

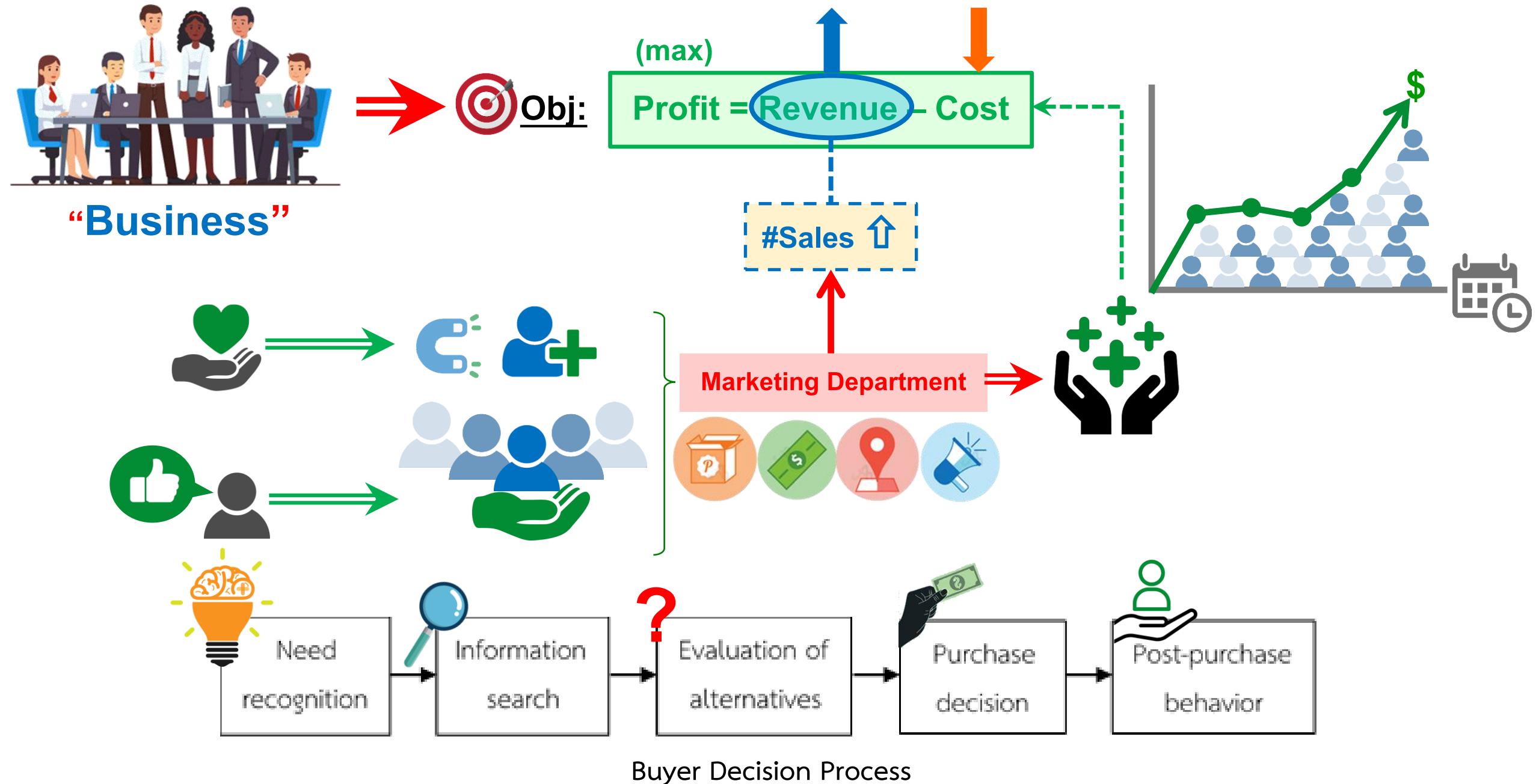


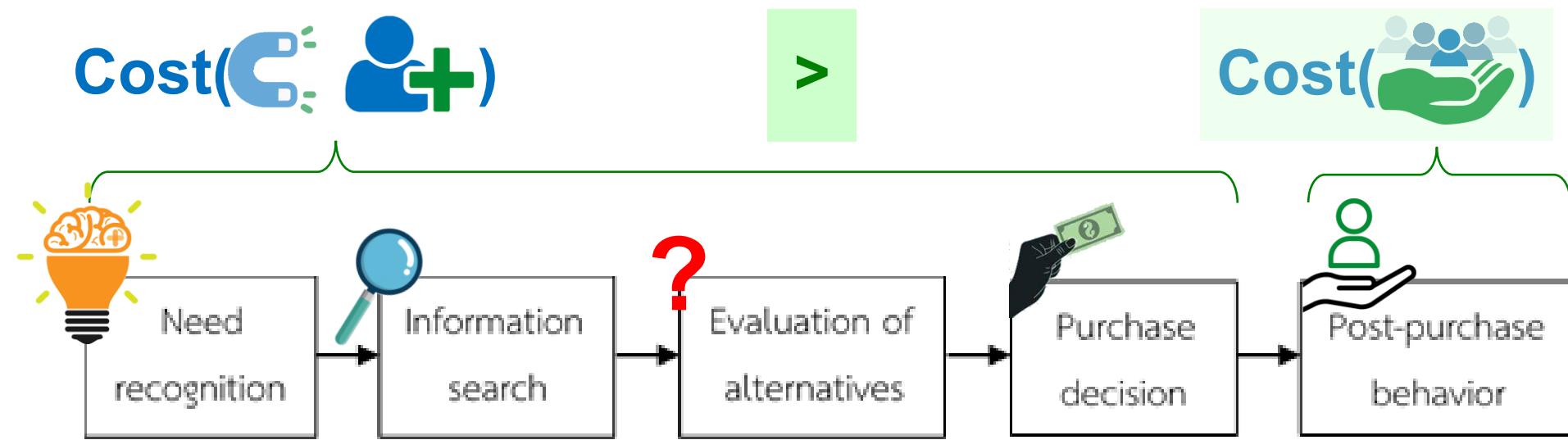
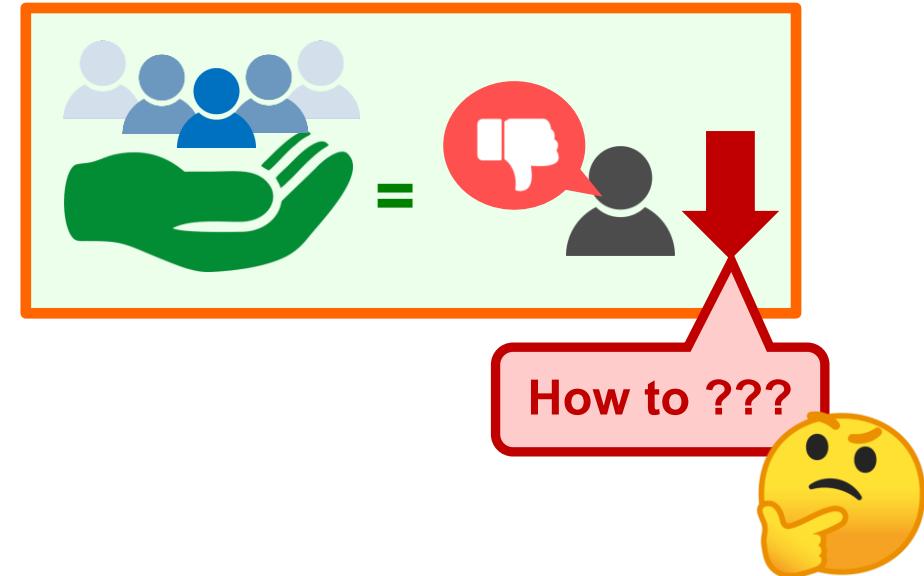
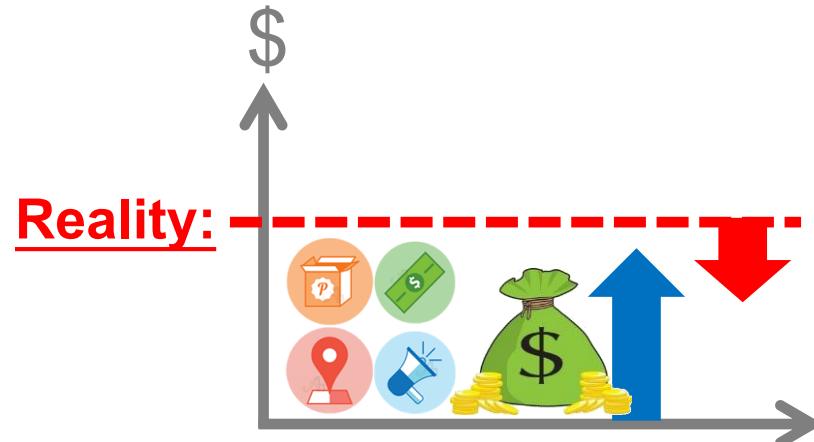
# Customer Churn Prediction & Retention Strategies for Telecommunication Business



**Presented by Group 23:**

Thatchakarn	Chariyasethapong	6130239421
Teethavat	Techaphatipong	6130252521



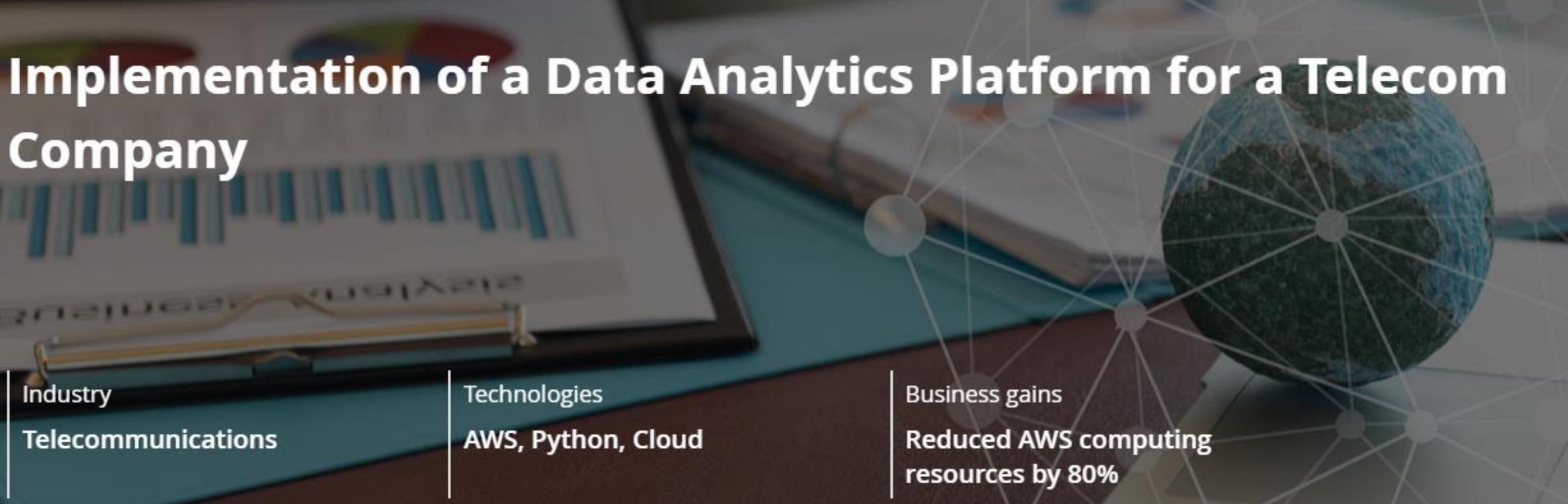


Buyer Decision Process



Home > Case Studies > Implementation of a Data Analytics Platform for a Telecom Company

# Implementation of a Data Analytics Platform for a Telecom Company



Industry

Telecommunications

Technologies

AWS, Python, Cloud

Business gains

Reduced AWS computing resources by 80%

## Customer

---

The Customer is a Texas-based telecom company participating in the federal Lifeline Support Program and providing pre-paid cell phones and service packages to low-income individuals.



Faculty of Engineering, Chulalongkorn University



Department of Industrial Engineering

# Results

With ScienceSoft's [big data services](#), the Customer was able to:

- Measure the engagement and identify the preferences of a particular user.
- Spot trends in the users' behavior.
- Make predictions about how users would behave.
- Invoice advertisers based on their calculated share.
- Benefit from insightful data analytics (for example, daily earnings, number of new users, customer service data and more).

The use of Amazon Spot Instances allowed the Customer to reduce the costs of AWS computing resources by 80%.

# A Business Problem in Telecom company



# Telecom Customer Churn Prediction

Abhijit Sundararajan

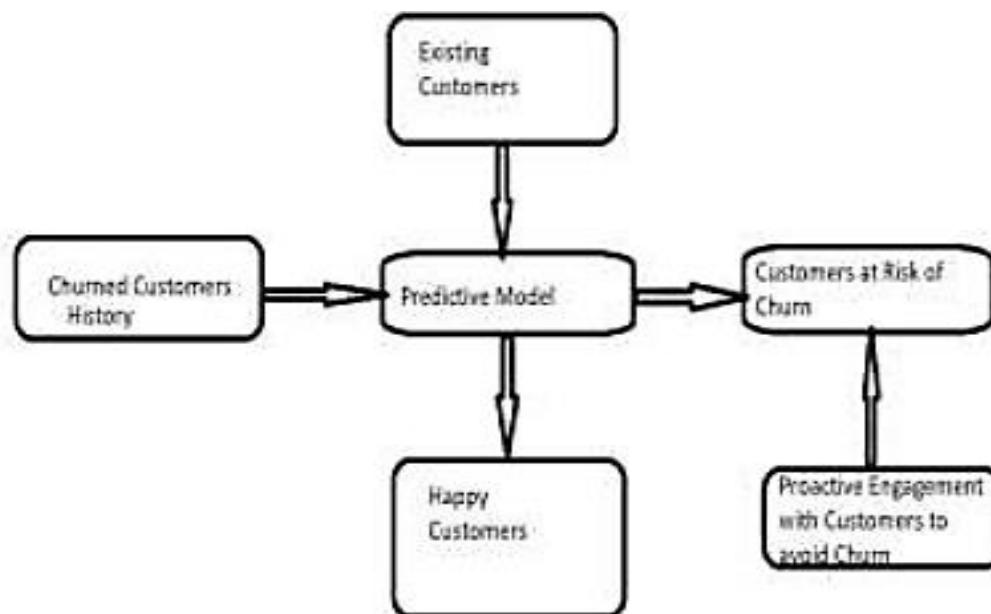
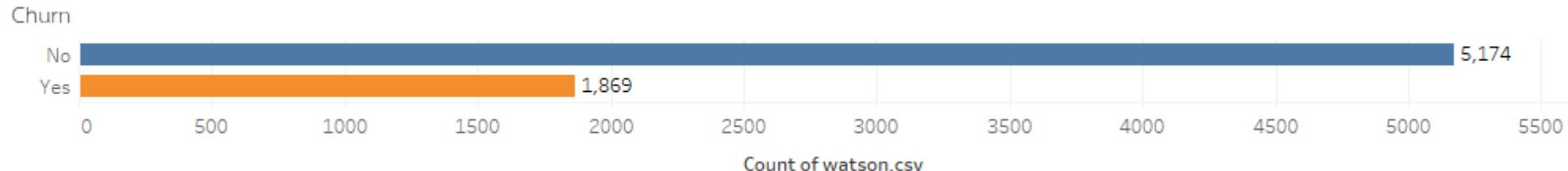
Department of MSIS, Rutgers University

E-mail: [abhijit.sundararajan@rutgers.edu](mailto:abhijit.sundararajan@rutgers.edu)

Kemal Gursoy

Department of MSIS, Rutgers University,

E-mail: [kgursoy@business.rutgers.edu](mailto:kgursoy@business.rutgers.edu)



## 5. Comparison of Methods

Models	Accuracy Score
Random Forest Classifier	0.9355
SVM	0.8192
Decision Tree	0.762
Logistic Regression	0.7894



# Scope & Objective

<https://www.kaggle.com/blastchar/telco-customer-churn>

**Telco Customer Churn**  
Focused customer retention programs

BlastChar • updated 4 years ago (Version 1)

Data Tasks (1) Code (758) Discussion (14) Activity Metadata Download (978 kB) New Notebook

Usability 8.8 License Data files © Original Authors Tags business

## Description

## Context

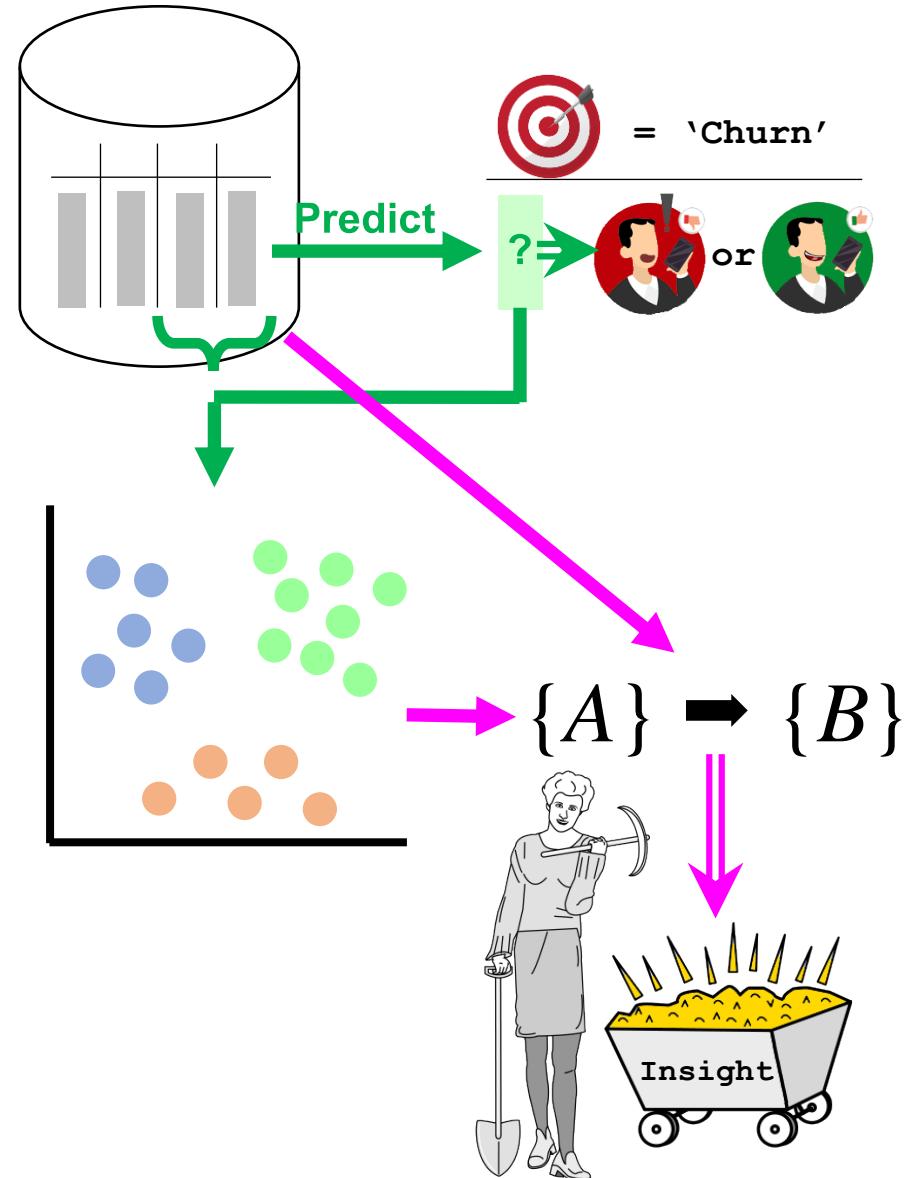
"Predict behavior to retain customers. You can analyze all relevant customer data and develop focused customer retention programs." [IBM Sample Data Sets]

## Content

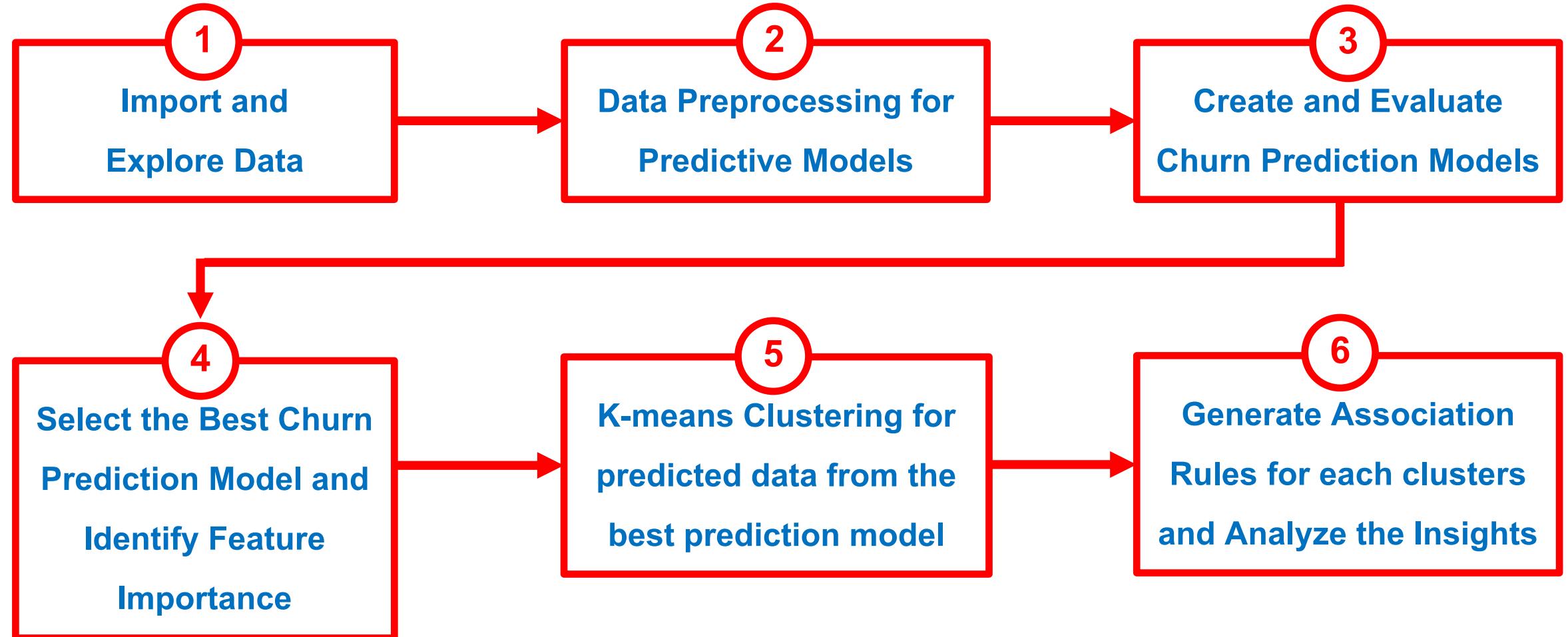
Each row represents a customer, each column contains customer's attributes described on the column Metadata.

The data set includes information about:

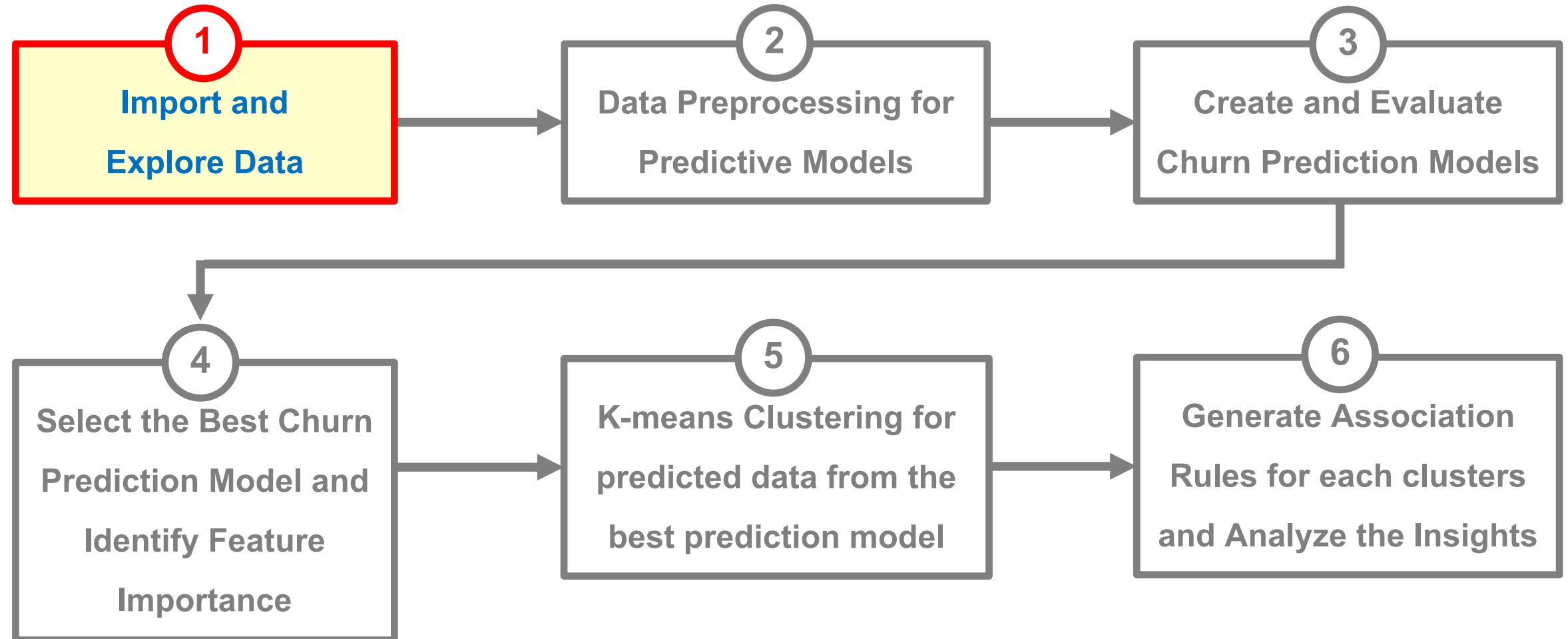
- Customers who left within the last month – the column is called Churn
- Services that each customer has signed up for – phone, multiple lines, internet, online security, online backup, device protection, tech support, and streaming TV and movies
- Customer account information – how long they've been a customer, contract, payment method, paperless billing, monthly charges, and total charges
- Demographic info about customers – gender, age range, and if they have partners and dependents



# Procedure



# Procedure



## 1

# Import and Explore Data

```
✓ 2s   ➜ !gdown https://drive.google.com/uc?id=1qXvzqhXaclx6dqKAQKYXnHeUZEZffnLE
```

```
✓ 0s   ➜ df = pd.read_csv('watson.csv', header=0, delimiter=',')
df
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	...	Churn
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	...	No
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	...	No
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	...	Yes
3	7795-CFOOW	Male	0	No	No	45	No	No phone service	DSL	...	No
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	...	Yes
...	...	...	...	...	...	...	...	...	...	...	...
7038	6840-RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	...	No
7039	2234-XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	...	No
7040	4801-JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	...	No
7041	8361-LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	...	Yes
7042	3186-AJIEK	Male	0	No	No	66	Yes	No	Fiber optic	...	No

7043 rows × 21 columns



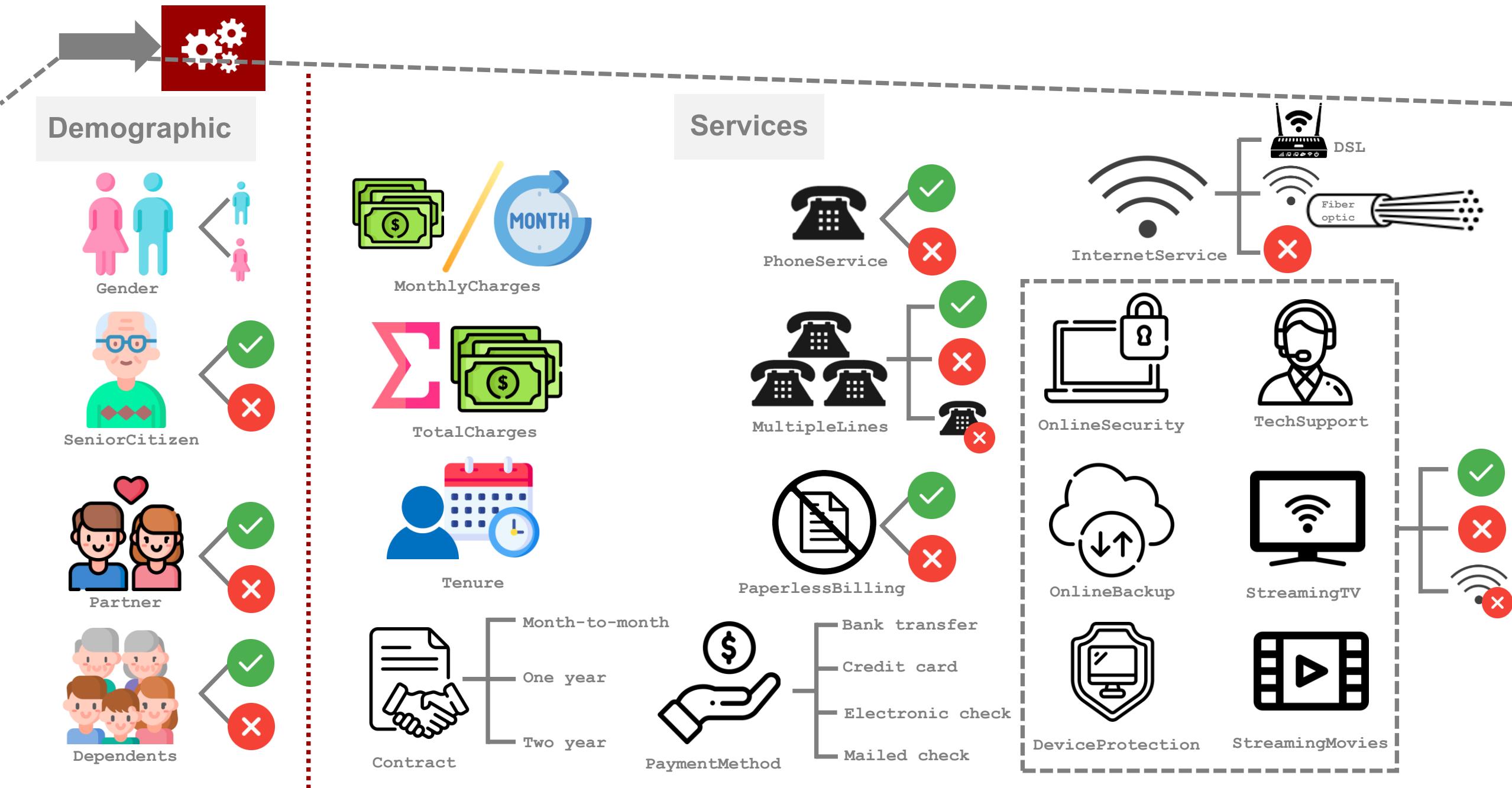
0s

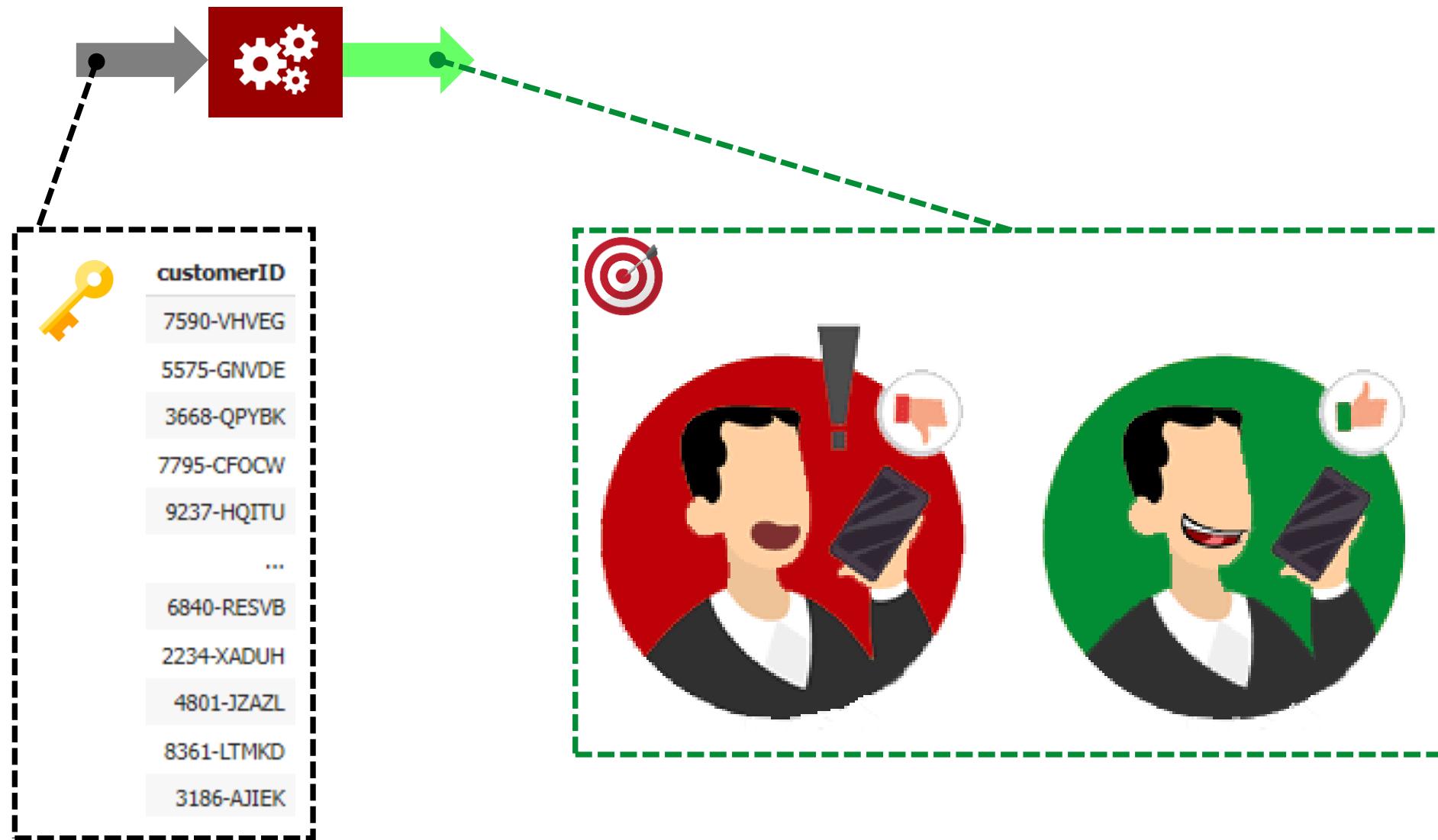


df.info()

```
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 21 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   customerID      7043 non-null    object 
 1   gender          7043 non-null    object 
 2   SeniorCitizen   7043 non-null    int64  
 3   Partner         7043 non-null    object 
 4   Dependents     7043 non-null    object 
 5   tenure          7043 non-null    int64  
 6   PhoneService    7043 non-null    object 
 7   MultipleLines   7043 non-null    object 
 8   InternetService 7043 non-null   object 
 9   OnlineSecurity  7043 non-null   object 
 10  OnlineBackup    7043 non-null   object 
 11  DeviceProtection 7043 non-null  object 
 12  TechSupport    7043 non-null   object 
 13  StreamingTV    7043 non-null   object 
 14  StreamingMovies 7043 non-null  object 
 15  Contract        7043 non-null   object 
 16  PaperlessBilling 7043 non-null  object 
 17  PaymentMethod   7043 non-null   object 
 18  MonthlyCharges 7043 non-null   float64
 19  TotalCharges   7043 non-null   object 
 20  Churn           7043 non-null   object 
dtypes: float64(1), int64(2), object(18)
```







✓

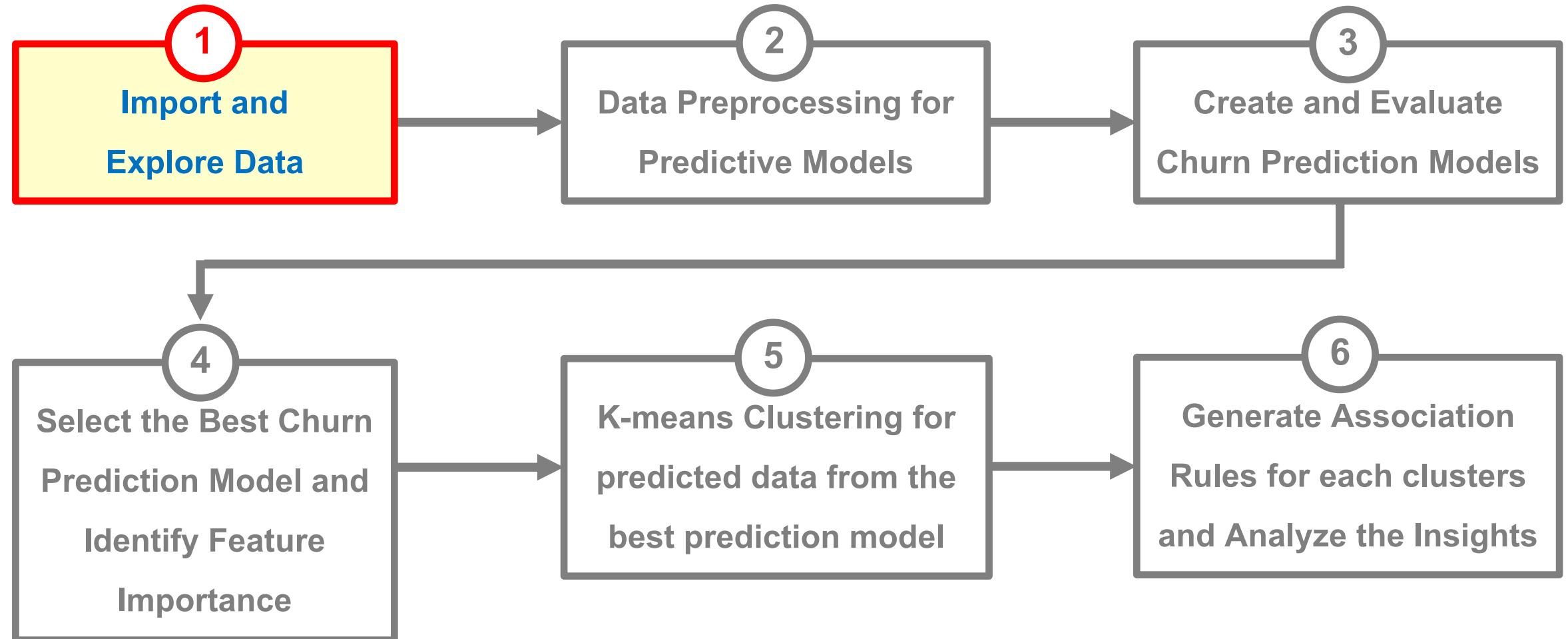


df.info()

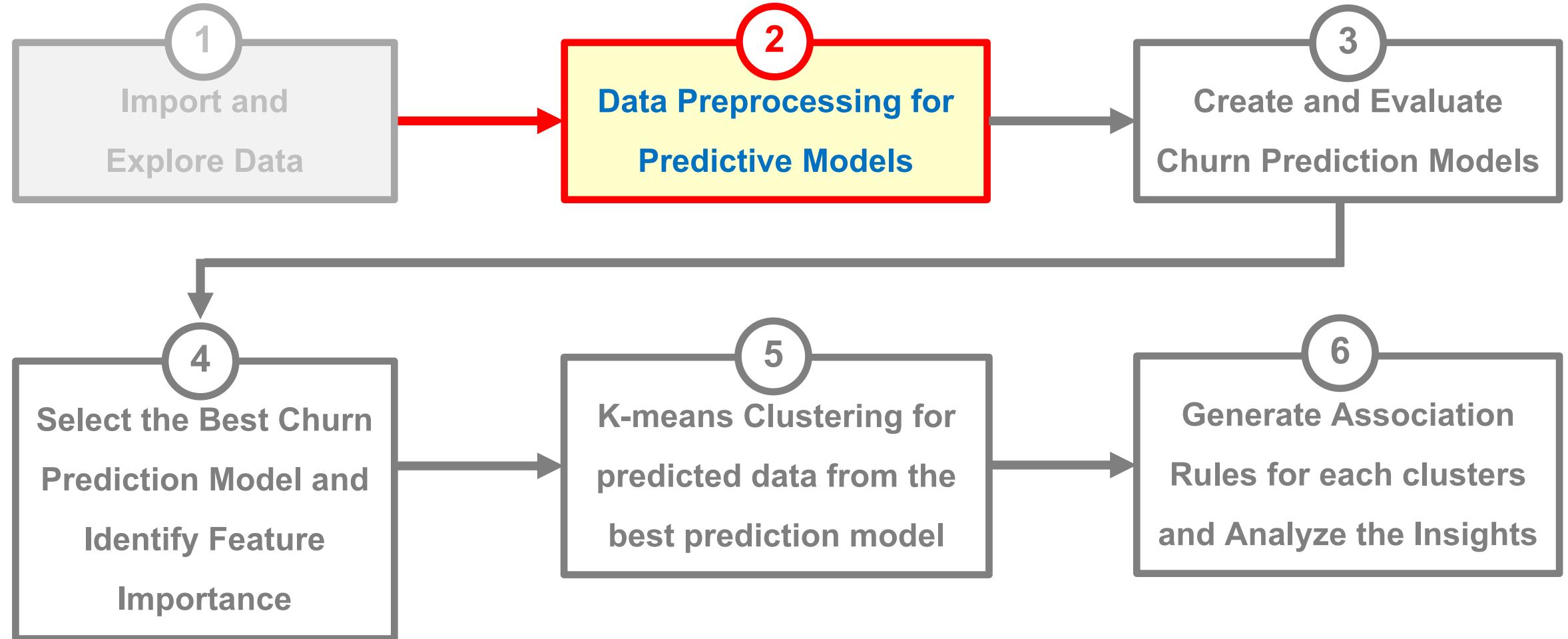
```
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 21 columns):
 #   Column           Non-Null Count  Dtype  
--- 
 0   customerID      7043 non-null    object  
 1   gender          7043 non-null    object  
 2   SeniorCitizen   7043 non-null    int64  
 3   Partner         7043 non-null    object  
 4   Dependents     7043 non-null    object  
 5   tenure          7043 non-null    int64  
 6   PhoneService    7043 non-null    object  
 7   MultipleLines   7043 non-null    object  
 8   InternetService 7043 non-null   object  
 9   OnlineSecurity  7043 non-null   object  
 10  OnlineBackup    7043 non-null   object  
 11  DeviceProtection 7043 non-null  object  
 12  TechSupport     7043 non-null   object  
 13  StreamingTV     7043 non-null   object  
 14  StreamingMovies 7043 non-null   object  
 15  Contract        7043 non-null   object  
 16  PaperlessBilling 7043 non-null  object  
 17  PaymentMethod   7043 non-null   object  
 18  MonthlyCharges 7043 non-null   float64 
 19  TotalCharges    7043 non-null   object  
 20  Churn           7043 non-null   object  
dtypes: float64(1), int64(2), object(18)
```

**Continuous (float)**

# Procedure

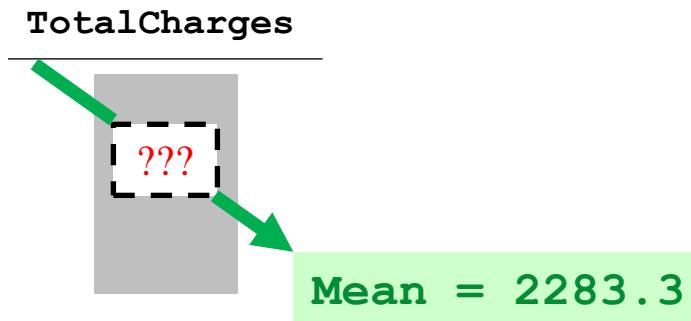


# Procedure



## Data Preprocessing for Predictive Models

### Data Cleansing



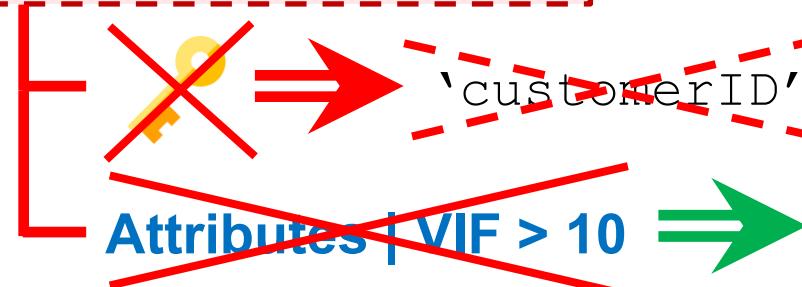
### Data Transformation

One-hot encoding  $\Rightarrow$  ABC  $\Rightarrow$  0 1

E.g. Contract = {Month-to-month, One year, Two year}

Contract	Contract_Month-to-month	Contract_One year	Contract_Two year
Month-to-month	1	0	0
One year	0	1	0
Two year	0	0	1

## Attribute Selection



[Only for Logistics Regression Model]

```
attribute_collinear =  
    ['MonthlyCharges', 'TotalCharges', 'gender_Female', 'gender_Male', 'Partner_No', 'Partner_res', 'Dependents_No',  
     'Dependents_Yes', 'PhoneService_No', 'PhoneService_Yes', 'MultipleLines_No', 'MultipleLines_No', 'phoneservice',  
     'MultipleLines_Yes', 'InternetService_DSL', 'InternetService_Fiber optic', 'InternetService_No', 'OnlineSecurity_No',  
     'OnlineSecurity_No', 'internet service', 'OnlineSecurity_Yes', 'OnlineBackup_No', 'OnlineBackup_No', 'internet service',  
     'OnlineBackup_Yes', 'DeviceProtection_No', 'DeviceProtection_No', 'internet service', 'DeviceProtection_Yes',  
     'TechSupport_No', 'TechSupport_No', 'internet service', 'TechSupport_Yes', 'StreamingTV_No',  
     'StreamingTV_No', 'internet service', 'StreamingTV_Yes', 'StreamingMovies_No', 'StreamingMovies_No', 'internet service',  
     'StreamingMovies_Yes', 'Contract_Month-to-month', 'Contract_One year', 'Contract_Two year', 'PaperlessBilling_No',  
     'PaperlessBilling_Yes', 'PaymentMethod_Bank transfer (automatic)', 'PaymentMethod_Credit card (automatic)',  
     'PaymentMethod_Electronic check', 'PaymentMethod_Mailed check']
```

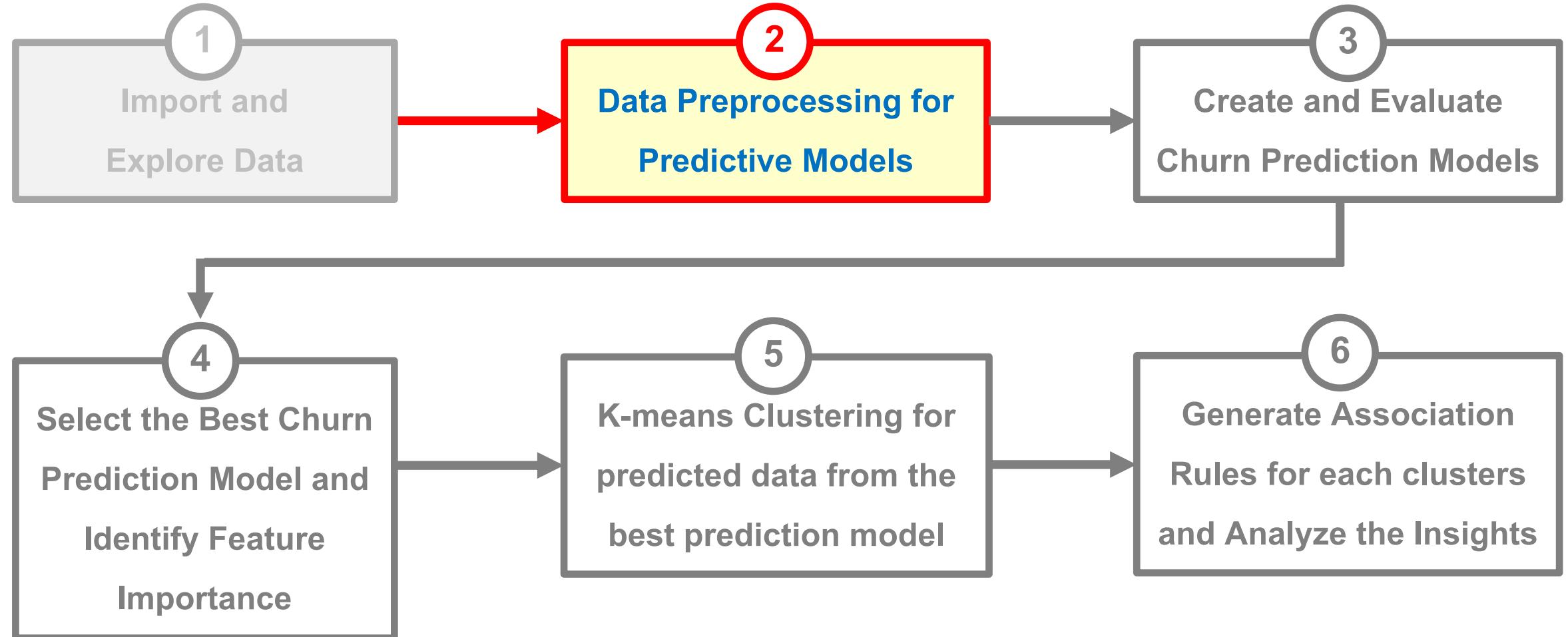
✓ 0s len(attribute\_collinear)

□ 43

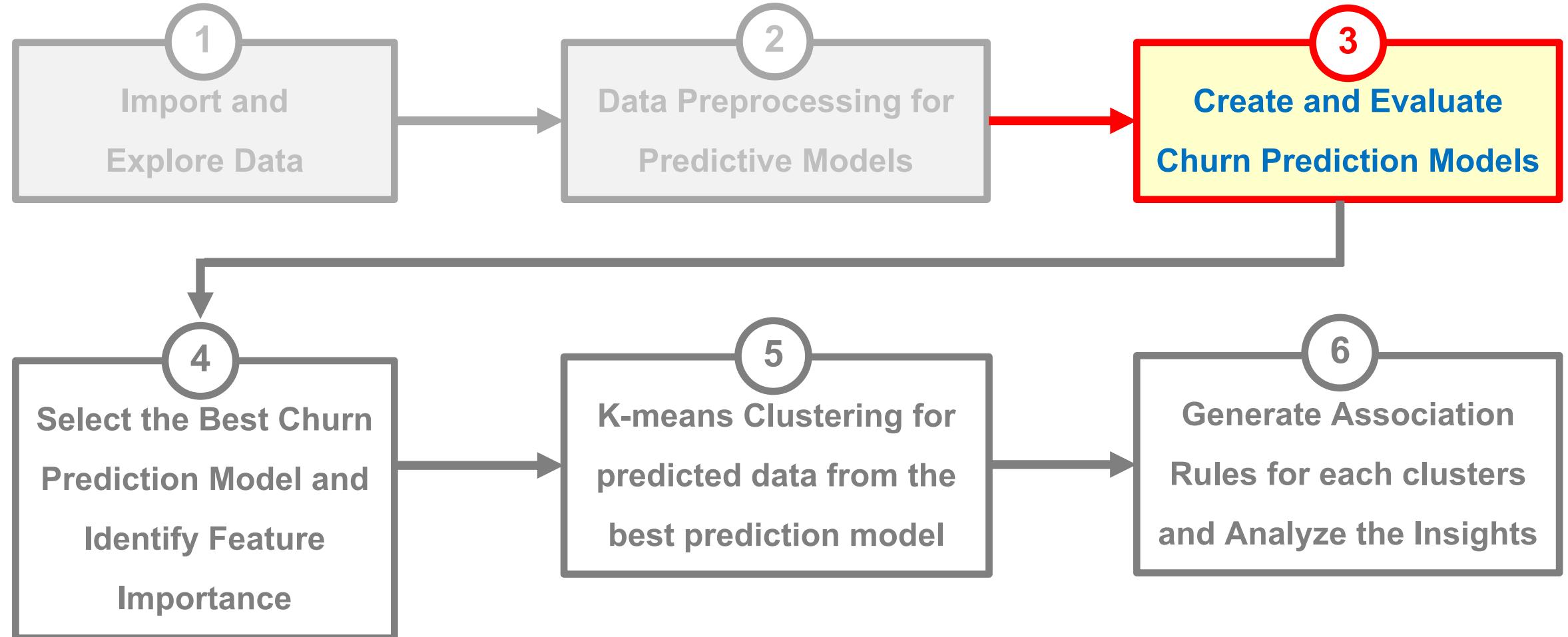
Logistics  
Regression  
Assumption



# Procedure



# Procedure

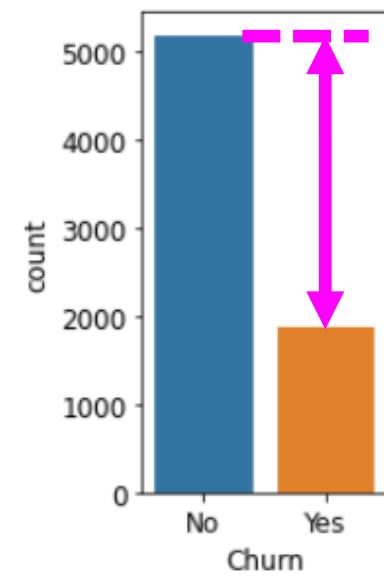
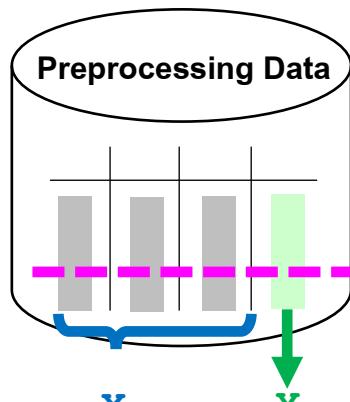


3

### Create and Evaluate Churn Prediction Models

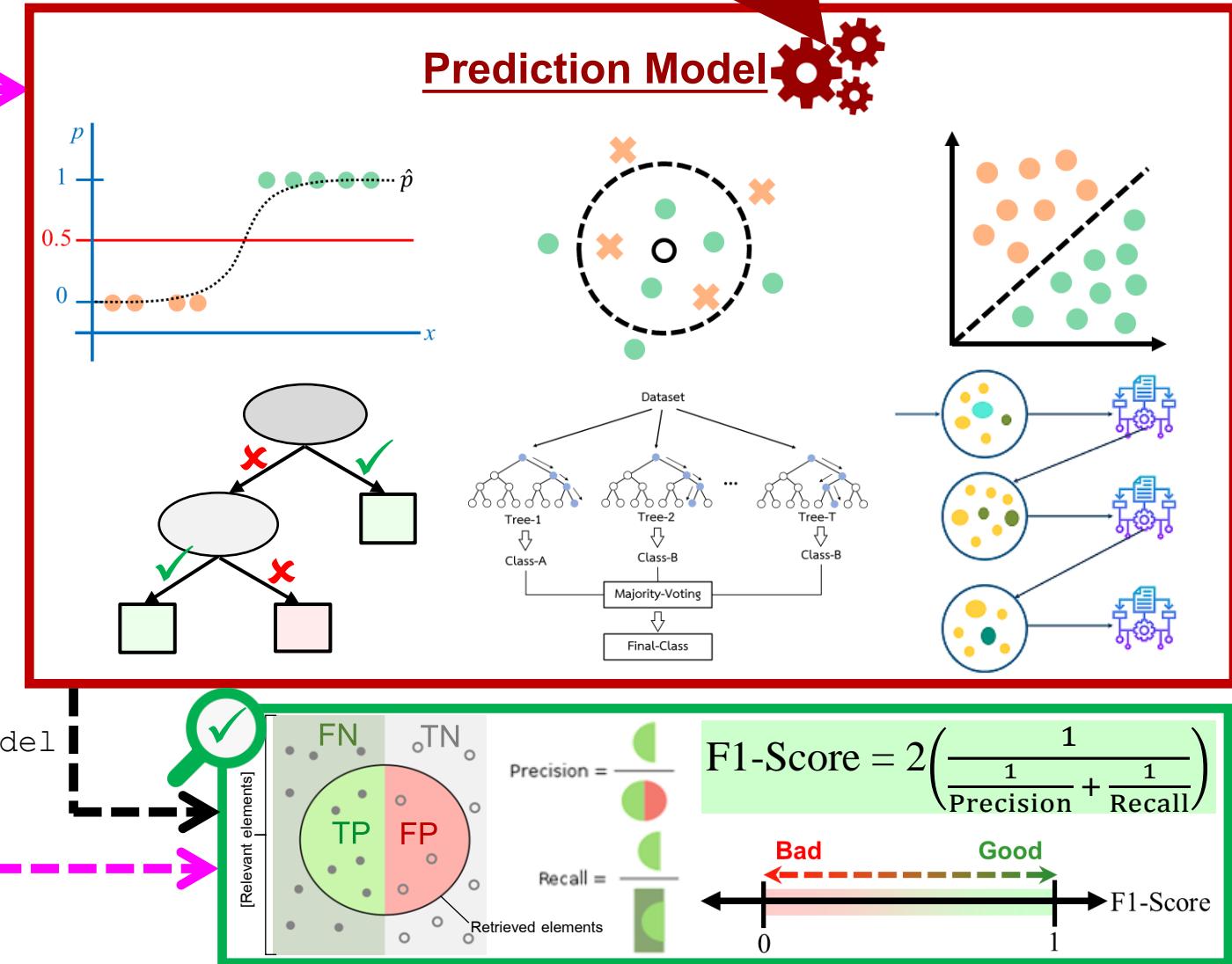
Hyperparameter Tuning with

```
.GridSearchCV(estimator, param_grid,
scoring = f1_score, n_jobs = -1, cv = 10)
```



80% Training set

20% Testing set



# F1-Score

Source: <https://towardsdatascience.com/micro-macro-weighted-averages-of-f1-score-clearly-explained-b603420b292f>

No	Actual	Predicted	Match
1	Airplane	Airplane	✓
2	Car	Boat	✗
3	Car	Car	✓
4	Car	Car	✓
5	Car	Boat	✗
6	Airplane	Boat	✗
7	Boat	Boat	✓
8	Car	Airplane	✗
9	Airplane	Airplane	✓
10	Car	Car	✓



		Predicted		
		Airplane	Boat	Car
Actual	Airplane	2	1	0
	Boat	0	1	0
	Car	1	2	3

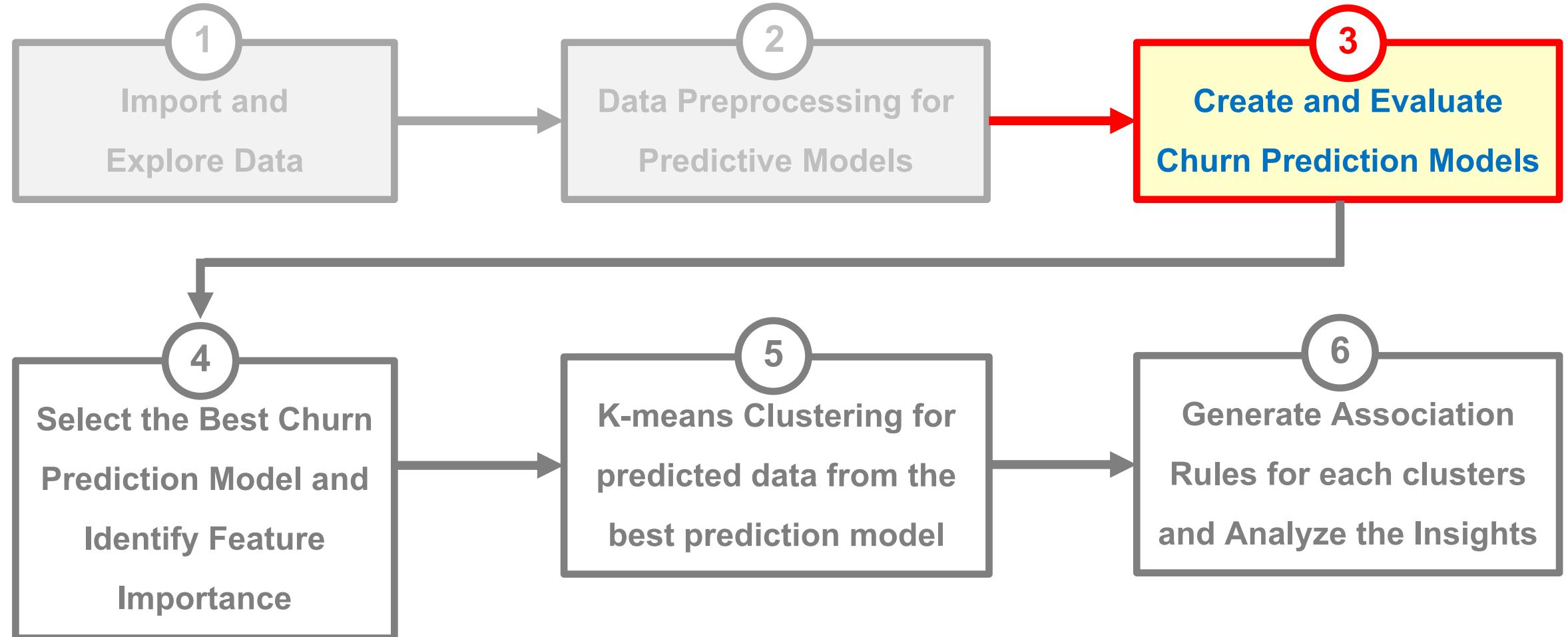
Label	True Positive (TP)	False Positive (FP)	False Negative (FN)	Precision	Recall	F1 Score
Airplane	2	1	1	0.67	0.67	$2 * (0.67 * 0.67) / (0.67 + 0.67) = 0.67$
Boat	1	3	0	0.25	1.00	$2 * (0.25 * 1.00) / (0.25 + 1.00) = 0.40$
Car	3	0	3	1.00	0.50	$2 * (1.00 * 0.50) / (1.00 + 0.50) = 0.67$

	precision	recall	f1-score	support
Aeroplane	0.67	0.67	0.67	3
Boat	0.25	1.00	0.40	1
Car	1.00	0.50	0.67	6
accuracy				
macro avg	0.64	0.72	0.60	10
weighted avg	0.82	0.60	0.64	10
↓ Average F1 scores				

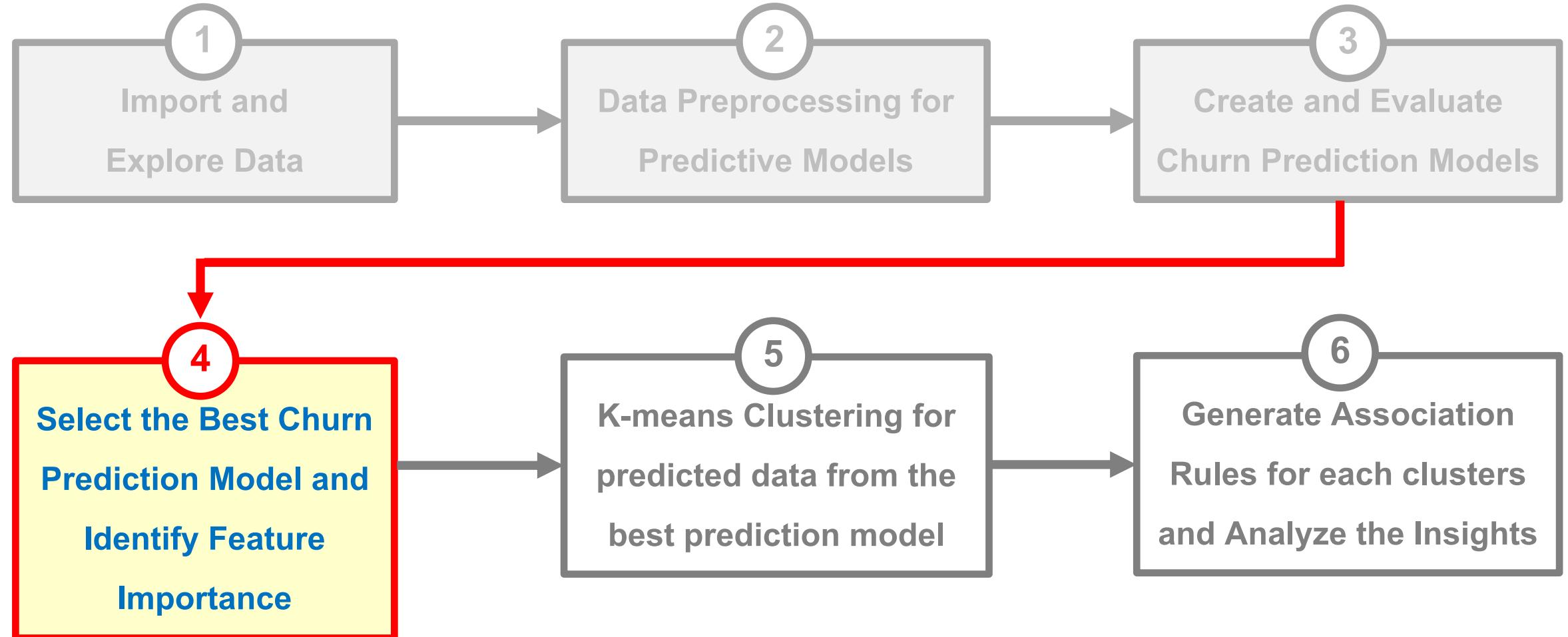
$$\begin{aligned} \text{Macro-Averaged F1 Score} &= \frac{0.67 + 0.40 + 0.67}{3} = 0.58 \\ \text{Weighted Average F1 Score} &= (0.67 * 0.3) + (0.40 * 0.1) + (0.67 * 0.6) = 0.64 \\ \text{Micro-Averaged F1 Score} &= \frac{\frac{TP}{TP+\frac{1}{2}(FP+FN)}}{6} = \frac{6}{6+\frac{1}{2}(4+4)} = 0.60 \end{aligned}$$



# Procedure



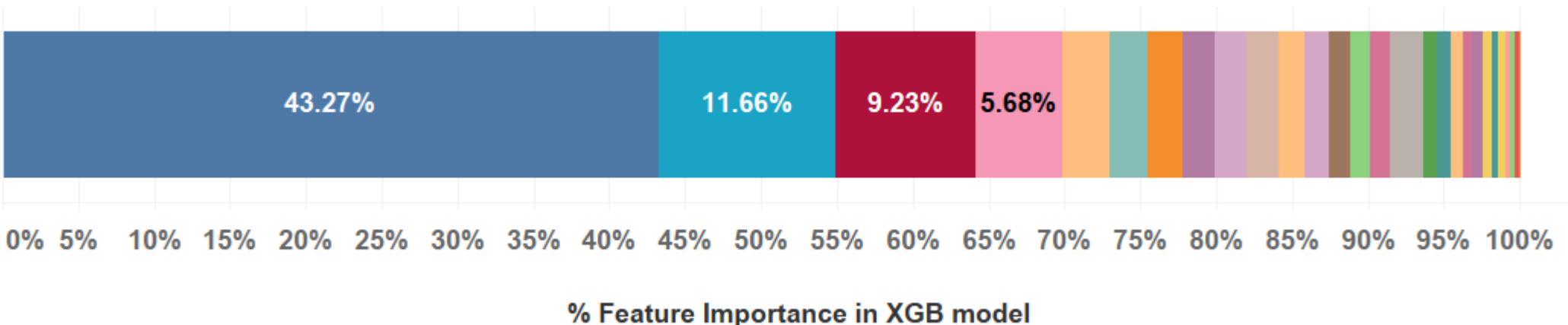
# Procedure



4

## Select the Best Churn Prediction Model and Identify Feature Importance

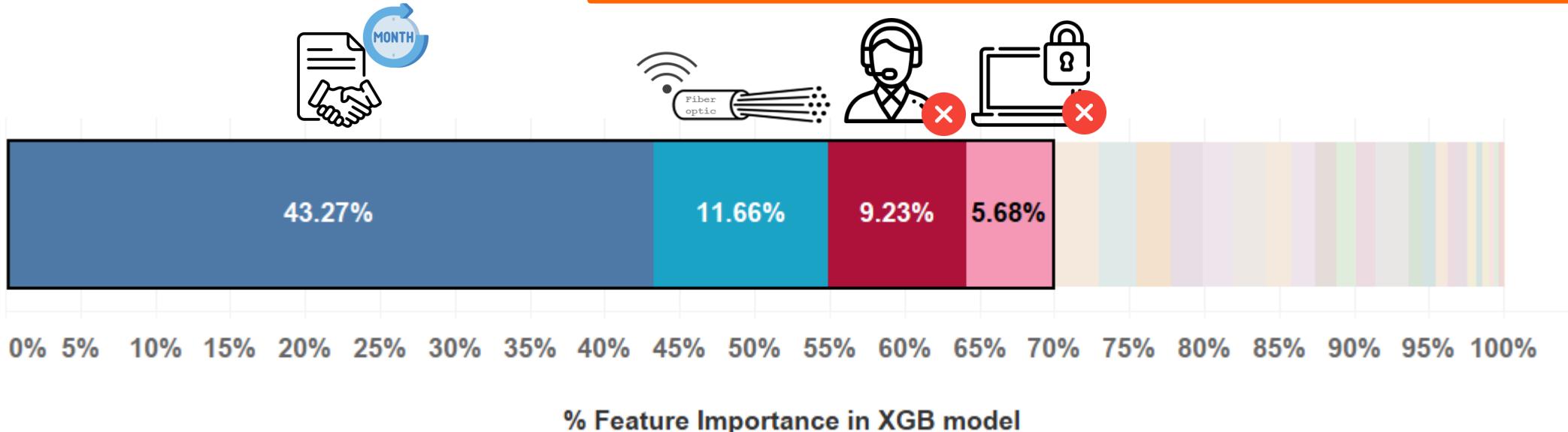
The Best Churn Prediction Model =



4

## Select the Best Churn Prediction Model and Identify Feature Importance

The Best Churn Prediction Model =



# Comparing Prediction Model Result

Prediction Model	Fine-tuned hyperparameter(s)	Total tuned combinations	Total runtime [seconds]	Average runtime per combination [seconds/combination]	Testing set performance	
					Accuracy	F1-score (macro)
Logistics Regression	C=0.01, penalty='none', solver='newton-cg'	100	47.486	0.475	74.521%	54.852%
Decision Tree	max_depth=6, min_samples_split=40	585	157.723	0.270	78.070%	71.359%
Random Forest	max_depth=10, min_samples_split=10, n_estimators=1000, max_features = 'auto'	480	4425.416	9.220	79.631%	71.396%
XGBoost	max_depth = 6, learning_rate = 0.3, subsample=1.0, tree method='auto'	180	1906.808	10.593	80.199%	72.657%
K-Nearest Neighbors	n_neighbors=21, weights='distance'	100	120.616	1.206	77.573%	67.478%
Support Vector Machine	'kernel'='rbf', C=100, gamma=0.0001	64	910.459	14.226	78.211%	69.788%

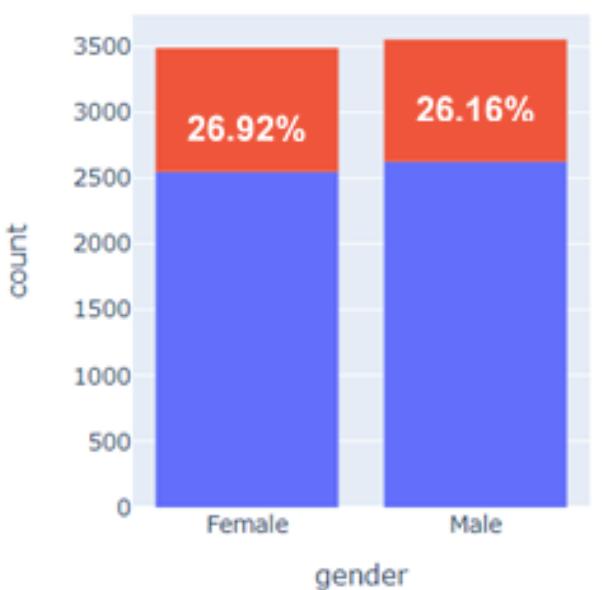
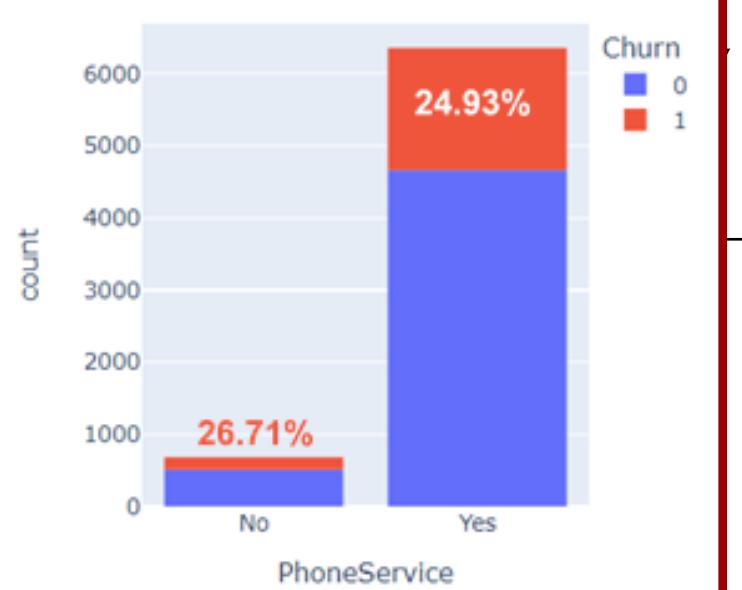
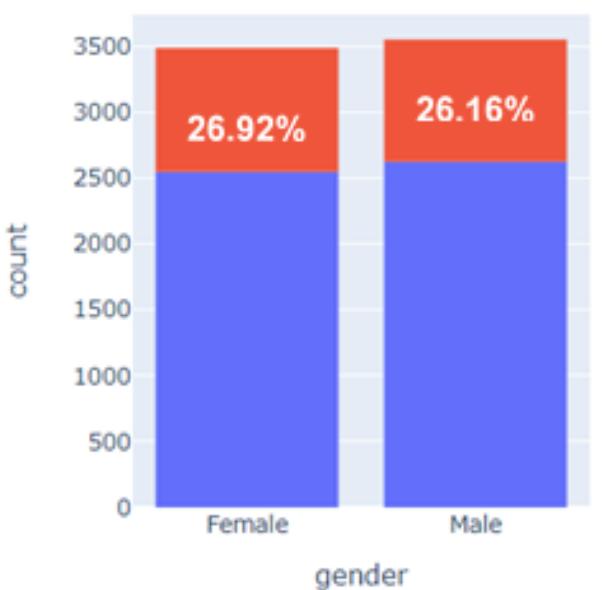
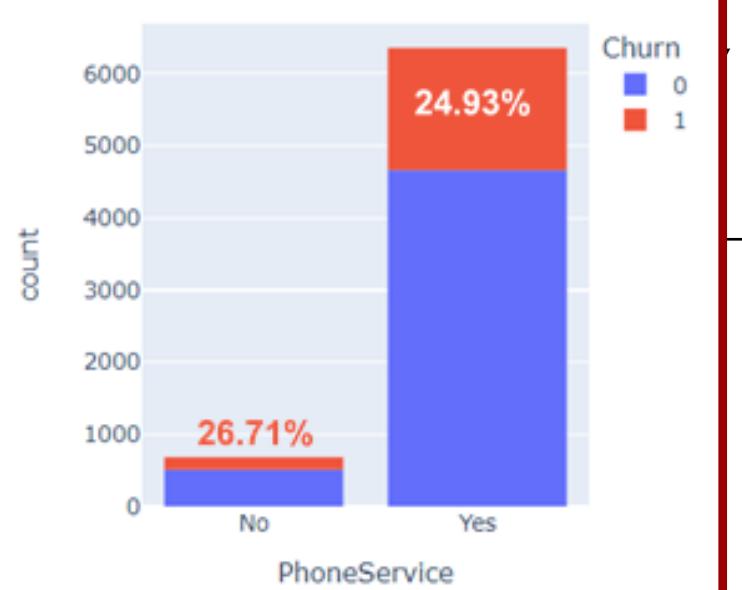


# Comparing Prediction Model with Kaggle user's

Rank	Title	Method	Algorithm	Author	Accuracy	F1-score
1	XGBoost & LightGBM & Catboost - Imbalanced Data	<ul style="list-style-type: none"> <li>- Manual hyperparameters tuning</li> <li>- Transform the data "No internet service" to "No" in the attribute, 'TechSupport', 'StreamingTV', 'StreamingMovies', 'OnlineSecurity', 'OnlineBackup', and 'DeviceProtection'</li> <li>- Drop the attributes, 'Gender' and 'Phone Service' from equivalent calculated churn rate within the attribute's class</li> </ul>	XGBClassifier(random_state=0)	 Kaan Boke	81.62%	75.22%
2	XGBoost Tuned With Random Search	<ul style="list-style-type: none"> <li>- Random Search with n_iter=10</li> <li>- Substitute missing data in 'TotalCharges' with 0</li> </ul>	XGBClassifier(colsample_bytree=0.6, gamma=2, learning_rate=0.02, max_depth=5, min_child_weight=10, n_estimators=200, nthread=-1, silent=True, subsample=0.6)	 Graeme Keleher	81.41%	74.81%
3	Customer Churn With Oversampling Techniques	<ul style="list-style-type: none"> <li>- Oversampling with BorderlineSMOTE</li> <li>- Substitute missing data in 'TotalCharges' with median</li> <li>- Drop the attribute 'CustomerID' and then eliminate the duplicate datasets</li> <li>- Min-max normalization</li> <li>- Eliminate the attribute which have VIF &gt; 10</li> <li>- Hyperparameter tuning with Random search</li> </ul>	XGBClassifier(learning_rate_init= 0.001, max_depth= 6, n_estimators= 300, n_jobs=-1, random_state = 0, min_child_weight= 2, colsample_bytree= 0.8, gamma= 2)	 Arezoo Dahesh	78.22%	73.14%



# Comparing Prediction Model with Kaggle user's

Rank	Title	Method	Algorithm	Author	Accuracy	F1-score								
1	XGBoost & LightGBM & Catboost - Imbalanced Data	<ul style="list-style-type: none"> <li>- Manual hyperparameters tuning</li> <li>- Transform the data "No internet service" to "No" in the attribute, 'TechSupport', 'StreamingTV', 'StreamingMovies', 'OnlineSecurity', 'OnlineBackup', and 'DeviceProtection'</li> <li>- Drop the attributes, 'Gender' and 'PhoneService' from equivalent calculated churn rate within the attribute's class</li> </ul>	XGBClassifier(random_state=0)	 Kaan Boke	81.62%	75.22%								
2	XGBoost Tuned With Random Search	<ul style="list-style-type: none"> <li>- Ra</li> <li>- Su</li> </ul>	 <p>Churn</p> <table border="1"> <tr> <td>0</td> <td>26.92%</td> </tr> <tr> <td>1</td> <td>26.16%</td> </tr> </table>  <p>Churn</p> <table border="1"> <tr> <td>0</td> <td>26.71%</td> </tr> <tr> <td>1</td> <td>24.93%</td> </tr> </table>	0	26.92%	1	26.16%	0	26.71%	1	24.93%	 Graeme Keleher	81.41%	74.81%
0	26.92%													
1	26.16%													
0	26.71%													
1	24.93%													
3	Customer Churn With Oversampling Techniques	<ul style="list-style-type: none"> <li>- Ov</li> <li>- Su</li> <li>- Dr</li> <li>- dup</li> <li>- Mi</li> <li>- El</li> <li>- Hy</li> </ul>	 	 Arezoo Dahesh	78.22%	73.14%								



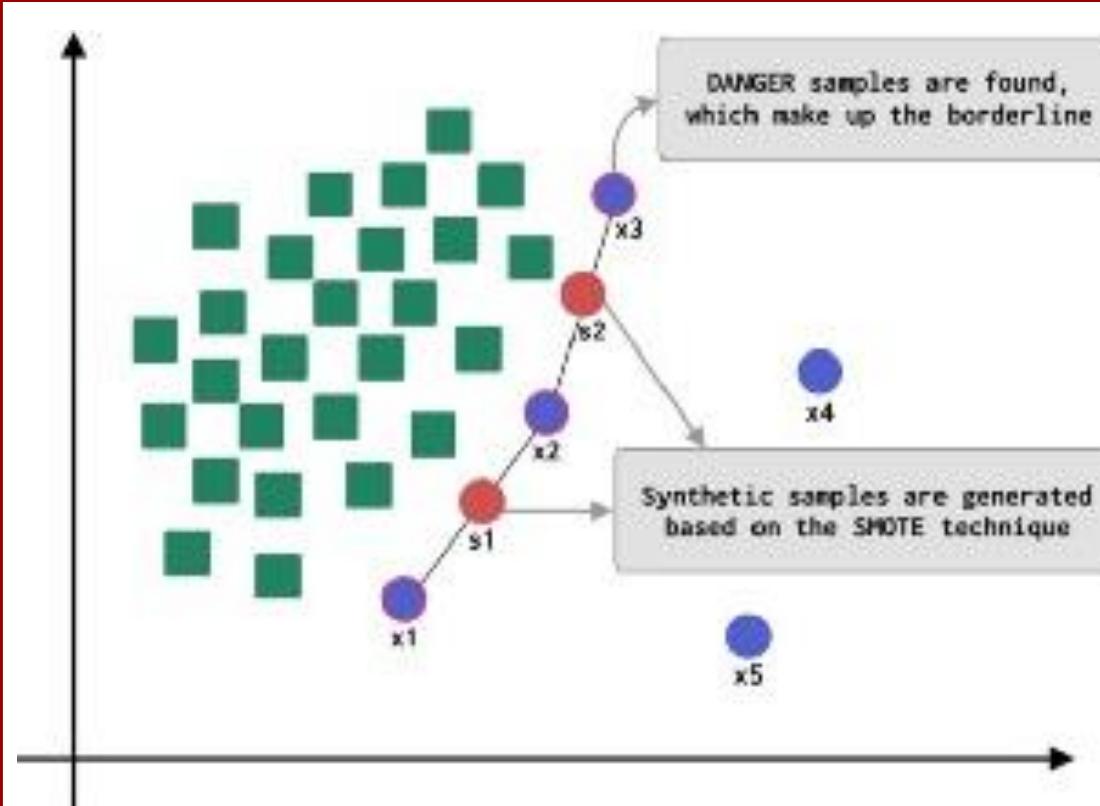
# Comparing Predictions

Rank	Title	Notes	Grid Search	Random Search	Accuracy	F1-score	
1	XGBoost & LightGBM & Catboost - Imbalanced Data	<ul style="list-style-type: none"> <li>- Manual hyperparameter tuning</li> <li>- Transform the dataset attribute, 'TechSupport', 'StreamingMovies', 'OnlineBackup'</li> <li>- Drop the attribute 'CustomerID' and then eliminate the duplicate datasets</li> <li>- Min-max normalization</li> <li>- Eliminate the attribute which have VIF &gt; 10</li> <li>- Hyperparameter tuning with Random search</li> </ul> <p>• Find absolute best way for tuned hyperparameter</p> <p>• Use exhaustive search (Full range of possibilities)</p> <p>= Time consuming + Chance for overfitting</p>		<p>sampled with fixed numbers of iterations given by n_iter</p>	<p>Source: <a href="https://betterprogramming.pub/comparing-grid-and-randomized-search-methods-in-python-81.62%-75.22%-cd9fe9c3572d">https://betterprogramming.pub/comparing-grid-and-randomized-search-methods-in-python-81.62%-75.22%-cd9fe9c3572d</a></p>		
2	XGBoost Tuned With Random Search	<ul style="list-style-type: none"> <li>- Random Search with n_iter=10</li> <li>- Substitute missing data in 'TotalCharges' with 0</li> </ul>		<pre>XGBClassifier(colsample_bytree=0.6, gamma=2, learning_rate=0.02, max_depth=5, min_child_weight=10, n_estimators=200, nthread=-1, silent=True, subsample=0.6)</pre>	<p>Graeme Keleher</p>	81.41%	74.81%
3	Customer Churn With Oversampling Techniques	<ul style="list-style-type: none"> <li>- Oversampling with BorderlineSMOTE</li> <li>- Substitute missing data in 'TotalCharges' with median</li> <li>- Drop the attribute 'CustomerID' and then eliminate the duplicate datasets</li> <li>- Min-max normalization</li> <li>- Eliminate the attribute which have VIF &gt; 10</li> <li>- Hyperparameter tuning with Random search</li> </ul>		<pre>XGBClassifier(learning_rate_init= 0.001, max_depth= 6, n_estimators= 300, n_jobs=-1, random_state = 0, min_child_weight= 2, colsample_bytree= 0.8, gamma= 2)</pre>	<p>Arezoo Dahesh</p>	78.22%	73.14%



# Comparing Predictive Models

Rank	Title	
1	XGBoost & LightGBM & Catboost Source: <a href="https://towardsdatascience.com/smote-synthetic-data-augmentation-for-imbalanced-data-1ce28090debc">https://towardsdatascience.com/smote-synthetic-data-augmentation-for-imbalanced-data-1ce28090debc</a>	<ul style="list-style-type: none"> <li>- Manual hyperparameter tuning</li> <li>- Transform the data using the 'get_dummies' attribute, 'TechSupport'</li> </ul>



(XGBClassifier(learning\_rate\_init= 0.001, max\_depth= 6, n\_estimators= 300, n\_jobs=-1, random\_state = 0, min\_child\_weight= 2, colsample\_bytree= 0.8, gamma= 2))

Reference: Nagarajan, 2017, p.9)

2	XGBoost Tuned With Random Search	<ul style="list-style-type: none"> <li>- Random Search with grid search</li> <li>- Substitute missing data</li> </ul>
---	----------------------------------	---

3	Customer Churn With Oversampling Techniques <ul style="list-style-type: none"> <li>- Oversampling with BorderlineSMOTE</li> <li>- Substitute missing data in 'TotalCharges' with median</li> <li>- Drop the attribute 'CustomerID' and then eliminate the duplicate datasets</li> <li>- Min-max normalization</li> <li>- Eliminate the attribute which have VIF &gt; 10</li> <li>- Hyperparameter tuning with Random search</li> </ul>	<pre>XGBClassifier(learning_rate_init= 0.001, max_depth= 6, n_estimators= 300, n_jobs=-1, random_state = 0, min_child_weight= 2, colsample_bytree= 0.8, gamma= 2)</pre>	<p>Arezoo Dahesh</p>	78.22%	73.14%
---	--	---	----------------------	--------	--------

Generate synthetic data by considering only samples that make up the border that divides one class from another

- Get higher accuracy than SMOTE based on border decision regions
- Tend to evade some important (extreme) examples in the minority class → bias in model

(Nagarajan, 2017, p.9)



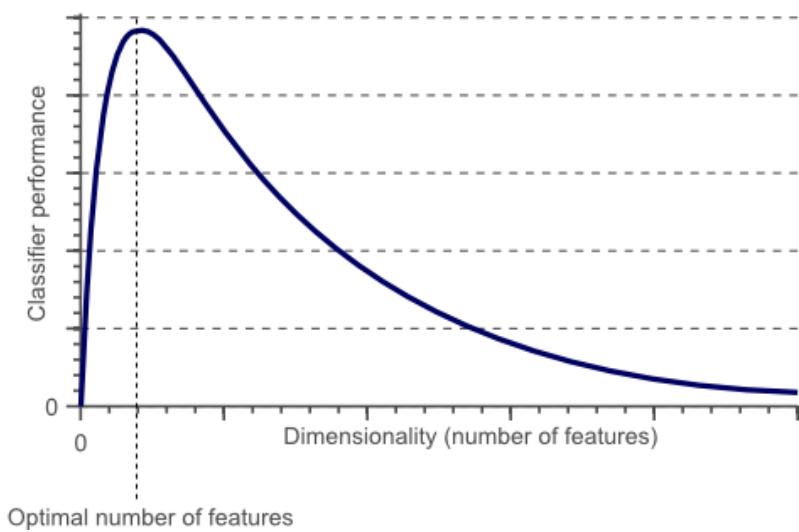
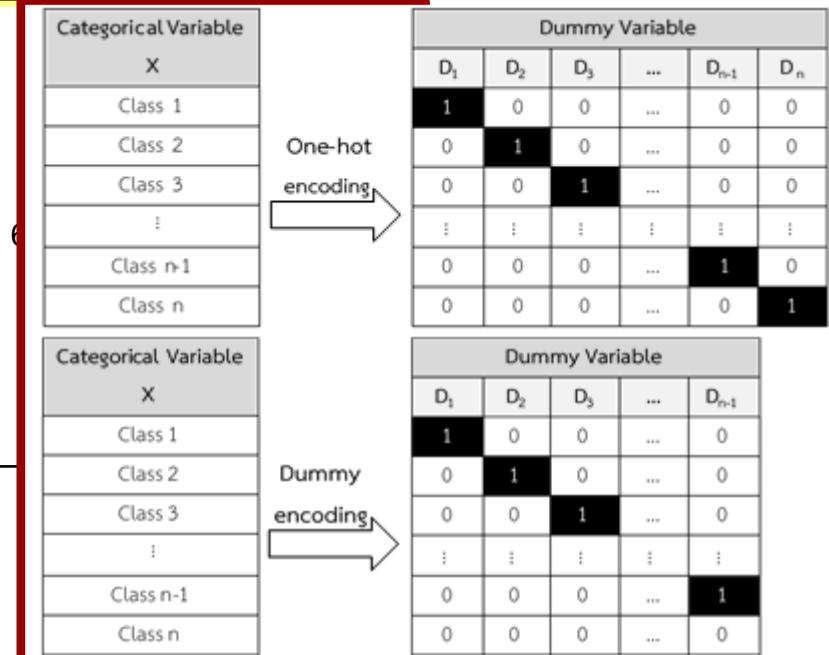
# Comparing Prediction Model with Kaggle users'

Rank	Title	Method	Algorithm	Author	Accuracy	F1-score
4	Senior Project	<ul style="list-style-type: none"> <li>- One-hot encoding for categorical attributes</li> <li>- Hyperparameter tuning with GridSearchCV()</li> <li>- Substitute missing data in 'TotalCharges' with mean</li> </ul>	XGBClassifier(max_depth = 6, learning_rate = 0.3, subsample=0.5, tree_method='hist')	Thatchakarn, Teethavat	79.77%	72.46%
5	Telco Customer Churn Prediction	<ul style="list-style-type: none"> <li>- Substitute missing data in 'TotalCharges' with mean</li> <li>- Dummy encoding for categorical attributes</li> </ul>	RandomForestClassifier()	 Iching Wang	79.13%	71.79%
6	Customer Churn Prediction Using ANN	<ul style="list-style-type: none"> <li>- Eliminate the datasets having a missing value</li> <li>- Transform class "No internet service" to "No" in the attribute, 'TechSupport', 'StreamingTV', 'StreamingMovies', 'OnlineSecurity', 'OnlineBackup', and 'DeviceProtection'</li> <li>- Transform class "No Phoneservice" to "No" in the attribute, 'MultipleLines'</li> <li>- Min-max normalization</li> </ul>	<pre>model = keras.Sequential([     keras.layers.Dense(27,                       input_shape=(27,), activation='relu'),     keras.layers.Dense(15,                       activation='relu'),     keras.layers.Dense(1,                       activation='sigmoid'))]  model.compile(optimizer='adam',               loss='binary_crossentropy',               metrics=['accuracy'])  model.fit(X_train, y_train, epochs=100)</pre>	 Jatin	79.25%	70.98%



# Comparing Prediction Model with Kaggle users'

Rank	Title	Method	Algorithm	Author	Accuracy	F1-score
4	Senior Project	<ul style="list-style-type: none"> <li>- One-hot encoding for categorical attributes</li> <li>- Hyperparameter tuning with GridSearchCV()</li> <li>- Substitute missing data in 'TotalCharges' with mean</li> </ul>	XGBClassifier(max_depth = 6, learning_rate = 0.3, subsample=0.5, tree_method='hist')	Thatchakarn, Teethavat	79.77%	72.46%
5	Telco Customer Churn Prediction	<ul style="list-style-type: none"> <li>- Substitute missing data in 'TotalCharges' with mean</li> <li>- Dummy encoding for categorical attributes</li> </ul>	RandomForestClassifier()	 Iching Wang	79.13%	71.79%



Curse of Dimensionality

*“As the dimensionality increases, the classifier’s performance increases until the optimal number of features is reached. Further increasing the dimensionality without increasing the number of training samples results in a decrease in classifier performance.”*



# Comparing

Rank	Title	
4	Senior Project	- - -
5	Telco Customer Churn Prediction	- S - D

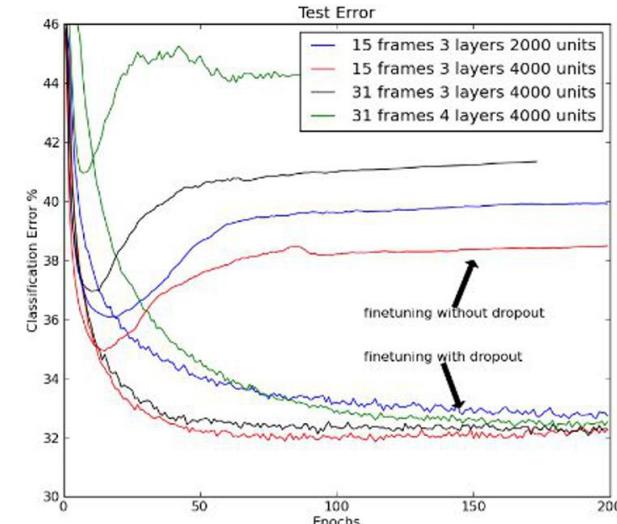
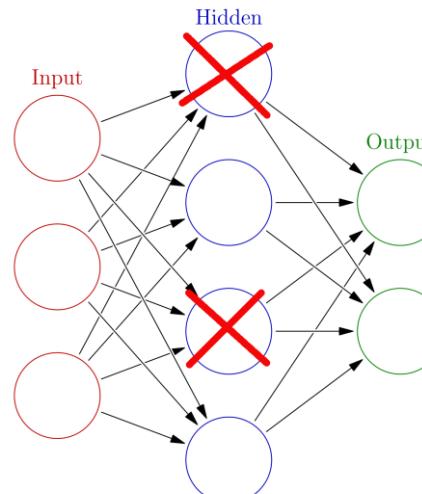
No regularization = Lack of generalization = Tend to overfit ↑

L1 regularization: Regularizing term is a sum

- Original loss +  $C \sum |w_i|$

L2 regularization: Regularizing term is a sum of squares

- Original loss +  $C \sum w_i^2$



- Eliminate the datasets having a missing value
- Transform class "No internet service" to "No" in the attribute, 'TechSupport', 'StreamingTV', 'StreamingMovies', 'OnlineSecurity', 'OnlineBackup', and 'DeviceProtection'
- Transform class "No PhoneService" to "No" in the attribute, 'MultipleLines'
- Min-max normalization

```
model = keras.Sequential([
    keras.layers.Dense(27,
    input_shape=(27,), activation='relu'),
    keras.layers.Dense(15,
    activation='relu'),
    keras.layers.Dense(1,
    activation='sigmoid'))]

model.compile(optimizer='adam',
    loss='binary_crossentropy',
    metrics=['accuracy'])

model.fit(X_train, y_train, epochs=100)
```



Jatin

79.25%

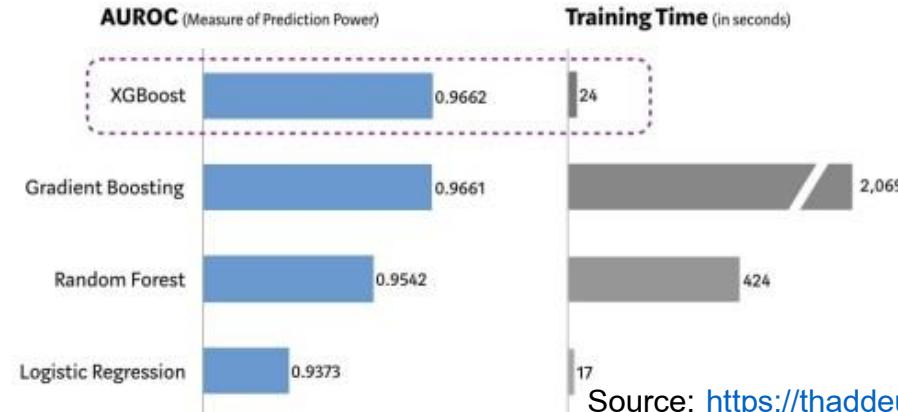
70.98%



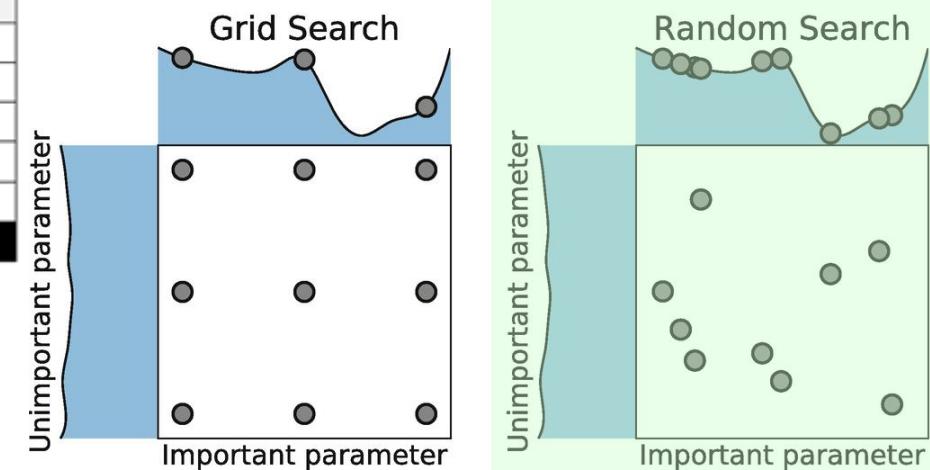
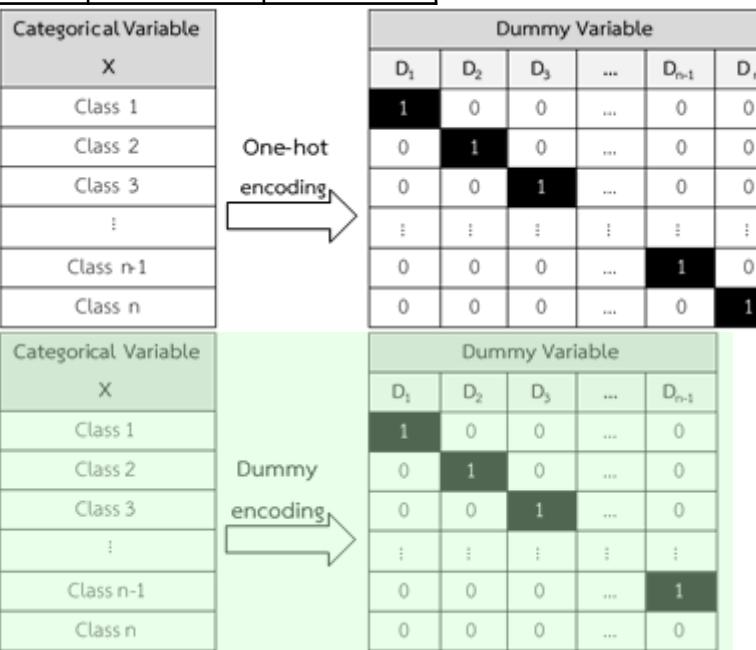
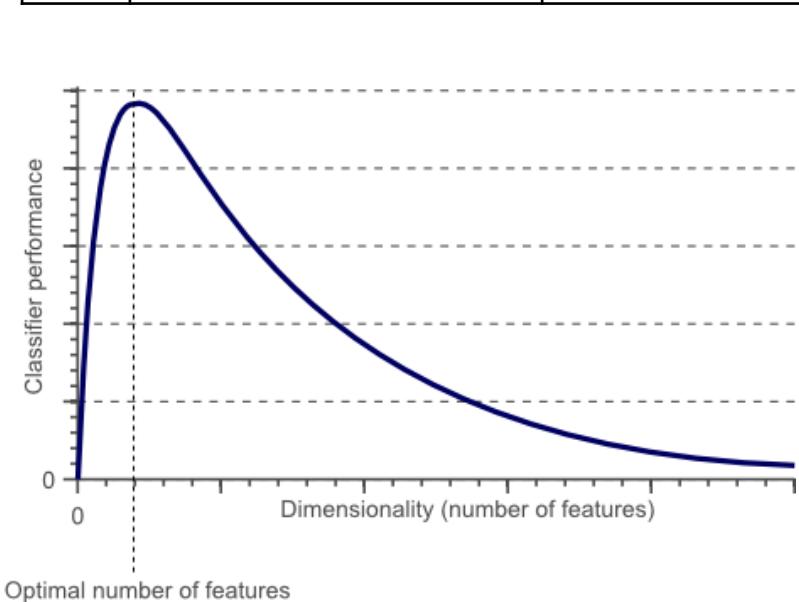
# Discussing Result from model comparison with Kaggle users'

Rank	Algorithm	Author	Accuracy	F1-score
1	XGB Classifier	Kaan Boke	81.62%	75.22%
2	XGB Classifier	Graeme Keleher	81.41%	74.81%
3	XGB Classifier	Arezoo Dahesh	78.22%	73.14%
4	XGB Classifier	Thatchakarn, Teethavat	79.77%	72.46%
5	RandomForest Classifier	Iching Wang	79.13%	71.79%
6	Neural Network: Sequence Model (RNN)	Jatin	79.25%	70.98%

Performance Comparison using SKLearn's 'Make\_Classification' Dataset  
(5 Fold Cross Validation, 1MM randomly generated data sample, 20 features)



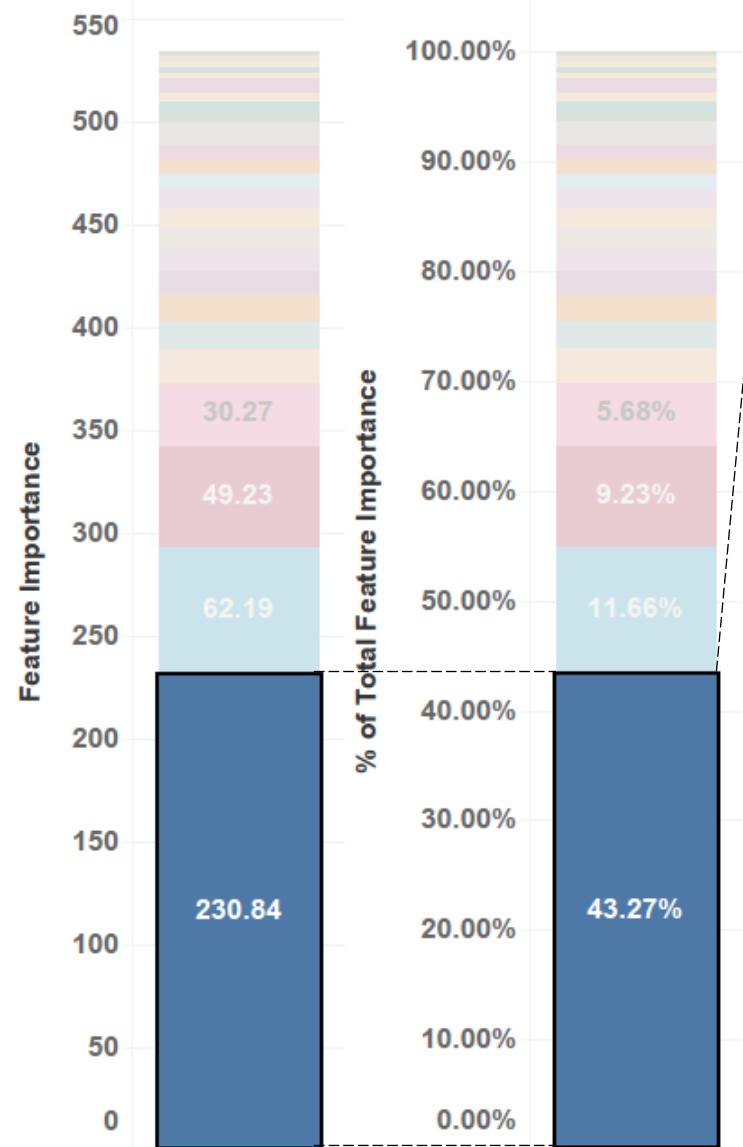
Source: <https://thaddeus-segura.com/xgboost/>



# Feature Importance



# Feature Importance



Feature Importance

% of Total Feature Importance

80%  
70.00%  
60.00%  
50.00%  
40.00%  
30.00%  
20.00%  
10.00%  
0.00%

60.72%  
39.28%

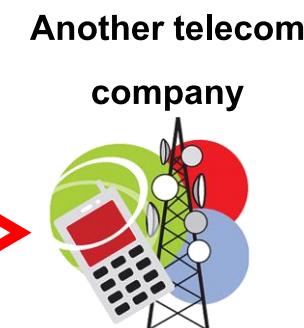
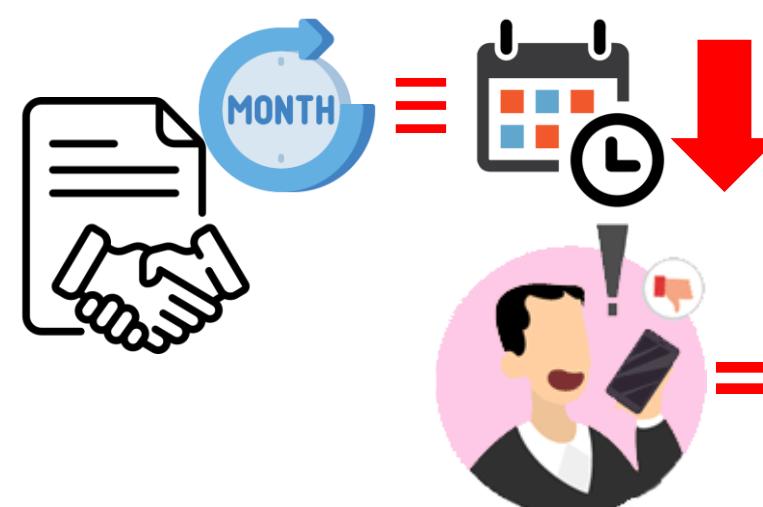
Month-to-month

## Contract



Predicted Churn

- Loyal
- Churn



Another telecom  
company

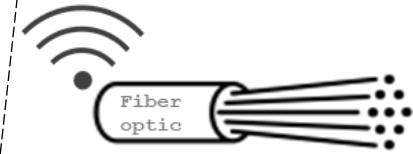
IBM Watson



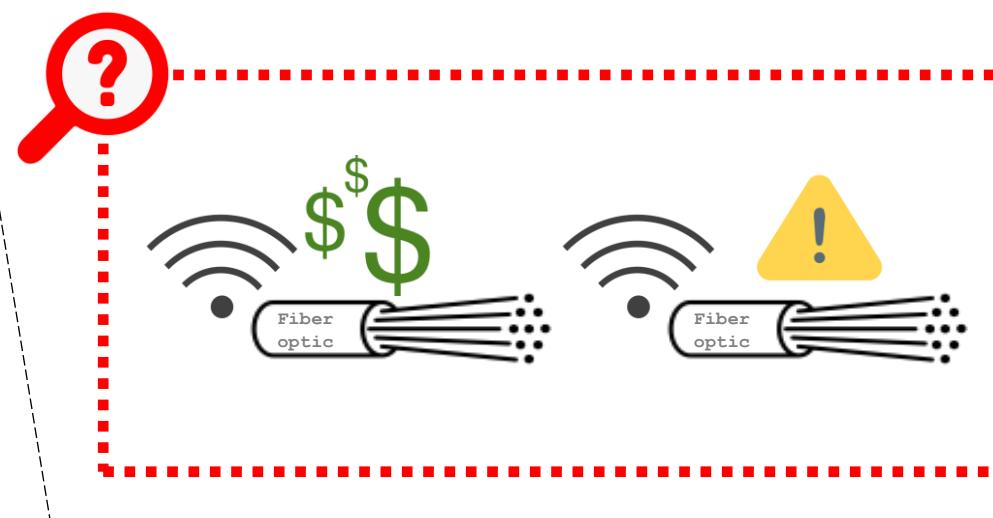
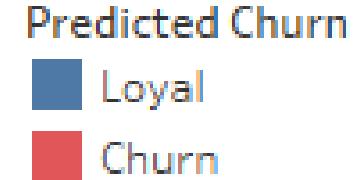
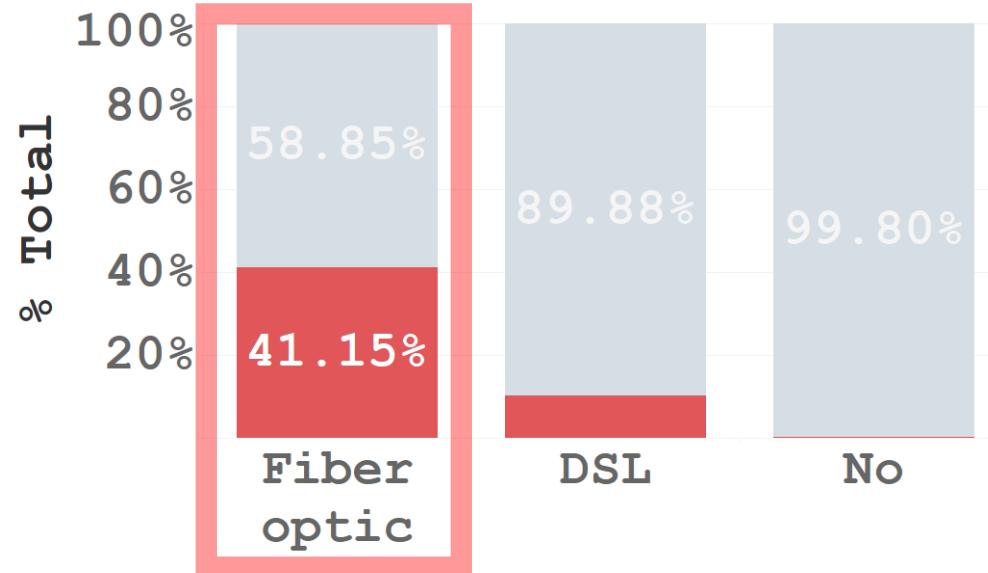
Shift



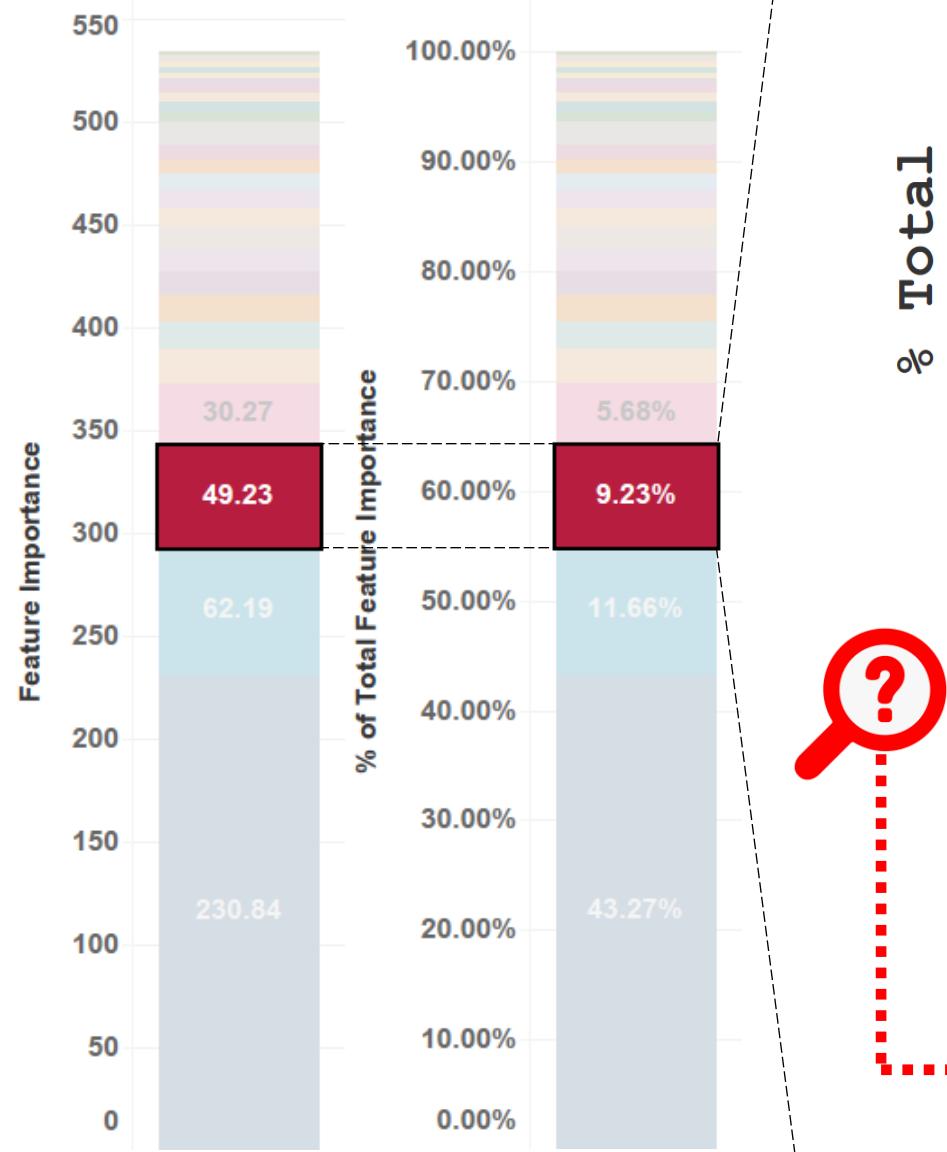
# Feature Importance



## Internet Service

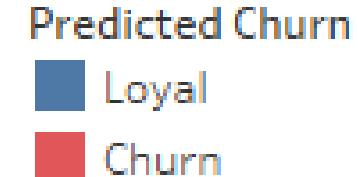
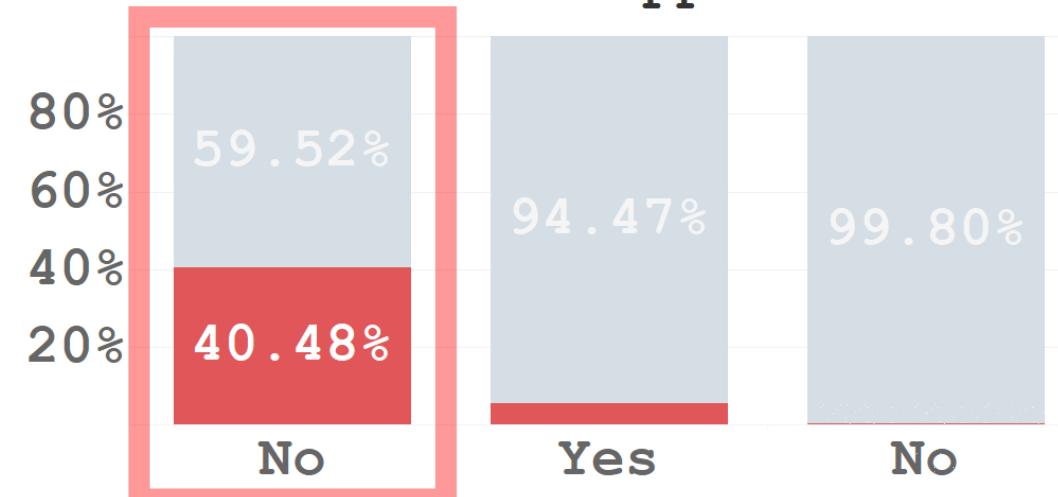


# Feature Importance

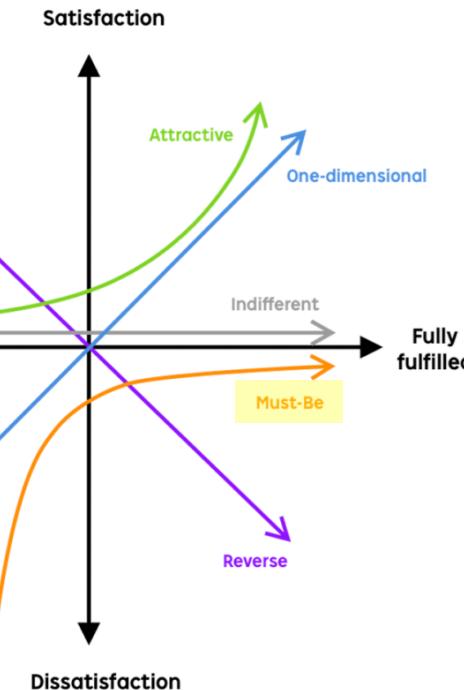
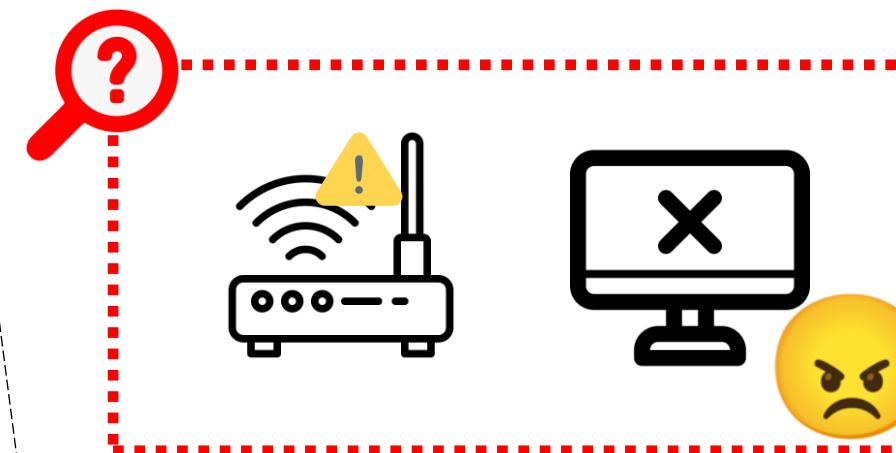


Total  
%

## Tech Support



No  
internet  
service



# Feature Importance

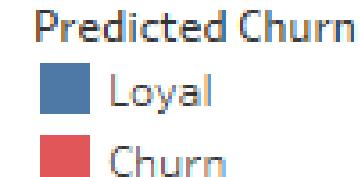
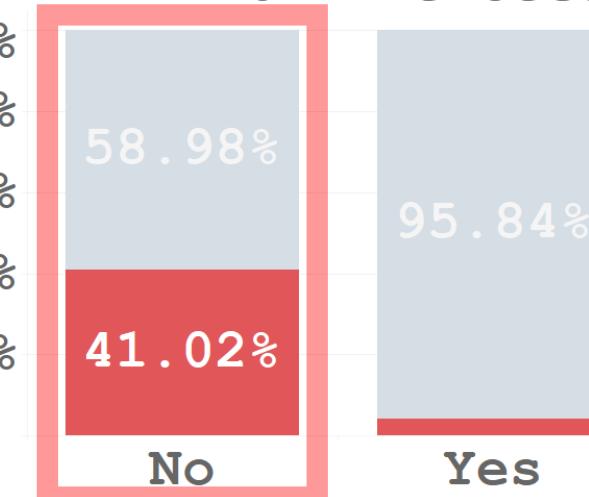


Feature Importance

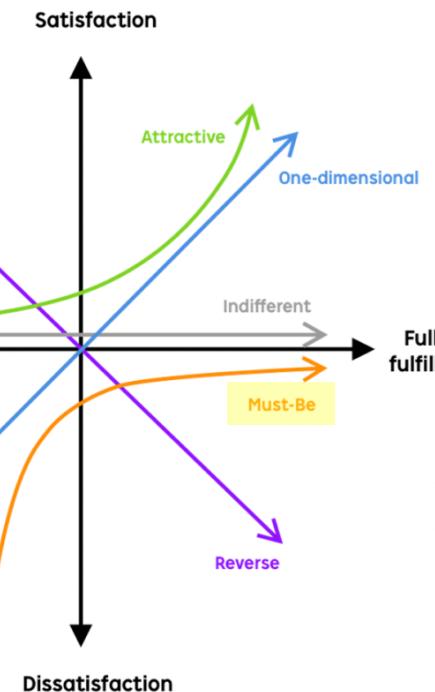
% of Total Feature Importance

## Online Security

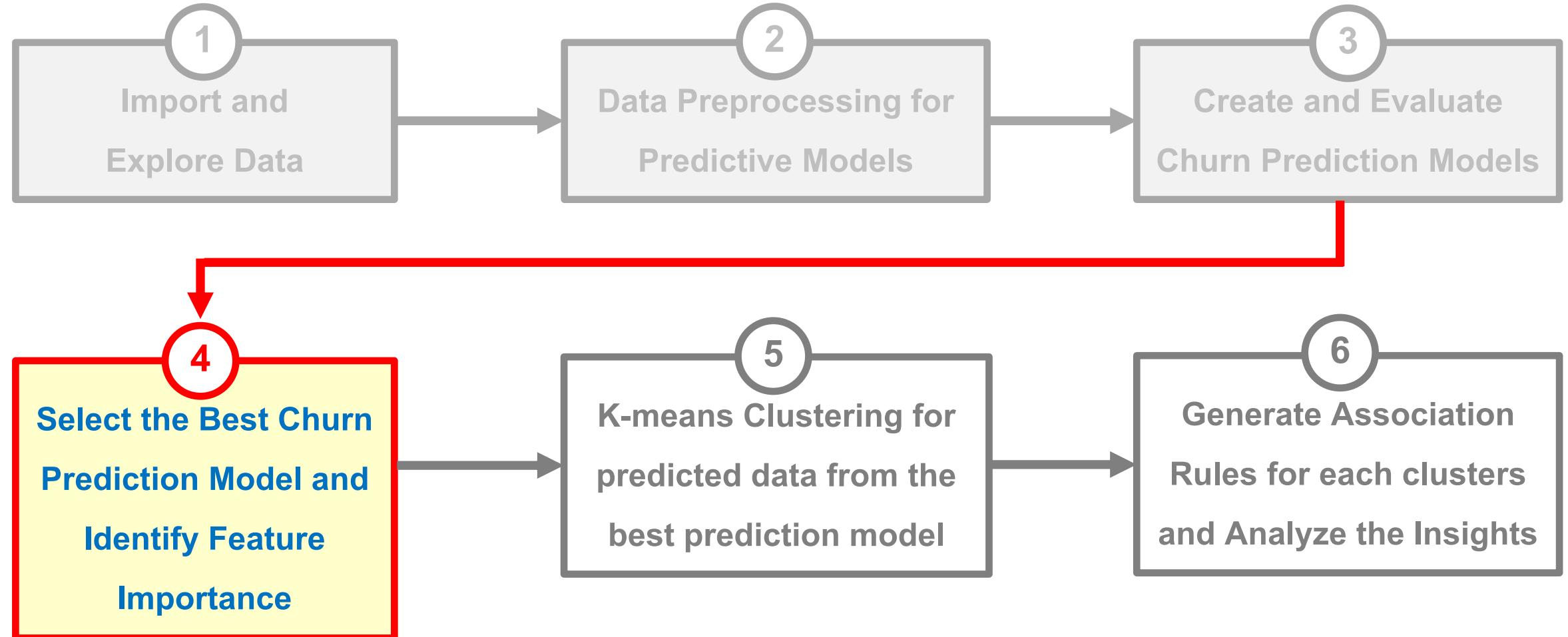
% Total



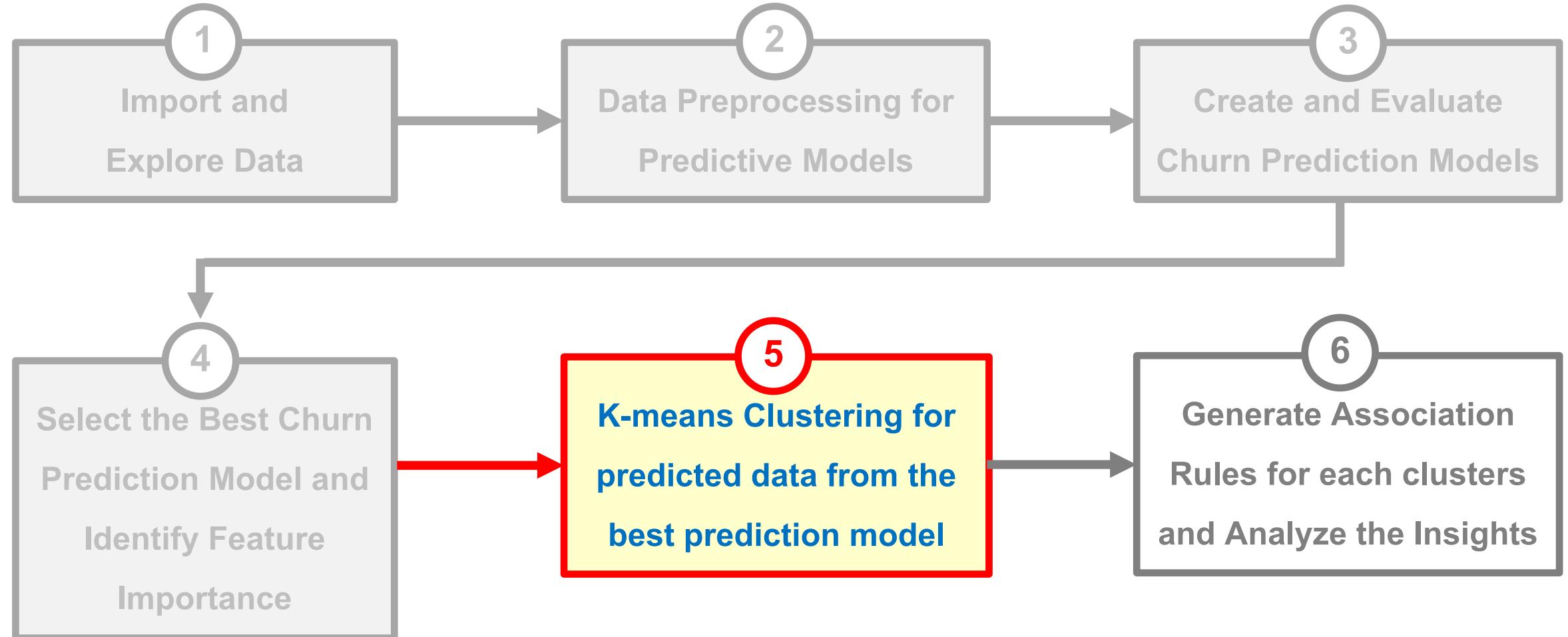
No  
internet  
service



# Procedure

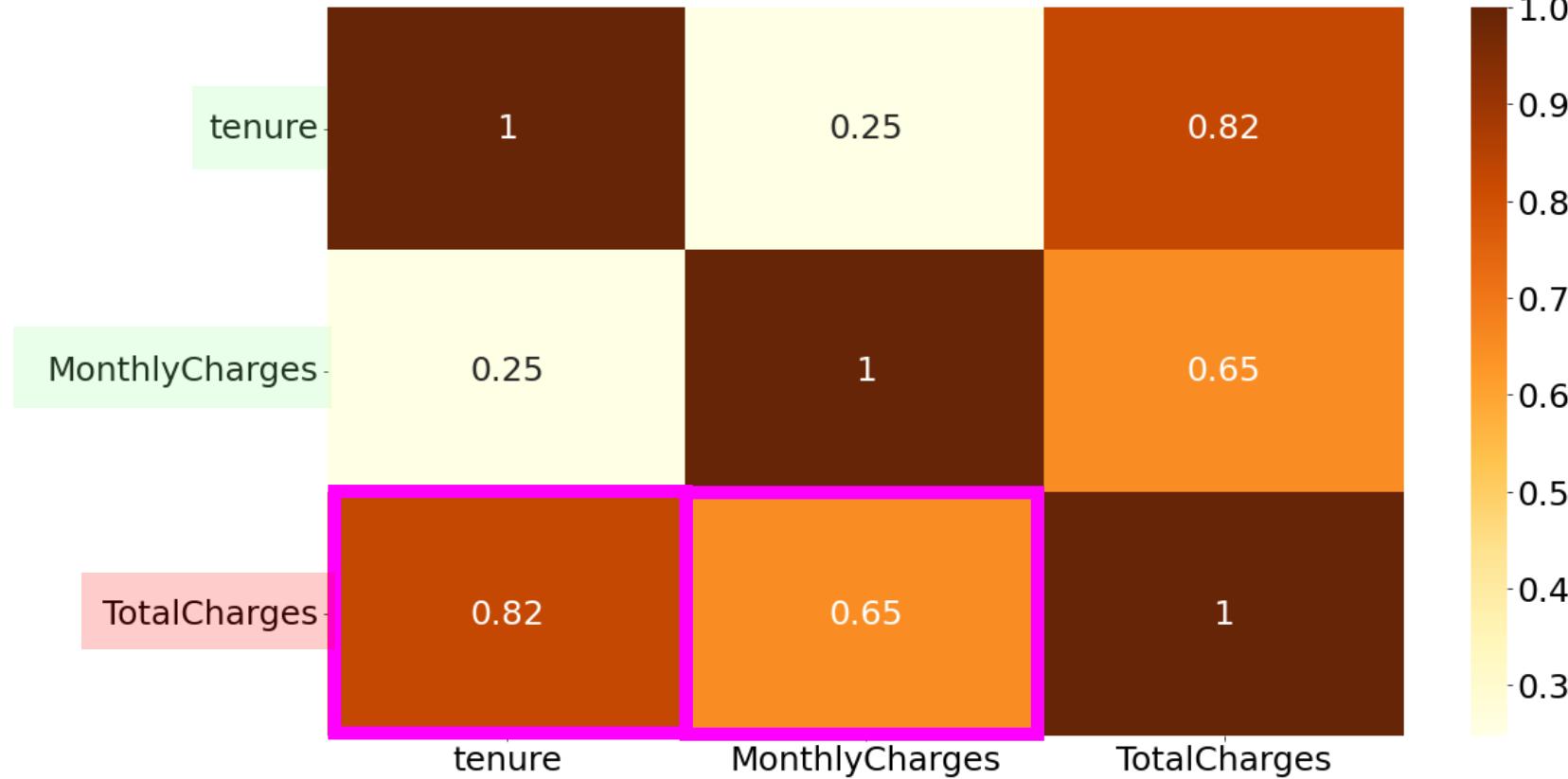
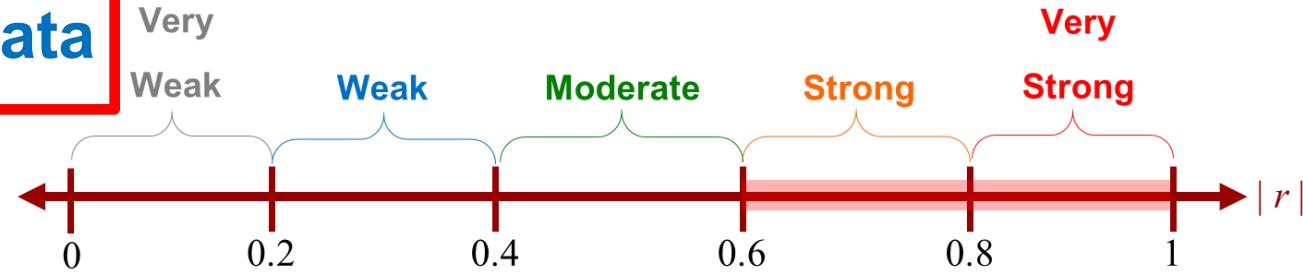


# Procedure

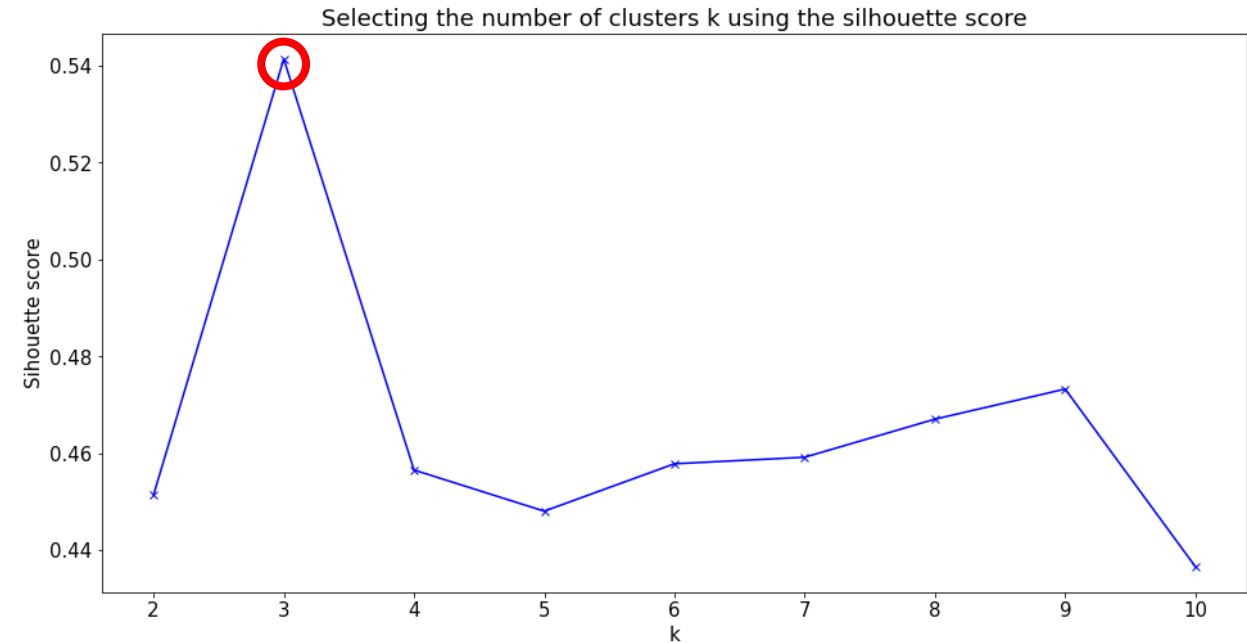
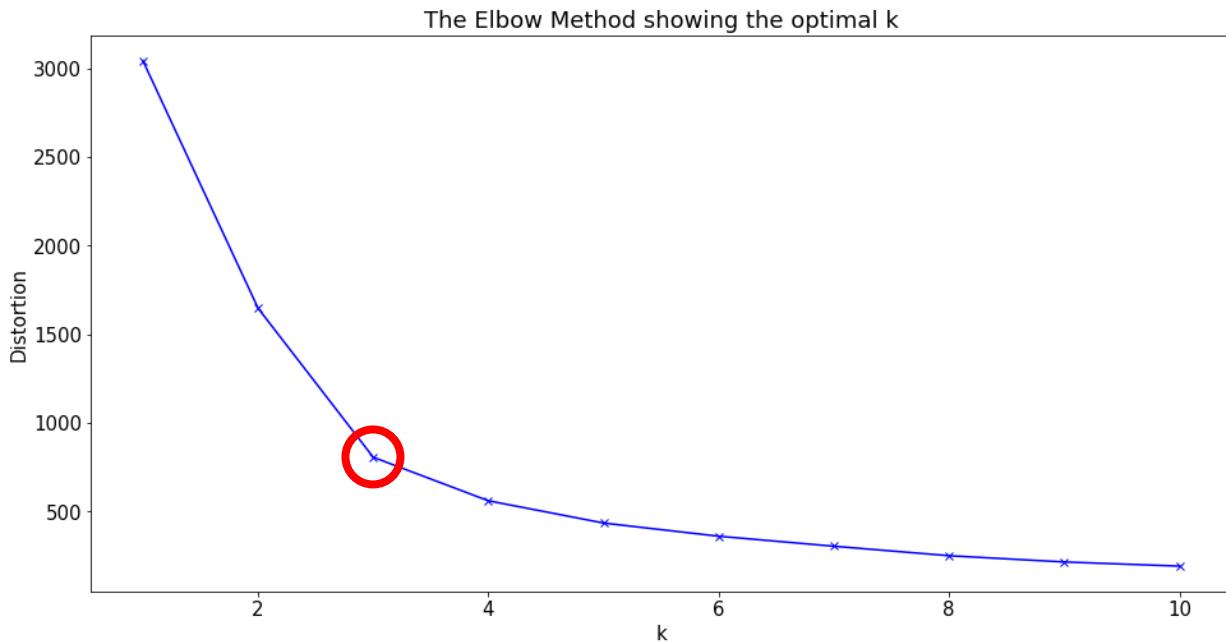


## 5 K-means Clustering for predicted data

### Attribute Selection

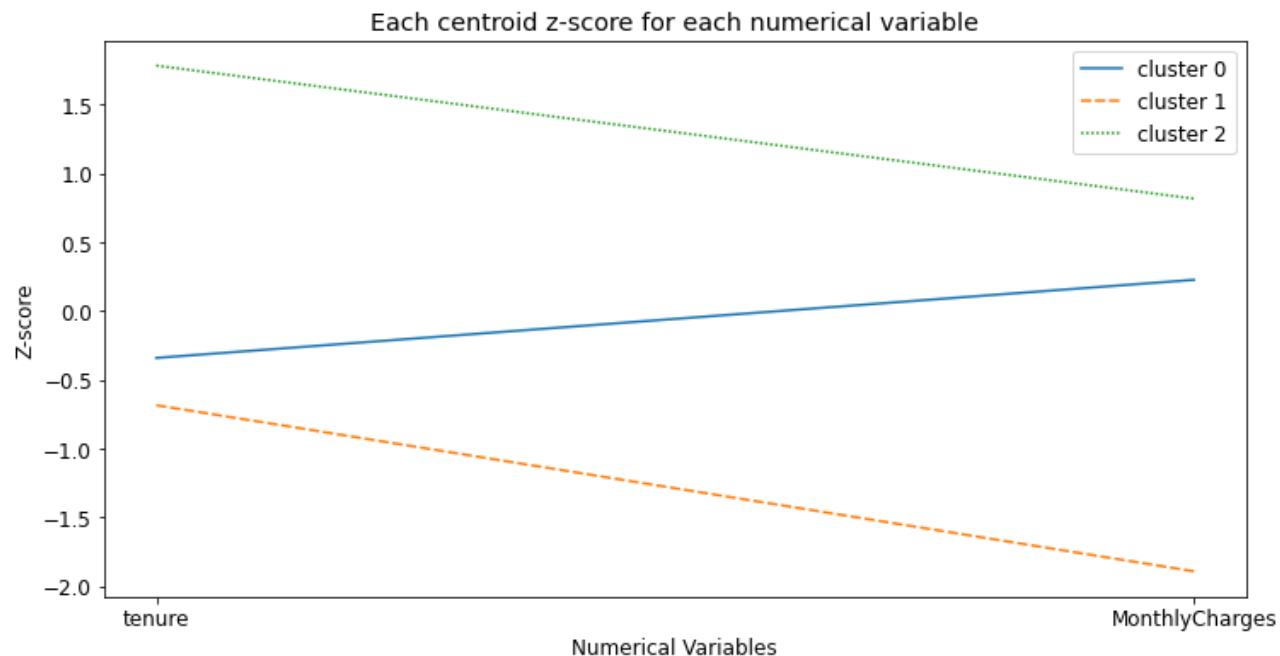
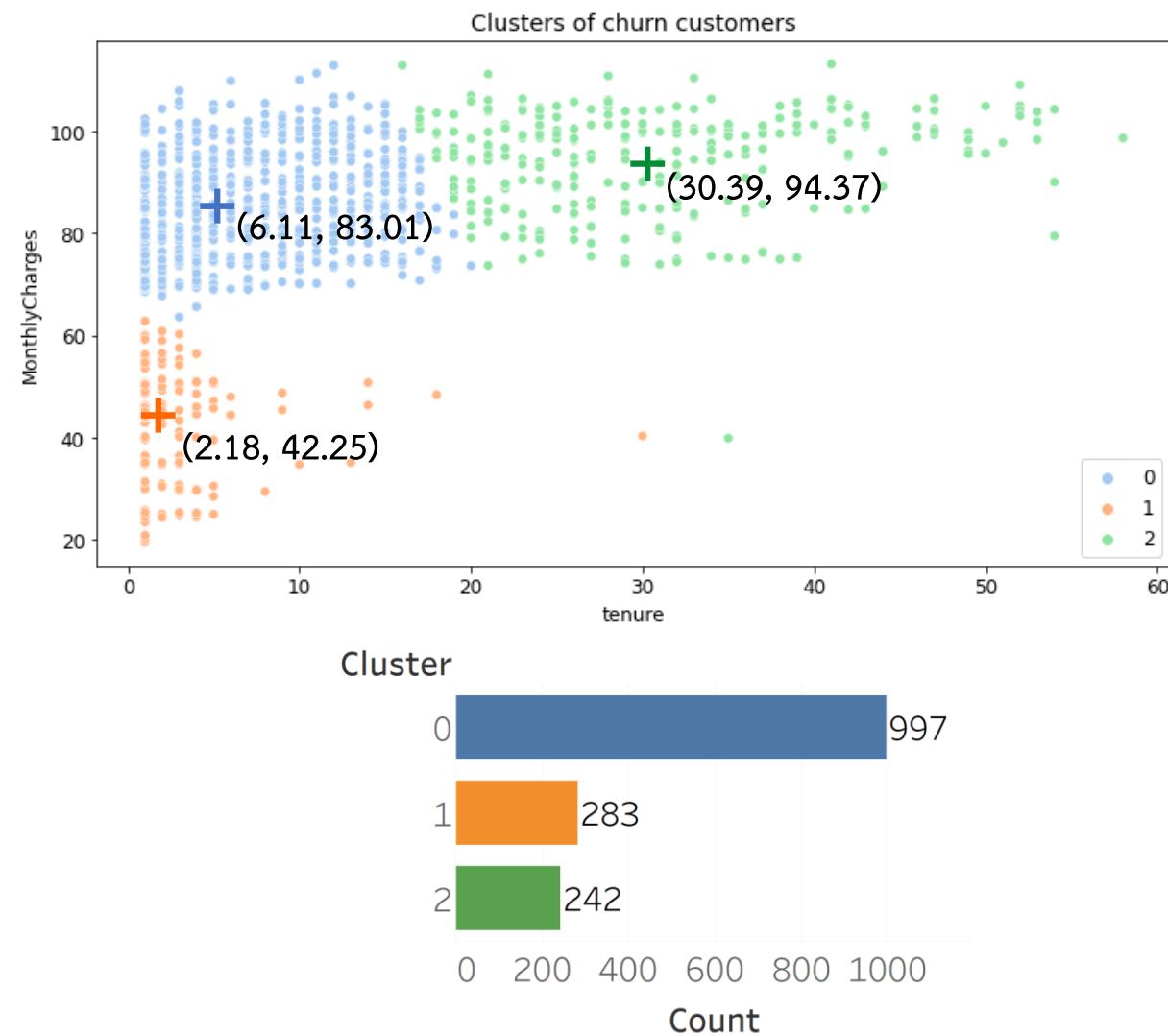


# Selecting optimal number of cluster (k)



# Creating Clustering Models

High intra-class similarity, low inter-class similarity



$$Z = \frac{x - \bar{x}}{s}$$

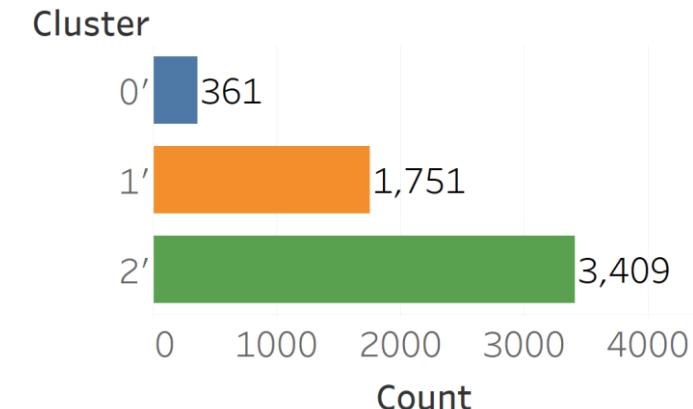
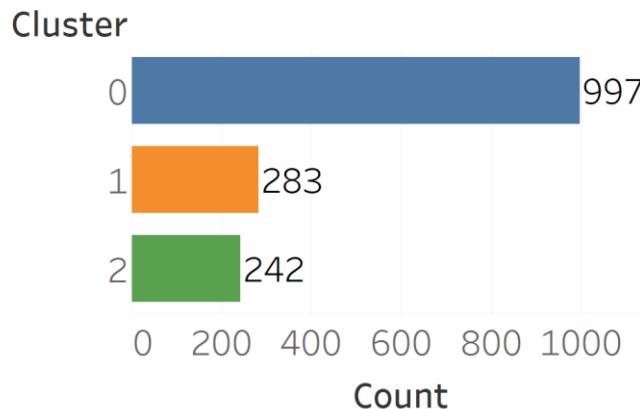
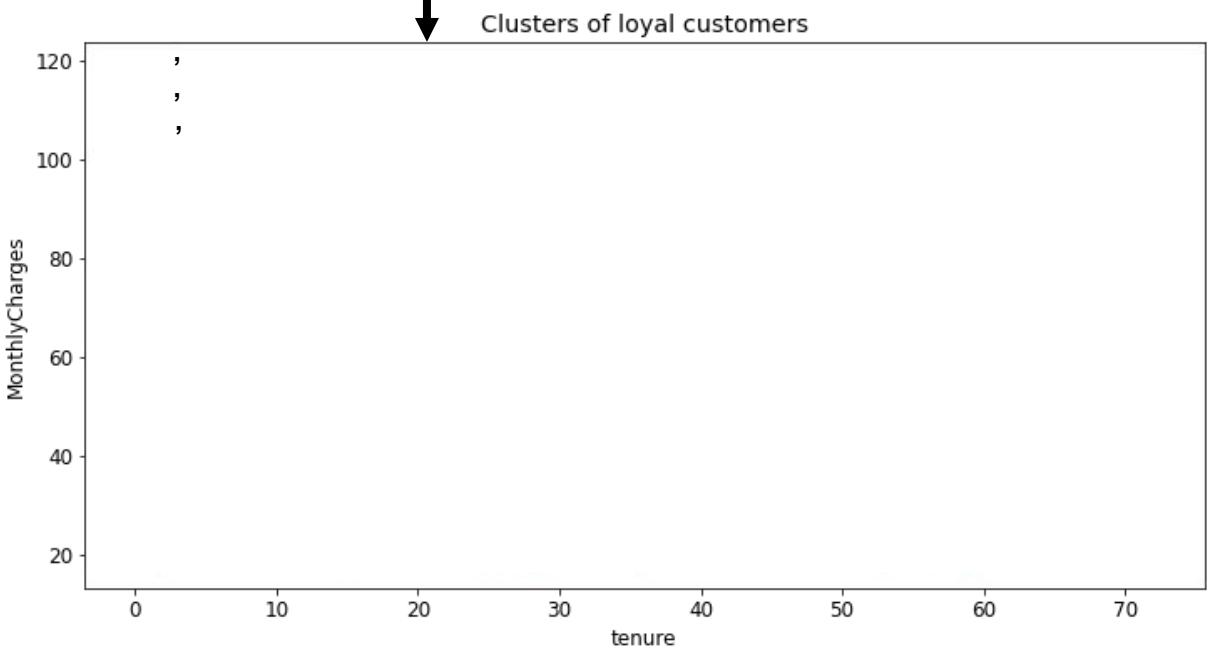
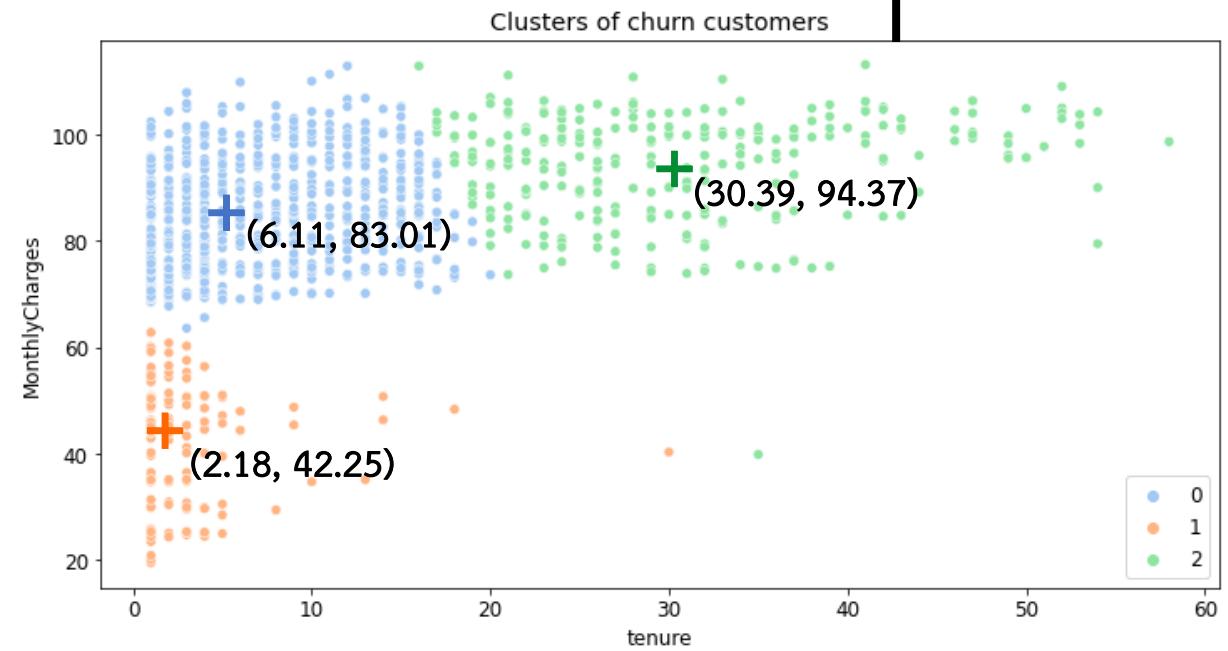
$Z$  = standard score  
 $x$  = observed value  
 $\bar{x}$  = mean of the sample  
 $s$  = standard deviation of the sample



# Creating Clustering Models

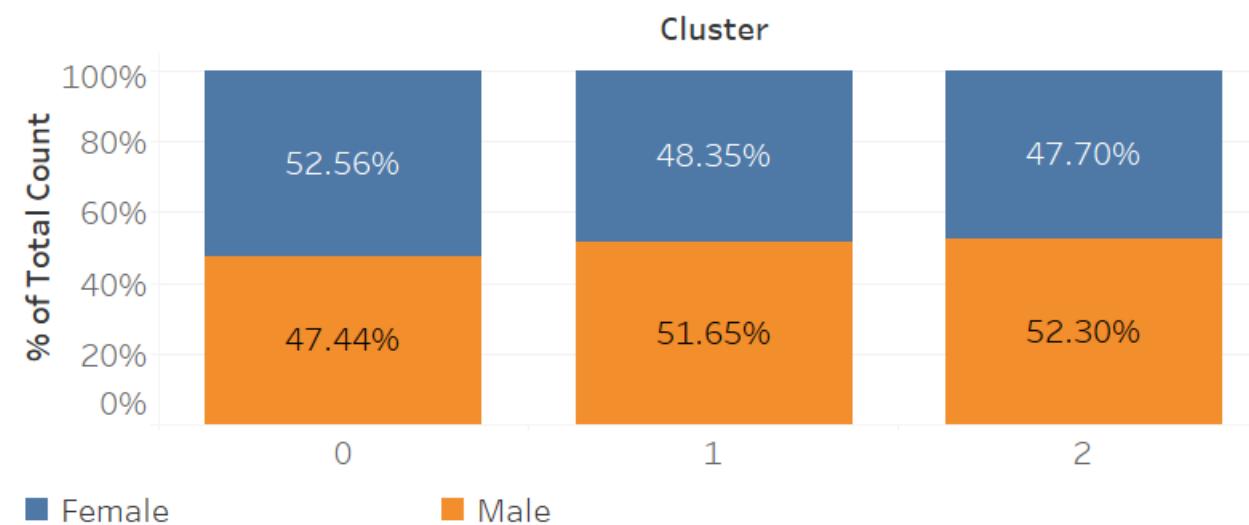
High intra-class similarity, low inter-class similarity

Use the same centroids as in Churn clusters

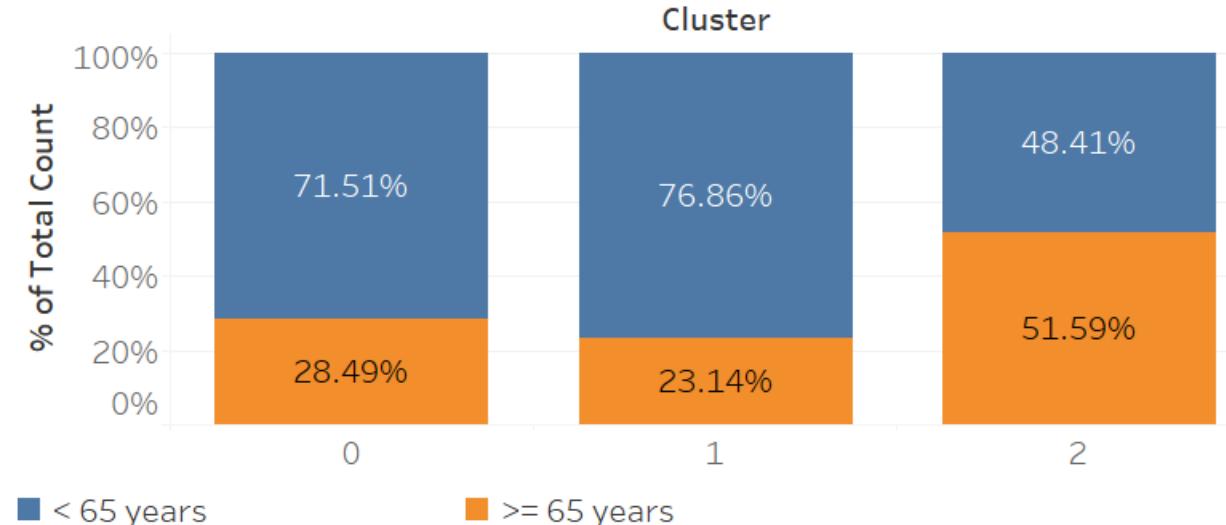


# Demographics Data of Churn Cluster

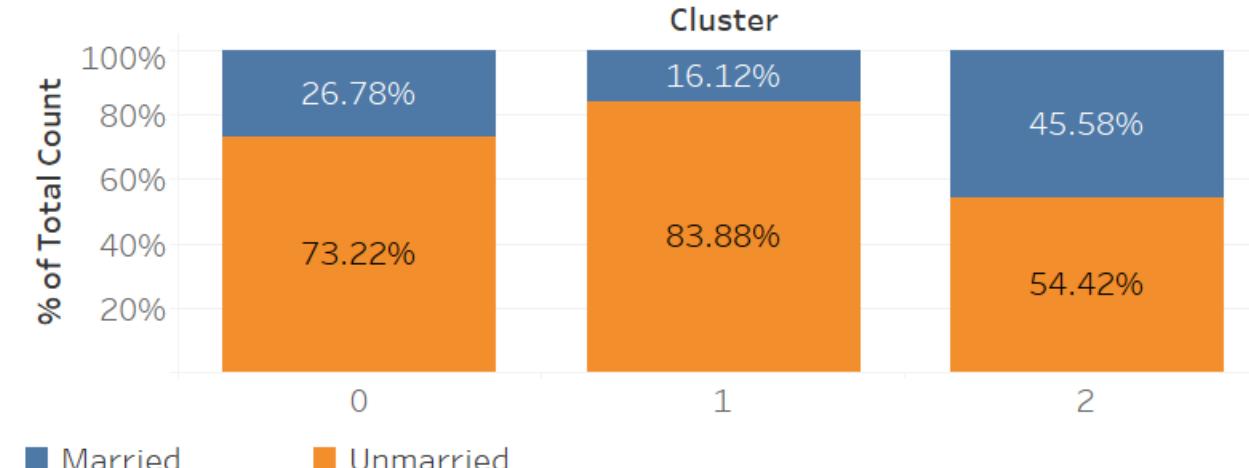
Gender



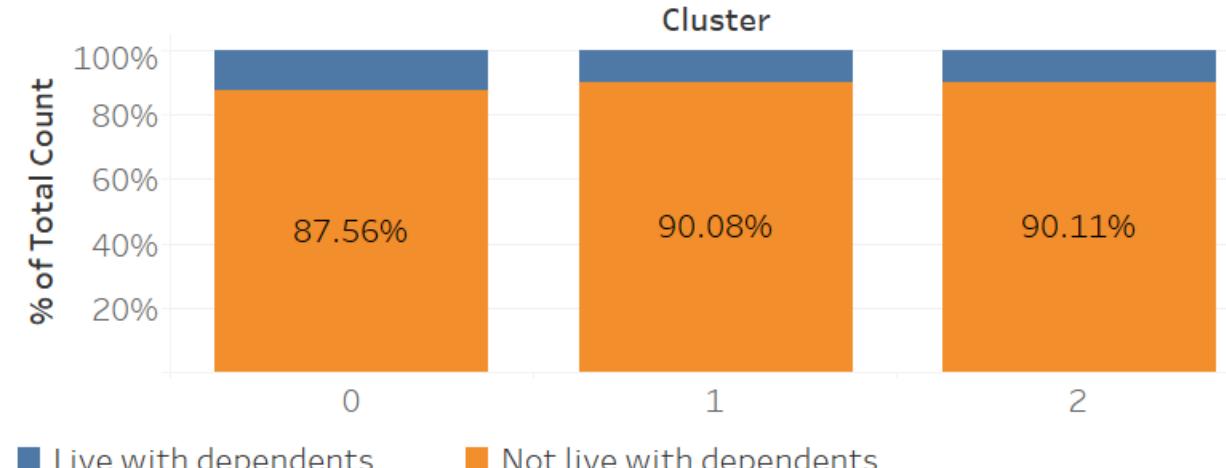
Senior Citizen



Partner

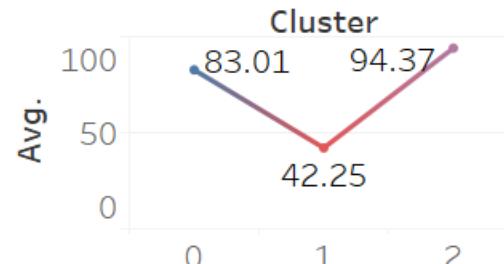


Dependent

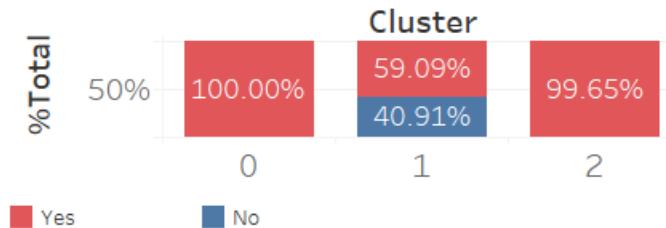


# Service Data of Churn Cluster

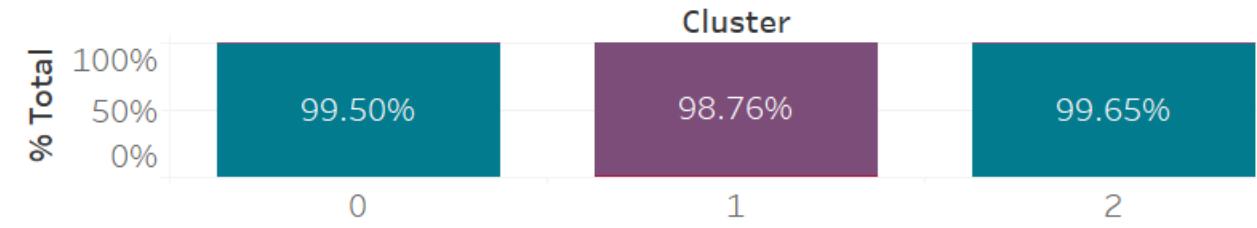
Monthly Charges



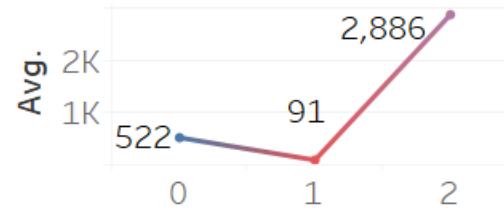
Phone Service



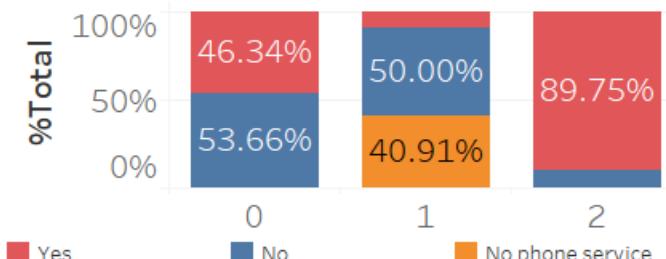
Internet Service



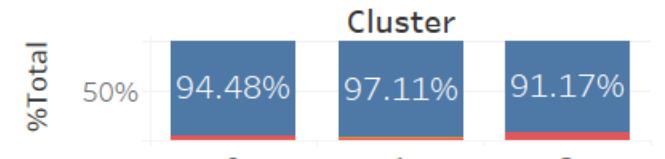
Total Charges



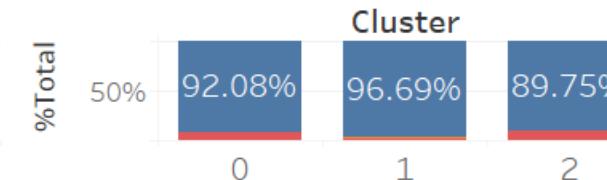
Multiple Line



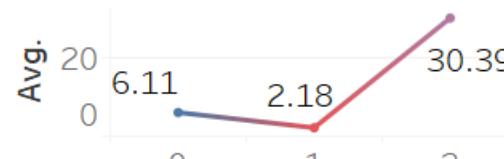
Online Security



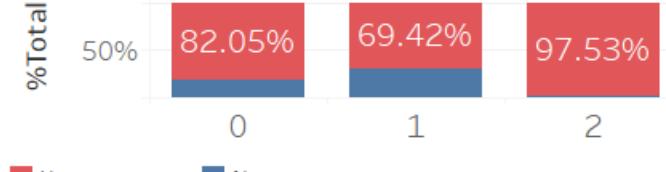
Tech Support



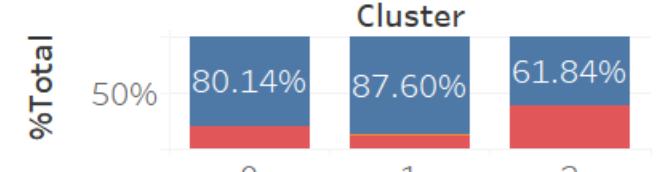
Tenure



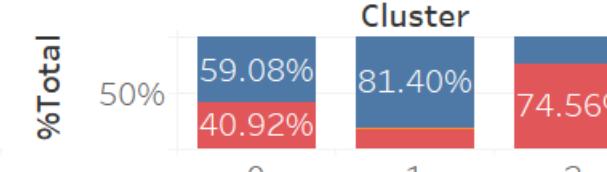
Paperless Billing



Online Backup



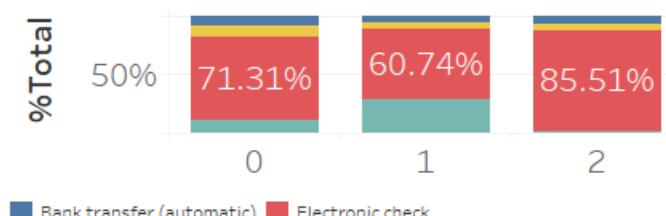
Streaming TV



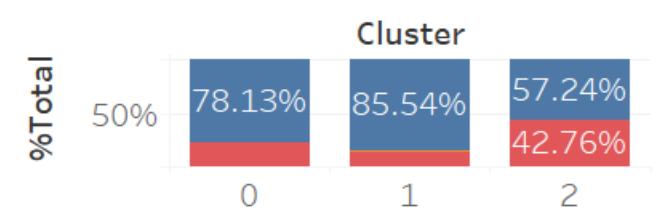
Contract



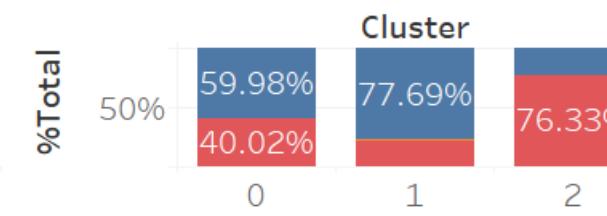
Payment Method



Device Protection

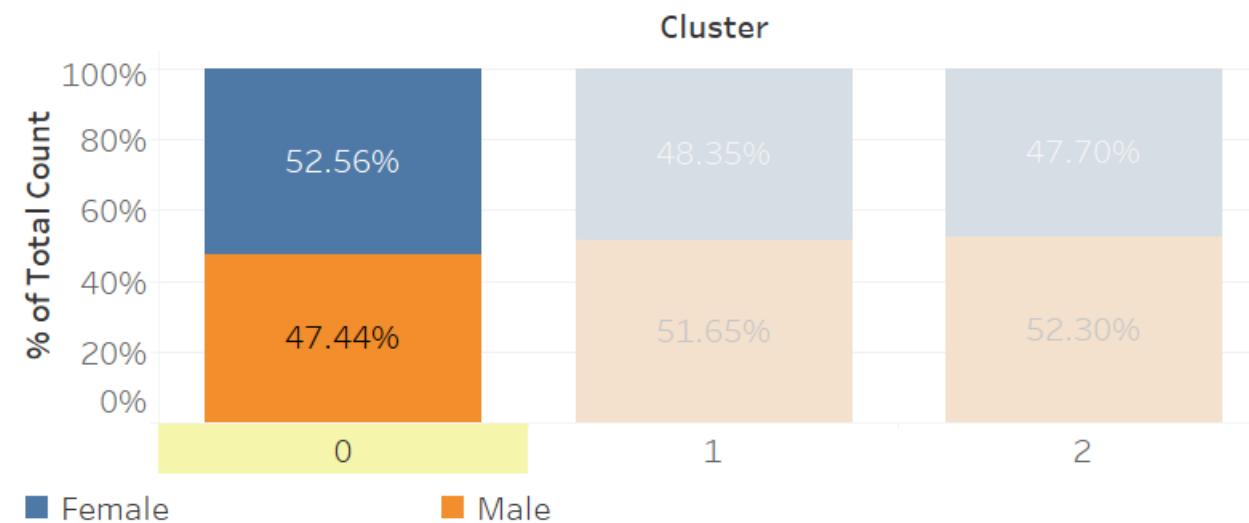


Streaming Movies

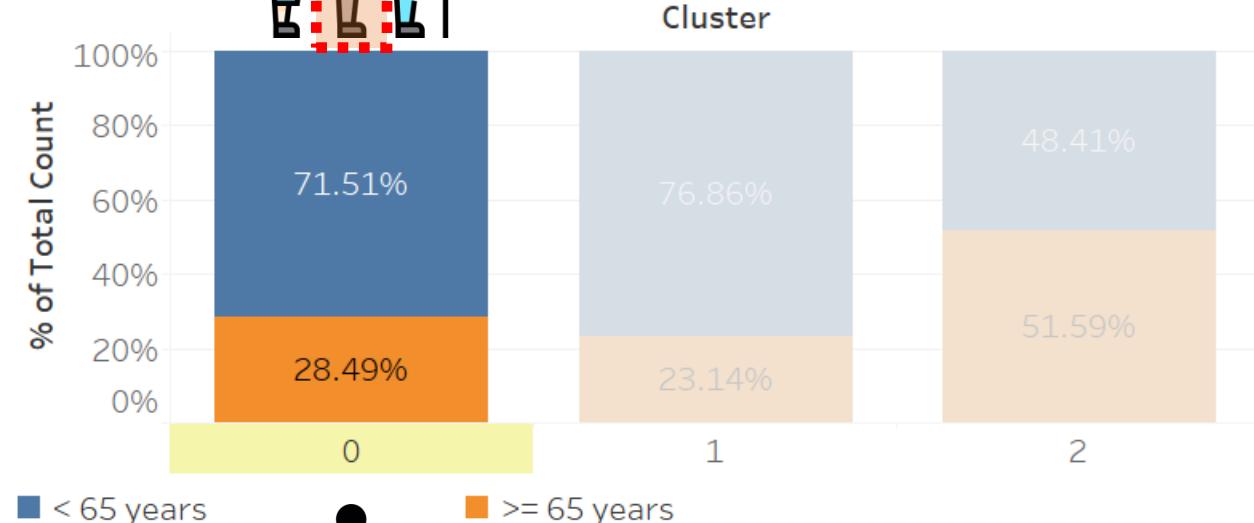


# Demographics Data of Churn Cluster

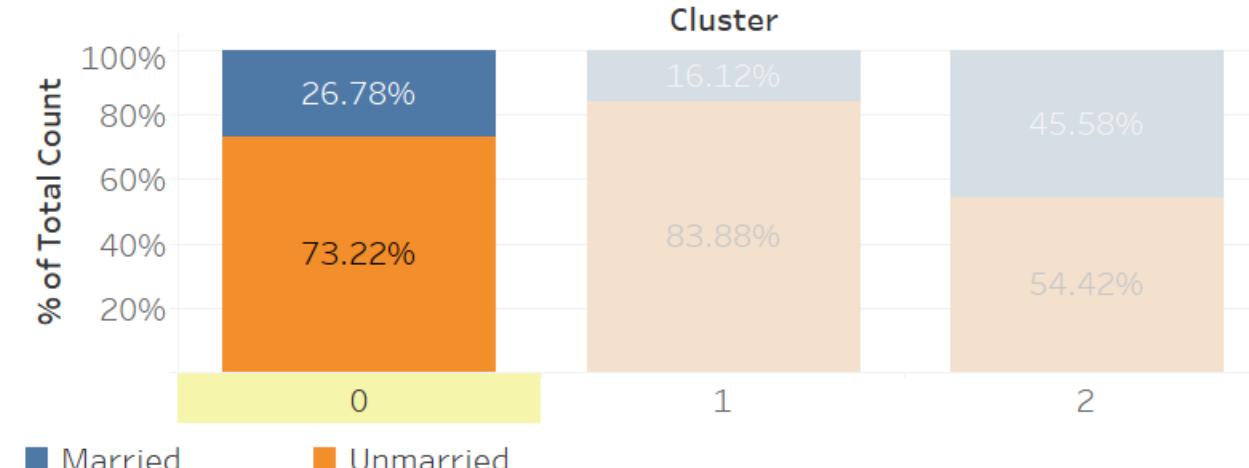
Gender



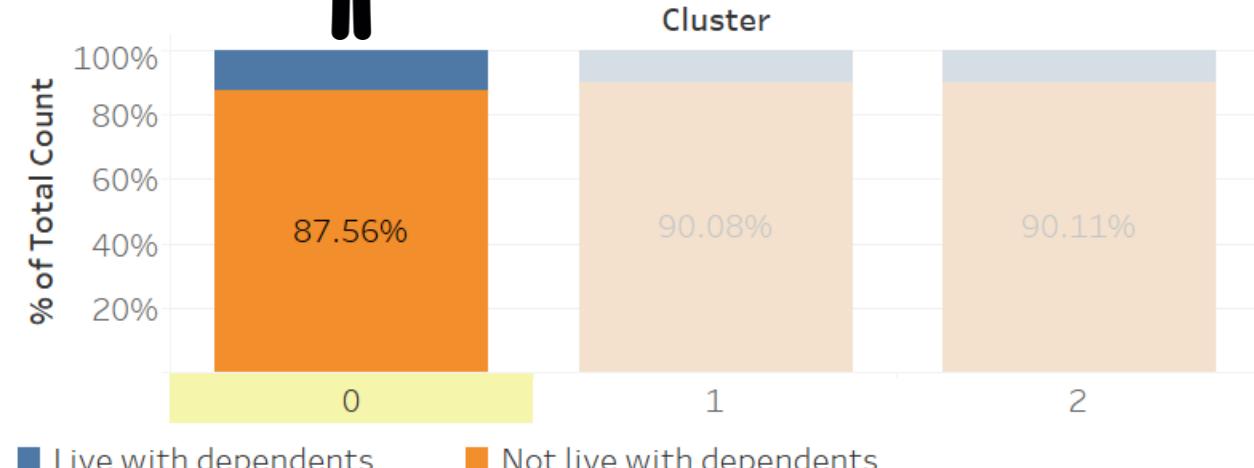
Senior Citizen



Partner

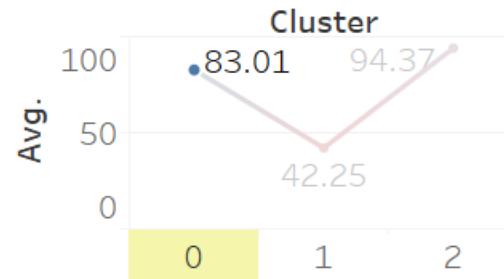


Dependent

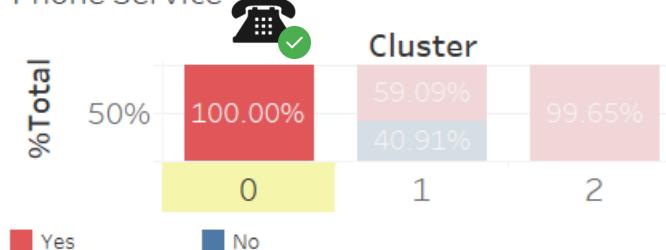


# Service Data of Churn Cluster

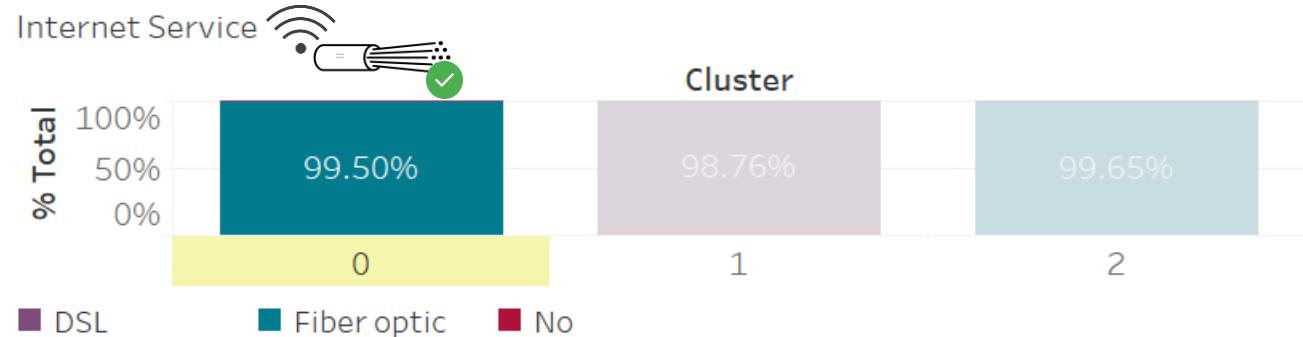
Monthly Charges



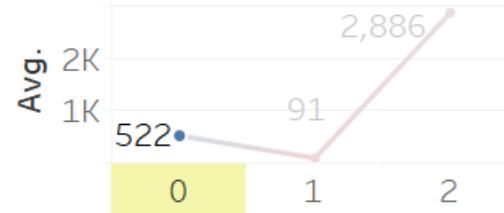
Phone Service



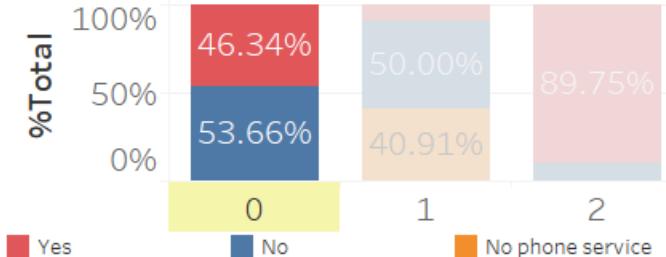
Internet Service



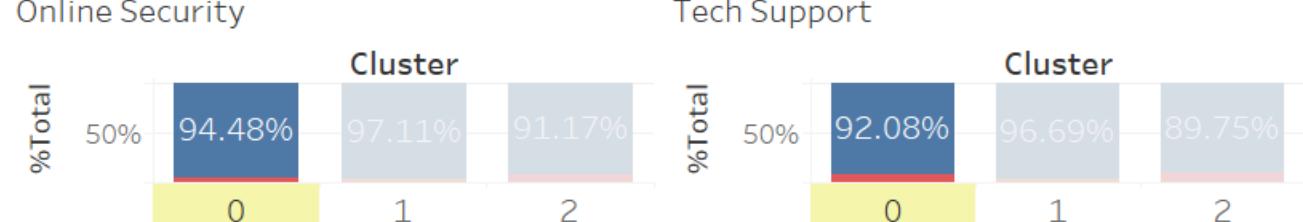
Total Charges



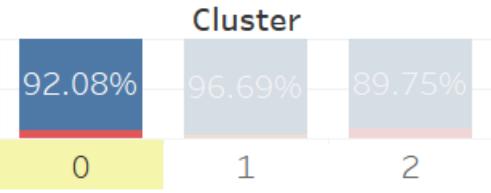
Multiple Line



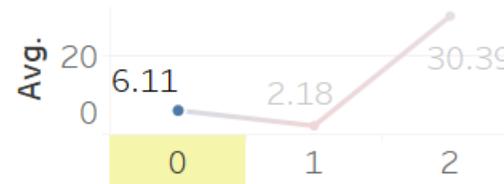
Online Security



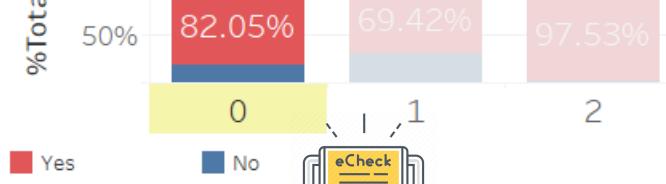
Cluster



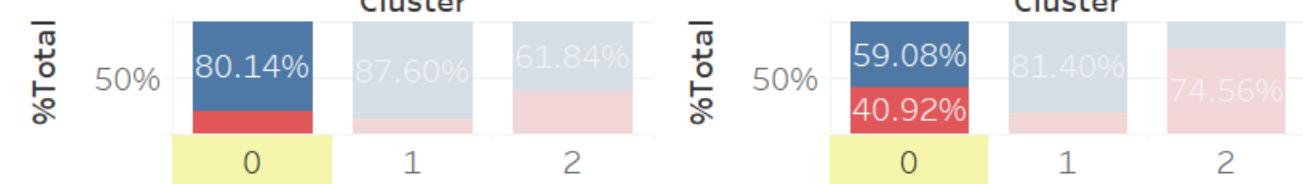
Tenure



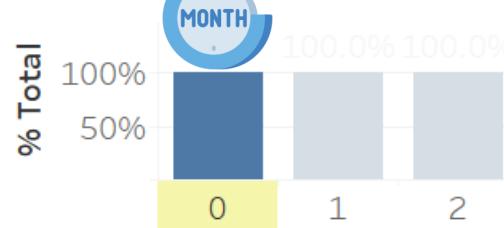
Paperless Billing



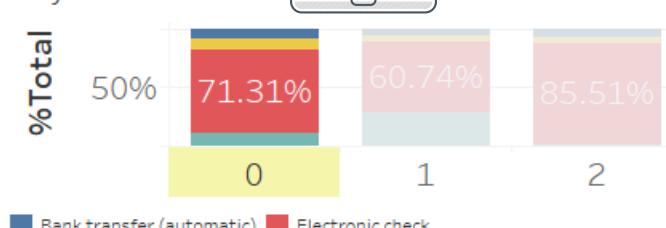
Online Backup



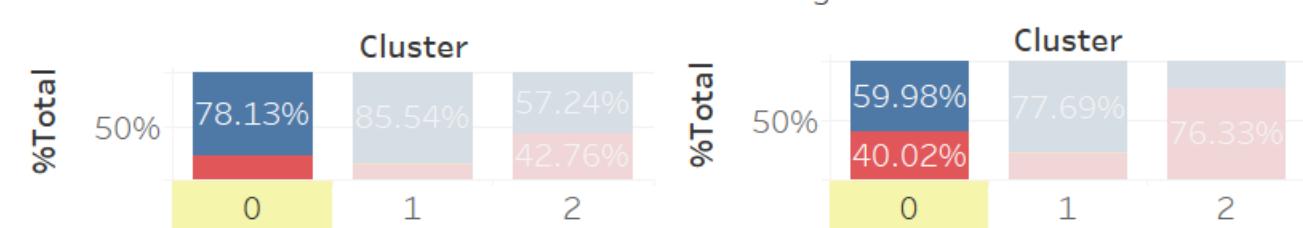
Contract



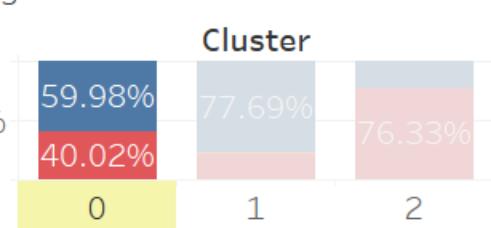
Payment Method



Device Protection



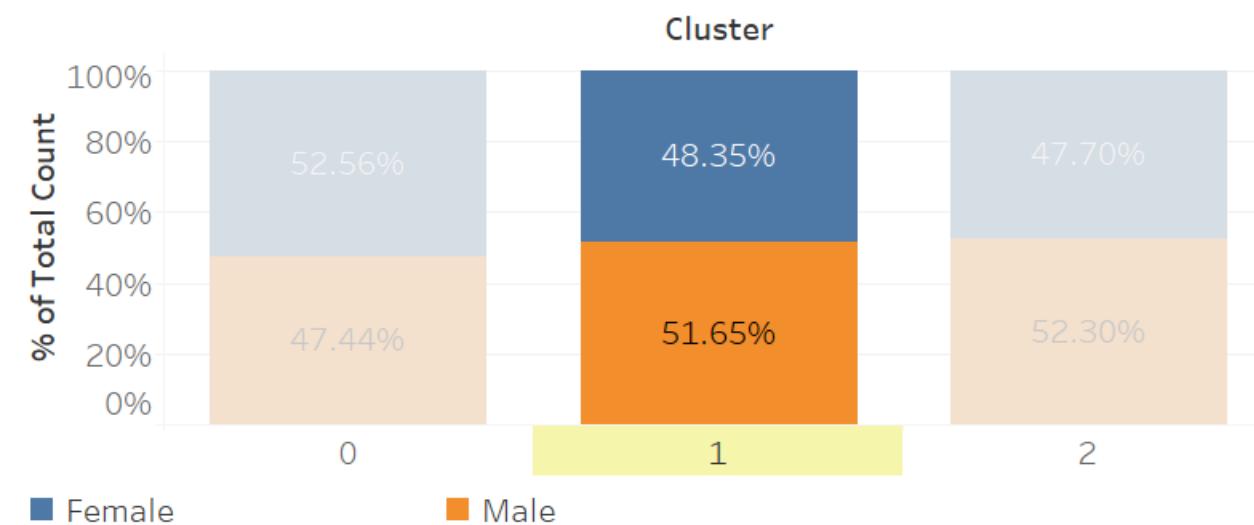
Cluster



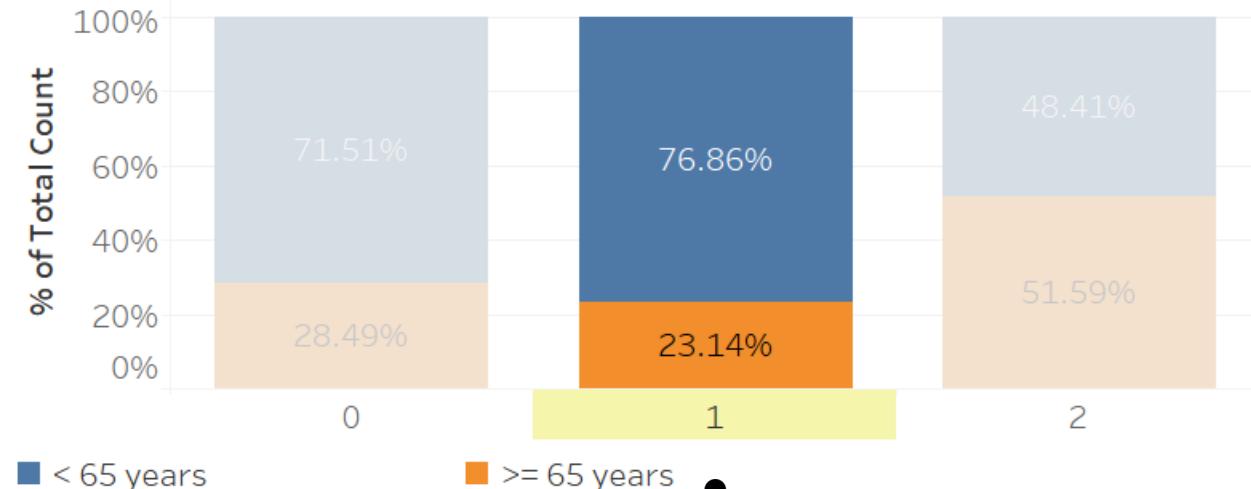
# Demographics Data of Churn Cluster



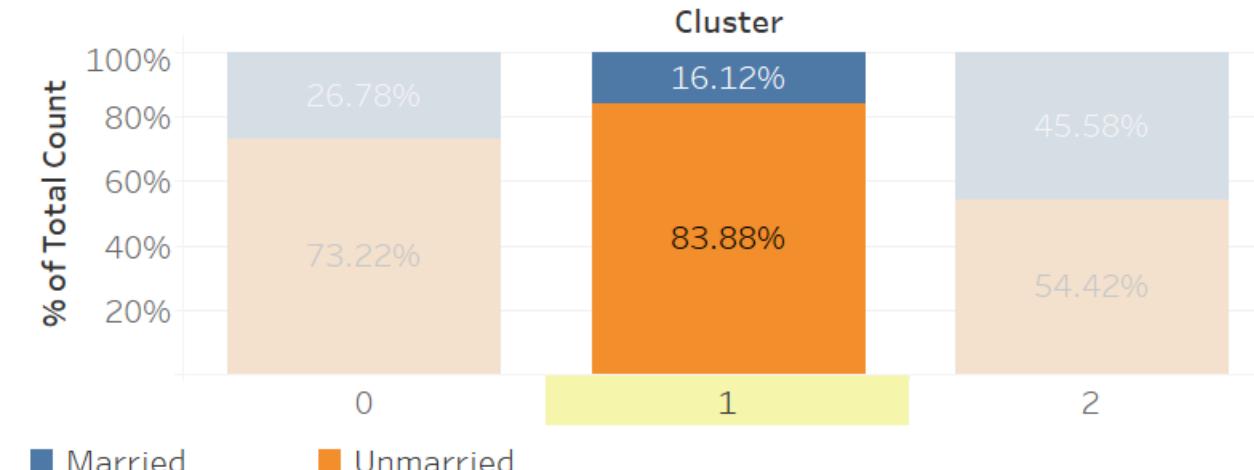
Gender



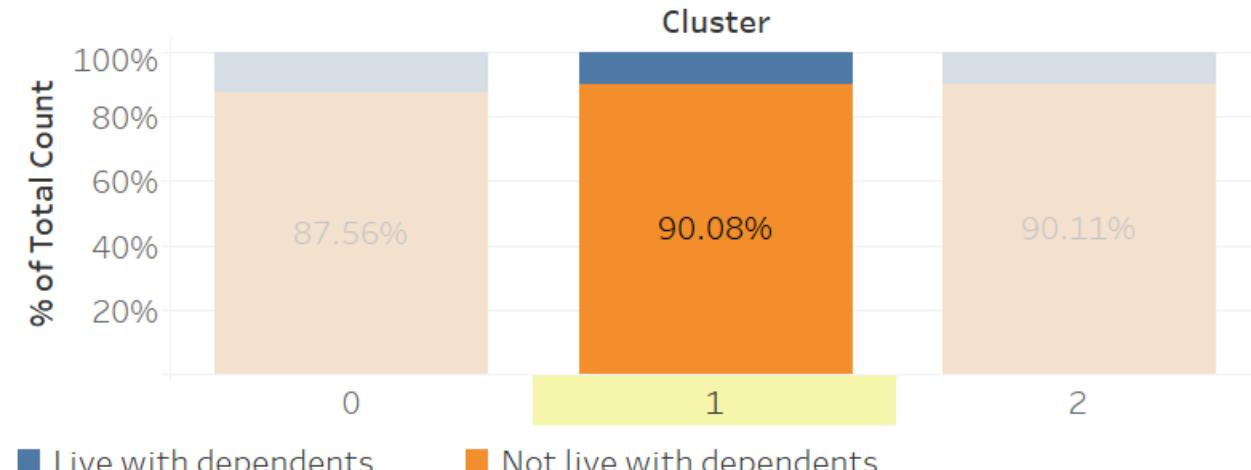
Senior Citizen



Partner

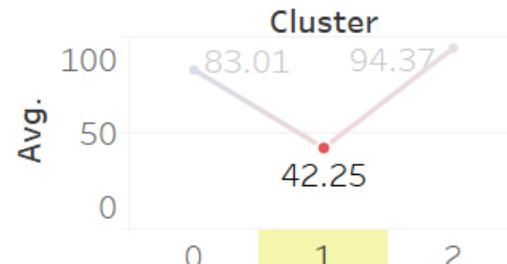


Dependent



# Service Data of Churn Cluster

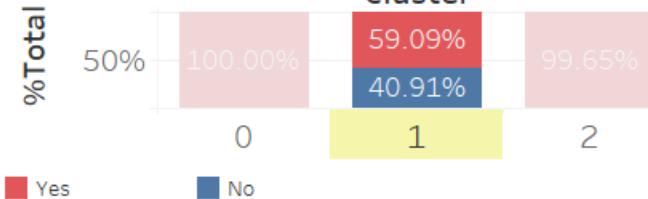
Monthly Charges



Phone Service



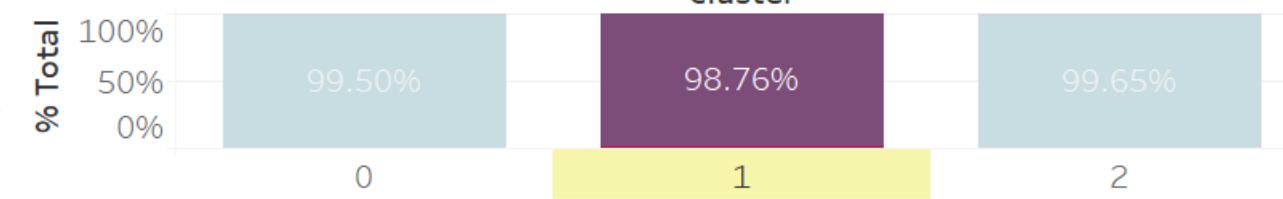
Cluster



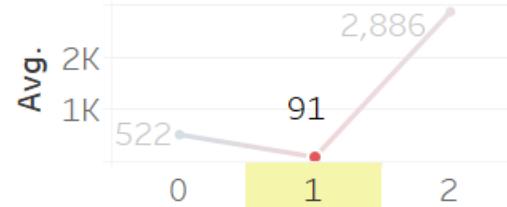
Internet Service



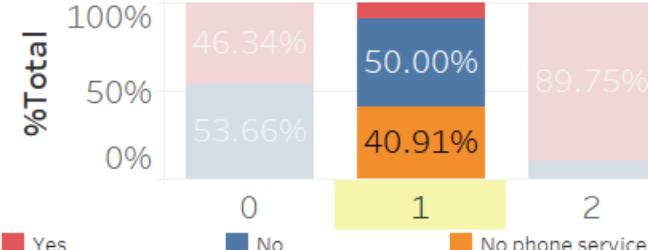
Cluster



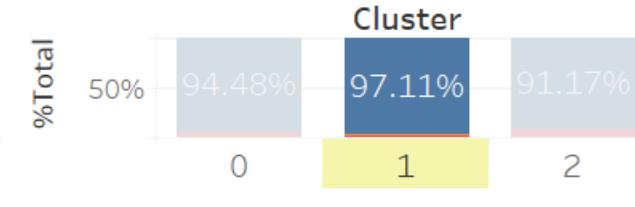
Total Charges



Multiple Line



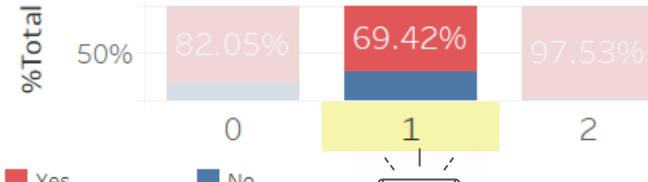
Online Security



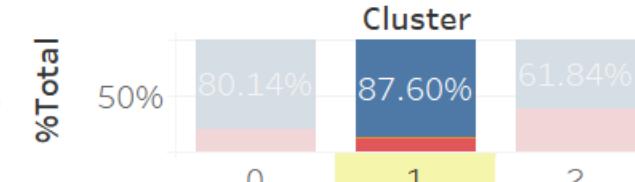
Tenure



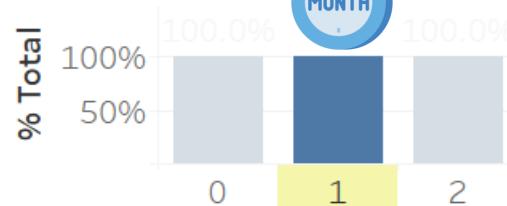
Paperless Billing



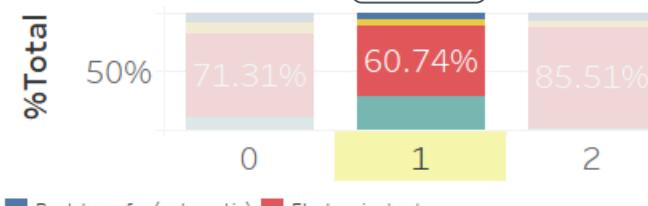
Online Backup



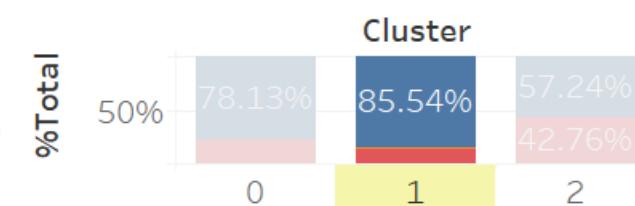
Contract



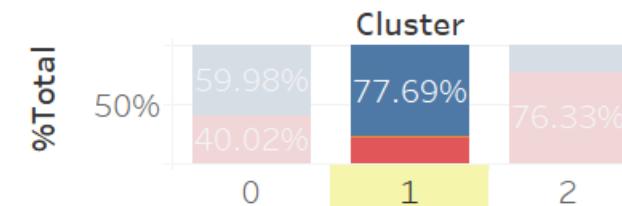
Payment Method

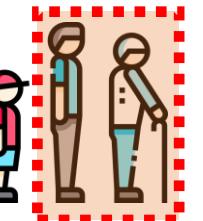


Device Protection



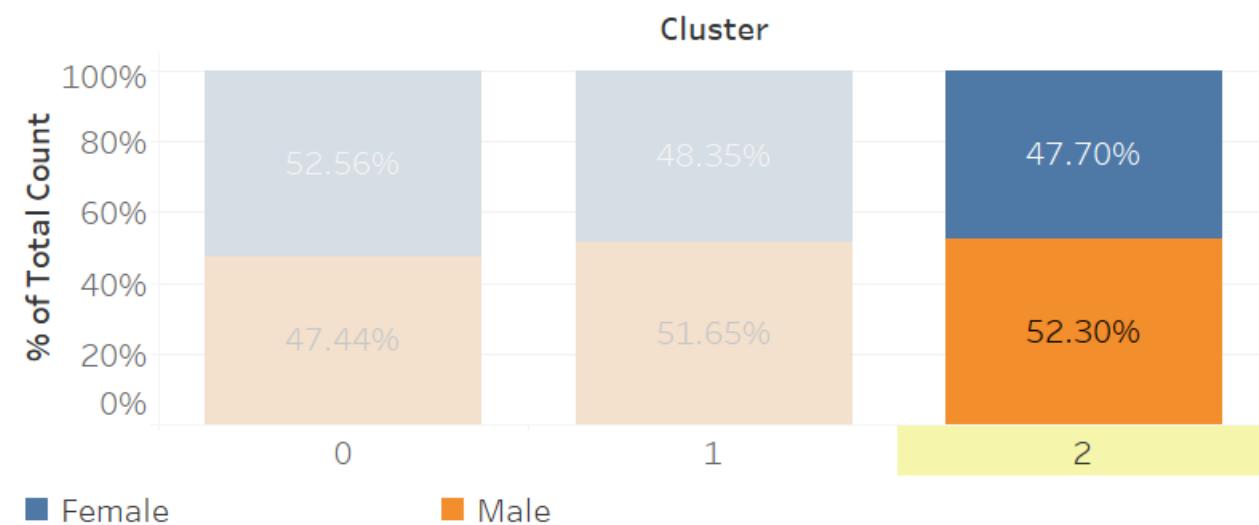
Streaming Movies



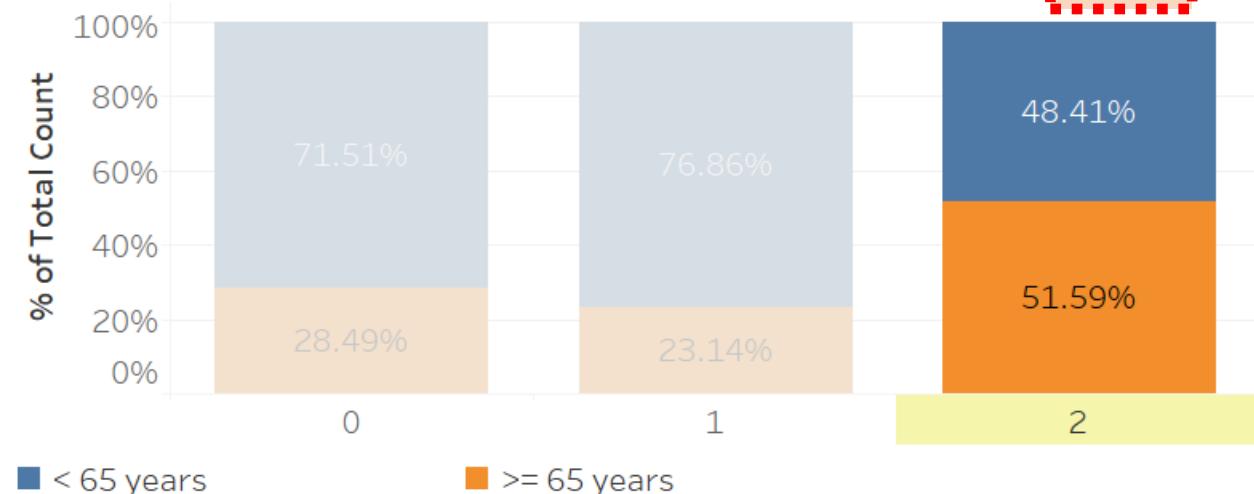


# Demographics Data of Churn Cluster

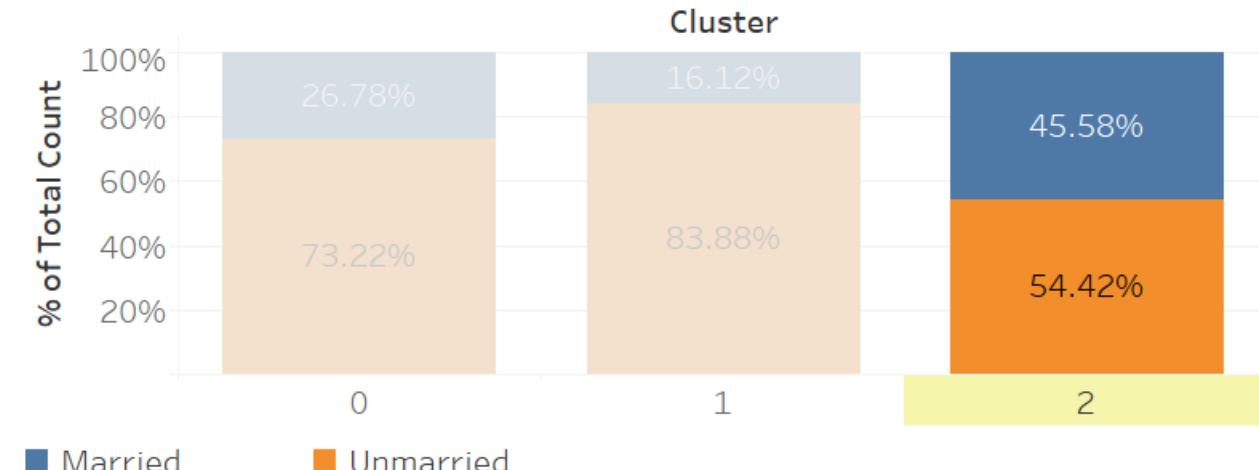
Gender



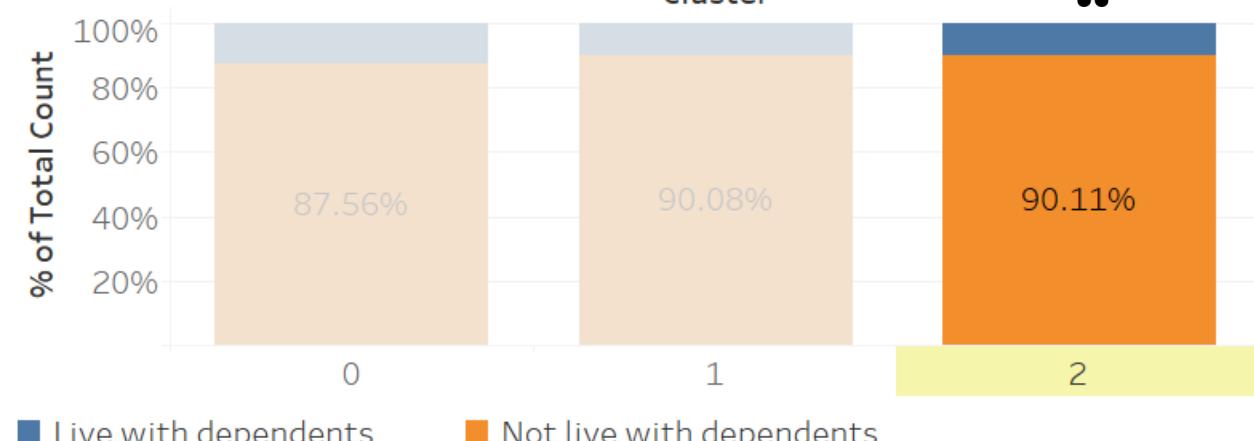
Senior Citizen



Partner

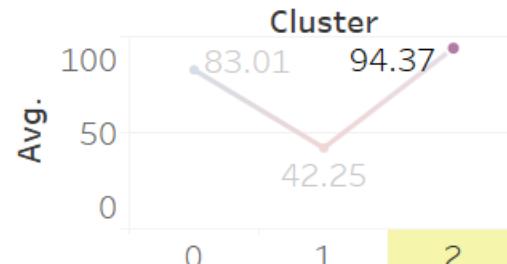


Dependent

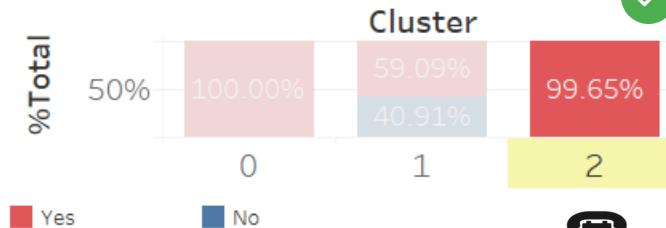


# Service Data of Churn Cluster

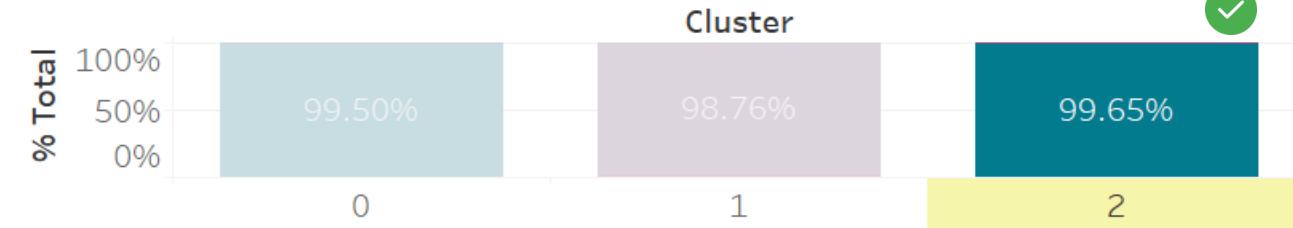
Monthly Charges



Phone Service

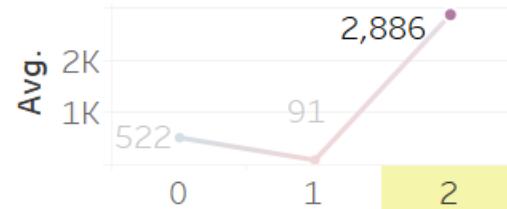


Internet Service

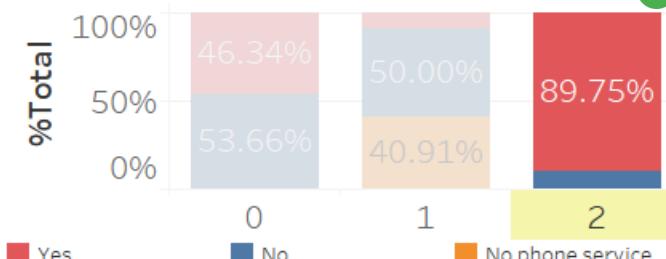


Cluster

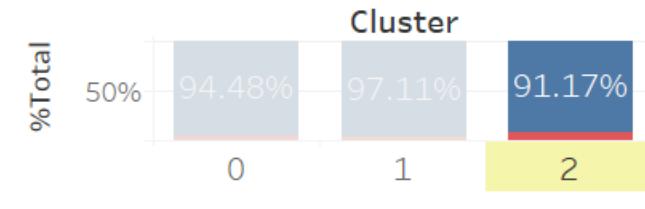
Total Charges



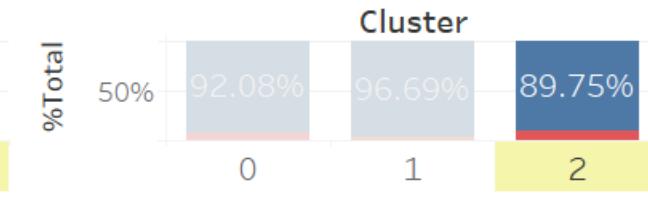
Multiple Line



Online Security



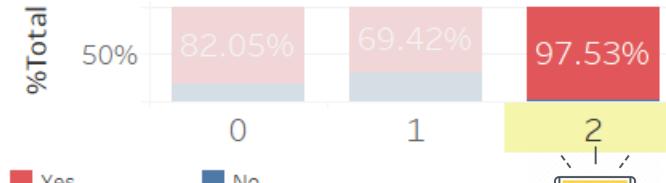
Tech Support



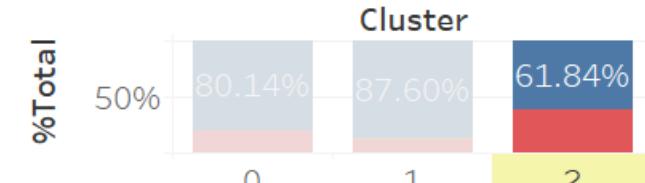
Tenure



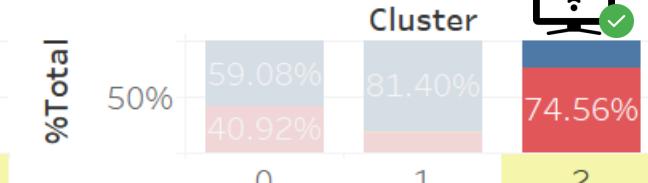
Paperless Billing



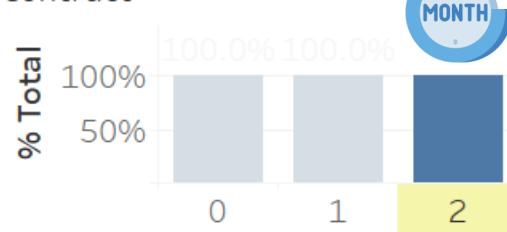
Online Backup



Streaming TV

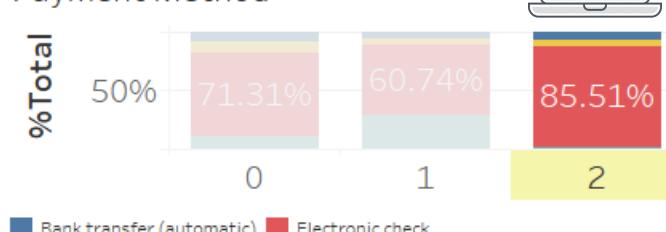


Contract

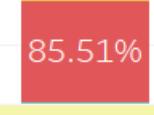
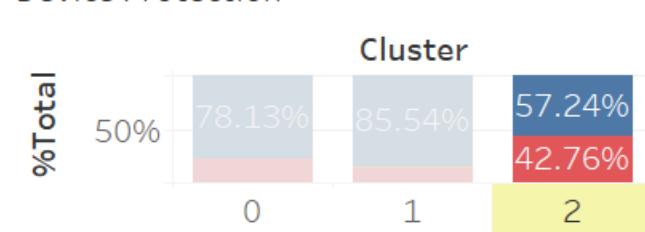


MONTH

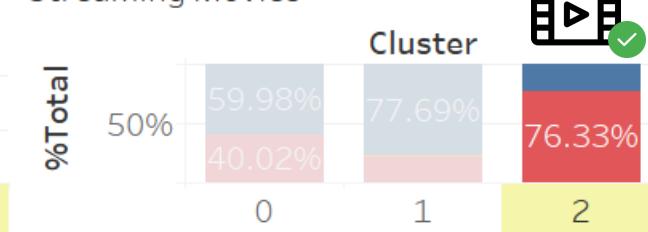
Payment Method



Device Protection



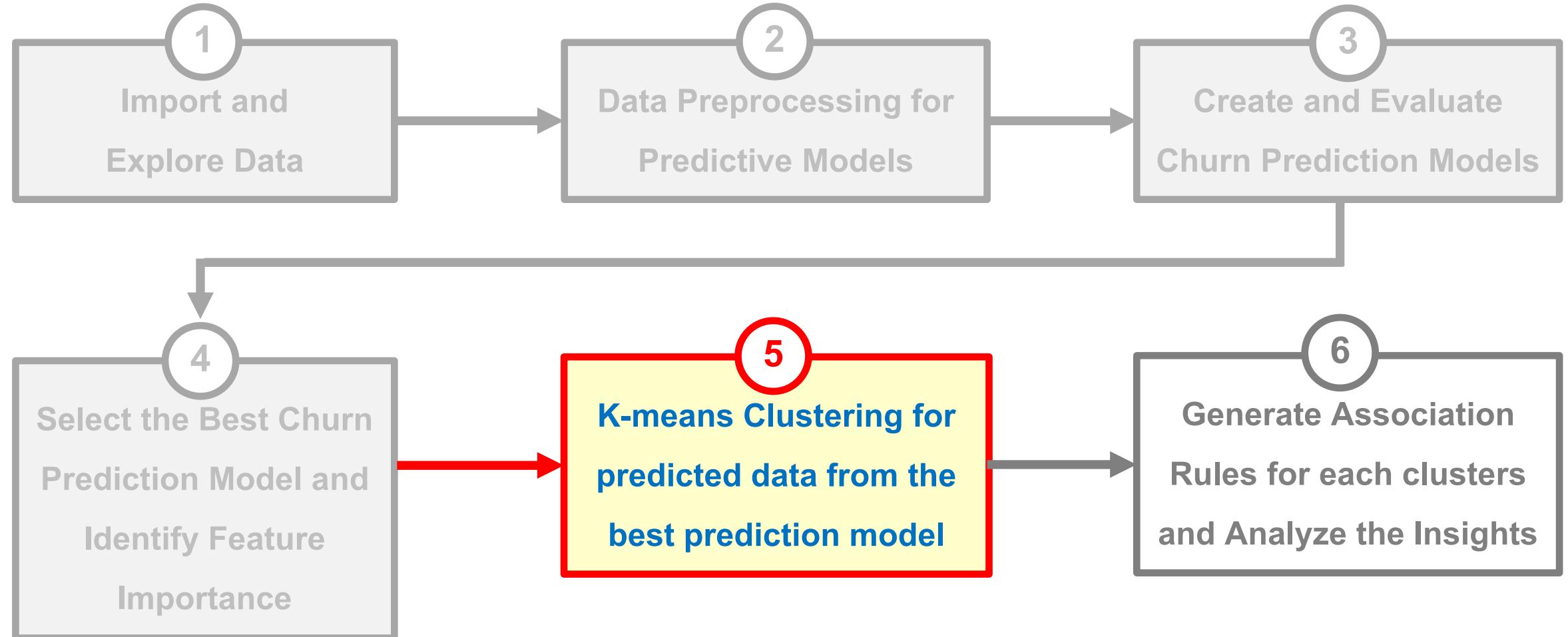
Streaming Movies



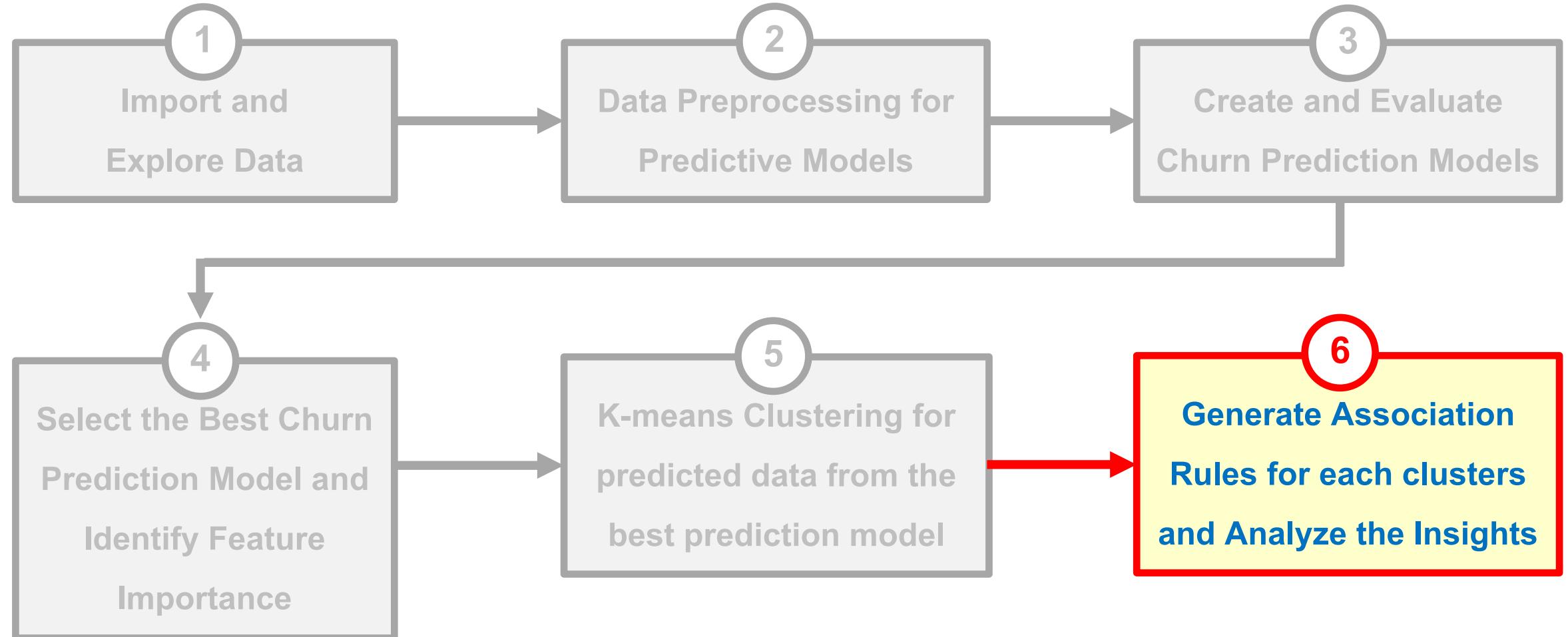
Cluster



# Procedure



# Procedure

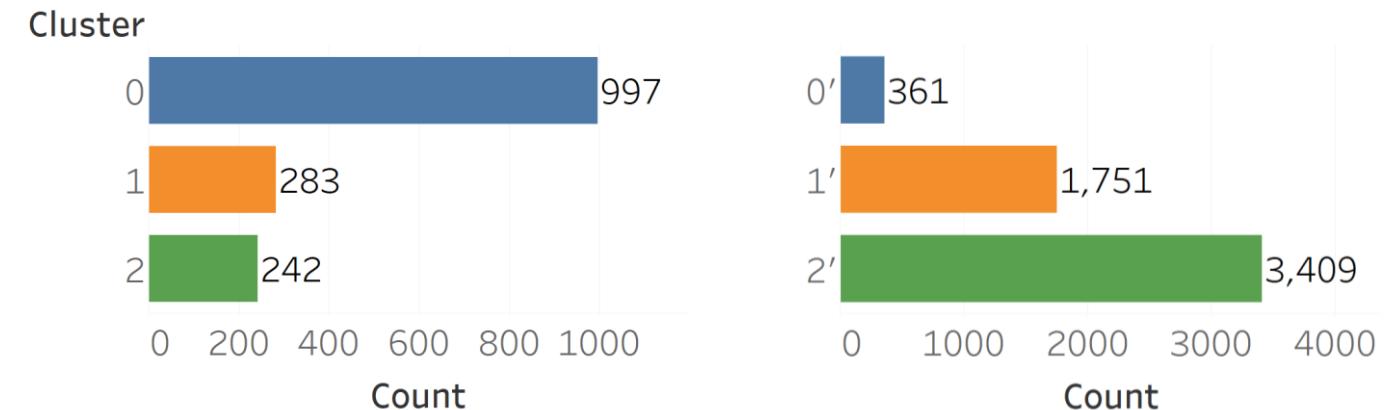


## Assigning minimum thresholds

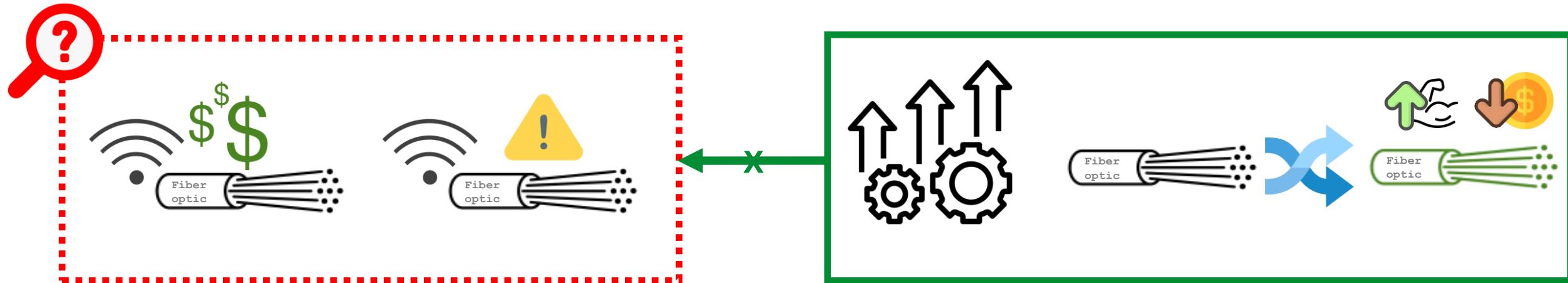
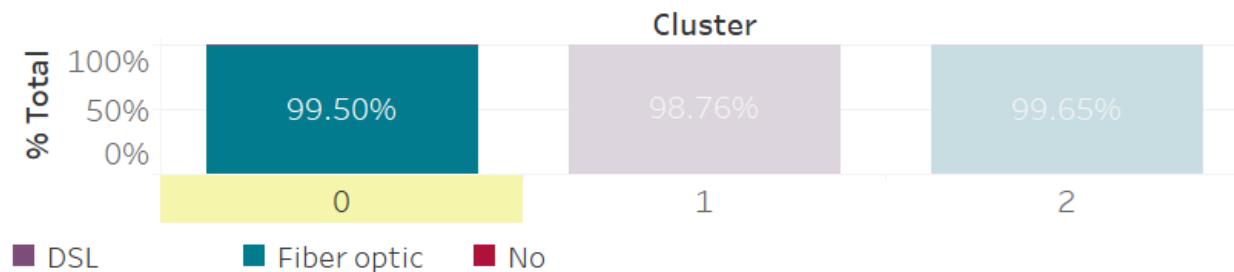
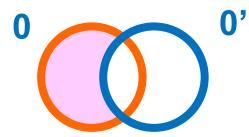
- Minimum Support:

#Customers in cluster	Minimum Support
<= 150	0.9
151 – 300	0.8
301 – 450	0.7
451 – 600	0.6
601 – 750	0.5
751 – 900	0.4
> 900	0.3

- Minimum Confidence = 0.8
- Lift > 1



## Cluster 0



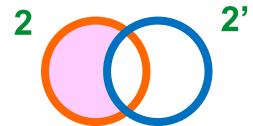
## Cluster 1



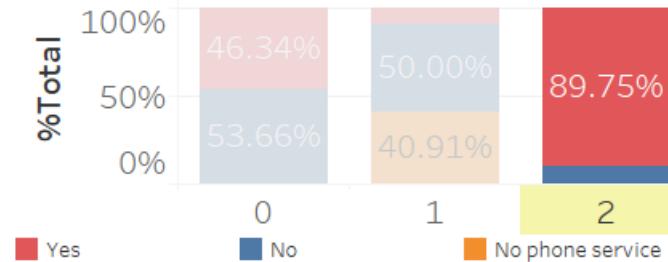
**Not found any association rules**



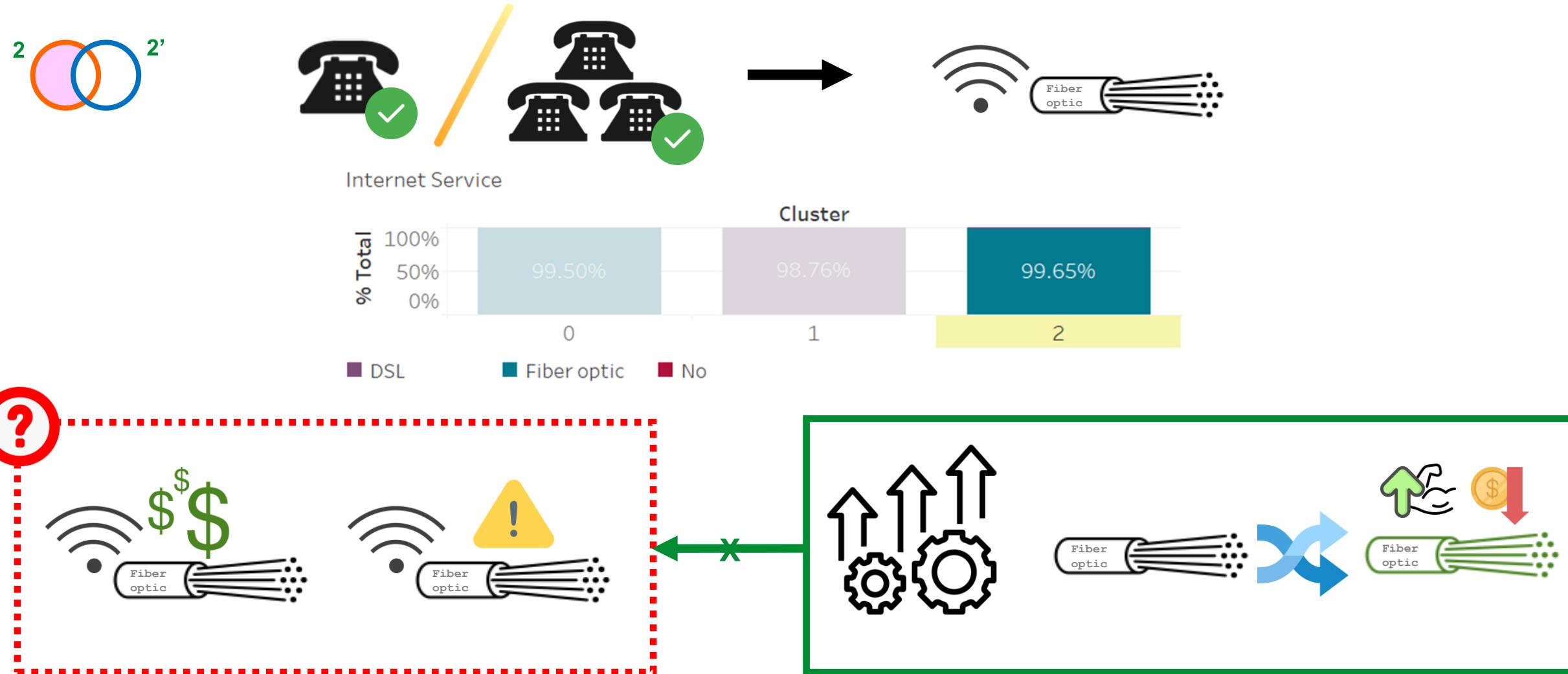
## Cluster 2



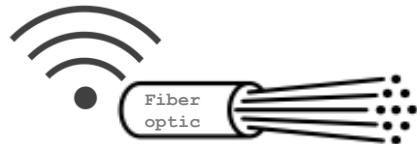
Multiple Line



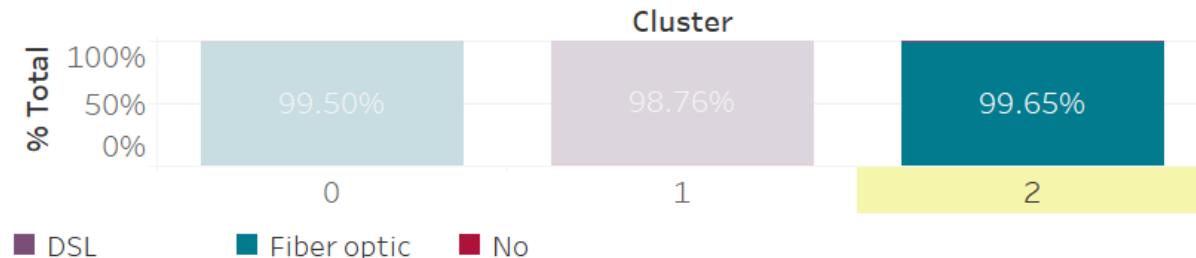
## Cluster 2



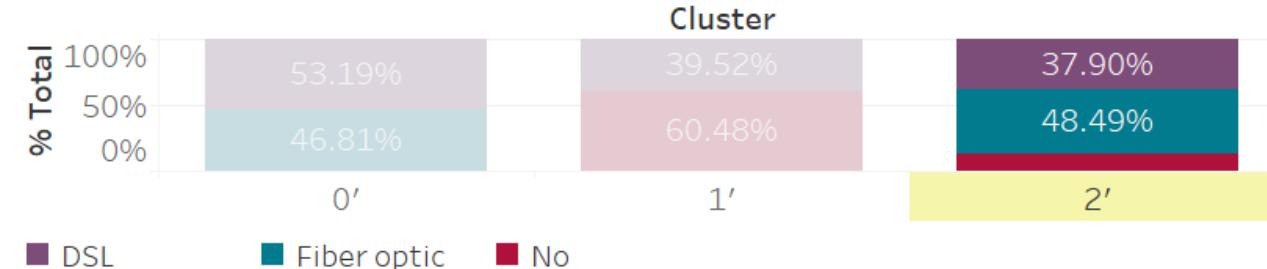
## Cluster 2



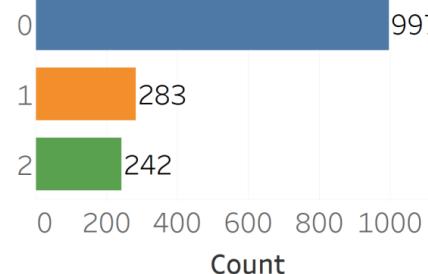
Internet Service



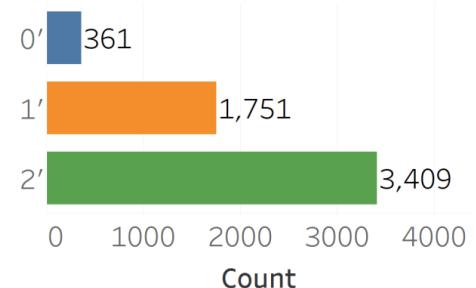
Internet Service



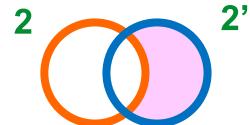
Cluster



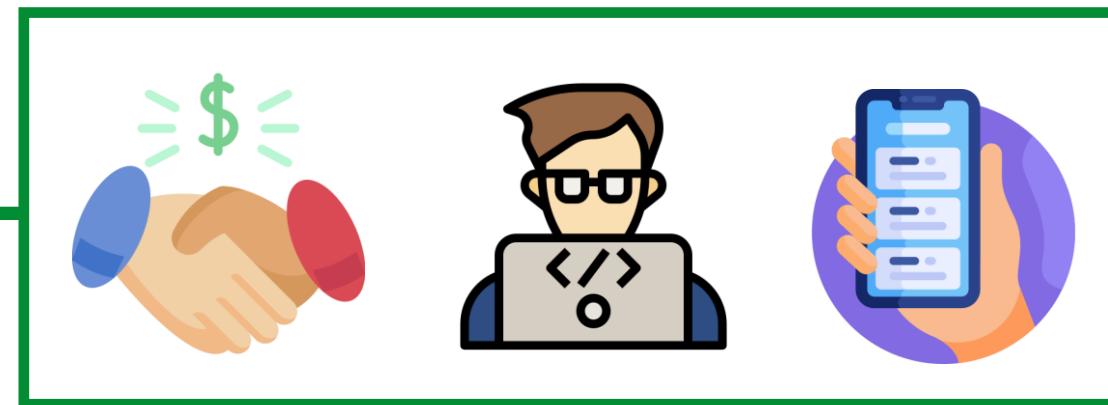
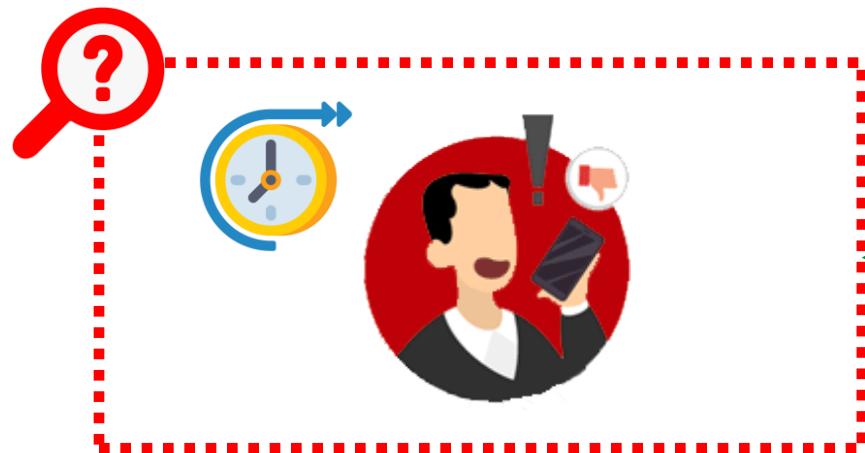
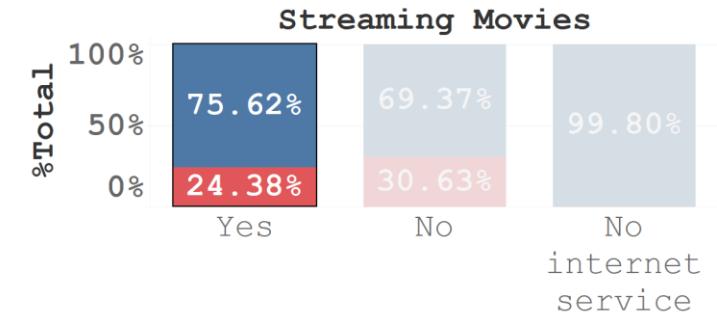
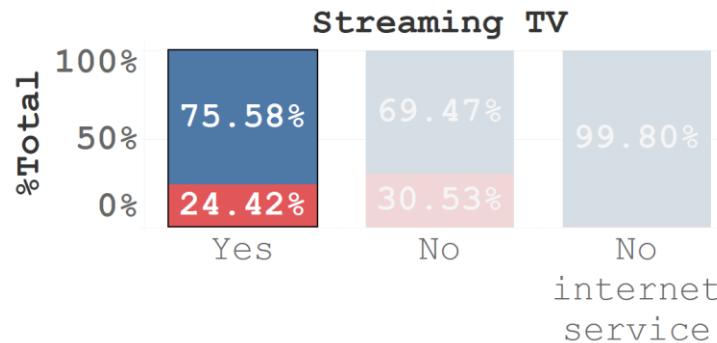
Cluster



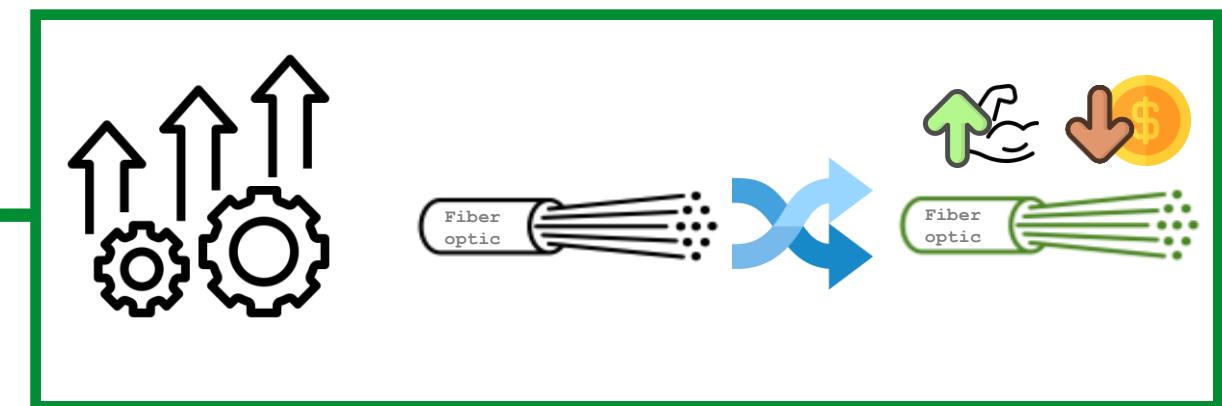
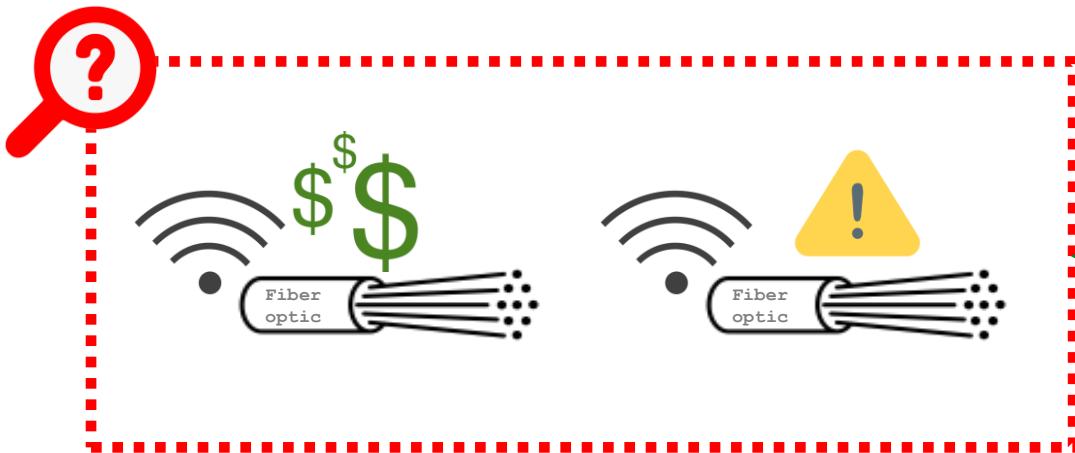
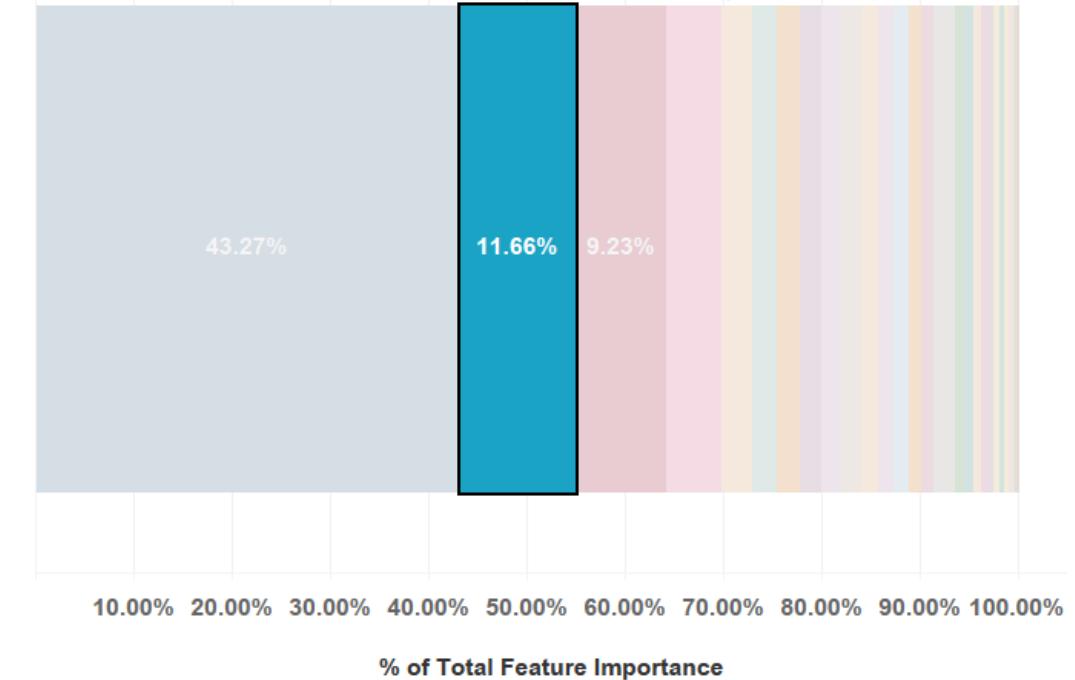
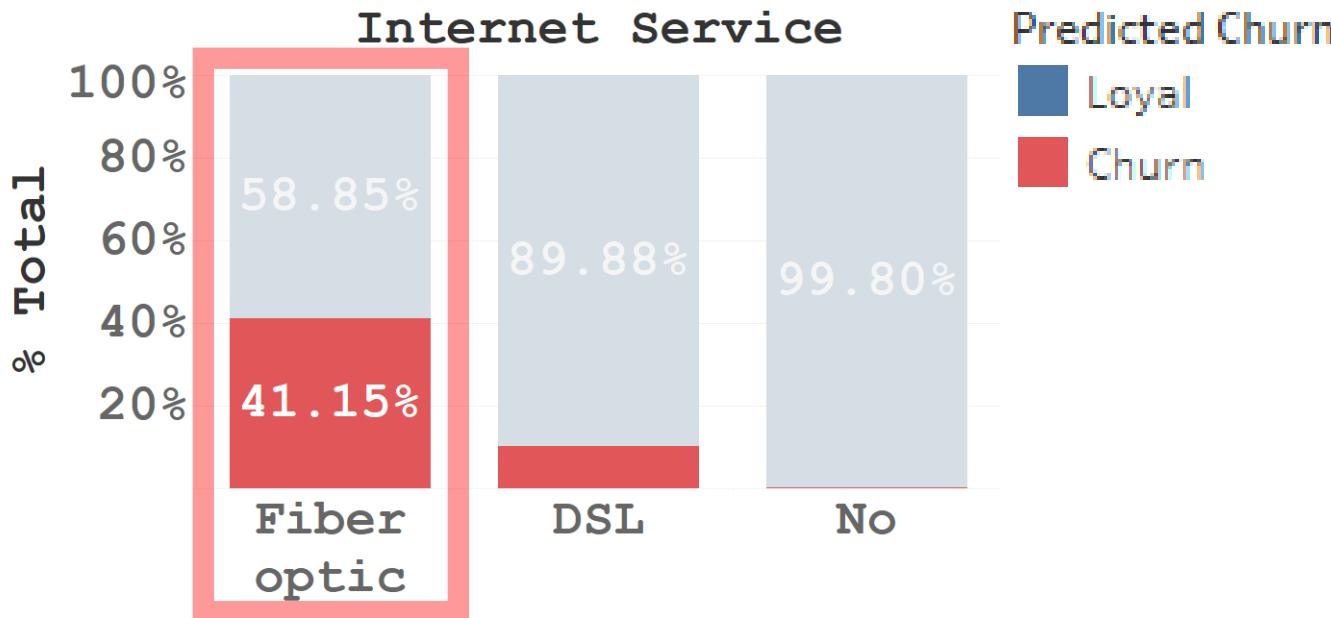
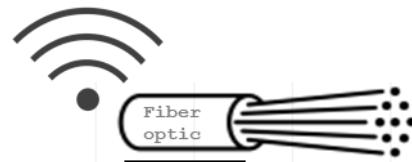
## Cluster 2



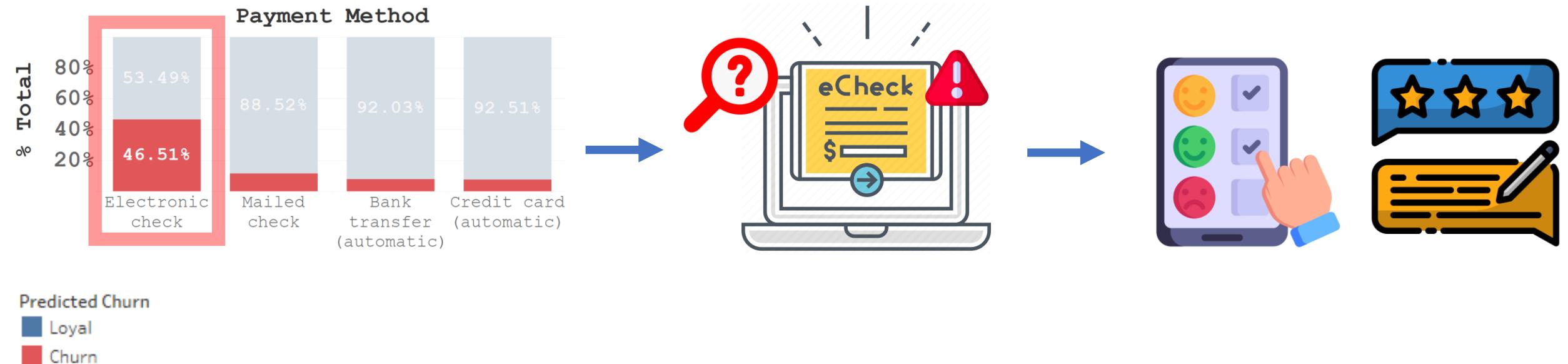
Predicted Churn  
Loyal  
Churn



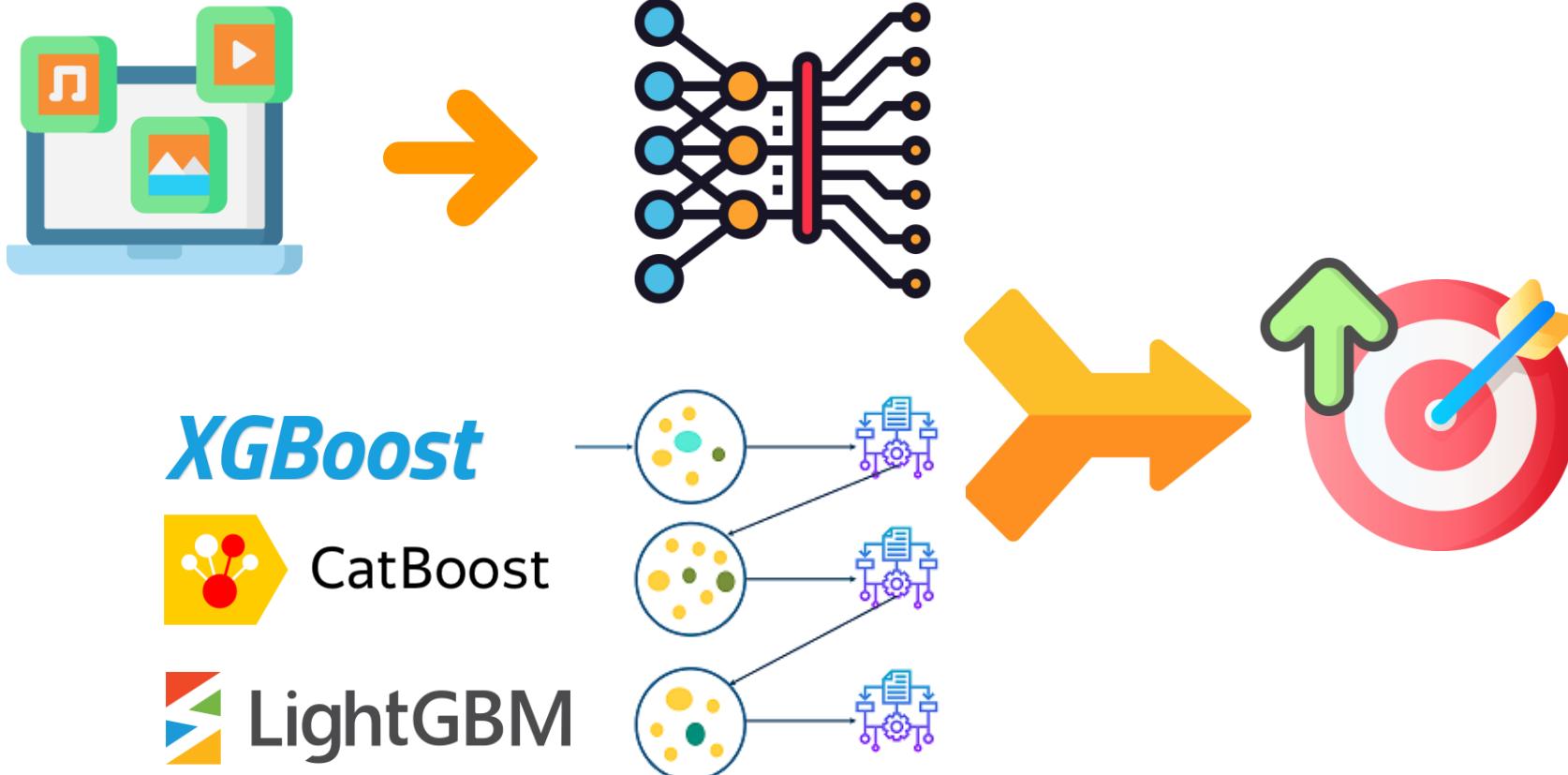
# Wrap-up: Retention Strategies



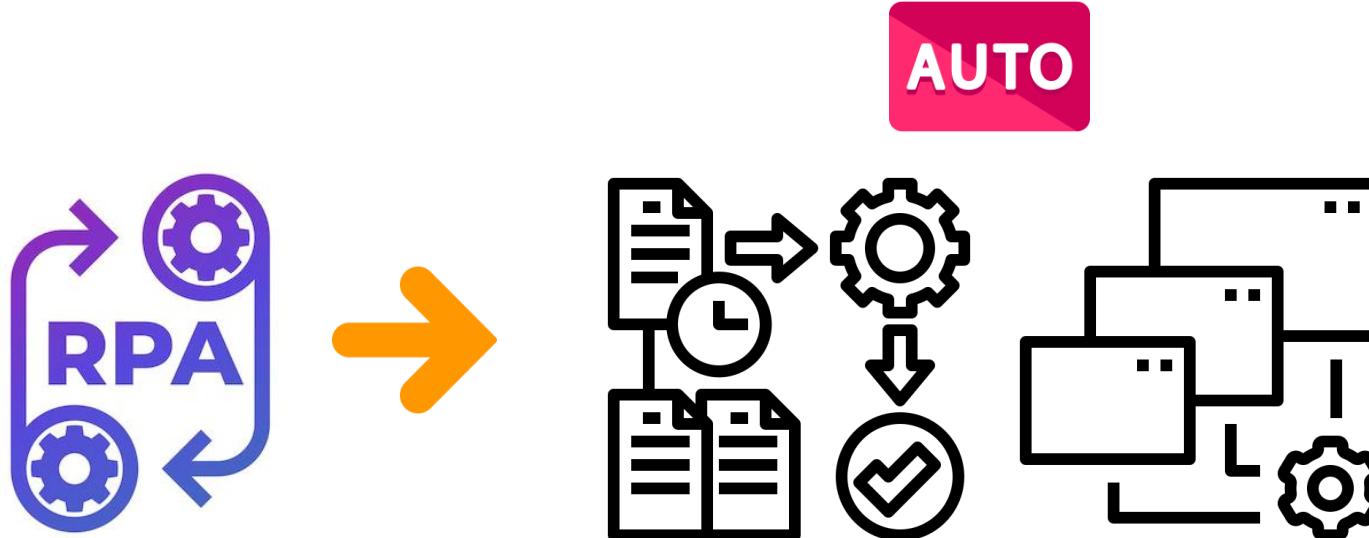
# Wrap-up: Retention Strategies



# Future Work



# Future Work



# CHULA ΣENGINEERING

Innovation toward Sustainability | ΑCTNOW