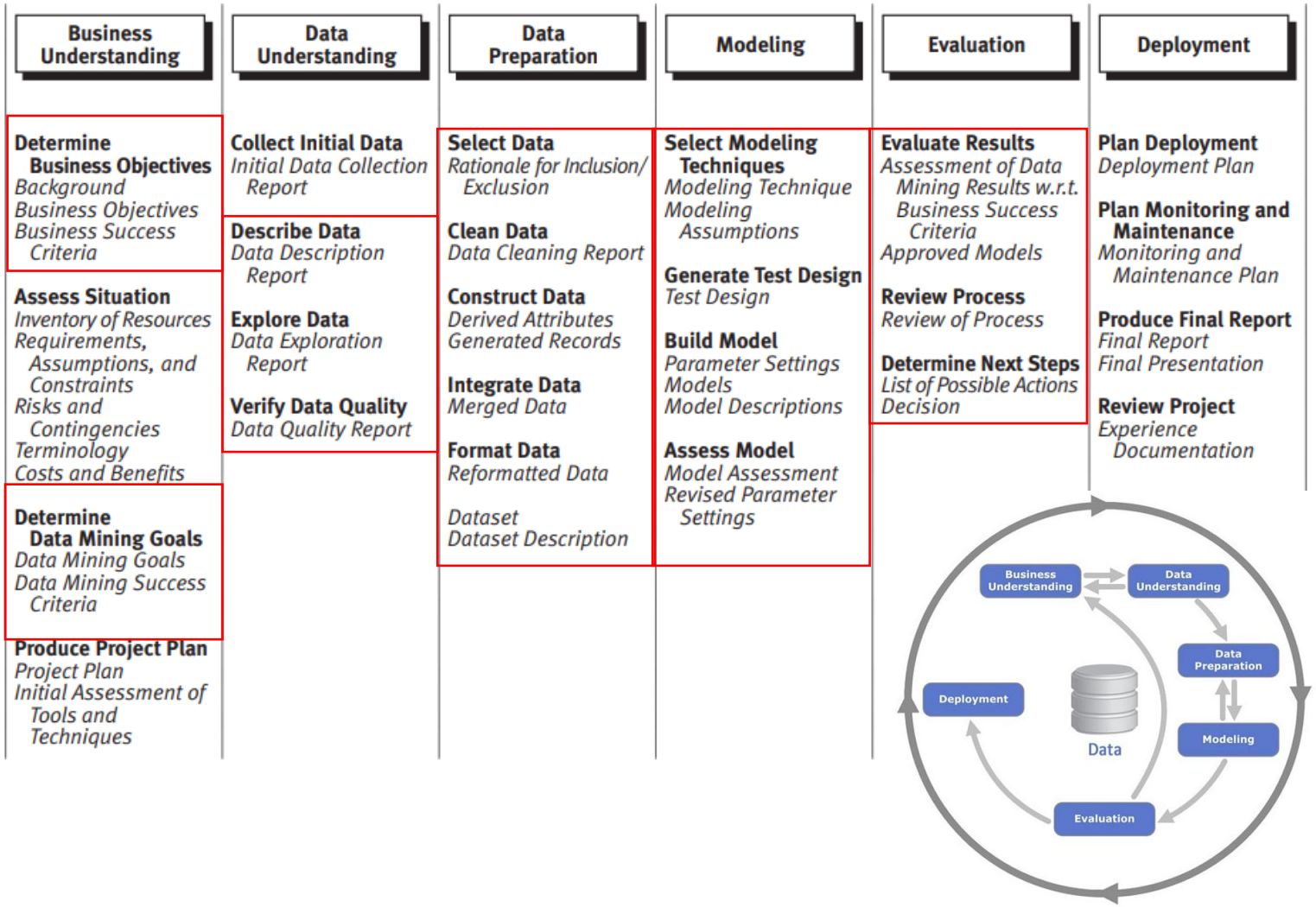
Outline

- Problem Understanding : Churn in Telco Company
  - Main problem
  - Get to know our customer
  - Get to know Telco Company services
  - Breaking down main problem
  - Causality Analysis
- Solution : Machine Learning Solution
  - Problem formulation
  - Data
  - Modeling
  - Evaluation : goal and success criteria
  - Action



Background

With the rapid development of telecommunication industry, the service providers are inclined more towards expansion of the subscriber base. To meet the need of surviving in the competitive environment, the retention of existing customers has become a huge challenge. It is stated that the cost of acquiring a new customer is far more than that for retaining the existing one. Therefore, it is imperative for the telecom industries to use advanced analytics to understand consumer behavior and in-turn predict the association of the customers as whether or not they will leave the company.

Point problem in fictional Telco company

- Huge challenge in the retention of existing customer
- Lack of competitive ability
- Cost acquisation > cost retention, better retention to keep the one you have happy
- Predict customer that will leave so we can retain more revenue

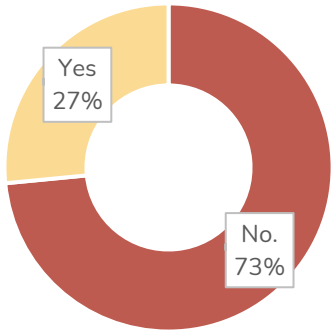
*data source : https://community.ibm.com/community/user/blogs/steven-macko/2019/07/11/telco-customer-churn-1113

High Churn Rate is What Happen in Q3

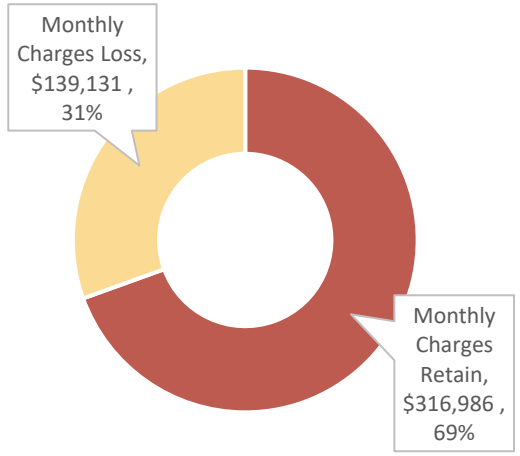
NUMBER OF USER

7043

Churn ?



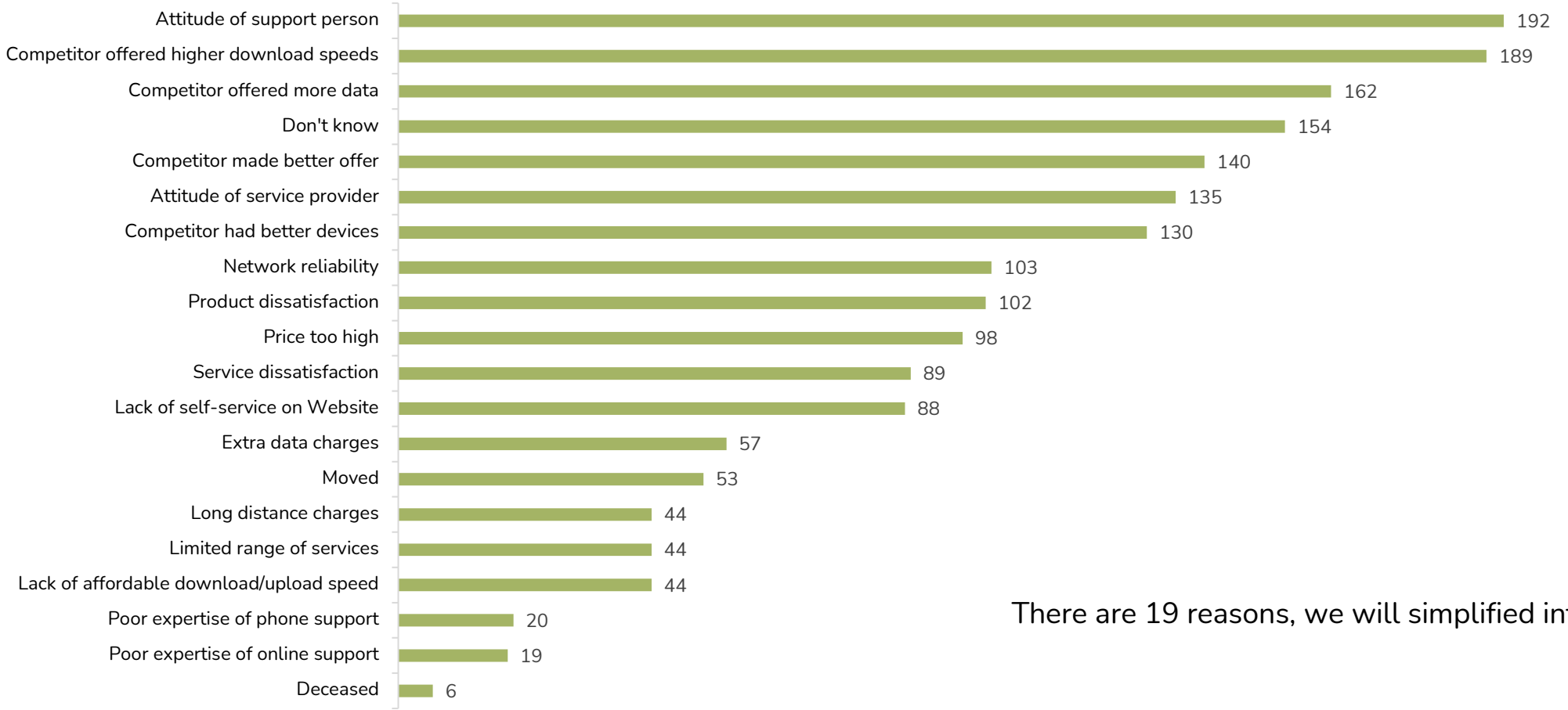
Total



Telco service provider In the state of california :

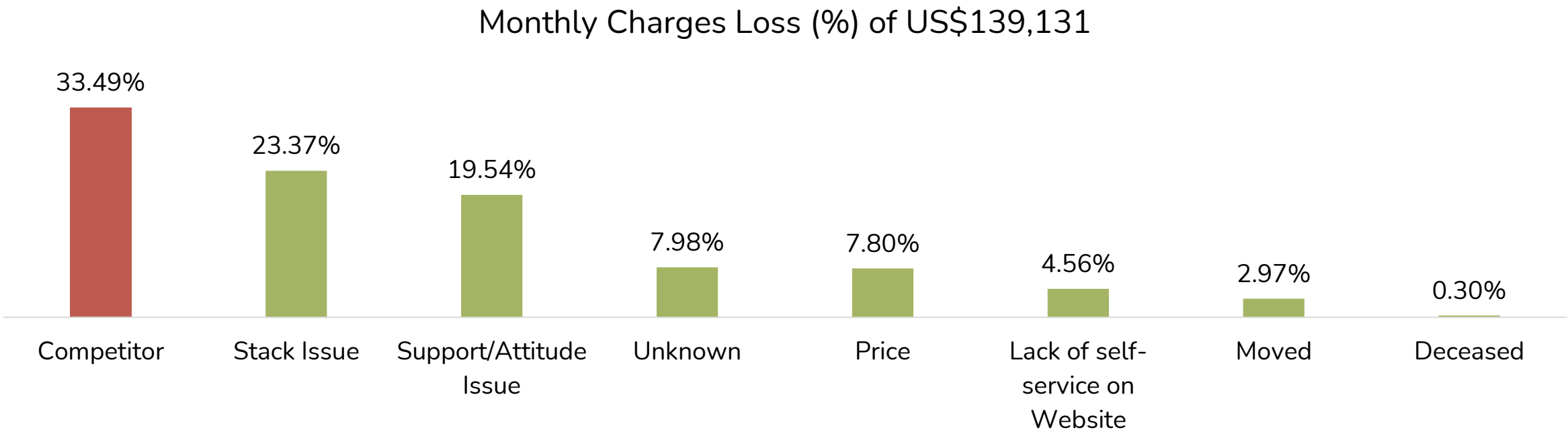
- 27% churn rate in the last quartal
- According to industry standard this is very high (above 20%)
- Revenue loss estimates is 31% or around U$139,131 based on current monthly charges

Why Customer are Leaving ???



There are 19 reasons, we will simplified into 8 categories

Customer Leave Mostly because of Better Offer

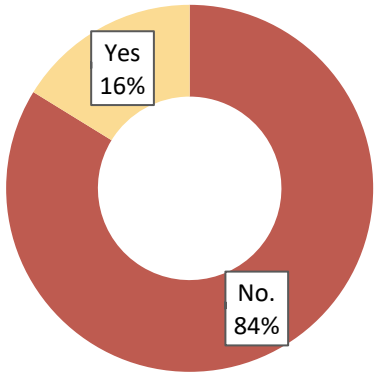


- We were trying to break down monthly loss per churn reason
- Most loss came from user that switch to competitor
- Top 3 issue are better offer from competitor, service perfomance and suppor issue

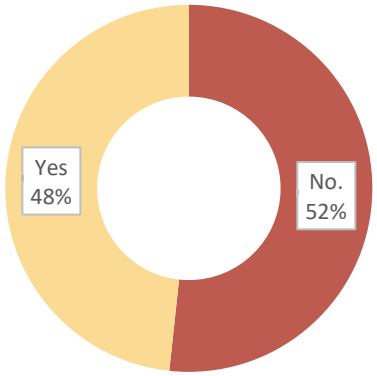
Telco Company User Profile

Dominated by Junior Citizen and People without Dependents

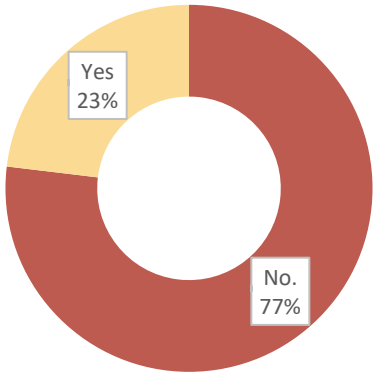
Senior Citizen ?



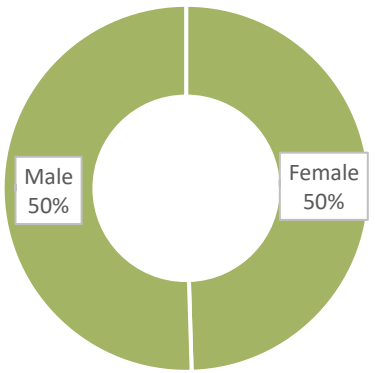
Partner ?



Dependents ?

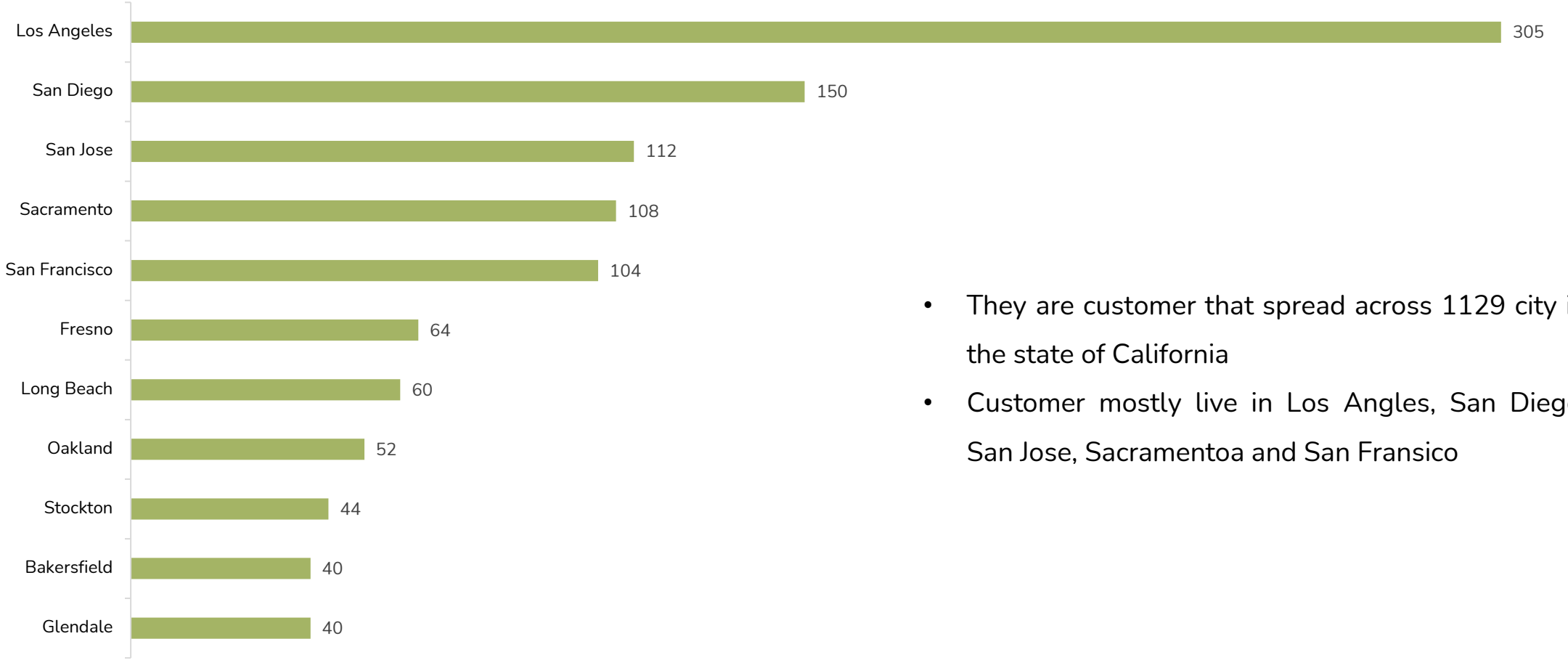


Gender ?



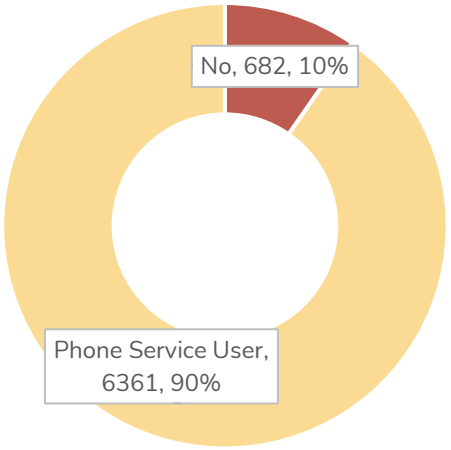
- Most User are Junior Citizen and also Without Dependents
- Half of them are whether with or without partner
- Half of them are whether male or female

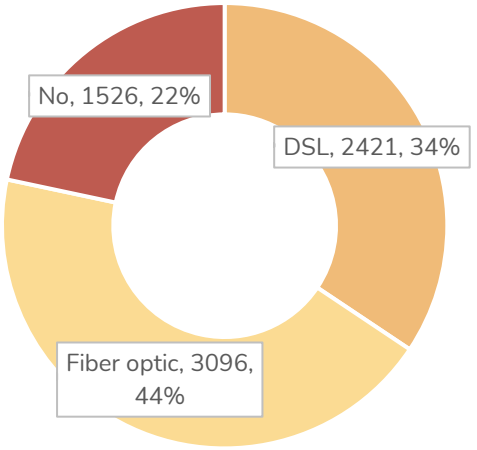
Where do They Live Mostly ?



- They are customer that spread across 1129 city in the state of California
- Customer mostly live in Los Angles, San Diego, San Jose, Sacramentoa and San Fransico

What Services are They Using ?

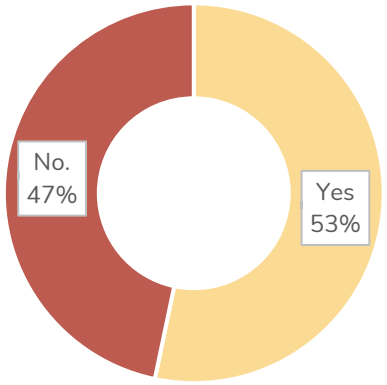




- In general, there are two products Phone Service and Internet Service
- 90% user using phone service
- 78% user using internet service with variance of product (44% Fiber Optic Internet and 34% DSL Internet)

What is More Telco Company can Offer ?

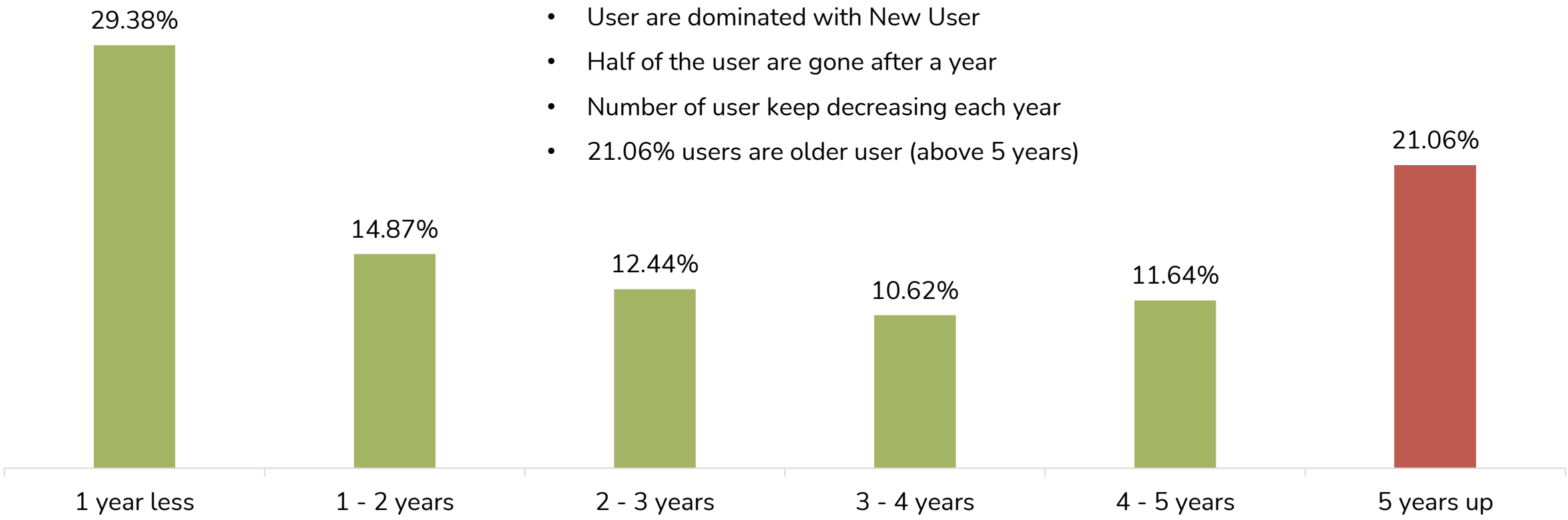
Multiple Line Services ?



Internet Additional Services User

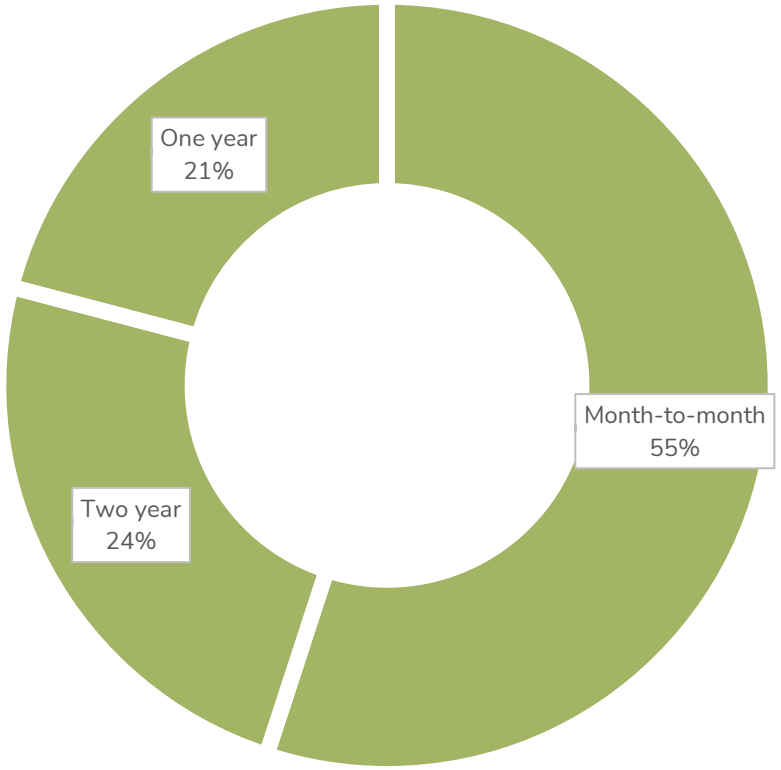
- Each services has additional service
- Phone Service : multiple lines
- Internet Service in order Streaming Movies, Streaming Online Backup, Device Protection, Tech Support and Online Security

How Long Have They been Subscribe ?



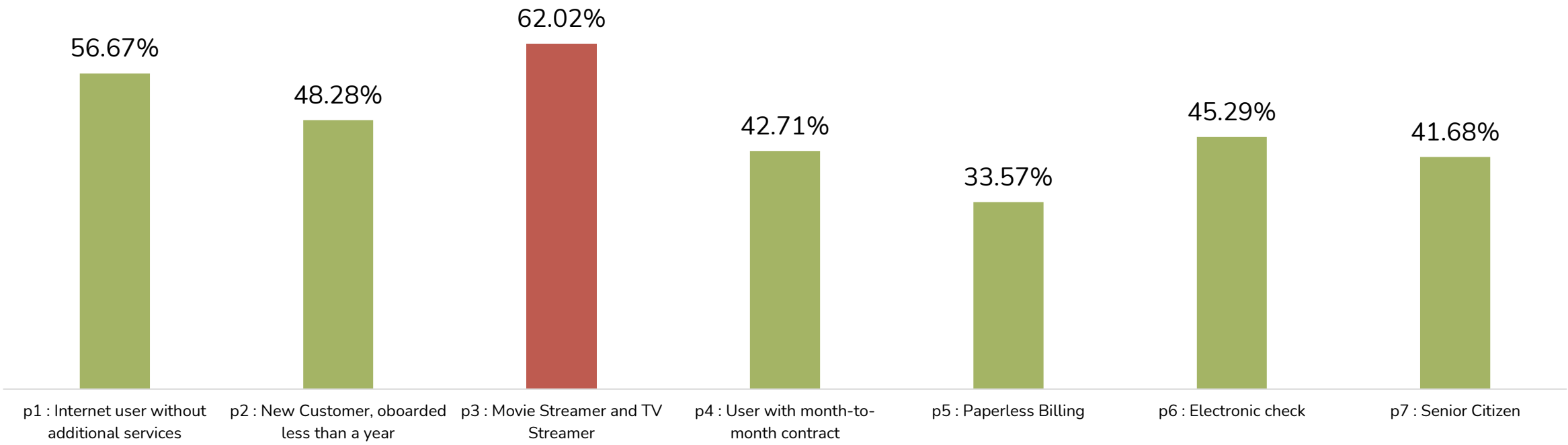
- User are dominated with New User
- Half of the user are gone after a year
- Number of user keep decreasing each year
- 21.06% users are older user (above 5 years)

How are They Subscribe to This Telco Company ?



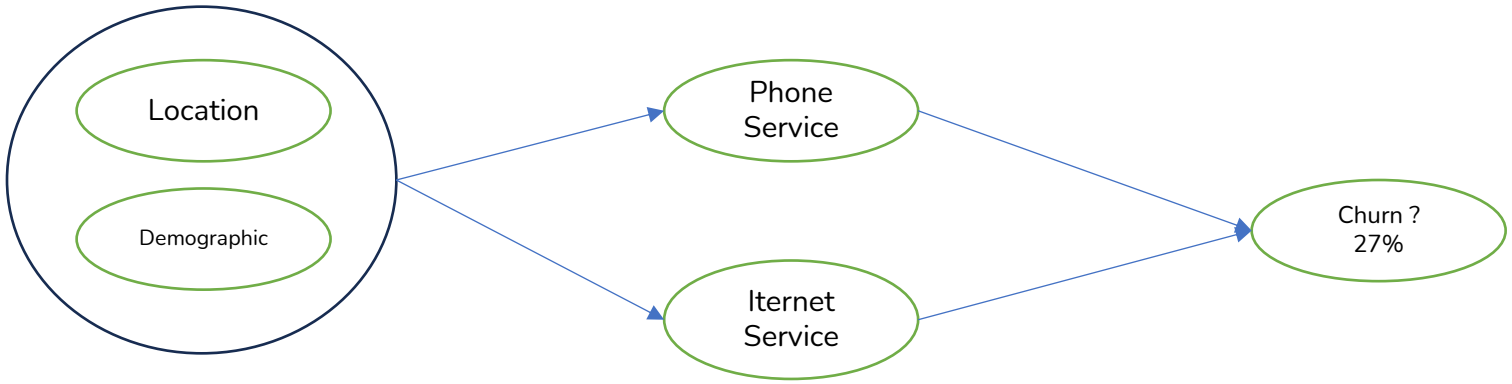
- Most Customer are using monthly subscription
- Another options are one year subscription and two year subscription

Breaking Down 27% Churn Rate

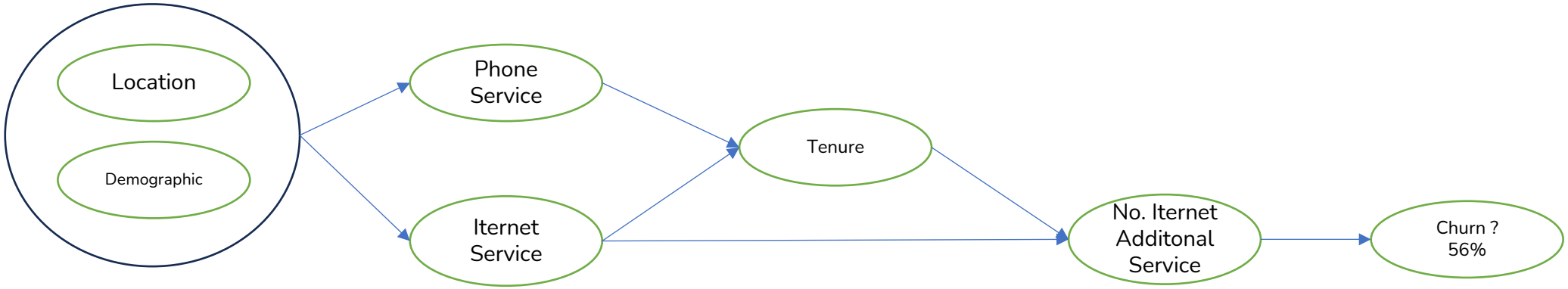


After doing some intense data exploratory. We found out 7 intersected segment that contribute to churn rate and loss. These segment combined cover 98.4% churner and 29.96% revenue loss out of 30.5% revenue loss.

Start of Journey for Every Customer



P1 : Internet User without Additional Services

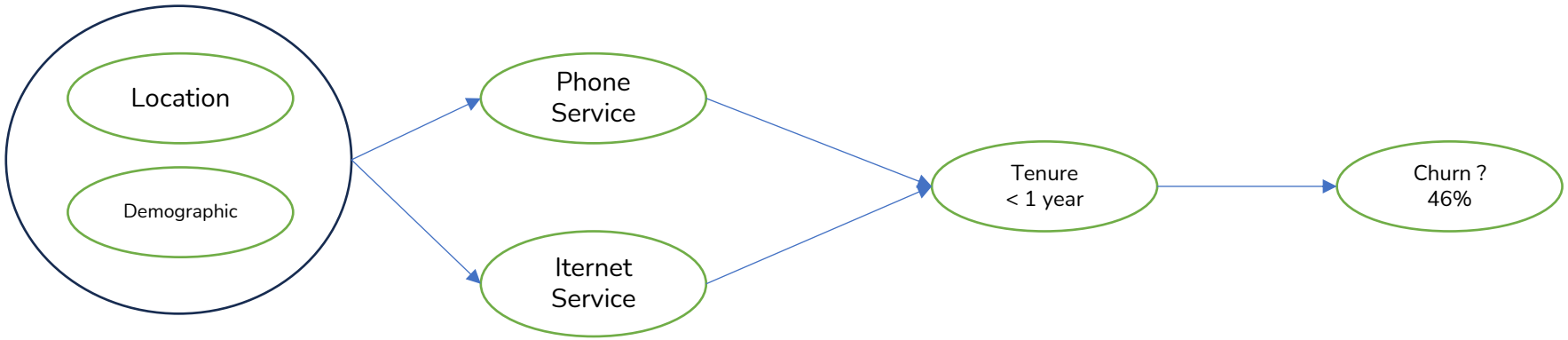


Typical User

- Low trust
- Low emotionnal bond
- Didn’t aware of the benefit and what they need yet

Solution : reframe into customer’s daily need and make product bundling, gamification and reward, free trial, education to customer, testimonies, give them attention

P2 : New Customer, Onboarded < 1 Year

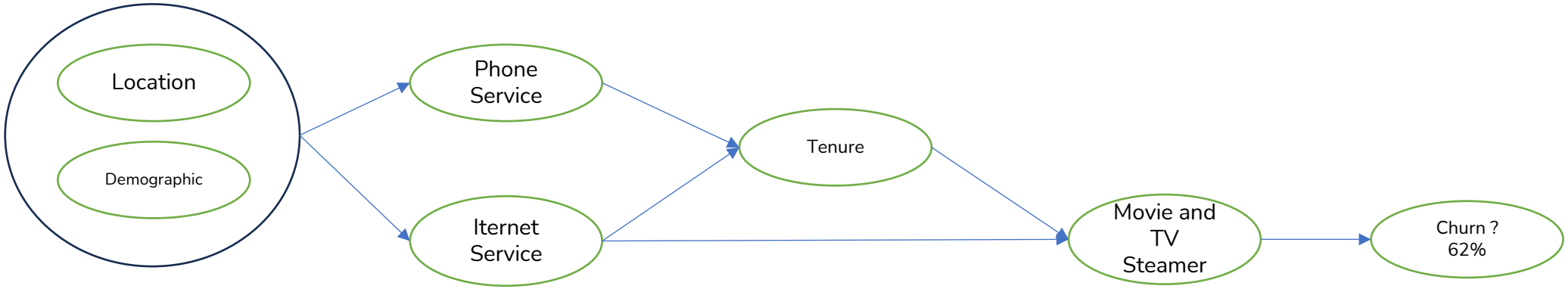


Typical User

- didn’t get value
- didn’t bound emotionally with the brand

Solution : focus on repairing a potentially disappointing initial experience and rebuilding trust, perhaps there are some problem in onboarding process

P3 : Movie and TV Streamer Only

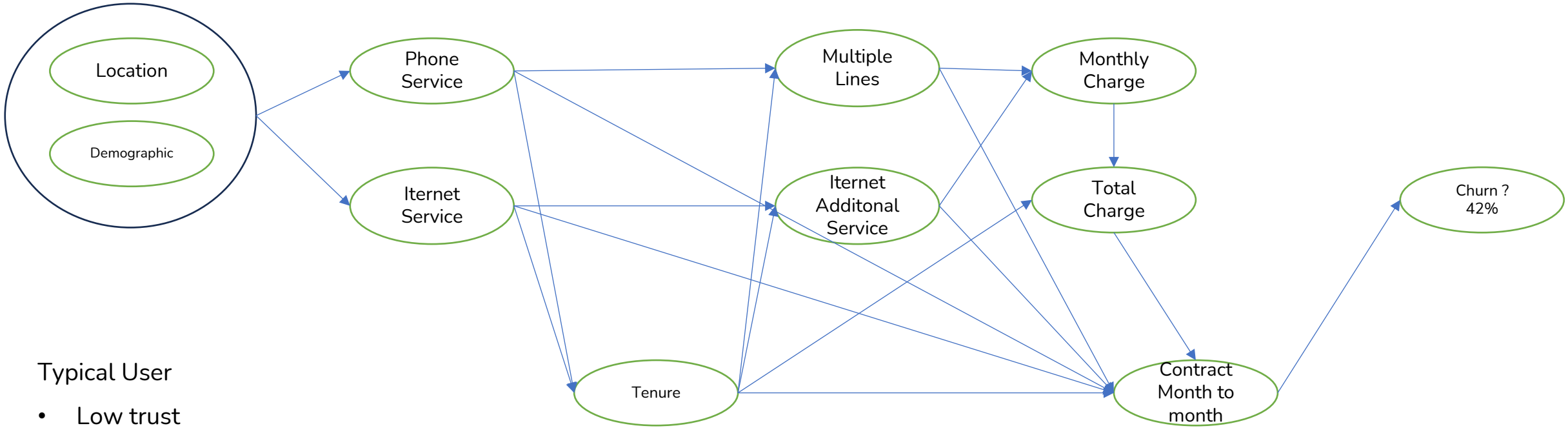


Typical User

- Need high speed internet
- Need support
- Movie and TV update

Solution : Focus on build better performance and optimization for streaming, over streaming bundling, early detection for churn

P4 : User with Month-to-month Contract

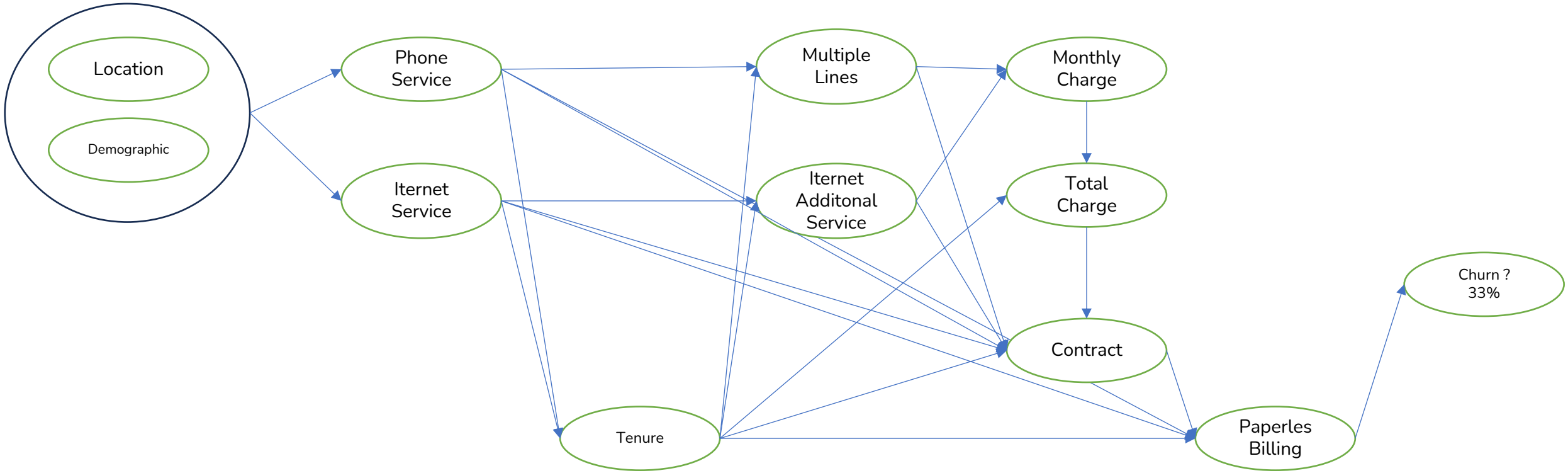


Typical User

- Low trust
- Dissatisfaction

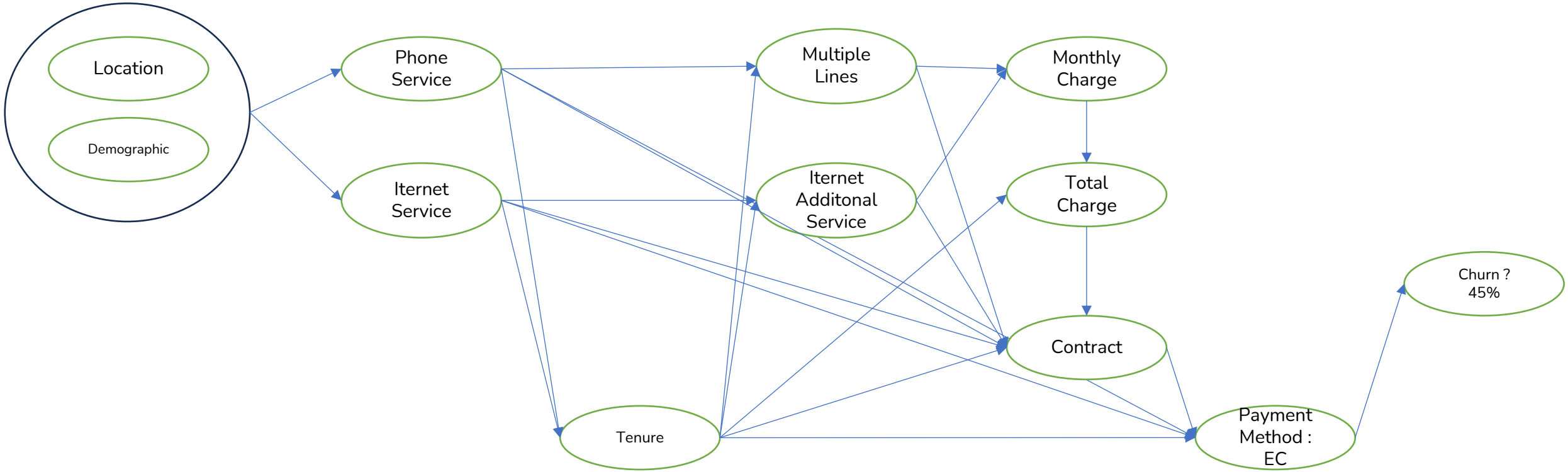
Solution : focus on repairing a potentially disappointing initial experience and rebuilding trust, premium and exclusive addtion, discount for additional internet service, discout for streaming, propect stable customer

P5 : Paperless Billing



- Paperless billing itself is not a cause of churn, but it can be a signal (proxy) of the type of customer who is more at risk of churn.
- Technology literate, proactive to search about similar service
- Prone to forgetfulness, lateness

P6 : Electronic Check



- Electronic Check itself also is not a cause of churn, but it can be a signal (proxy) of the type of customer who is more at risk of churn
- Prone to forgetfulness, lateness, technical churn
- Low trust or temporary user

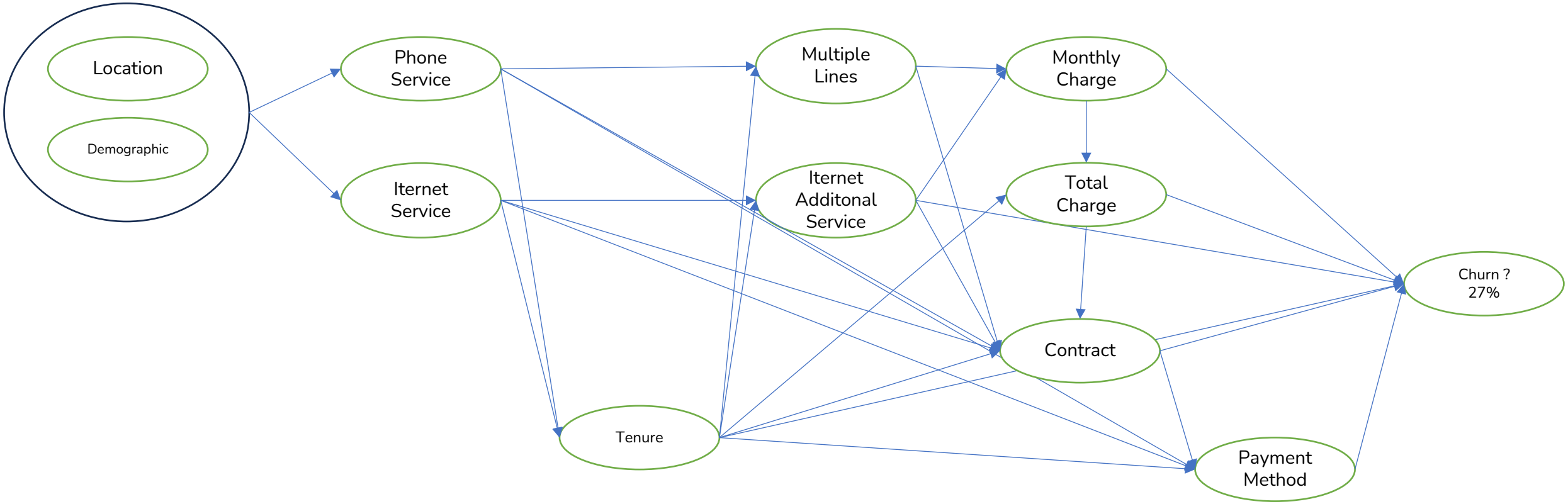
P7 : Senior Citizen

Typical User

- Tend to technologically illiterate
- Prioritize security and convenience
- Vulnerable to scams or phishing

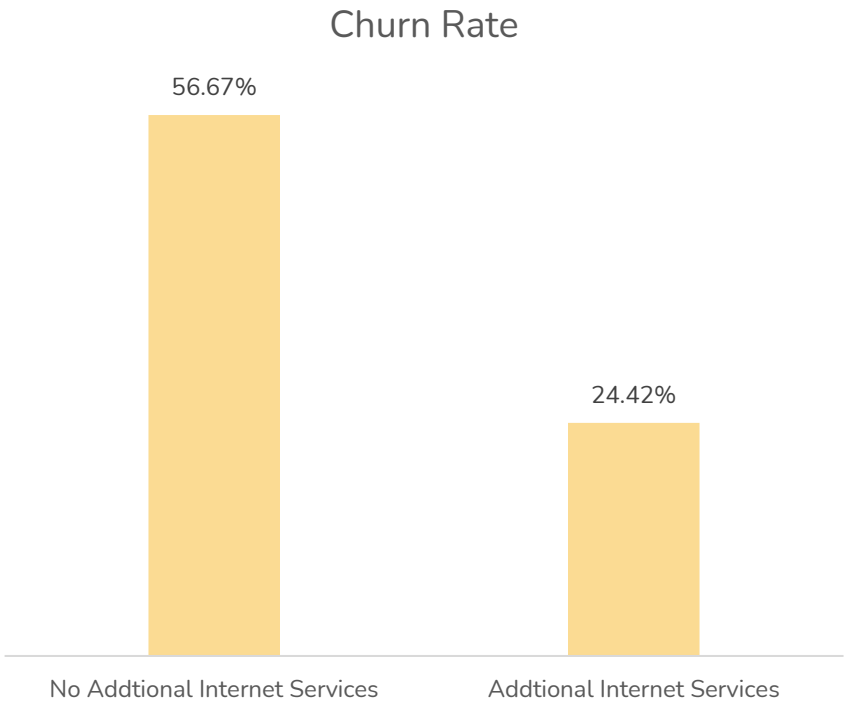
Solution : Offer Comfortable & Safe product bundling Tech Support Priority and Online Security

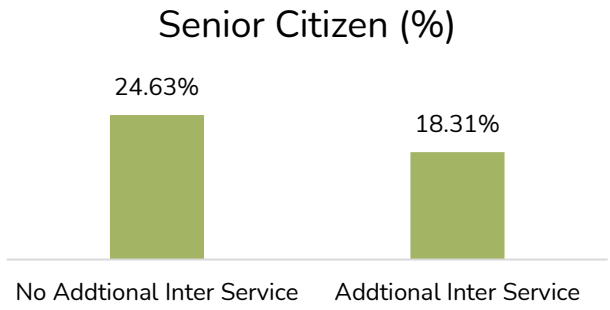
Full Journey

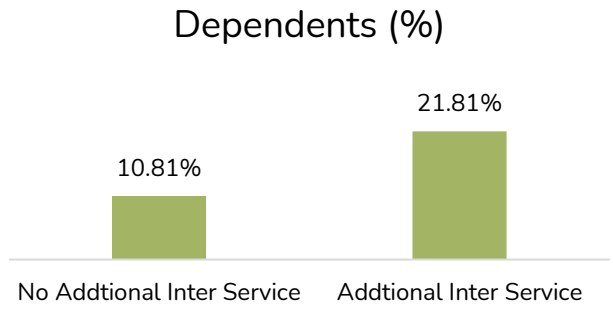


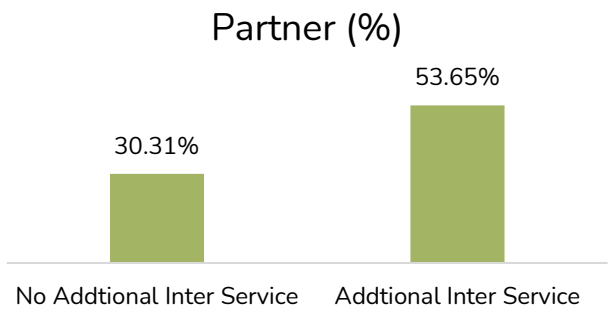
Causality Analysis : Additional Internet Services

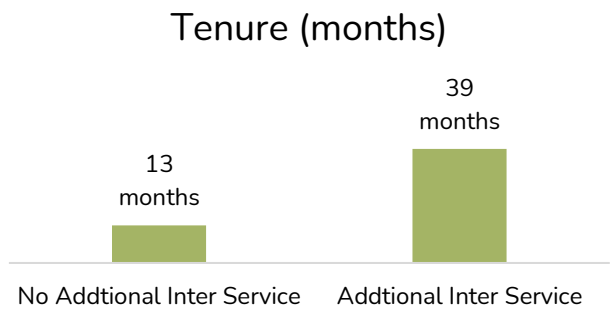
Encouraging user to use this service may reward nicely to reduce churn rate











- Additional Internet service is a direct cause of churn. This analysis will show you why we should get more user to use this services
- Data composition is not homogenous
- Additional internet service user and without online scurity user has different profile in : senior citizen, partner, dependent, tenure
- We will use propensity score matching to address this issue
- Senior citizen, partner, dependents and tenure will be used as comfound variable

Additional Internet Services Effect

Churn Rate in Internet Service User

31.82%

Churn Rate in Add. Internet Service

24.41%

Churn Rate in Diffent on Data

-7.41%

Estimated of True effect on churn rate

-7.80%

- Churn Different in 31.82 – 24.41 % = 7.41%
- True Effect based on Propensity Score Matching : - 7.80%
- Additional intenet service usage can significantly (p-value < 0.05) decrease churn rate on internet user around 7.80%
- Recommendation : encourage internet user to use additional internet service

Problem Understanding Conclusion

- Churn rate 27%
- Contribute to Revenue Loss (Monthly Charges) 30.5%
- We are breaking down the churners into 7 segment
  - p1 : nternet user without additional services
  - p2 : New Customer, oboarded less than a year
  - p3 : Movie Streamer and TV Streamer
  - p4 : User with month-to-month contract
  - p5 : Paperless Billing
  - p6 : Electronic check
  - p7 : Senior Citizen
- Causality analysis : additional internet service usage significantly decrease churn rate. Encourage internet service user to use addtional internet service.
- Further sugestion : each segment has higher churn rate and can be treated diffrently to retain more customer

ML Solution Approach

PROBLEM FORMULATION:

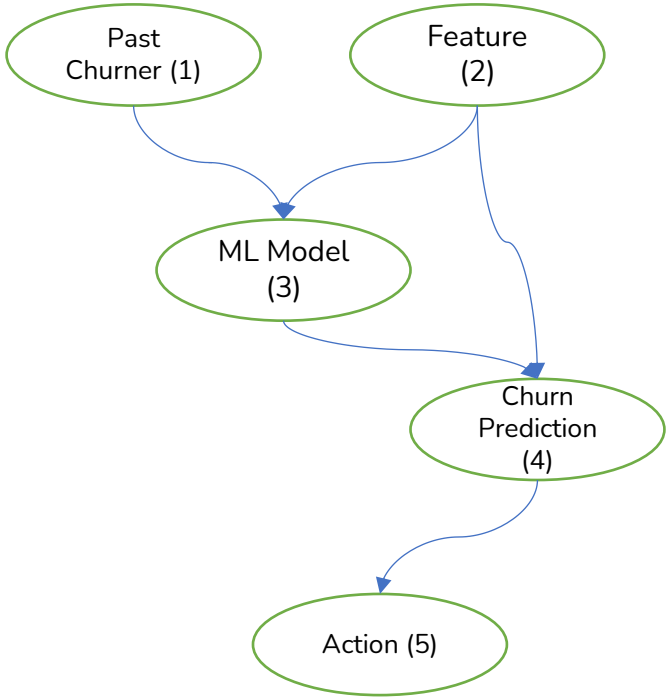
"How migh we predict customer that will leave company so we can retain more revenue, at least 40% of revenue loss ?"

WHY Machine Learning ???

- Preventive action (churn use case of course needs this)
- Efficient effort, focus on going after high churn risk customer
- Able to adapt to new pattern
- Able to learn complex pattern

HIGH LEVEL STEP BY STEP APPROACH

1. Identify Past Churner
2. Define variables to identify future churner (Feature)
3. Develop ML model to learn the pattern
4. Identify customer that will leave (churn prediction)
5. Action : Reach out the customer to prevent them from leaving



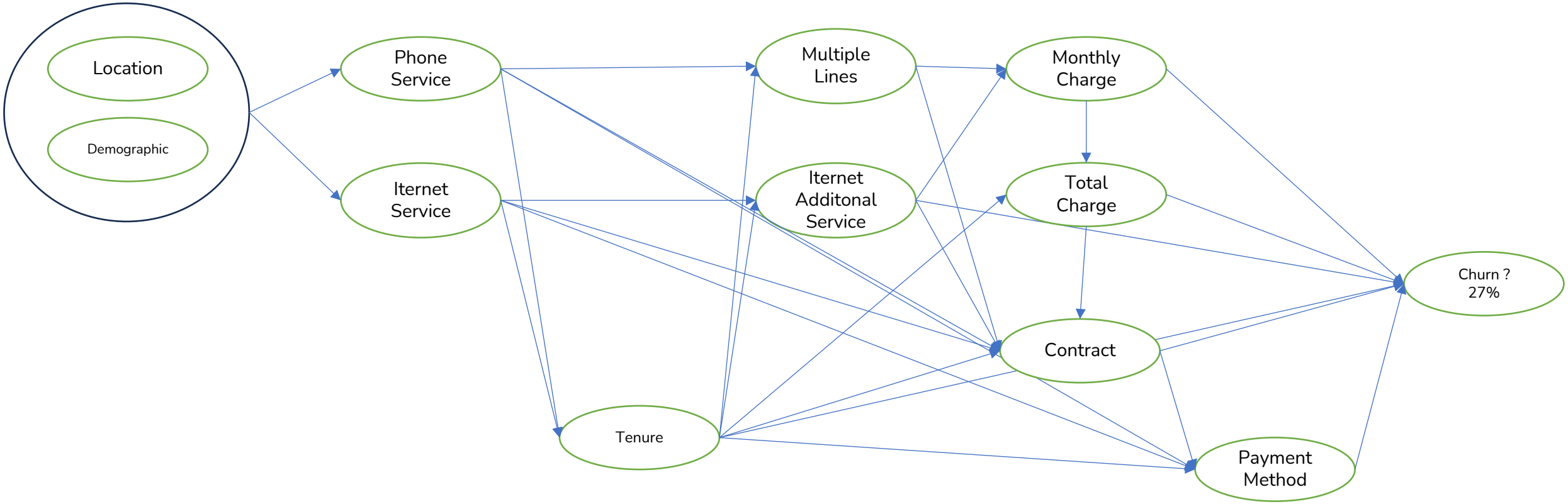
Data

We need historical that to predict what will happen in the future, such as

- Past Churner
- Demographic and Location
- Service Usage : Internet Service or Phone Service
- Additonal Service usage : Multiple Lines, Online Security, Online Backup, Device Protection, Technical Support, Streaming TV, Streaming Movies
- Payment Method : credit card, bank transfer, electronic check, mailed check, Paperless Billing
- Contract Type : Month-to-Month, 1 Year, 2 Year
- Transaction : Monthly Charge and Total Charge
- Tenure : Monthly, Yearly

*more complete dataset description : https://www.kaggle.com/datasets/yeanzc/telco-customer-churn-ibm-dataset

So, Basically These Data We Analyze Before



We don’t use “Churn Reason” because we can’t predict future using something that we don’t know yet

Illustration

Cust ID	Senior Citizen	...	Internet Service	Online Security	Streaming TV	...	Tenure (month)	Churn Score (%)	Churn Risk
123	No	...	Yes	No	Yes	...	6	75%	high
124	No	...	Yes	No	No	...	7	71%	High
...	...	...	...	...	...	...	...	...	...
991	No	...	Yes	Yes	No	...	24	23%	low

Historical data

Predicted Result :
Churner for next quarter

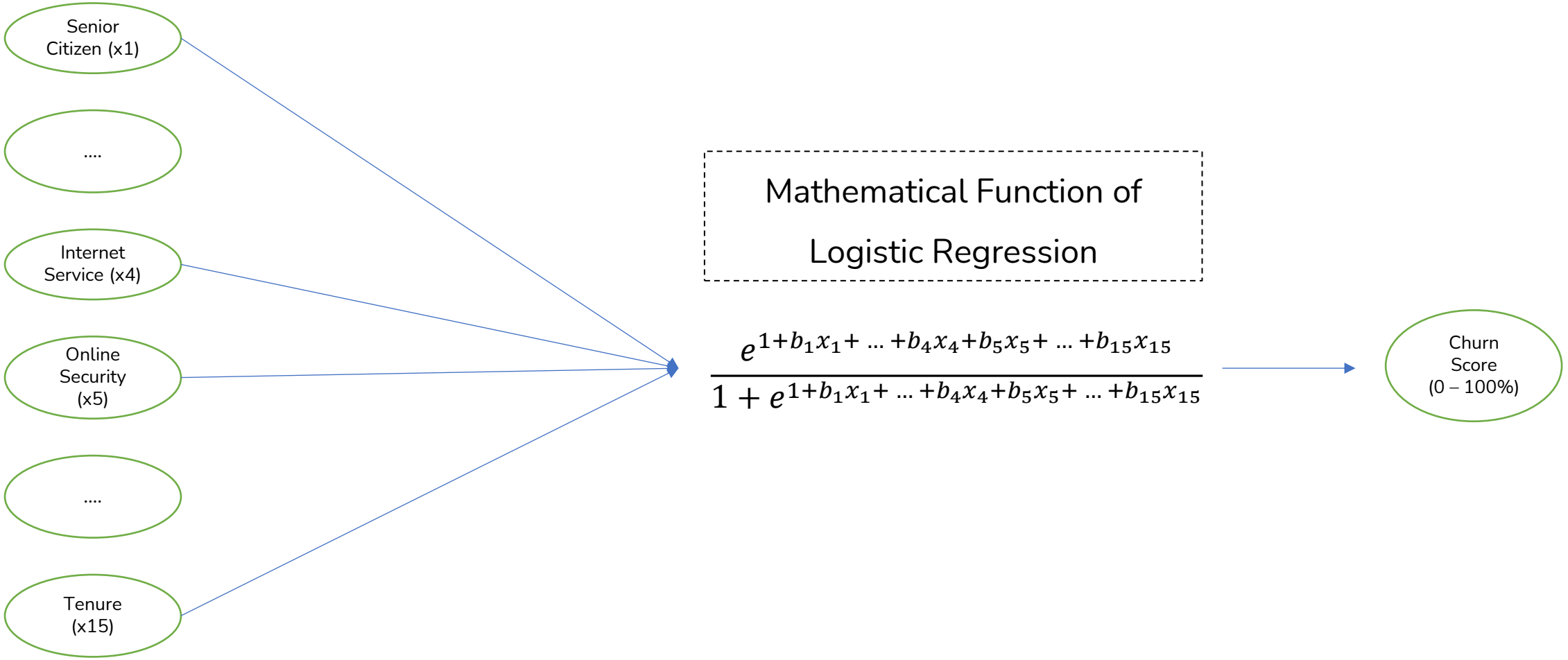
Needs Retention :
New customer (< 1 year)
and streamer

Needs Retention : New customer (<1 year) and No. additional internet service

No Retention / soft maintanance

Actions

Modeling : Logistic Regression



*let's say this model is a benchmark beacuse there might be following post about model comparisson

Evaluation : Modeling Succes Criteria

True	Prediction (Churn Score 0 – 100%)	
	Low risk churn (Churn Score > threshold)	High risk churn (Churn Score > threshold)
Stay	Staying Customer	Inacurate Retention Effort
Churn	Undetected Churner	Detected Churner

DATA MINING GOALS:

We must chose threshold to Minimize (Undetected Churner + Inacurate Retention Effort)

DATA MINING SUCCES CRITERIA:

We can minimize both “Undetected Churner” and “Inacurate Retention Effort” using classification metrics f1 score. By using this criteria in the end we hope we can prevent revenue loss as much as possible

Evaluation : Modeling Succes Criteria Result

threshold	f1 (minimize both)	recall (minimize Undetected churner)	precision (minimize wasted retention effort)
20%	0.626	0.833	0.502
25%	0.650	0.793	0.551
30%	0.640	0.712	0.580
35%	0.640	0.666	0.616
40%	0.639	0.622	0.657
45%	0.611	0.559	0.673
50%	0.584	0.495	0.712
55%	0.566	0.458	0.741

Threshold we chose for churn score is 35%. It minimize both undetected churner and wasted retention effort:

- Recall : 0.666 of 1 (the bigger the value smaller undetected churner)
- Precision : 0.616 of 1 (the bigger the value less retention effor)
- F1 – score : 0.640 of 1 (the bigger the value lee both)

Evaluation : Business Succes Criteria

True	Prediction (Churn Score 0 – 100%)	
	Low risk churn (Churn Score > threshold)	High risk churn (Churn Score > threshold)
Stay	Staying Customer	Wasted Retention Cost
Churn	Undetected Revenue Loss	Potentially Saved Revenue

Definition :

Revenu Loss = Undetected Revnue Loss + 100% Potentially Saved Revenue
Wasted Retention Cost = sum of (customer in “high risk” and staying customer) x (5% of monthly charge per customer (assumption))
Undetected Revenue Loss = sum of (number of customer in “low risk” and churning customer) x monthly charge
Potentialy Saved Revenue = sum of (number of customer in “high risk” and churning customer) x monthly charge per customer x retention campaign succes rate (assumption)

Business Goals and Succes Criteria :

Succesfully decrease revenue loss 40% of total revenue loss by minimizing (Undetected revenue loss + wasted retention cost)

Revenue Loss Decrease 52.31%

So we can decrease 30.5% revenue loss up to 15.95%

- Given : Retention cost = 15% of revenue
- Assumption : Campaign effectivity Target = 80% succes rate of retention campaign
- Revenue Loss = Undetected Revnue Loss + 100% Potentially Saved Revenue = US$136,616
- Wasted Retention Cost = sum of (customer in “high risk” and staying customer) x (5% of monthly charge per customer (assumption)) = US$8,754
- Undetected Revenue Loss = sum of (number of customer in “low risk” and churning customer) x monthly charge =US$41,372
- Potentialy Saved Revenue = sum of (number of customer in “high risk” and churning customer) x monthly charge per customer x retention campaign succes rate) = US$73,910
- **Potentialy Saved Revenue (%) = (US$73,910 - US$8,754)/ US$136,616 = 47.69%**
- **Revenue Loss Decrease (%) = 100% - 47.69% = 52.31%**

What Actions to Achieve Revenue Loss Decrease

In general, according to churn reason Telco Company should do:

1. Support person training
2. Get better internet peroformance, better speed especially for streaming
3. Formulate better offer for customer
4. Review the acquisition process. Since there are so many new user leave

Specifically, Telco Company should approcah customer according to :

1. churn risk prediction (high/low) for the next quartal
2. which segment they are in (p1, p2, p3, p4, p5, p6, p7)

Example 1

Cust ID	Senior Citizen	...	Internet Service	Online Security	Streaming TV	...	Tenure (month)	Churn Score (%)	Churn Risk
123	No	...	Yes	No	Yes	...	6	75%	high
...	...	...	...	...	...	...	...	...	...

P1 P3 P2

Cust ID	Senior Citizen	...	Internet Service	Online Security	Streaming TV	...	Tenure (month)	Churn Score (%)	Churn Risk
123	No	...	Yes	Yes	Yes	...	14	31%	Low
...	...	...	...	...	...	...	...	...	...

Needs Retention : High Churn Risk

- No additional internet service (p1)
- Segment New customer (< 1 year) (p2)
- Segment TV and Movie Streamer (p3)

Possible Action from p1, p2, p3 :

- P1: reframe into customer's daily need and make product bundling, gamification and reward, free trial, education to customer, testimonies, give them attention
- P2: focus on repairing a potentially disappointing initial experience and rebuilding trust
- P3: Focus on build better performance and optimization for streaming, over streaming bundling, early detection for churn

Example 2

Cust ID	Senior Citizen	...	Internet Service	Online Security	Streaming TV	...	Tenure (month)	Churn Score (%)	Churn Risk
123	Yes	...	Yes	No	No	...	7	75%	high
...	...	...	...	...	...	...	...	...	...

P1 and P7

P2

Cust ID	Senior Citizen	...	Internet Service	Online Security	Streaming TV	...	Tenure (month)	Churn Score (%)	Churn Risk
123	Yes	...	Yes	Yes	No	...	13	29%	Low
...	...	...	...	...	...	...	...	...	...

Needs Retention : High Churn Risk

- No additional internet service (p1)
- Segment New customer (< 1 year) (p2)
- Senior Citizen (p7)

Possible Action from p1, p2, p7 :

- P1 and P7: Offer Comfortable & Safe product bundling, Tech Support Priority and Online Security
- P2: focus on repairing a potentially disappointing initial experience and rebuilding trust

Machine Learning Solution Conclusion

- We aim to solve the problem using machine learning, with problem statement : “How migh we predict customer that will leave company so we can retain more revenue, at least 40% of loss revenue ?”
- Prediction will be done based on historical data : demographic, location, service usage, additonal internet service, streaming TV and Movies. payment method, contract type. transaction and tenure
- Machine learning model we used is Logistic Regression
- In term of model performance : With adjusting threshold to 35% We achieve f1 score 0.640, recall 0.666, precision 0.616
- In term of business succes criteria : With certain assumption we can decrease 30.5% revenue loss to 15.95% revenue loss
- We can acvhive these after doing some following action
    - In general, according to churn reason Telco Company should do: support person training; get better internet peroformance, better speed especially for streaming; formulate better offer for customer and review the acquisition process. Since there are so many new user leave
    - Specifically, Telco Company should approcah customer according to churn risk (high/low) prediction for the next quartal n and which segment they are in (p1, p2, p3, p4, p5, p6, p7)
    - Total Retention cost 15% of total revenue
    - Achieve 80% success rate of retention campaign (to acieve this, it will be a sepparate work)

Disclaimer Post

- In, CRIS-DM there are many points. In this post, I'm just willing to share certain point for the shake of analytic, storyline, proble solving and limited information about the data.
- Another model indeed can be used. May be there will be follow up post focusing on getting better performance.
- Action can be depend on company budget and another situation assesment. In this case perhaps they want to focus on bulding customer trust and worry less about Movie and TV streaming
- In reality, Better offer action for customer should be discuss further with many party such as stakeholder and business team.
- Assumption needed because we can't figure out a few things right away. Especially information about data that I get.

Referencess

- https://hbr.org/2014/10/the-value-of-keeping-the-right-customers
- https://www.datascience-pm.com/crisp-dm-2/
- https://web.archive.org/web/20220401041957/https://www.the-modeling-agency.com/crisp-dm.pdf
- https://www.kaggle.com/datasets/yeanzc/telco-customer-churn-ibm-dataset
- https://builtin.com/data-science/propensity-score-matching
- https://community.ibm.com/community/user/blogs/steven-macko/2019/07/11/telco-customer-churn-1113