions
- *Cloud Computing Resources*: Cloud servers, storage, databases, and content distribution networks (CDN)
- *Business Intelligence (BI)*: Tools that allow users to query and visualize data of a variety of types
- *Security products*

This short list doesn't do justice to the wide variety of SaaS products that are used by businesses today. Nearly every modern SaaS company is dependent on a wide array of other SaaS products to provide the software running the *non-core* portions of the operation. A typical SaaS company uses its in-house software engineers to create only what's unique about the service, while almost every other part of the operation is run using software provided by another SaaS venture, and always paid on subscription basis. This includes the standard applications listed above that are used by all types of companies, but there are also many SaaS products designed for use only by companies in a specific vertical:

- Software to manage comments and discussions are used by most online news or media services but that service isn't created by those companies, which focus on creating their content.
- Information services that support specific verticals. This used to be common only in the finance and legal industries but now many other industries have vertical specific information services such as real estate, energy, manufacturing, and agriculture.
- Services that manage invoices and accounts payable functions that are specific to a single industry or vertical.

These are a few examples from a large and heterogeneous category.

1.7.3 *Ad supported media and apps*

Since the early days of the internet, one of the most common businesses models (if not *the* most common) has been providing free media content (reading material, videos, music) but with ads displayed before, during, and/or after. You have no formal subscription to use such a product, so the scenario has several important differences from figure 1.1. Although there's no "subscription" database, another database probably tracks profile information about the users for the purpose of selecting ads. While the means of capturing value from the consumer is different, these services share the same discretionary nature as regular paid consumer services and the same obvious concern with churn:

the services want their users (ad viewers) to keep coming back. As will be explained, churn can be defined without a subscription to mean that a customer disappears for an extended period of time. And this type of product can also benefit from the churn analysis techniques taught in this book, as long as there's a form of event data warehouse that tracks users across multiple sessions. Tracking user events across multiple sessions is the one minimum requirement to use the techniques in this book.

1.7.4 Consumer feed subscriptions

Another novel type of subscription is a subscription to free information services to be delivered in a feed, such as YouTube subscriptions or email updates. This is a variant of the ad supported media model mentioned in the last section in which the content is still advertising supported, but the provider has created the option to upgrade the experience by "subscribing" for free. The subscription often means the consumer has given the product the right to update them with new information pushed to their inbox or video feed. In this case the structure of the relationship exactly fits the typical scenario of figure 1.1, with the only difference being that the subscription has no fixed payments but instead has advertising revenue derived from the events.

1.7.5 Freemium business models

A Freemium offering refers to any subscription service with both a free version of services as well as a paid, or premium, level of service. For certain services, the free version may be time limited, giving users a chance to "try before you buy". For other services, it may be possible to use the free level of service in perpetuity. Another common variety is to have a free version with advertisements (as described in section 1.4.3) and a paid version that's ad free. As far as churn is concerned, Freemium services are like services without a free level except there are two distinct types of churn for the same service: churn from the premium service as well as churn from the free level of service. The free level is analyzed using techniques described for subscriptionless, activity-based, churn analysis. Also, there's also the transition from the free to the paid service level, known as free trial conversion, which can be analyzed with exactly the same techniques as churn (see section 1.4: Non-subscription churn use cases.)

1.7.6 In-app purchase models

A variant on the freemium model are products that are free to use in perpetuity (or which require only a relatively small one-time payment to use in perpetuity) but which have a variety of ways to upgrade the experience by making one-time payments during use. This model is becoming the predominant model in online gaming. To play the game is free, but if you want a cool-looking skin for your avatar, it's going to cost you! Or a better weapon, or a shortcut to a higher level, and so on. This is another type of scenario that fits the model of figure 1.1, but without the subscriptions. Instead of subscriptions, you have a transactional database of the one-time purchases (and the original purchase of the app, if there's an initial fee.) This is another scenario where churn

can be defined as a user (customer) going inactive and all of the usual techniques in this book apply, as long as user behavior can be tracked across sessions of use.

1.8   Non-subscription churn scenarios

The focus of this book is on explicit churns from subscriptions, but the same approach works in a variety of other common business scenarios. I'll briefly explain those here, but then teach the techniques, mainly focusing on churn from subscriptions for simplicity. After you master the techniques you should have no trouble repurposing what you've learned for these other scenarios.

1.8.1   Inactivity as churn

User inactivity can be seen as churn for the free tier of a Freemium service, and the same applies for apps or ad-supported products with no explicit subscription. You choose a time window in which a user must engage with the service (that is, one month, three months) and churn is defined as users who go inactive for that long. In contrast to the typical scenario in figure 1.1, there's no subscription transaction database—only an event data warehouse. But the techniques in the book can still be used as long one key requirement is satisfied: user behaviors must be tracked consistently across different episodes or sessions of activity. For explicit subscriptions where every subscriber has an account in the subscription database, tracking across sessions is generally a given, and any app that requires a signup with an email should have the same capability. But this may not be the case for apps without a user sign-in or websites, where it will depend on the web analytic system in place.

1.8.2   Free trial conversion

As described above, a Freemium model is one with both a free and a paid level of service. Because similar behavioral data is available at the free level as in the paid subscription, it's as easy to analyze the *conversion* of subscribers from free service to the paid service as it is to analyze churn from the paid service. This is essentially the opposite of churn, and the scenario still looks like figure 1.1, so the exact same analytic techniques can be used.

1.8.3   Upsell/downsell

Adding new services or moving to a higher cost plan is known as an *upsell*, while removing services or moving to a lower cost plan (without churning) is known as a *downsell*. As with churn, behavioral data and user characteristics can be analyzed to determine what scenarios are most likely to lead to an upsell or a downsell. The additional challenge is that there may be many different possible upsell/downsell options to analyze. However, in practice an additional analysis for upsell/downsell is often unnecessary: most of the time the customers who are most likely to purchase some form of upsell are those who are the best customers identified by the churn analysis, while those customers most like to downsell are also those most likely to churn. Also,

upsells are often related to crossing a specific threshold in usage, such as the number of users (seats) sold for an enterprise software license or the amount of GB of data in a mobile phone service. In that case, an in-depth analysis is unnecessary: You know who's a candidate for an upsell simply by looking at a single relevant metric of usage.

1.8.4   Other yes/no (binary) customer predictions

Customer churn is an example of a prediction problem where with a yes or no answer. In statistics and data science this is referred to as a binary outcome, in reference to the two possibilities. The methods of this book can be applied without modification to virtually any situation involving predicting a future customer state that can be framed as a yes/no question. Examples include whether any kind of insurance policy will result in a claim (medical, auto, and so on) and whether a borrower will default on a loan. The one caveat here is that this book tends to focus on *rare* outcomes (which also happen to be the case for insurance claims and loan default.) Slightly different methods may be used if a "yes" outcome is equally common as a "no", and I'll note the differences when appropriate.

1.8.5   Customer activity predictions

Assuming a subscriber continues a service, it's reasonable to want to analyze what a subscriber is likely to do in the future. This is especially important for behaviors that are revenue generating. This could be how much the subscriber will use a pay-as-you go feature, or how much content the user will consume and the resulting ad revenue generated for an ad-supported service. Most of the techniques in this book also work for this sort of analysis, but with the major caveat that for churn we use techniques for modeling two-state, or "binary" outcomes (churn vs. continue); for activity prediction, you use techniques for modeling numeric outcomes.

But here are a few reasons why a separate activity analysis often not necessary: predicting a future real valued outcome with a high degree of accuracy is even harder than predicting simple yes/no outcomes such as churn vs. continuation. The best prediction for the amount of a future activity by a subscriber is usually the value observed for that subscriber in the recent past. For example, how much content a user will view next month is likely to be about as much as was viewed last month, so a decent prediction is made by creating the relevant behavioral metric for the churn analysis.

1.8.6   Use cases that aren't like churn

One use case that's different from churn and won't be covered in this book is product recommendation systems. Those are scenarios where customers have a wide variety of products or content to choose from, and the goal is to recommend the most suitable ones based on previous choices. The techniques in this book do apply for customers choosing between a small number of products, such as a basic, standard, and premium plans on a subscription. But for large catalogs of physical or media products, you should consult a book specifically about recommender systems.

1.9   *Customer behavior data*

Given the huge variety of products and services that include recurring customer/user interactions, those interactions can take an even larger variety of possible forms. One section in one chapter cannot possibly give an exhaustive listing of all the possibilities, so consider this only an introduction. Almost anything that happens in software or that can be tracked by software can be considered an interaction or event.

> **DEFINITION**   An *event* in the context of fighting churn is any user interaction or result on the product that's tracked in the data warehouse. Events are time stamped and pertain to a single account or user.

1.9.1   *Customer events in common product categories*

I'll make the discussion of data concrete by listing typical customer events for common product categories. What they all have in common is that they refer to individual events that can happen to one customer or user at a point in time. For certain events, the simple fact that the event took place for a certain user at a certain time may be the only information available, and for other events there will be details tracked along with the event.

- *Software*: Refers to any software product (SaaS) and can also refer to other types of products with software interfaces.
  - *Logins:* Logging into the application is usually tracked as an event.
  - *User Interface (UI) Interaction:* Almost any click or typing in the user interface may be tracked as events. The event will usually include a detailed reference to what part of the UI.
  - *Document/Record Actions:* These include creating, editing, updating, and deleting records or documents that are tracked in the application database. The event may include information about what type of document and specifically what field of the document, when appropriate.
  - *Batch Processing:* Many applications include processes that users run periodically. Every item processed may be seen as an event, as can the batch job itself.
- *Social Networks*: Not only dedicated social networks, but also products that have social functions.
  - *Liking:* Indicating that the user likes something they see is one of the most ubiquitous interactions with a social network.
  - *Posting:* Sharing any type of media supported by the network.
  - *Share:* Usually a specialized post that refers to another user's post.
  - *Connect:* Connecting with other users is often the most important form of engagement, because it enriches the future user experience.
- *Telecommunications*: Or Telco, for short, refers to providers of mobile or fixed location telecommunication products and services.
  - *Call:* Making a voice or video call. Typically call events are tracked with the duration and the type of call.
  - *Data:* Data usage is often tracked with the amount of data.

- *App:* Using an application and specifically which one.
- *Add/Remove Device:* Updating devices is an important event in the lifecycle of using the service.

■ *Internet of Things (IoT)*: Refers to products consisting of connected devices.
- *Geospatial:* Events about movement (of devices that can move) include location and speed.
- *Sensor:* Data received by a sensor can include almost any type of additional information that was received from the sensor.
- *Device:* As with sensors, device activity can refer to almost any kind of activity and includes device specific information.

■ *Media*: Any product that provides any type of pre-recorded or live-streamed media, including video, audio, images, and text, not only for enjoyment but also for education and professional training.
- *View/Play:* Playing a media is the most common event on a media specific service and often includes details about what media was played and how much was played. This includes articles or pages read for news and books.
- *Dwell Time:* Viewing a page or other content and explicitly capturing the time spent.
- *Like:* Indicating preferences on media by liking (or giving the thumbs down) is an important event for media.

■ *Gaming*: Any product that's a game.
- *Play:* Many events are typically generated during playing a game and may include information about exactly what parts were played, for how long, and so on.
- *Levels and Score:* Many games include points or other forms of "leveling" and achieving these is often tracked as events.

■ *Retail*: Shopping websites or services that allow purchase of individually selected items, which may be either physical or digital products.
- *View:* Viewing products may be tracked as an event, along with details of what product.
- *Search:* Searching a product catalog may be tracked as an event, along with keywords used for the search.
- *Add to Cart:* Adding products to a shopping cart may be tracked as an event.
- *Returns:* Returning products is also tracked as an event, along with product details.

■ *Box Delivery*: Refers to services where curated selections of (mostly physical) products are delivered to customers periodically.
- *Delivery:* Successful delivery of each box is an important event, as well as any failures or difficulties in delivery. Tracked information typically includes the type of box, when there are more than one, and the time it took to arrive.

- *Returns:* Many box services allow items to be returned, an important event indicating user dissatisfaction.
- *Retail:* Most box delivery products also include retail options so all retail events are also relevant.

There are also categories of events that occur across a wide variety of different kinds of products and services. Here are several examples:

- *Financial*: Financial events occur on all products and services that aren't free.
  - *Recurring Payments:* These are so common they're often non-events, but they're tracked and often more important when they don't occur than when they do.
  - *Non-Recurring Purchases:* These are all sales on retail sites and include any extra or in-app purchases made on games or subscription services.
  - *Overage Charges:* Charges that apply when users exceed a threshold.
- *Support*: Support interactions refers to whenever the customer turns to the service for help, whether through a call, email, chat, or even searching support/help documentation.
  - *Ticket:* A support ticket or case, usually tracked with an opening and closing time and a wide variety of details.
  - *Call/Email/Chat:* Any kind of interaction between the customer and support representatives, possibly including the full text of the interaction.
  - *Documentation:* Use of online documentation resources may be tracked as UI events.
- *Plan*: Subscription plan related events occur for any product or service that has an actual subscription.
  - *Plan Change:* The time/date of plan changes may be tracked as events.
  - *Billing Change:* These are events liking changing the credit card or other payment method, as well as switching billing details such as monthly vs. annual billing.
  - *Cancel:* Yes, cancelling the product or service is also tracked as an event. But note that when we talk about cancellation as events, we're talking about the date and time the cancellation change is entered in the system but not necessarily the time when the service contract ends, whenever the user has time remaining. *For this reason, a cancellation event isn't the same as a churn.* The churn occurs when the subscriber completes the current term without signing up for a new one, often allowing a short grace period. Consequently, *a cancellation event doesn't necessarily mean a churn will occur, because the customer still might resign before much time has passed.* This happens often enough that the act of cancellation should be considered an event that suggests churn is likely, but it's still not a certain conclusion.

Note also that many products or services will contain events from multiple categories. For example, any product with software user interface will collect software type events, while many products have social network features (for example games), even if they aren't exclusively or primarily a social network product.

1.9.2   *The most important events*

With all this discussion of types of events you're probably already wondering which type are the most important, and with good reason—with so much variety it's going to be important to stay focused. We have no hard rules for this, but we can give general guidelines to frame the discussion for the rest of the book. And to be clear: figuring out which events are the most important is one of the main points of doing the analyses in this book, because it's always different for every product service. This is a preview of things we'll look at in depth later.

The bottom line is this: the most important events are the events that are closest to the customer achieving the goal or purpose of the service. That's vague, but these examples should make it clear:

- Software products have a goal. For example writing documents. In this example, creating documents is more important than logging in. In general, login events are much less important than the events that are directly involved with achieving the goals of the product.
- Many B2B software products are used for making money, so if you can measure how much money is likely to be made from the events, then those are the most important. For example, if a product is a Customer Relationship Management system (CRM) used to track sales deals, then closed deals and their value is probably the most important type of event. Often a product isn't that close to the money business customers make, but you should still focus on events associated with commercial success. For example, if the product is an email marketing tool, opened emails are important events.
- For most media services, the purpose is to enjoy the media. Playing the content is generally important, and more specifically, indicators of enjoying the content are watching the whole thing, giving it a like, or sharing it. But you can never directly measure enjoyment because it's a subjective state of the users.
- For a dating service, the purpose for the user is to go on dates, so actual meetings are probably more important than tasks such as searching, viewing profiles, or interacting online. That presents a challenge because success on the service is well-defined but the actual events occur off the platform.
- For gaming, the purpose is to have fun. As with Media, subjective feelings are hard to measure, so the most important events could be achieving scores and levels or social interactions with friends.

Many important caveats go along with this point, and we've noted a few, but the following takeaway is important.

TAKEAWAY   Look for events that are as close as possible to the value created by using the service, even when that value cannot be measured directly.

The rest of the book is all about bringing rigor to this simple intuition.

1.10   Companies that fought churn with data and won

Now it's my pleasure to introduce you to a few companies that will show up in the case studies throughout the book. They all used data in some way to address their churn problem in a way that made a difference. This section introduces them and makes the data discussion of the last section more concrete with examples.

1.10.1   Klipfolio

Klipfolio is a data analytics cloud app for building and sharing real-time business dashboards and reports on web browsers, TV monitors, and mobile devices. Klipfolio helps companies stay in-the-know and in control of their business by giving visibility into the KPIs (Key Performance Indicators) and metrics that matter most. Klipfolio believes in empowering people to use and understand their data, anytime and anywhere, eliminating the unknown, making them more competitive every day. Subscriptions to Klipfolio's online app are sold to businesses. Like most B2B SaaS products, the price depends on the number of users and a variety of extra features. Most subscriptions bill either monthly or annually and continue until the subscriber cancels.

As an extremely data driven company (its products are all about data!), Klipfolio was enthusiastic in using data to fight churn and increase customer engagement. Klipfolio learned early on that luring customers to stay with downsells and discounts wasn't worth it—customers tend to churn anyway as soon as the discount expired. By analyzing usage and churn patterns, Klipfolio discovered that customers were at risk of churn if only one person in an organization used the product, so a key metric for them became the number of active users per account. They also found there was a high risk of churn early in the subscription if the customer didn't fully adopt the product, so they instituted onboarding calls by their support group and free support for the first three months. They also realized that the lifetime value of certain customers was too low to justify. To address the situation, they reconfigured their pricing and packaging to more profitable versions of their plan and features.

We'll see more about how Klipfolio used their data to achieve those outcomes throughout the book. For now, we start by taking a quick look at part of their event data. As the product is for making Klips and Dashboards the most common events are "View Dashboard" and various events around the Klips such as "Edit Klip" and "Save Klip". Also, there are events around social features related to the Klips, such as sharing. More than 80 different events were captured and the most common are listed in Table 1.2. Klipfolio's events also identify individual users within each company, and some of the events contain additional data such as the session duration.

Table 1.2   Most Common Klipfolio Customer Events

View Dashboard	Orientation Switch
Tab Switched	Account Active Today
Klip Editor	Klip Editor Exit
Edit Klip from Dashboard	Add Klip Overlay
Klip Saved in Editor	Reconfigure Data Source

1.10.2 *Broadly*

Broadly ensures that businesses looks great online and are found and chosen by potential customers. Broadly's review strategy makes it easy for customers to write great feedback on Google, Yelp, Facebook, and company's website with one click, and Broadly's Search Engine Optimization technology helps your business stand out. Like most B2B SaaS companies, Broadly sells subscriptions with different plans depending on the size of the buyer and things, such as the number of users. Subscriptions are sold on either a monthly or annual basis and include a variety of add on products to help companies with more advanced needs such as website design and consulting.

Because Broadly is a company whose products facilitate communications and connection with customers, they appreciated the importance of engaging with their own customers. Broadly has a team of Customer Success Managers (CSMs), and this team makes time every day to reach out to customers struggling with the product. The CSMs use metrics to guide the conversation that they have with the customers: the strengths and weaknesses suggested by the metrics become talking points. Another tactic Broadly uses is to focus on customers in the mid-range of customer engagement and risk: they don't try to help customers who haven't even logged in, but instead focus on customers showing signs of usage but below target usage levels. Broadly also found that one of the most important factors in influencing their customer retention was when a customer integrated Broadly with the company booking systems, so now the CSMs make sure to help with integrations whenever possible. Broadly also uses email campaigns to drive interest and adoption of new product features.

One of the most important aspects of Broadly's product is finding customers and persuading customers to review the business. As such, their customer data includes events for adding customers, asking customers to review, and determining whether customers review positively or negatively. More than 60 different events from the SaaS product were captured and the 10 most common are listed in Table 1.3.

Table 1.3   Most Common Broadly Customer Events

Transaction Added	Follow Up Email Sent
Ask Presented	Review Ask Decision
Customer Added	Customer Promoter
Thank You Email Sent	Ask Fulfilled
Affinity Updated	Page View (path: /add_customer)

1.10.3   Versature

Versature is disrupting the Canadian telecom industry with cloud-based unified communication solutions for businesses. Trusted by clients and partners across the country, Versature is an award-winning company that's raising the bar with superior, cost-effective technology and Canadian-based support. Businesses take advantage of Versature's unified communication packages with plans that depend on the number of users, volume of calls, and other online communications services. Most customers are subscribed to monthly renewing services.

The systematic study and fight against churn actually started at telecommunication companies back in the late 20th century, following deregulation. As a telecommunications provider born into the deregulated market, Versature has been focused on customer success from day one with an emphasis on churn prevention. Versature has experienced steady growth thanks to negative net churn year over year. The ability to identify at-risk customers early in their lifecycle has dramatically reduced controllable churn and allowed Versature's Customer Success Managers to have value-driven conversations before the customer reaches the point of no return.

Versature's service combines traditional telecommunications call features with additional digital features, such as a client-accessible administrative portal known as Sonar, a call data management offering known as Insights, and integration with digital products like Google Chrome and Salesforce. Their customer events are a combination of traditional telco calls with software events such as logins and page views. Versature's top 10 most common events are listed in Table 1.4.

Table 1.4   Most Common Versature Customer Events

Local Call	Sonar Login
Sonar Page View	International Call
Canada Call	Sonar Call Center Support
Toll Free Call	Conference Phone Number
US Call	Conference Call

1.10.4 *Churn case studies and privacy protection*

The companies described in this book have all generously allowed themselves to be profiled and have made their data available for demonstrating techniques in this book. However, these companies are all going concerns at the time of this writing and have a strategic need to protect information about their operations. For that reason, there will be many parts of the book where detailed information that might have been interesting to many readers is nevertheless withheld. In particular note:

- No churn rates of actual companies are reported in this book.
  - All numeric examples of churn calculation (like those in chapter 2) use randomly generated data for illustrative purposes.
  - All figures showing relationships with churn found through analysis (such as those in the next section and later chapters) will show relative and not actual churn rates.

Other types of information that will be obscured throughout the book include facts and figures that portray or can be used to derive information about the customer base or pricing whenever such information isn't already available through other public sources of information.

Also note: No personally identifiable information (PII) about the case study companies' customers was accessed at any point in the production of this book.

The books' emphasis is on anonymous behavioral analysis that doesn't depend on PII. Although certain PII, such as geographic information, may be useful in churn analysis (see chapter 10, Demographic and Firmographic), discussion of such use cases in this book are provided using simulated data for examples.

1.11 *Case studies in great customer metrics*

Before getting into technical details in chapter 2, I'm going to show you few examples of results from using the techniques in the book. I need to warn you that won't make findings like this immediately. You have to learn the techniques to prepare your data (chapters 2 through 4) before you start doing churn analysis in chapter 5.

As I mentioned in the discussion of events, the behaviors that are most closely related to the value delivered by the service are most important. But choosing the measurement to make is also crucial. Here are three metrics that I've found to be especially effective in the fight against churn:

- *Utilization*: Metrics that shows how much of the service the customer uses. If the service imposes limits on certain types of use, a utilization metric shows what percentage of the allowed amount the customer took advantage of.
- *Success*: Metrics that show how successful a user is in activities that have different outcomes.
- *Unit cost*: Metrics that relate the price the customer pays to the quantity of the service consumed or used.

Don't worry if you don't follow every detail in these case study examples—this is going to be a quick preview of what the rest of the book is about! The details are presented in later chapters.

1.11.1 *Utilization*

Introduced in the last section, Klipfolio is a data analytics cloud app for building and sharing real-time business dashboards. These dashboards can be created by multiple users, and a common metric for any product that allows multiple users on one subscription is the number of users that are active. Figure 1.2 is a demonstration of how the number of active users per month at a Klipfolio customer is related to churn.

Figure 1.2 uses a technique called *metric cohorts* to show the relationship between a behavior and churn. You'll see many of these plots in this book and learn how to create them, and for now I'll give a brief explanation of how it works. Given a pool of customers and a metric like the number of active users per month, the customers are organized into cohorts by their measurements on that metric. Typically, 10 cohorts are used, so the first cohort contains the bottom 10% of customers in terms of the metric, the second cohort contains the next 10%, on up to the final cohort which contains the top 10% of customers on that metric. Once the cohorts are formed, you calculate what percent of customers in each cohort churned. The result is displayed in a plot similar to figure 1.2: each point in the plot corresponds to one cohort, with the x-value of the point given by the average value of the metric for the customers in the cohort and the y-value of the point given by percentage of churns (that is, the churn rate) in the cohort.
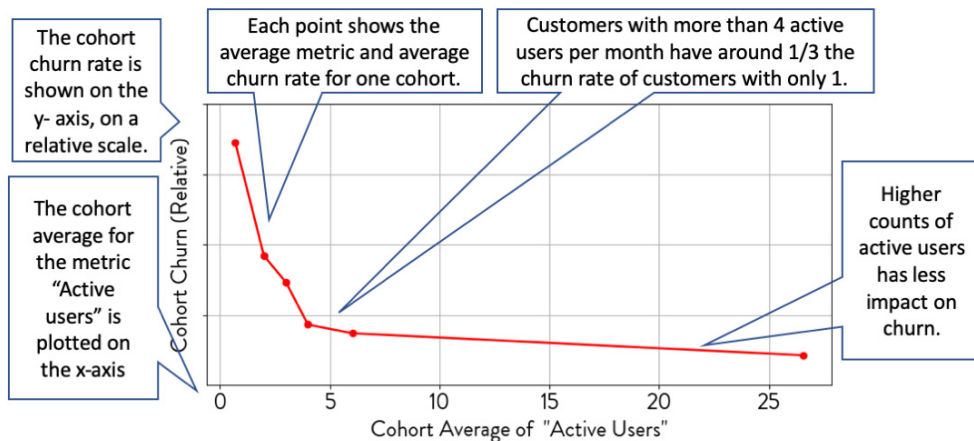

Figure 1.2   Klipfolio churn versus active users.

As mentioned in section 1.10.4 on privacy protections, the case study churn plots in this book don't show the actual churn rates, only the relative difference between the cohorts. However, the bottom of the metric cohort plots is always set to zero churn rate, so the distance of the points from the bottom of the cohort plot can be used for

comparing *relative* churn rates. For example, if one point is half as far from the bottom of the plot as another, that means the churn rate in that cohort is one half the other.

Turning to the details of figure 1.2 and what they mean, it shows that the lowest cohort has under one active user per month (an average over multiple months), and the highest cohort has an average of more than 25 active users per month. In terms of churn, the churn rate on the cohort with the lowest active users per month is approximately 8 times greater than churn in the cohort with the highest number of active users. At the same time, most of the differences in churn rates occurs between around one and five active users per month.

While measuring the number of active users is a good metric for fighting churn, an even better one is shown in figure 1.3. This is the license utilization metric calculated by dividing the number of active users by the number of seats the user has purchased. Many SaaS products are sold "by the seat" meaning the number of users allowed (this is called the licensed number of seats). If the number of active users is divided by the number of seats licensed, the resulting metric measures the percentage utilization of the seat license by the customer.

The result in figure 1.3 shows that license utilization is an effective metric for fighting churn: the lowest cohort in license utilization has an average utilization above zero, and the highest cohort has license utilization around 1.0. The lowest cohort has around seven times the churn rate as the highest cohort and the churn rate varies more or less continuously across the cohorts. In contrast to figure 1.2 (showing churn and the number of active users per month), there isn't a level at which having higher utilization no longer makes a difference. This makes license utilization more effective for understanding customer health than active users alone. As will be explained further in later chapters, the reason is that the active users per month conflates two different underlying factors related to churn: how many seats were sold to the customer, and how often a typical user is active. Utilization is a measure of how active the users are on a relative basis that's independent of the number of seats sold. License utilization is often useful for segmenting customers with respect to their engagement and churn risk.
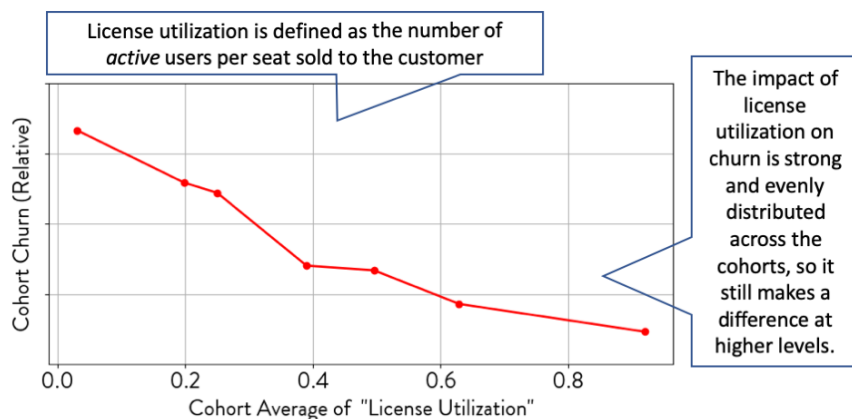

Figure 1.3   Klipfolio churn versus license utilization.

1.11.2 Success Rates

Introduced in the last section, Broadly is an online service that helps businesses manage their online presence including reviews. An important metric for Broadly's customers is the number of times the business is reviewed positively or promoted. Figure 1.4 shows the relationship between the number of promoters per month that a Broadly customer has and churn. In figure 1.4, the cohort with the fewest promoters per month (just above zero promoters on average) has a churn rate that's approximately four times higher than the cohorts with the most promoters; the reduction in churn mostly happens between 0 and 20 promoters per month. This is a clear relationship for an important event and it's easy to understand why customers who have promoters are more likely to stay with the Broadly service—receiving positive reviews is one of the main goals for a business using Broadly!
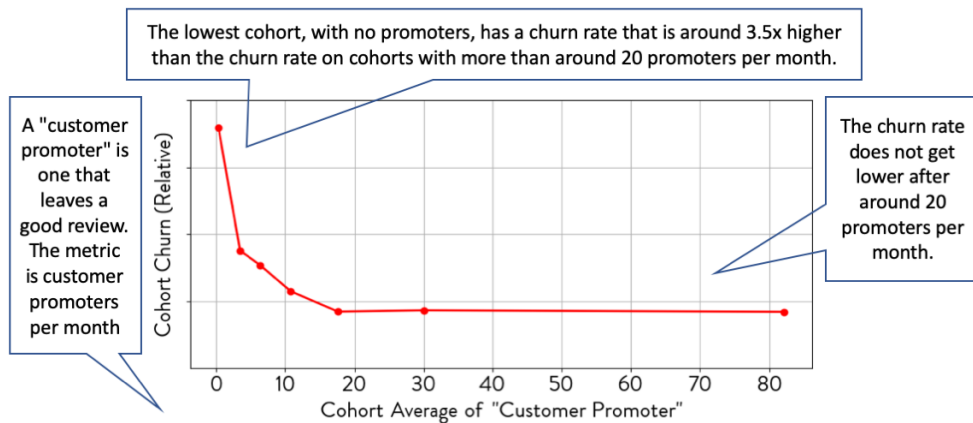

The lowest cohort, with no promoters, has a churn rate that is around 3.5x higher than the churn rate on cohorts with more than around 20 promoters per month.

A "customer promoter" is one that leaves a good review. The metric is customer promoters per month

The churn rate does not get lower after around 20 promoters per month.

Figure 1.4　Broadly churn vs customer promoters.

Another important event for Broadly's customers related to the number of promoters is the number of detractors or the number of times the business is reviewed negatively. Figure 1.5 shows the relationship between the number of detractors per month that a Broadly customer has and churns. The cohort with the fewest detractors per month (just above zero) has a churn rate that is approximately two times higher than the cohorts with the most detractors (average of under five detractors per month); the reduction in churn mostly happens between 0 and 1 detractors per month.

While this relationship looks much like the one for customer promoters shown in figure 1.4, doesn't it seem like something is wrong here? Getting negative reviews is a bad thing, and presumably not the result that Broadly's customers were looking for, so why is having negative reviews associated with reduced churn?
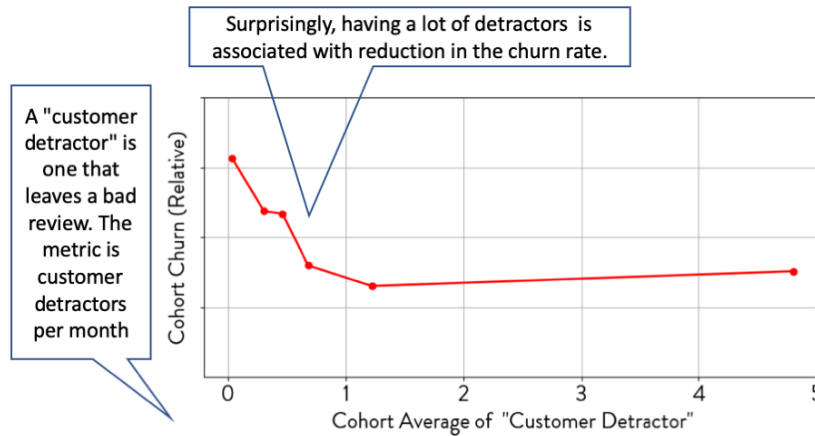
A "customer detractor" is one that leaves a bad review. The metric is customer detractors per month

Surprisingly, having a lot of detractors is associated with reduction in the churn rate.

Figure 1.5 Broadly churn vs customer detractors.

To understand why more of a bad thing like detractors may be associated with *less* churn, it helps to look at another, better metric for Broadly's customers. If you take the number of detractors and divide it by the total number of reviews (promoters plus detractors), then the result is the percentage of detractors, which I call the Detractor Rate. Figure 1.6 shows the relationship between churn and the detractor rate—this is probably more the kind of relationship you were expecting for a product event that's negative for the customer. The higher the detractor rate, the higher the churn, and in a significant way.

Why does the relationship to churn show that more detractors is good when you look at detractor count in figure 1.5, and that more detractors is bad when you look at the detractor rate in figure 1.6? The answer is that the *total* number of detractors in figure 1.5 is related to the total number of promoters shown in figure 1.4, because Broadly customers who receive many reviews overall are likely to receive more of both good and bad reviews because they have more reviews. When you look at the impact on the relationship between the number of detractors and churn in the simple way of figure 1.5, it conflates two underlying factors driving the metric: having many reviews (which is good), and having a high proportion of bad reviews (which is bad). When the proportion of bad reviews is analyzed alone, you get the more useful result shown in figure 1.6. This illustrates why success and failure rates can be so effective for understanding churn.

1.11.3 *Unit cost*

Introduced in the last section, Versature provides telecommunication services for businesses. As a unified communications provider, many of their most important events are voice calls that have a duration stored in a field attached to each event. Figure 1.7 shows the relationship between the total time spent on voice calls and churn
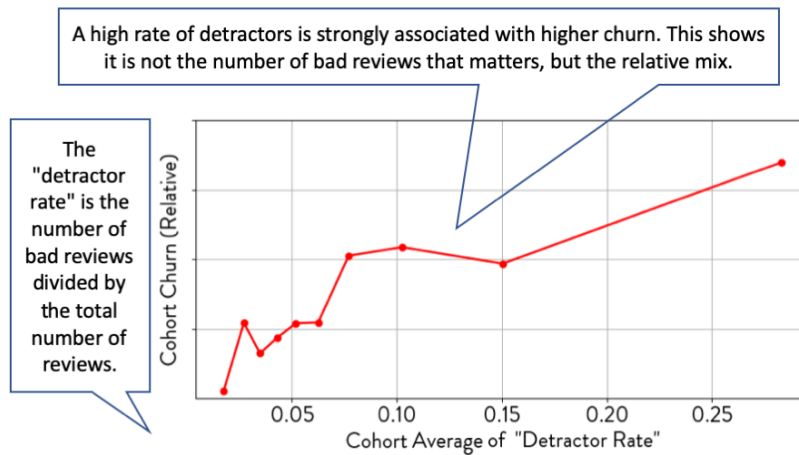
Figure 1.6   Broadly Churn vs detractor rate.

for Versature customers. The lowest cohort in terms of local calls has practically zero calls and a churn rate that's approximately three times higher than those cohorts of customers that have a local call time per month that's in the thousands.
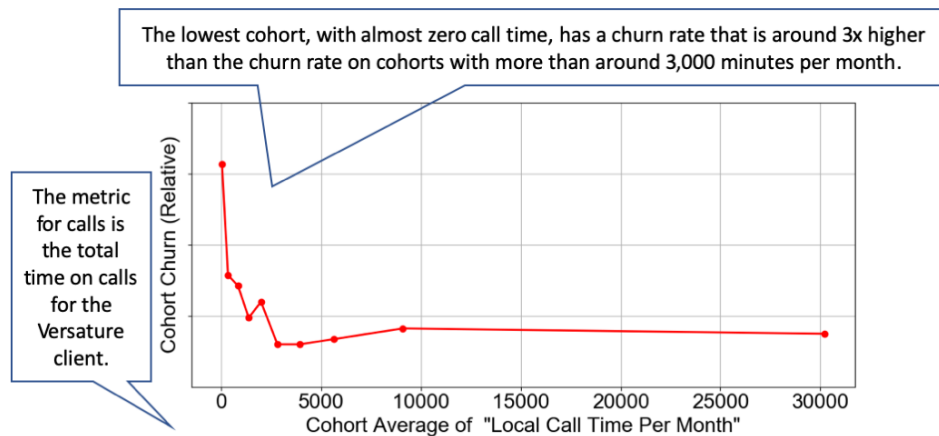

Figure 1.7   Versature churn vs local calls.

When trying to understand churn, it's important to consider not only the amount of a service customers use, but also how much they pay. Monthly Recurring Revenue (MRR) is a standard metric for calculating the amount a customer pays to use a subscription service: it's the recurring amount a customer pays each month to use a service, but not including any setup fees or irregular charges. (I'll say more about MRR and how to calculate it in chapter 3.) The amount customers pay can also be analyzed

with a metric cohort approach to look for a relationship with churn, which is shown in figure 1.8 for Versature.

The metric cohort plot in figure 1.8 does something new: rather than displaying the average MRR of the cohorts directly, it shows the average after every MRR measurement has been converted to a *score*. If you're familiar with the concept of "grading on a curve" then metric scores are the same idea: the measurements are converted from one scale to another, but the ordering remains the same. Given cohort like the bottom 10% on the metric is still the same set of customers if the metric is converted into a score, and the cohort has the exact same churn rate. Converting a metric into a score only affects how the cohorts are placed along the horizontal axis of the cohort plot, but not the vertical position of the points, which is the churn rate. Metrics are converted to scores when the re-scaling on the horizontal access makes the result easier to understand. I'll say more about metric scores and teach you how to calculate them in chapter 3.
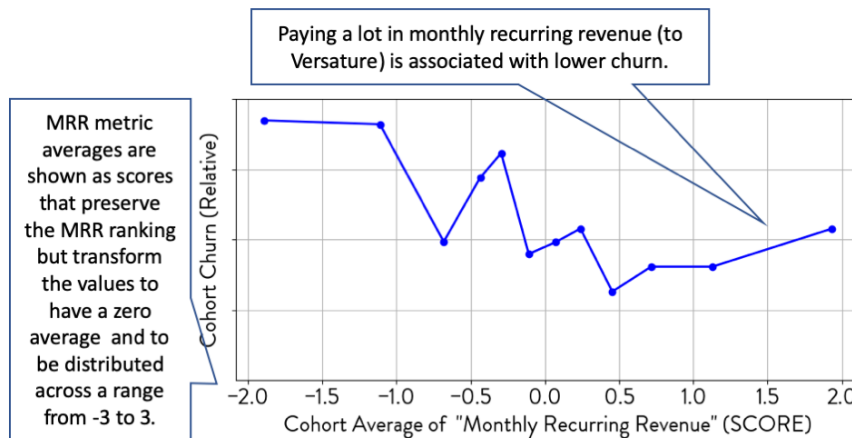

Figure 1.8 Versature churn vs monthly recurring revenue (price) scores.

The cohort churn rates in figure 1.8 show that MRR is also related to churn, though not as strongly as making calls. The churn rates in the different cohorts don't vary in a totally consistent way, and the lowest cohort churn rates are only about one half or one third less than the highest churn cohorts. But this is another case where it makes you stop and think about what it's saying: people who pay more churn less. Is that what you expected? This may be surprising, but it's quite common, especially in business products. That's because business products are sold with higher prices for bigger customers and bigger customers churn less for interrelated reasons. First, they have more employees. When it comes to product use like making calls or using software, customers who pay more for a product are often using it more too. The lower churn

for customers paying higher MRR shown in figure 1.8 is related to the lower churn for customers with more calls shown in figure 1.7.

Figure 1.9 shows a different metric for looking at how the amount customers pay relates to churn. The MRR metric is divided by the metric for the number of calls per month. This results in a metric that's the cost per call the customer makes. I call this a *unit cost metric*, because it explains how much of the service the customer receives for their money. Like in figure 1.8, figure 1.9 shows the cohort average as a score rather than in dollars. The metric cohort churn plot for cost per call shows that customers that pay more really do churn more, when the payment is measured in relation to the amount of the service used. The highest cohort in cost per call has a churn rate that's approximately six times higher than the cohorts with the lowest cost per call. Value metrics like this are a key weapon for understanding why customers churn and an important subject that will be explored fully in later chapters.
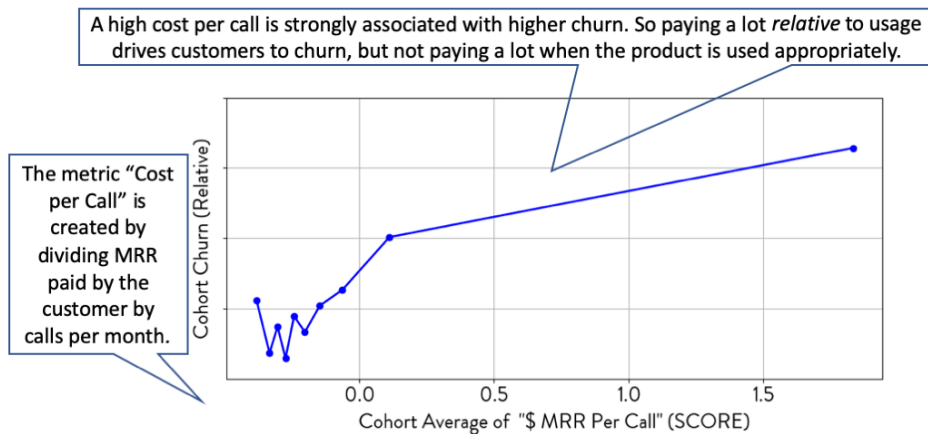

Figure 1.9   Versature churn vs monthly price per call.

Summary

- The term churn arose in the context of subscription products and means users quitting or canceling the service.
- Churn also applies to all products and services where customers or users repeatedly interact with the product over long periods of time, whether or not a formal subscription or any form of payment exists.
- Recent years have seen an explosion in the number of discretionary online services for consumers and businesses. The discretionary nature of these services means that churn is a constant problem that needs to be addressed.
- People skilled in data analysis are frequently called in to help understand what causes churn and what can be done to reduce it.

- Online products and services track a wide variety of interactions of the users with the service, generically referred to as events. The history of such events is one of the primary sources of data for fighting churn.
- Churn isn't a situation where a predictive model alone helps to achieve the goal (churn reduction) because churn reducing interventions depend on knowing the causes for churn.
- To help reduce churn, the data person should create the best customer metrics possible.
- Great customer metrics have these characteristics:
    - Easily understood by the business.
    - Clearly associated with churn and retention, so it's obvious what a healthy customer looks like.
    - Segments customers in a way useful for targeted interventions that increase engagement.
    - Useful to multiple different functions of the business (that is, product, marketing, support, and so on).