Fighting Churn With Data

Carl Gold, PhD
Chief Data Scientist @
zuora

www.fightchurnwithdata.com

zuora: The leading Subscription Management platform

www.fightchurnwithdata.com

Customer Case Studies

Broadly ensures that your business looks great online, and is found and chosen by potential customers.

Klipfolio is a data analytics cloud app for building and sharing real-time business dashboards.

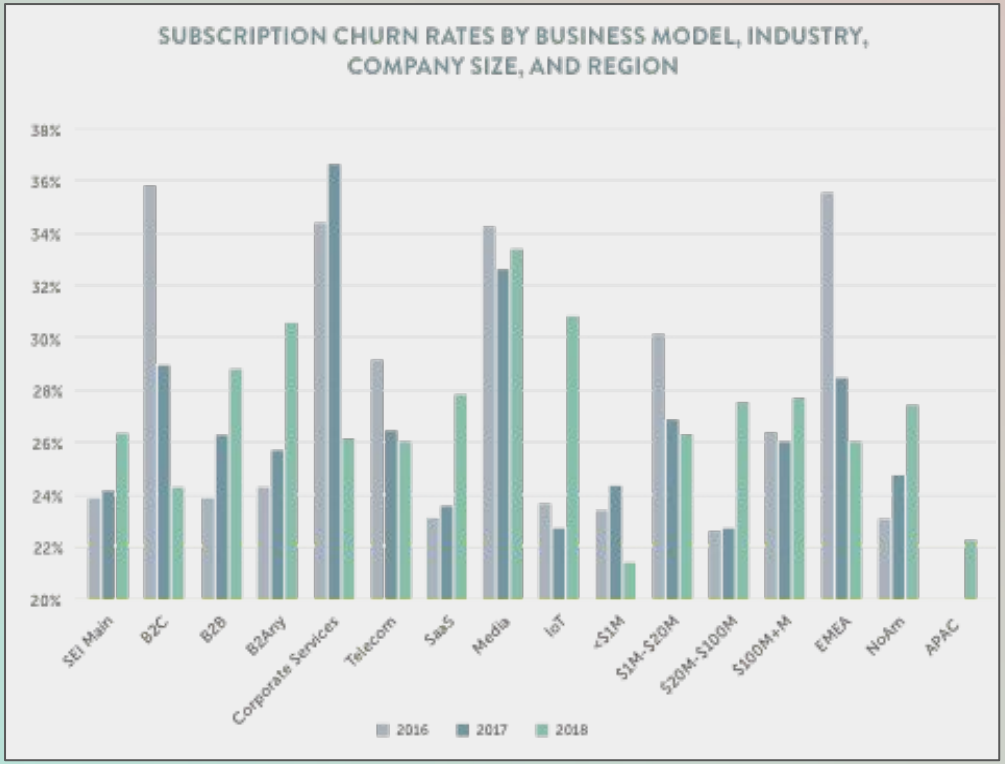
Versature is disrupting the Canadian telecom industry with Cloud-based business communication solutions.

What is Churn?

- Churn = cancellation of subscriptions
  - on a subscription product
- Generally: users quitting or leaving any product or service when you don't want them to
- The term originated from "Churn rate"
  - Proportion of customers quitting in a time period
- But now it is also:
  - A verb : "The customer churned"
  - A noun : "Make a list of all the churns last month"

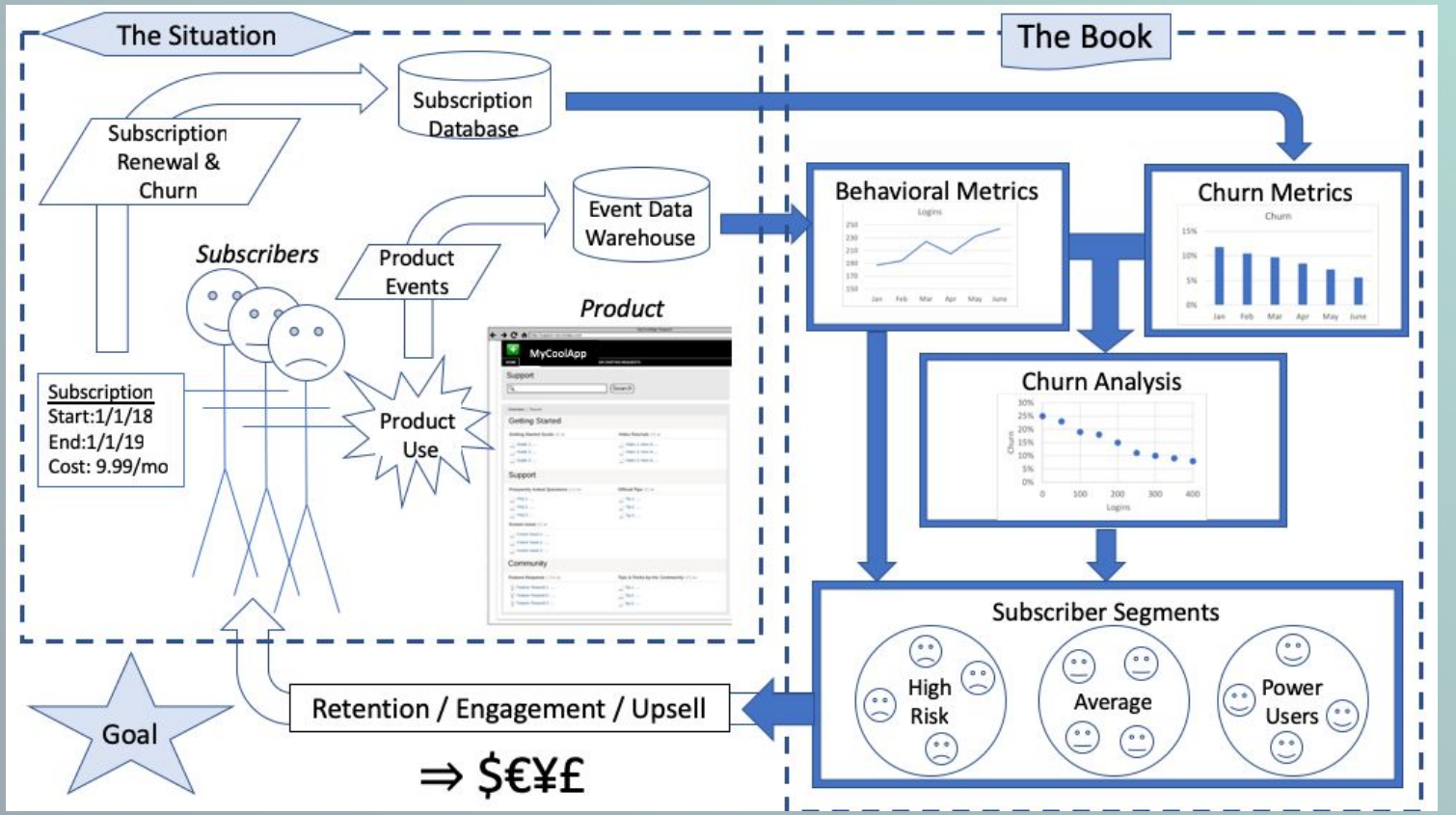
Churn Rates

www.zuora.com/resource/subscription-economy-index/



www.fightchurnwithdata.com

What is Fighting Churn with Data About?



www.fightchurnwithdata.com

Why churn is hard to fight...

1. HARD TO PREDICT | 2. HARDER TO PREVENT | 3. THE BUSINESS

1. Churn is hard to predict

- Important information is usually out of reach:
  - Ability to pay
  - Subjective Utility
  - Alternatives & Switching Cost
- Even when churn is obvious...
  - Timing is unpredictable
  - Depends on external factors

Why churn is hard to fight...

1. HARD TO PREDICT | 2. HARDER TO PREVENT | 3. THE BUSINESS

2. Churn is harder to prevent

- These people already know the product
- To reduce churn significantly:
  - ***You have to actually deliver more value (utility)***
- There are no "silver bullets"
  - Churn is a lead bullet situation
- Discounting is *not* a churn mitigation strategy

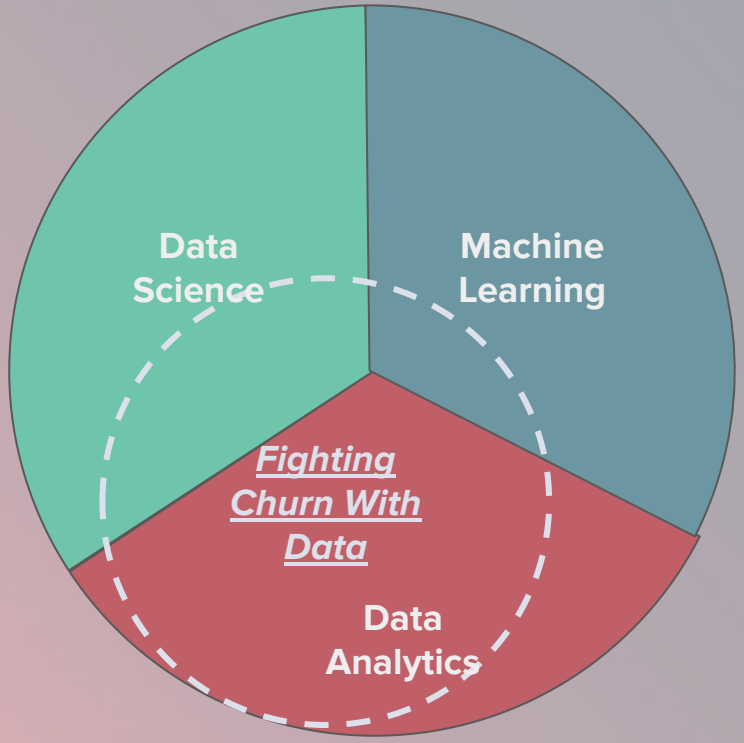
Why churn is hard to fight...

1. HARD TO PREDICT | 2. HARDER TO PREVENT | 3. THE BUSINESS

3. Preventing Churn is Owned by the Business

1. Product Creators
    - Make a more engaging, stickier product
2. Marketers
    - Engagement & Education campaigns
3. Customer Success & Support
    - Proactive & Reactive 1:1 interventions
4. Account Managers
    - Right Size Price/Plan

The role of data...



1. Design behavioral metrics
2. Test hypotheses
3. Explain the results
4. Help design segments
    - Maybe predict churn
5. Help monitor effectiveness

Metric Design (AKA Feature Engineering)

Your not so secret weapon:

Prove Interpretable Hypotheses

So the business gets the knowledge they need to act, and believes in it

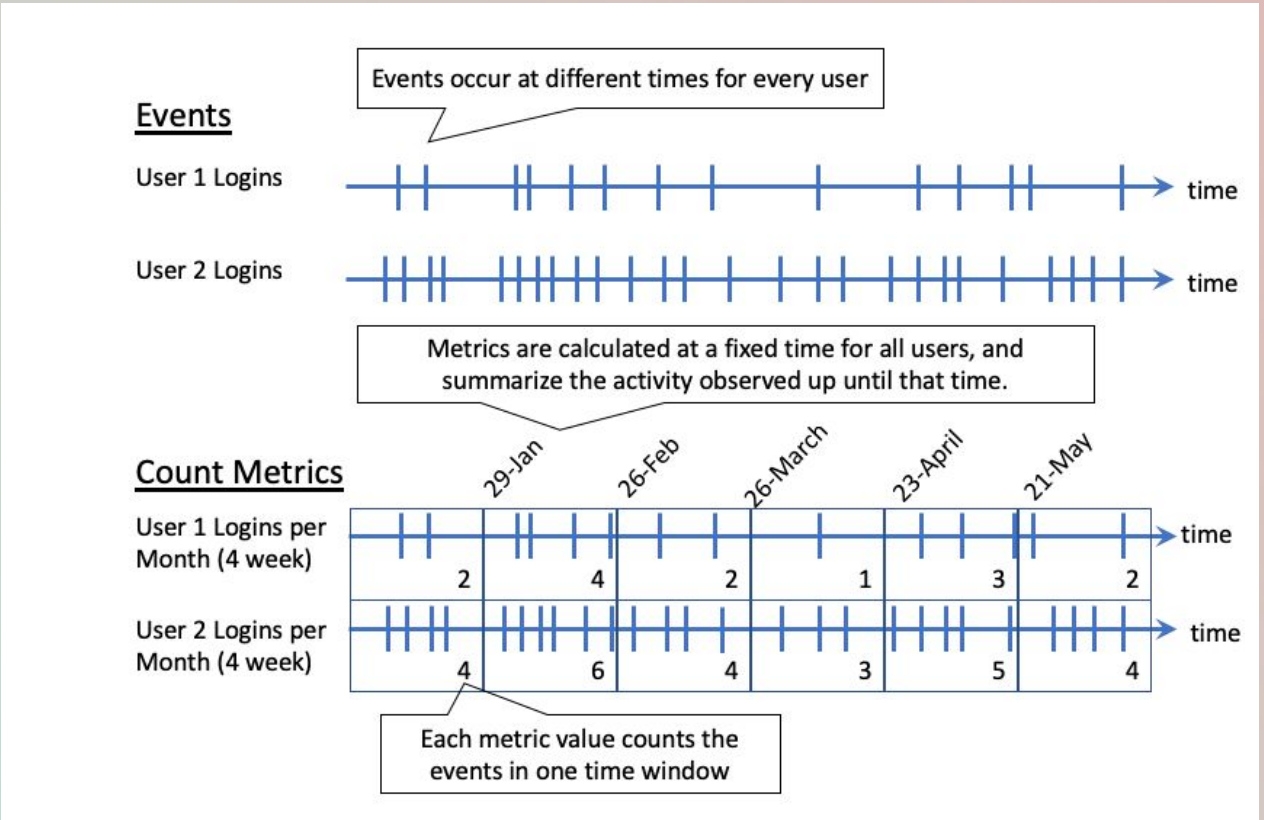
Perform Dimension Reduction

That increases business insight rather than confusion

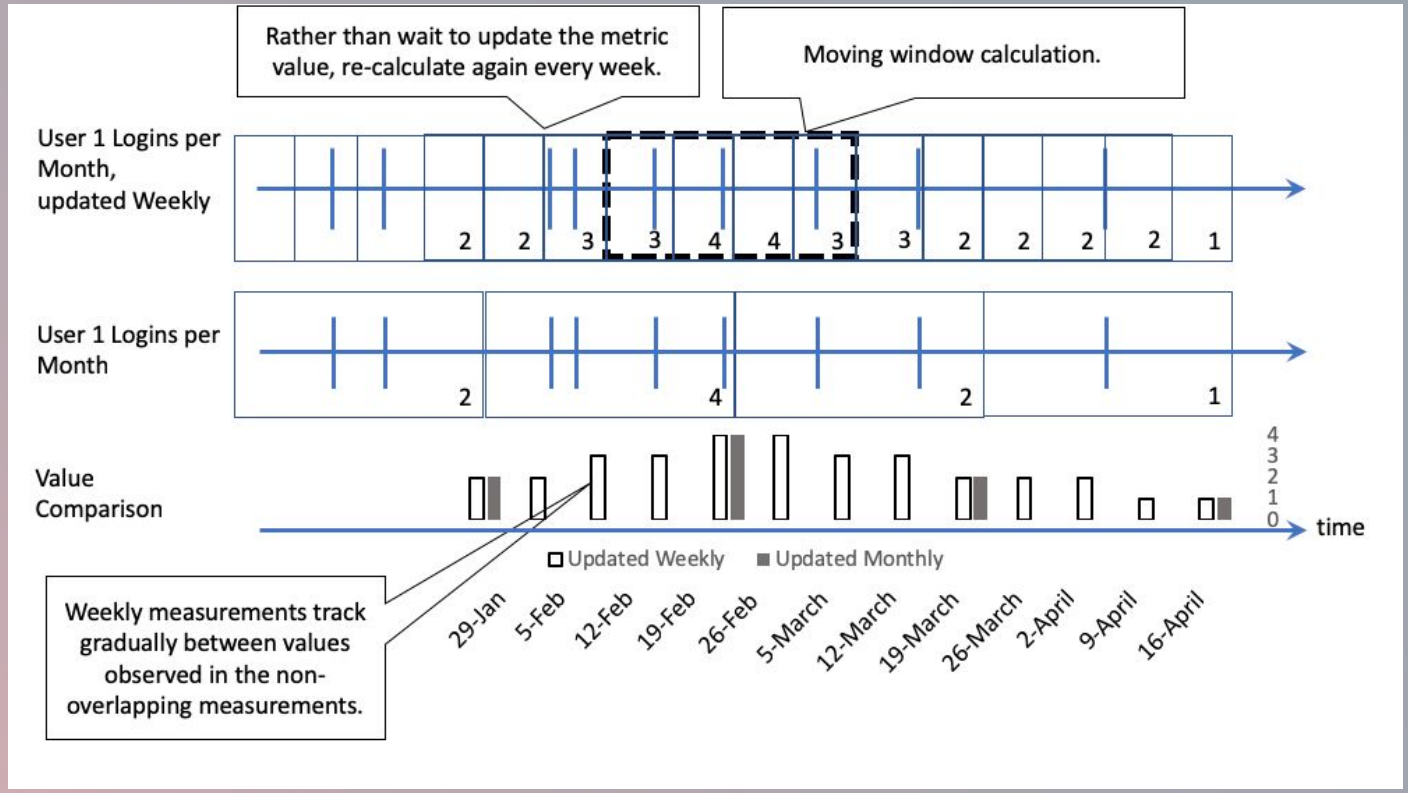
Accurately Predict with Any Model

Including interpretable linear models

Basic Count Metrics

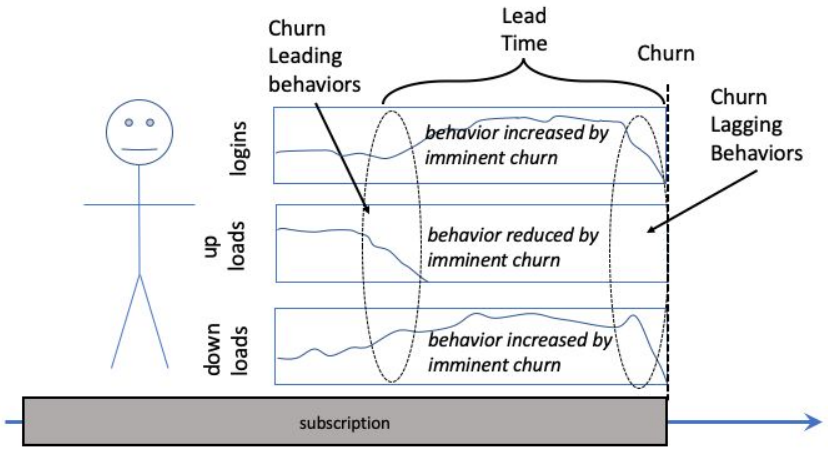


Staggered Metric Calculations



Data Set Formation

Form a Dat Set by compiling metric observations *in advance of* <u>both</u> Churn and Renewal events...

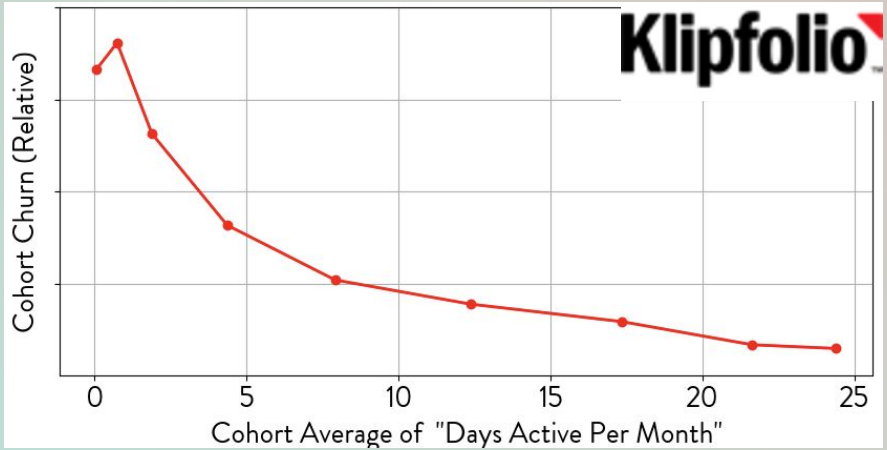


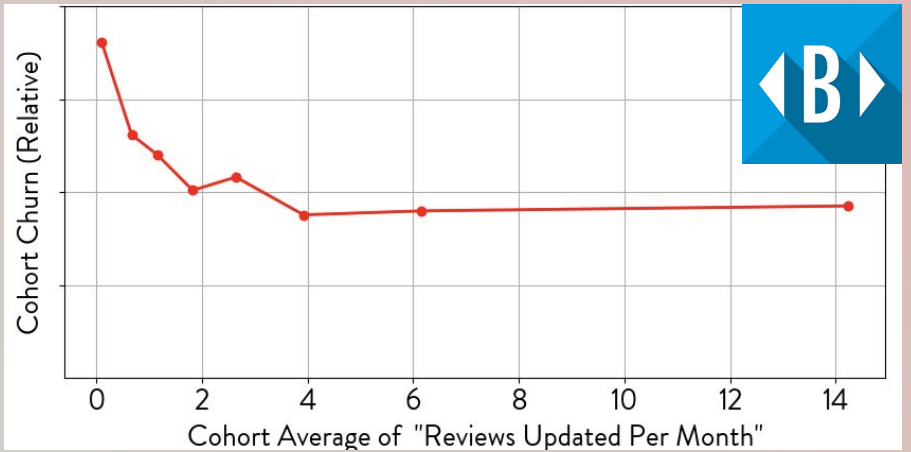
Account	Metric	Value
A	logins_per_month	5
A	posts_per_month	2
A	likes_per_month	1
A	shares_per_month	4
B	logins_per_month	8
B	...	...

Subscription Database

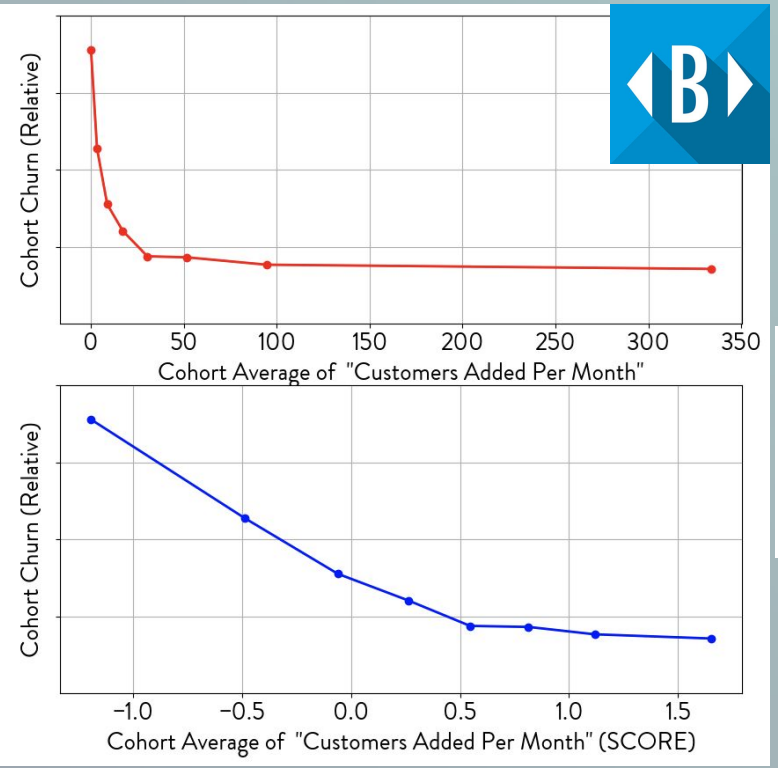
Account	logins_per_month	posts_per_month	likes_per_month	shares_per_month	is_churn
A	5	2	1	4	False
B	8	...			True
...	...				...

Behavioral Cohorts & Churn





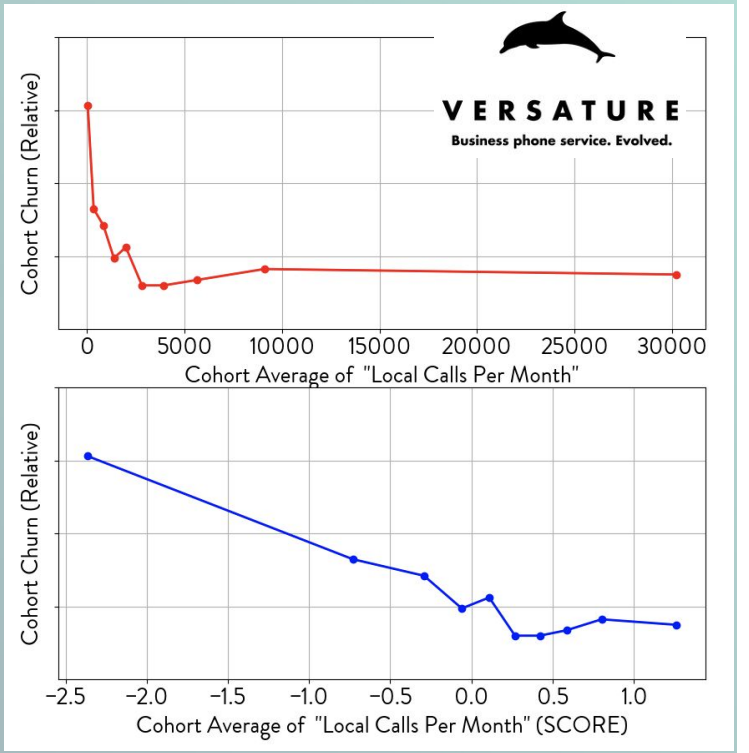
Scoring Skewed Behavioral Cohorts



Log Scale Scoring:

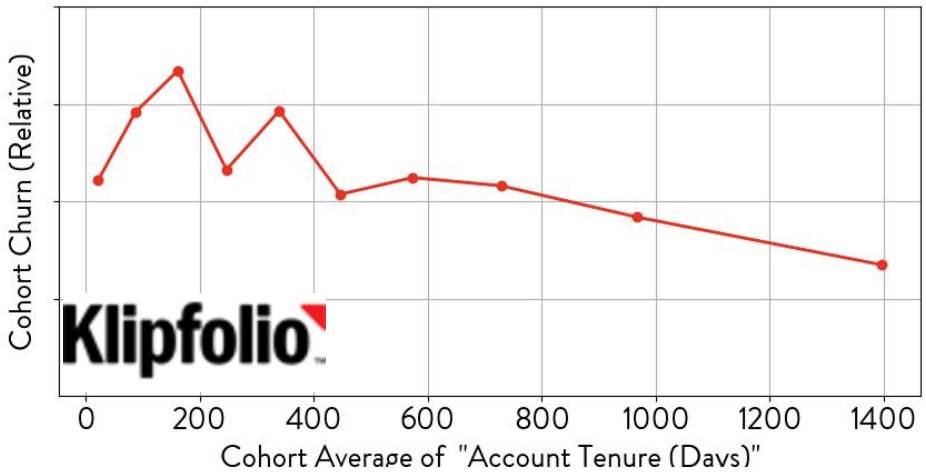
$$l = \log(1 + m)$$

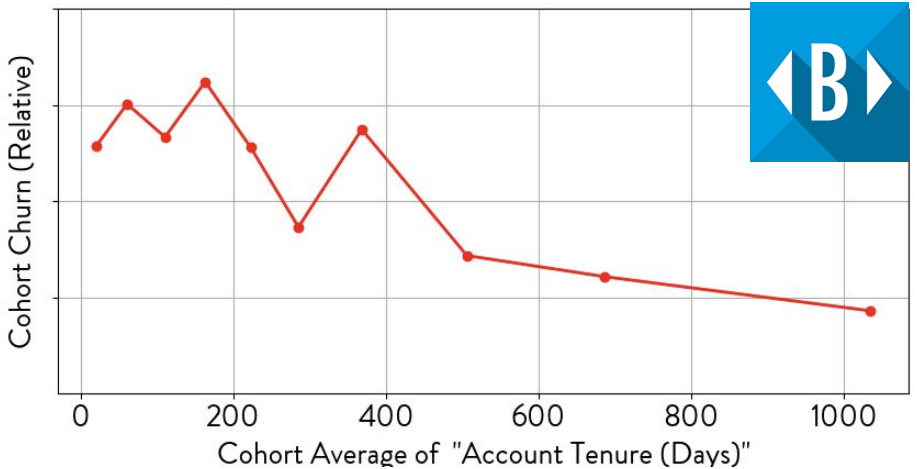
$$s = \frac{l - \bar{l}}{\sigma_l}$$



Account Tenure ("Age" on the Product)

- **Tenure is a standard for churn cohort analysis**
  - Calculate it as an account metric for unified analysis



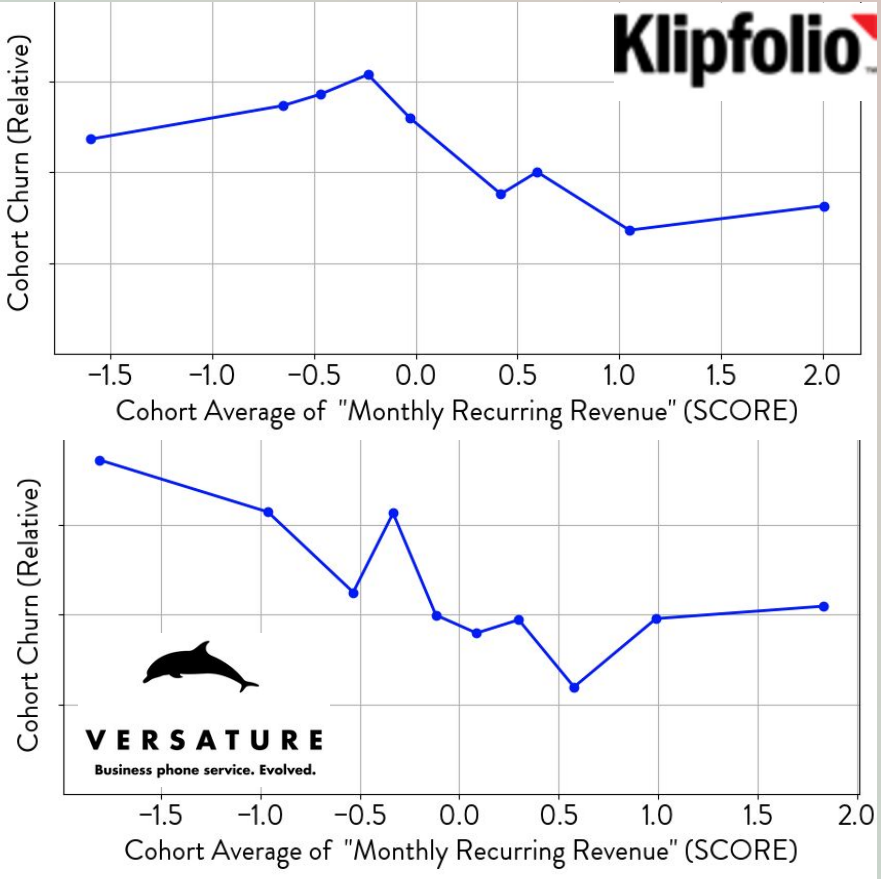


Monthly Recurring Revenue

- **MRR = Monthly Recurring Revenue**
  - A slowly changing dimension
  - Calculate it as an account metric for unified analysis

- **Question: Does paying more cause people to churn?**

Monthly Recurring Revenue and Churn



- Usually those who pay more churn less
- "Involuntary churn" = Churn by those who want to pay but can't
- Involuntary churn is less common among those paying more
- But it does not entirely explain Churn vs. MRR

Correlation in Churn Analysis

Many behaviors related to churn are correlated.

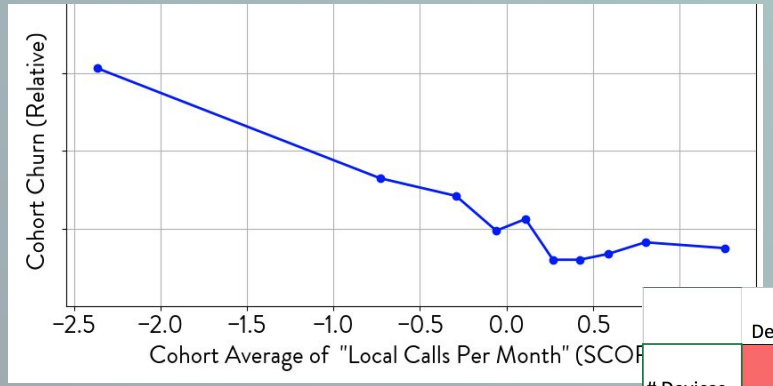
- Monthly Recurring Revenue
- # Devices
- Local Calls
- Domestic Calls

	# Devices	Domestic Calls	Local Calls	MRR
# Devices		48%	40%	76%
Domestic Calls	48%		75%	62%
Local Calls	40%	75%		54%
MRR	76%	62%	54%	

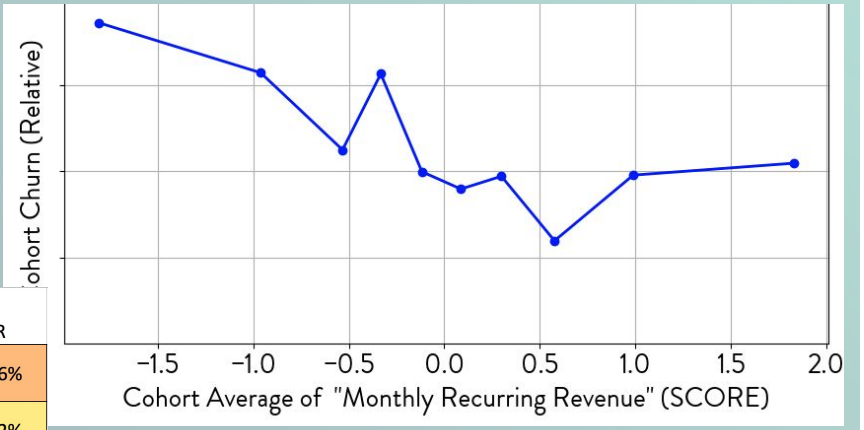
VERSATURE
Business phone service. Evolved.

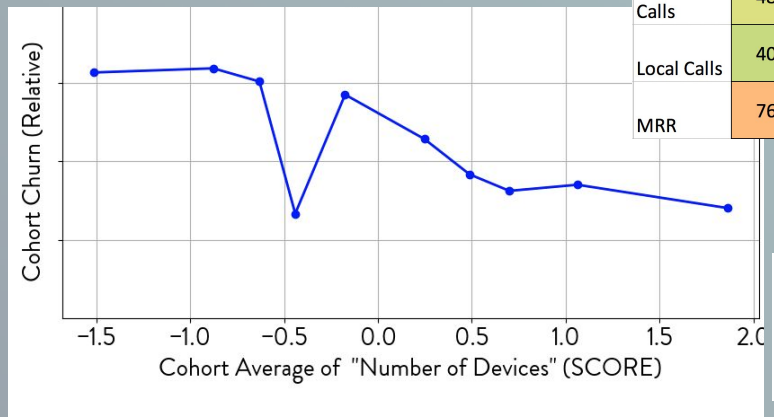
www.fightchurnwithdata.com

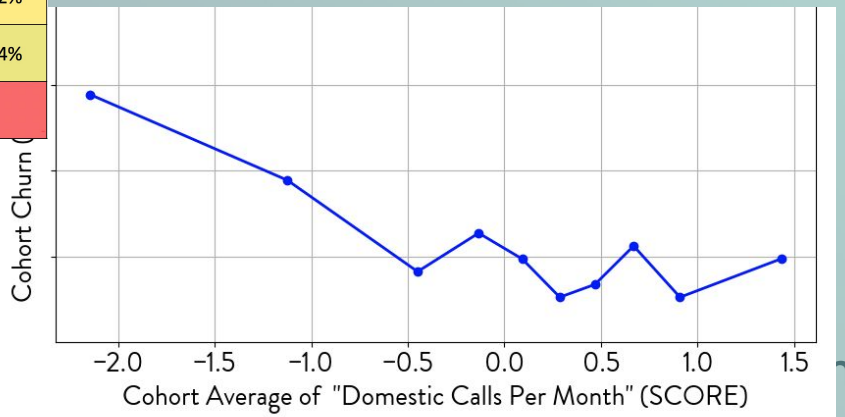
Churn & Correlated Behaviors



	# Devices	Domestic Calls	Local Calls	MRR
# Devices		48%	40%	76%
Domestic Calls	48%		75%	62%
Local Calls	40%	75%		54%
MRR	76%	62%	54%	



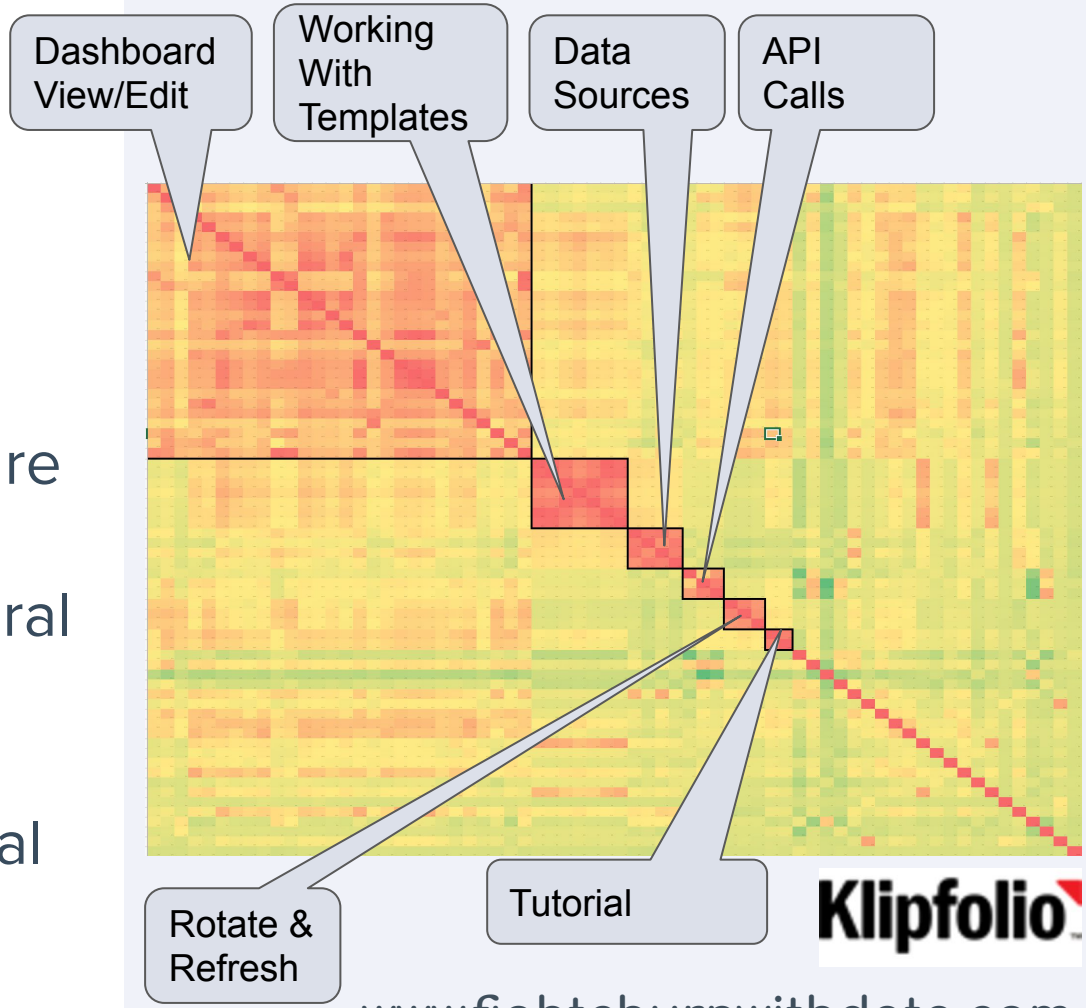




VERSATURE
Business phone service. Evolved.

Typical SaaS Behavioral Metric Correlations

- Many software features are used in tandem
- As a result many behavioral metrics for SaaS will be highly correlated
- Groups relate to functional areas of the product



Hierarchical (Agglomerative) Clustering

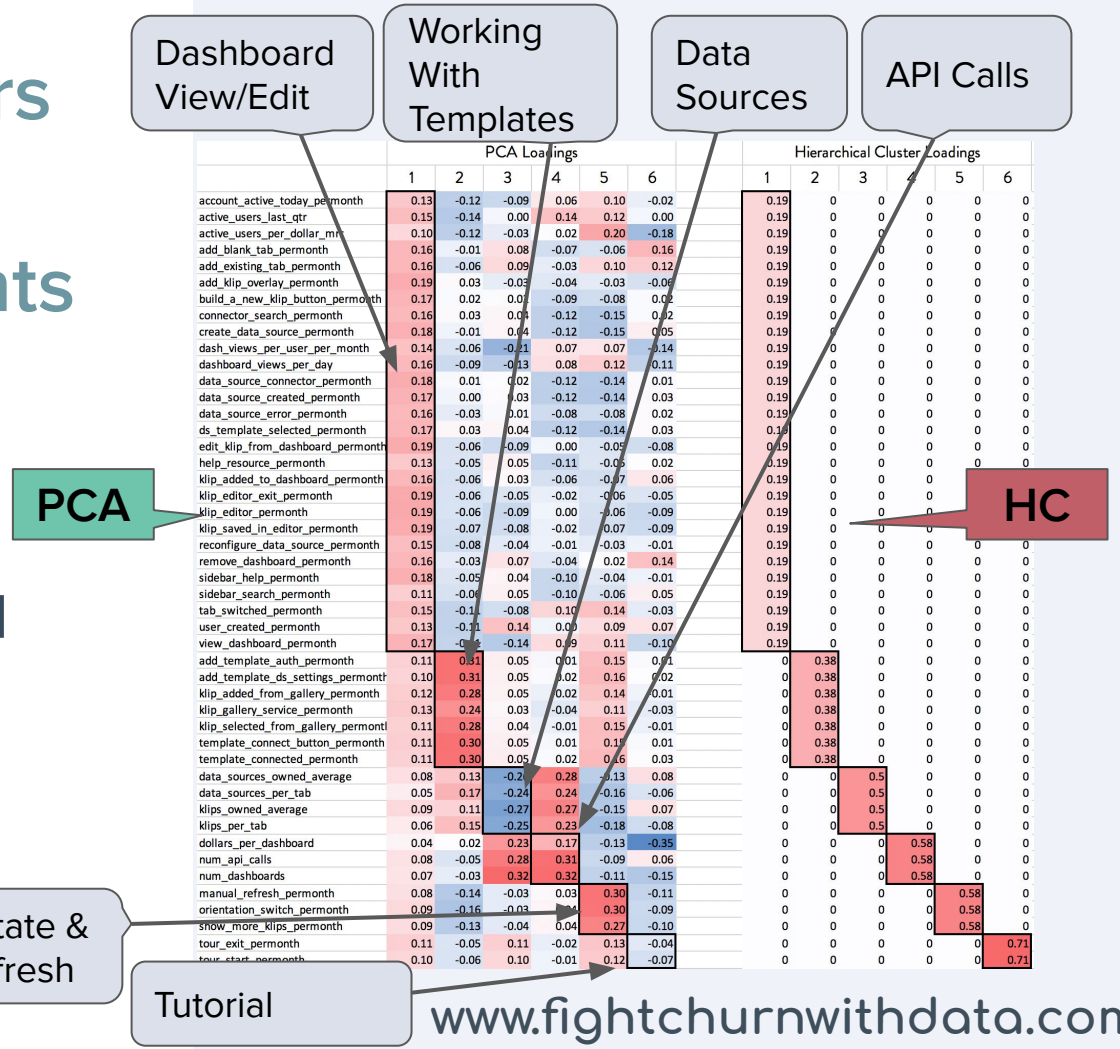
- **Dimension reduction is hard to explain**
- **Hierarchical Clusters are Understandable By The Business**

Algorithm:

1. **Merge two most correlated metrics by weighted average**
   - Merge operates on *Scores*, not un-normalized metrics
   - Sum of squares weighting preserves variance
2. **Re-Calculate Correlations**
3. **Repeat**
   - Until Remaining Correlations are below threshold, or Achieve a target number of groups

Hierarchical Clusters vs. Principal Components

- The clusters from HC capture similar groupings of correlated variance as PCA

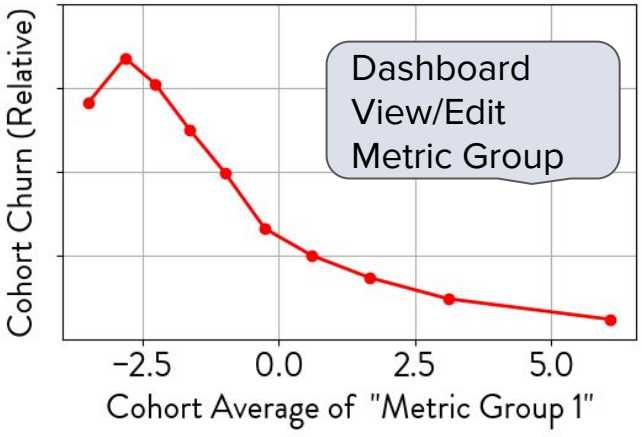


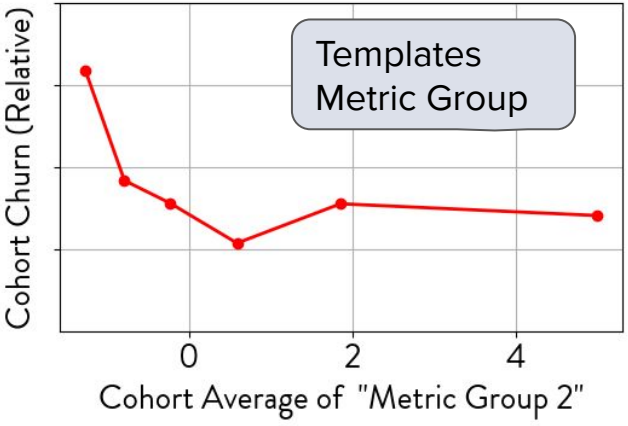
Klipfolio

www.fightchurnwithdata.com

Dimension Reduction For the Business

- Prepared, business people generally accept averages of scored metrics in this context
  - Name the groups intuitively
  - Show the Business people the heatmap
  - Do not mention "loadings", sum-of-square weights





Klipfolio

What about the Differences?

- PCA captures information about the relative values (differences) between underlying metrics
- Simple hierarchical clusters do not
- How can this information be captured in a way that is understandable?
- Take a page from the Wall Street playbook...

Company Analysis (Finance)

Many measures of a company:

1. **Share Price**
2. **Earnings**
3. **Dividends**
4. **Number of Shares**
5. **Value of assets and debts**
6. **Market Capitalization**

These measures are generally *correlated* in the following sense:

- **Big/successful companies have big numbers on all of them**
- **Small companies have small numbers**
- **All metrics scale with the size/success of the company being measured**

Stock Ratio Metrics

1. EPS = Earnings *per share*
2. P/E = Price *divided by earnings* (per share)
3. Dividend Yield = Dividend *divided by Price*
4. Book Values *per Share* = Total Assets / # of Shares
5. etc.

- These ratios make stocks of different size companies comparable
  - Cheap or expensive : Look at P/E, not price alone
- Divide one thing that scales with size by another
  - The result is less correlated with the underlying metrics

Intuitiveness of Ratios

- Ratios are very easy to for humans understand
  - Success Rate (Successes / Attempt)
  - Miles per Hour (Miles / hours)
  - $ per Gallon (gas prices)
  - Users per Seat (AKA License Utilization)
- Statistical multiplicative interactions are usually unintuitive
  - "Mile hour" (of miles * hours)
  - "Gallon dollar" (gallons * $)
  - "User Seat" (users * seats)

Key Ratios for Churn

Utilization

Amount used of a budgeted resource

Efficiency

Completion or Success rate on activities

Value

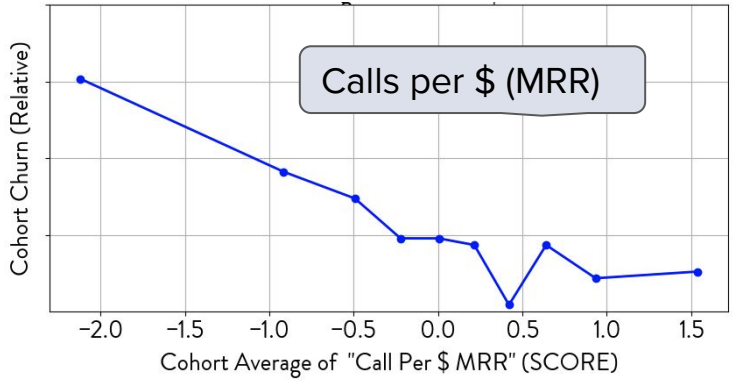
Cost / Use

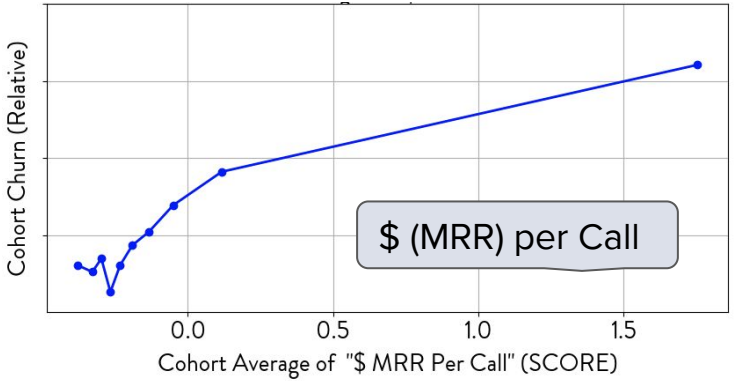
or

Use / Cost

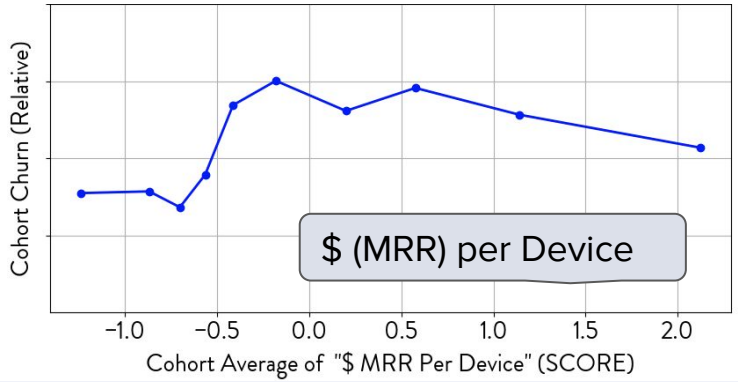
Value

VERSATURE
Business phone service. Evolved.

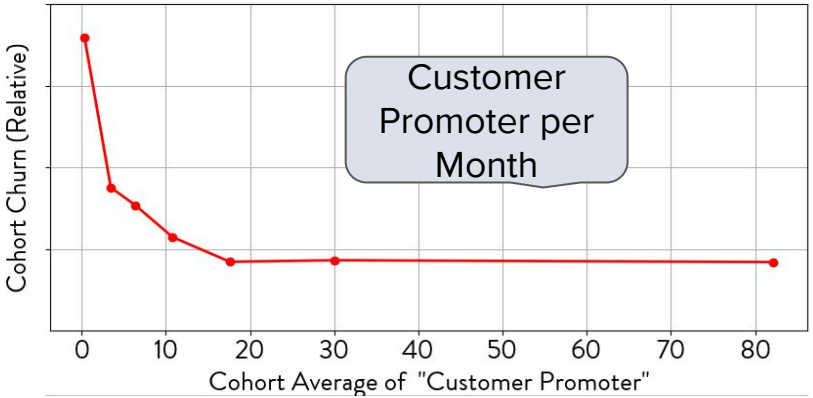


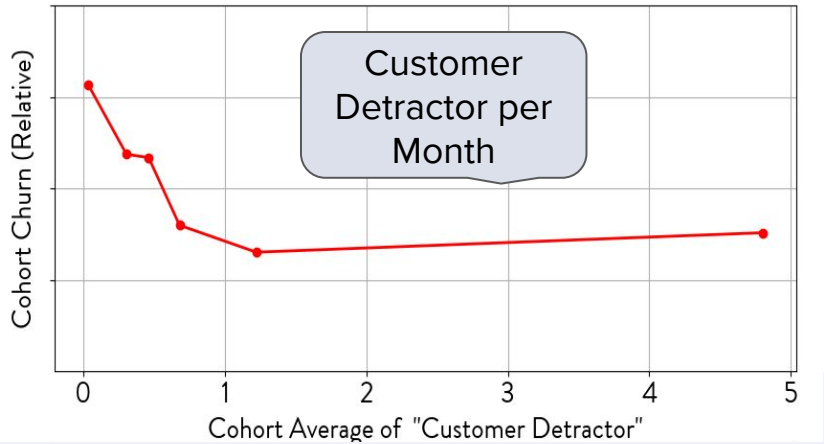


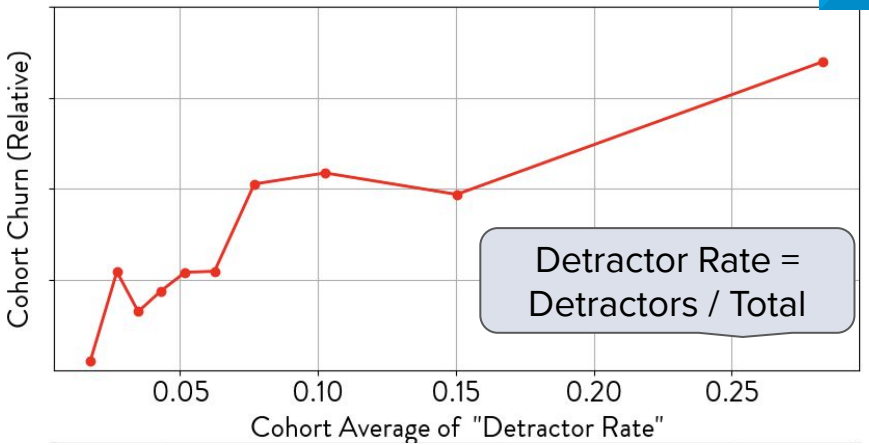
# Devices						
MRR (Monthly Recurring Revenue)	76%					
Total Calls Per Month	43%	57%				
Calls Per Device	25%	20%	65%			
Calls Per $ MRR	13%	17%	83%	65%		
$ MRR Per Call	-12%	-14%	-39%	-31%	-47%	
$ MRR Per Device	-25%	10%	6%	43%	0%	-2%
	# Devices	MRR	Calls Per Month	Calls Per Device	Calls Per $ MRR	$ MRR Per Call



Success / Failure

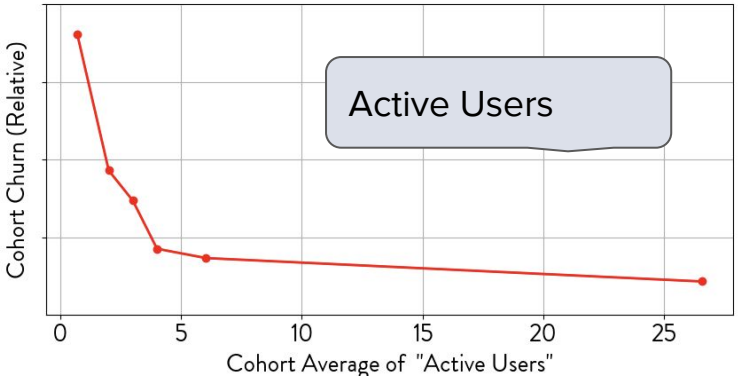


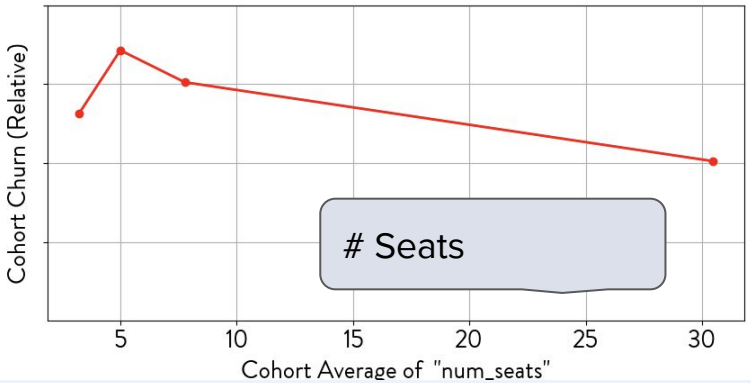


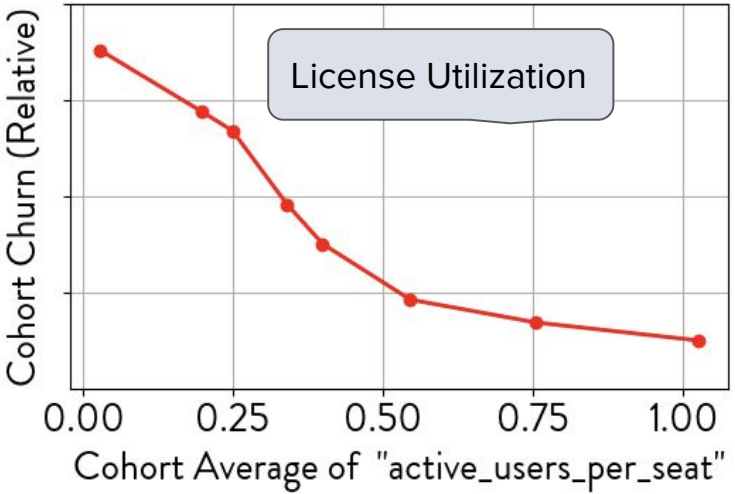


	Customer Promoter	Customer Detractor	Detractor Rate
Customer Promoter			
Customer Detractor	51%		
Detractor Rate	-56%	44%	
Promoter Rate	56%	-51%	-100%

Utilization







	Active Users	# Seats	License Utilization
Active Users			
# Seats	58%		
License Utilization	9%	28%	

Summary

- Fighting churn is not easy and requires data people to provide insight and understanding to the business
- Well designed metrics (features) allow you to effectively analyze and predict churn in an interpretable way
- Pro Tip: Use Ratios of simple metrics
    - Interpretable as Efficiency, Utilization & Value
    - Reveals interactions between correlated metrics without complex dimension reduction

THANK YOU!

Book available for early online access beginning in June


carl.gold@zuora.com


www.linkedin.com/in/carlgold/


@carl24k


github.com/carl24k/fight-churn

www.fightchurnwithdata.com

Things I don't have time to tell you about...

- How to calculate the appropriate churn rate measurements
- More advanced metric tricks
  - Percents of a total
  - Measuring change over time
  - Scaling metric measurements for new accounts
- How to prepare & QA your data for churn analysis
- Pitfalls of churn data set construction
- How to measure predictive model accuracy for churn
- How different predictive models compare
- Calculating customer lifetime value from churn predictions