



# Telco customers churn analysis

Kalvin Chan, Yuri Jeong, Ahmed Gherfal

Winter 2022

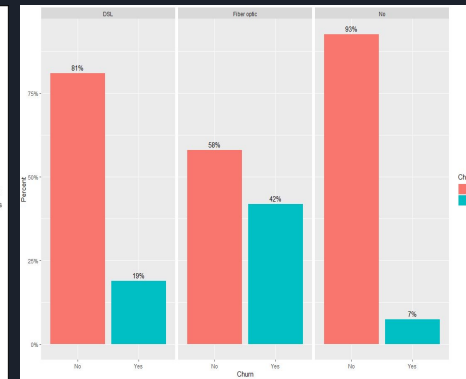
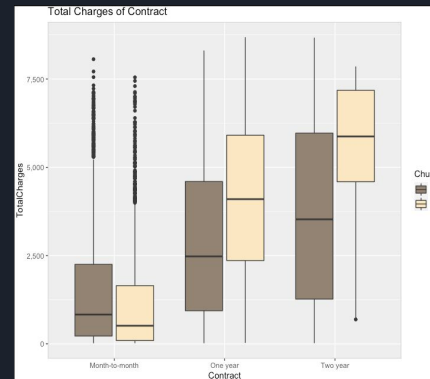
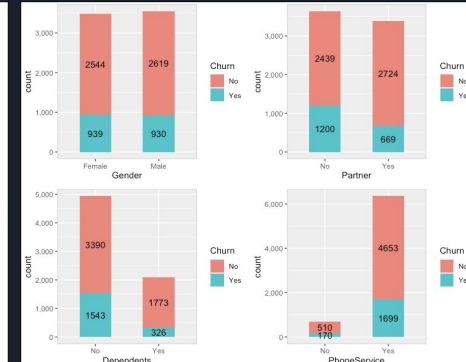
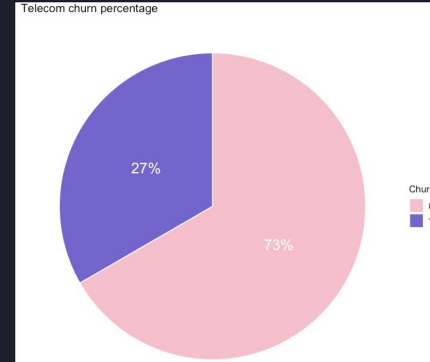
# Introduction

- How to retain customers in business environment
- Telecom Churn dataset
  - why customers were leaving the network provider?
- Exploratory Data Analysis(EDA)
- Correlation
- Feature Selection
  - Stepwise Selection
- Logistic Regression
  - ROC Curve, AUC value
- Evaluation
  - Confusion Matrix
- Conclusion



# Exploratory Data Analysis (EDA)

- Total 7,044 rows, 20 columns
- Churn - No 5,163(73%), Yes 1,869(27%)
- Contract and Total Charges
  - Month-to-month, one year, two year
  - \$18.8 ~ \$8,684.8
  - Longer contract → Higher charges
  - Total charges → Different churn range
- Gender(Female/Male)
  - Similar Churn
- Partner(No/Yes)
  - More Churn when people have no partner
- Dependents(No/Yes)
  - More Churn when people are independent
- Phone Service(No/Yes)
  - A lot more churn with phone service
- Internet Service(DSL/Fiber Optic/No)
  - Fiber Optic - Most Churn rate
- Churn → SMOTE() to balance the rate



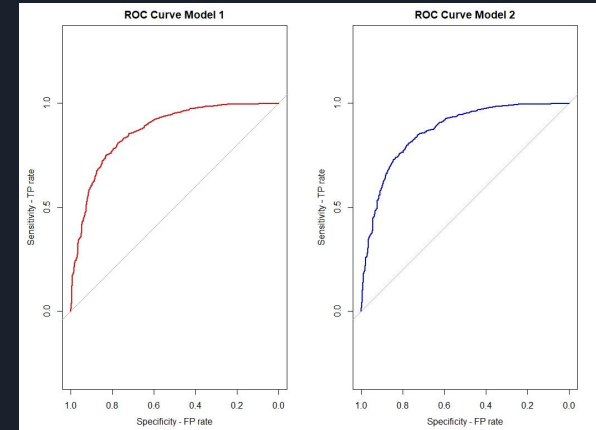
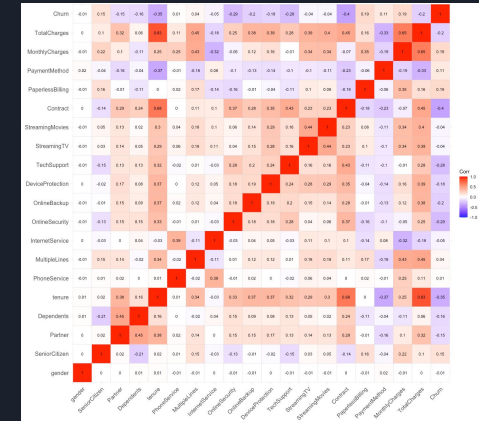
```
> table(balanced.data.train$Churn) > table(balanced.data.test$Churn)
```

	No	Yes
Count	2618	2618

	No	Yes
Count	1120	1120

# Logistics Regression Analysis

- Contract correlation -0.4
- Tenure correlation -0.35
- Feature Selection
  - Stepwise method
  - Model 1
    - Dependents, Tenure, multiple lines, internet service, online backup, device protection, streaming TV, streaming movies, contract, paperless billing, payment method, and total charges.
- Model 2
  - Statistically significant from original model
  - Similar predictor variables as Model 1 except online back up
- Comparison
  - Model 1 and Model 2
    - AIC
    - ROC
    - AUC
- Conclusion
  - Model 1 is significantly better than model 2





# Evaluation

## Model 1 Confusion Matrix

- True Negative = 847      False Negative = 203
- False Positive = 273      True Positive = 917
- Accuracy : 0.78
- Sensitivity : 0.818
- Specificity : 0.756
- Precision : 0.7706

## Confusion Matrix and Statistics

	Reference	
Prediction	No	Yes
No	847	203
Yes	273	917

Accuracy : 0.7875  
95% CI : (0.77, 0.8043)  
No Information Rate : 0.5  
P-Value [Acc > NIR] : < 2.2e-16

Kappa : 0.575

Mcnemar's Test P-Value : 0.001564

Sensitivity : 0.8187  
Specificity : 0.7562  
Pos Pred Value : 0.7706  
Neg Pred Value : 0.8067  
Prevalence : 0.5000  
Detection Rate : 0.4094  
Detection Prevalence : 0.5312  
Balanced Accuracy : 0.7875

'Positive' Class : Yes



# Conclusion

- EDA was conducted
  - Relative graphs were provided and interpreted
  - Audience have an understanding of the dataset
- Two variables have mild correlation to Churn
  - Contract
  - Tenure
- Solutions
  - Improve contracts policy, such as, activation fee and hidden fee
  - Give out free subscription services like Spotify and Netflix for contracted customers
  - Used phone trade in promotion to make customers stay longer or even extend the contract
  - VIP status, point earning program



# Reference

BlastChar. (2018, February 23). Telco customer churn. Kaggle. Retrieved February 1, 2022, from <https://www.kaggle.com/blastchar/telco-customer-churn>

How to put labels over geom\_bar for each bar in R with ggplot2. (2019, July 25). Intellipaat Community. <https://intellipaat.com/community/16343/how-to-put-labels-over-geombar-for-each-bar-in-r-with-ggplot2>

R Plot Mean in ggplot2 Barplot (Example) | Show Averages in geom\_bar. (2021, December 13). Statistics Globe. <https://statisticsglobe.com/plot-mean-in-ggplot2-barplot-r>

Datatricks, One-Hot Encoding in R: Three simple methods. Retrieved February 7, from <https://datatricks.co.uk/one-hot-encoding-in-r-three-simple-methods>

RDocumentation, Cor: Covariance and Correlation (Matrices), Retrieved February 7, from <https://www.rdocumentation.org/packages/DescTools/versions/0.99.44/topics/Cor>

Bevans, R. (2021, June 18). An introduction to the akaike information criterion. Scribbr. Retrieved February 9, 2022, from <https://www.scribbr.com/statistics/akaike-information-criterion/>