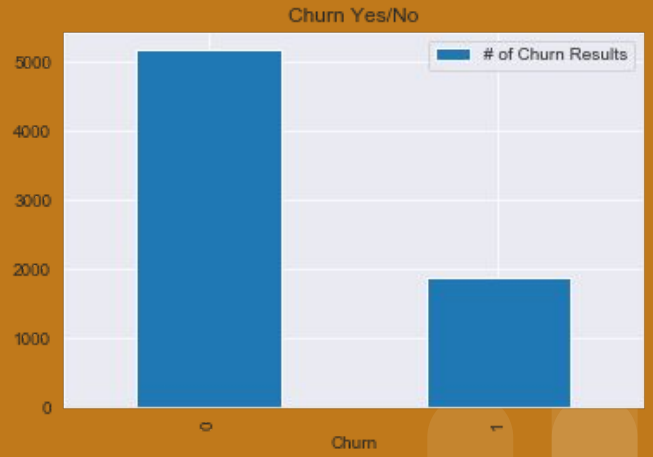
Telco Customer Churn Presentation

Flatiron School Module 3 Data Science Project

By: William Newton

Data and Methodology



Data and Methodology

- Data for this project is the Telco Customer Churn Data available on Kaggle
  - https://www.kaggle.com/blastchar/telco-customer-churn
- Data set contains the following information from an anonymous Telecom company
  - 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
- Used the OSEMN method to obtain, clean, explore, model, and evaluate data set

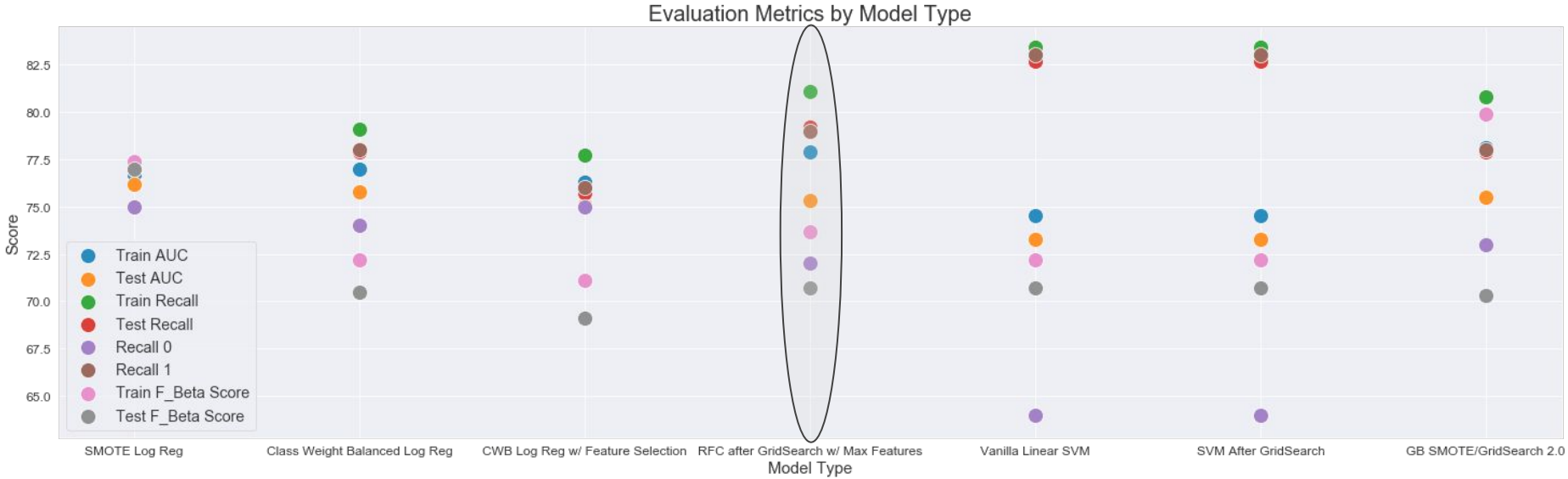
Goal

Use the data set to predict customer churn using Machine Learning models

Model Results

	Model #	Model Type	Train AUC	Test AUC	Train Recall	Test Recall	Recall 0	Recall 1	Train F_Beta Score	Test F_Beta Score	Notes	Train_Test_Difference
0	1	Vanilla Log Reg	73.2	70.2	57.1	51.9	89.0	52.0	58.7	53.6	Baseline performs poorly when attempting to pr...	3.0
0	2	SMOTE Log Reg	76.7	76.2	77.4	77.0	75.0	77.0	77.4	77.0	Model performed better, additional tuning needed	0.5
0	3	Class Weight Balanced Log Reg	77.0	75.8	79.1	77.9	74.0	78.0	72.2	70.5	Balanced class weights results in slightly bet...	1.2
0	4	CWB Log Reg w/ Feature Selection	76.3	75.1	77.7	75.7	75.0	76.0	71.1	69.1	Slightly worse model, moving on to different m...	1.2
0	5	Vanilla Random Forest	99.7	68.0	99.5	46.7	89.0	47.0	99.5	48.9	Significant Overfitting	31.7
0	6	Random Forest Classifier after GridSearch	80.9	74.5	82.9	74.4	75.0	74.0	76.7	68.1	Overfitting After GridSearch	6.4
0	7	RFC after Feature Selection	80.3	74.4	81.8	74.0	75.0	74.0	75.8	67.8	Overfitting After Feature Selection	5.9
0	8	RFC after GridSearch w/ Max Features	77.9	75.3	81.1	79.2	72.0	79.0	73.7	70.7		2.6
0	9	Vanilla Linear SVM	74.5	73.3	83.4	82.7	64.0	83.0	72.2	70.7	Highest Recall Score	1.2
0	10	SVM After GridSearch	74.5	73.3	83.4	82.7	64.0	83.0	72.2	70.7	No Change	1.2
0	11	Vanilla Gradient Booster	75.4	69.6	59.4	49.5	90.0	50.0	61.5	51.7	Overfitting	4.8
0	12	Gradient Booster SMOTE	85.4	74.5	87.8	69.0	80.0	69.0	86.9	65.6	Overfitting Worse	10.8
0	13	GB Feature Select	74.9	69.5	58.6	49.8	89.0	50.0	60.7	51.8	Lost Predictive Power	5.4
0	14	GridSearch GB Model	73.8	70.0	56.5	50.4	90.0	50.0	58.8	52.5	Overfit and Underpowered	3.8
0	15	GB SMOTE 2.0	82.3	75.0	86.3	75.1	75.0	75.0	85.0	68.7	Overfitting Worse	7.3
0	16	GB SMOTE/GridSearch 2.0	78.1	75.5	80.8	77.9	73.0	78.0	79.9	70.3		2.6

Ran 4 different types of models using the data set



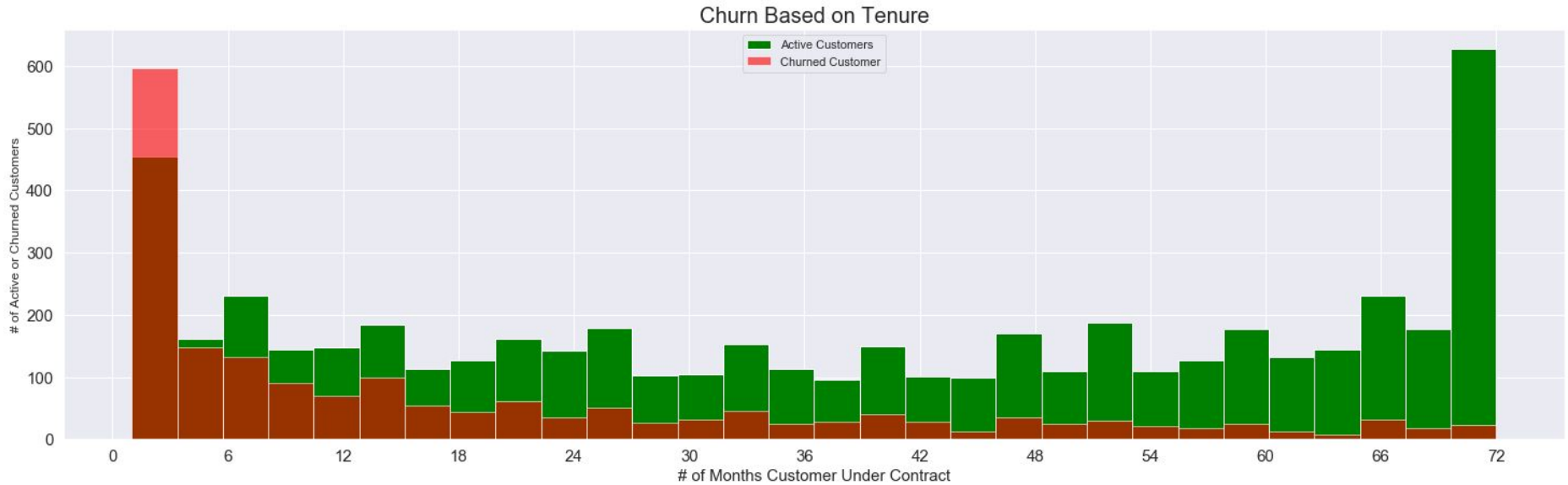
And visualized the evaluation metrics to determine which one performed the best

The 3 Best Predictors of Customer Churn Were...

- Tenure
  - How long has the customer been with the company?
- Contract
  - Is the customer month-to-month or 1-2 year contract?
- Internet Service Plan
  - Is the customer signed up for a DSL or Fiber Optic plan?

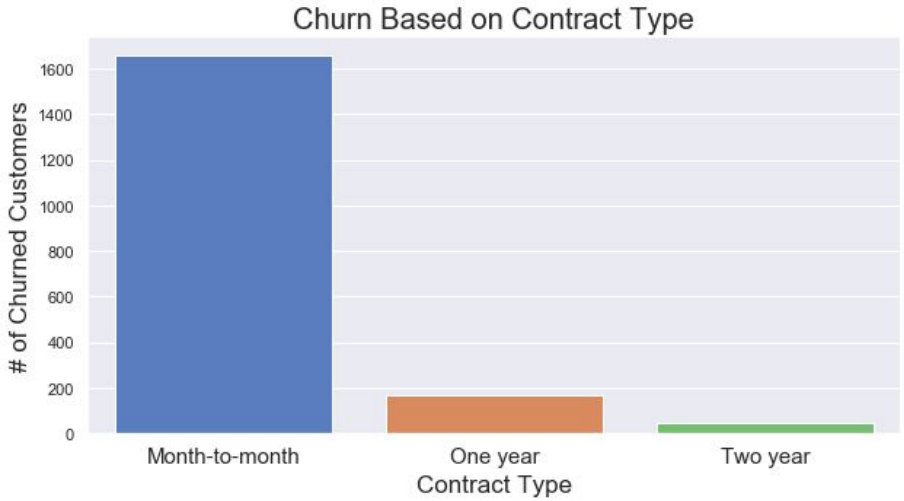
Tenure

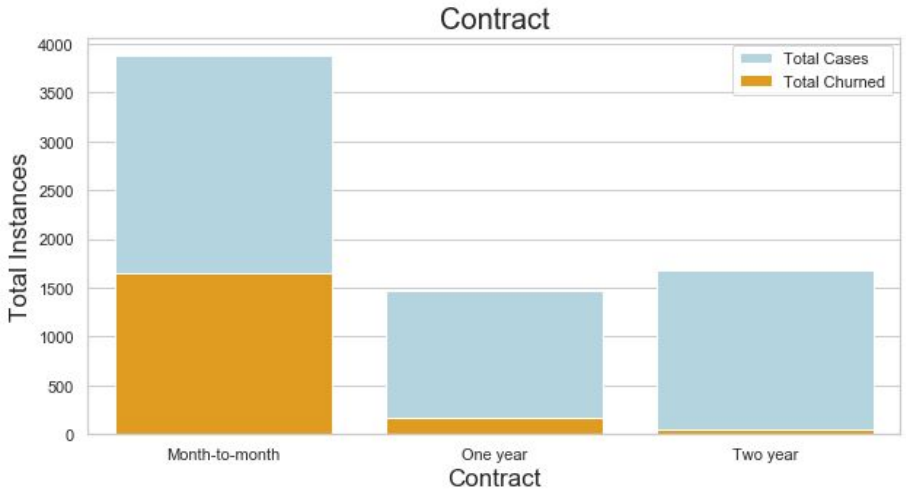
Happier Customers Tend to Stick Around



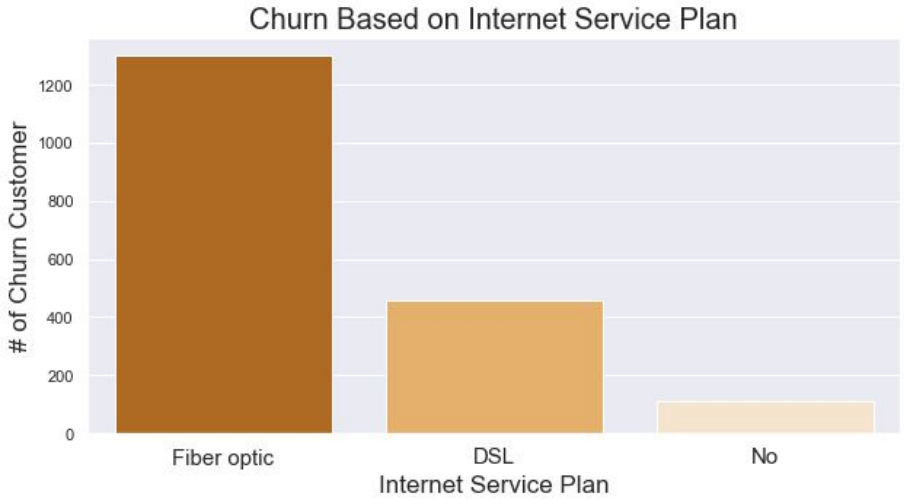
Contract

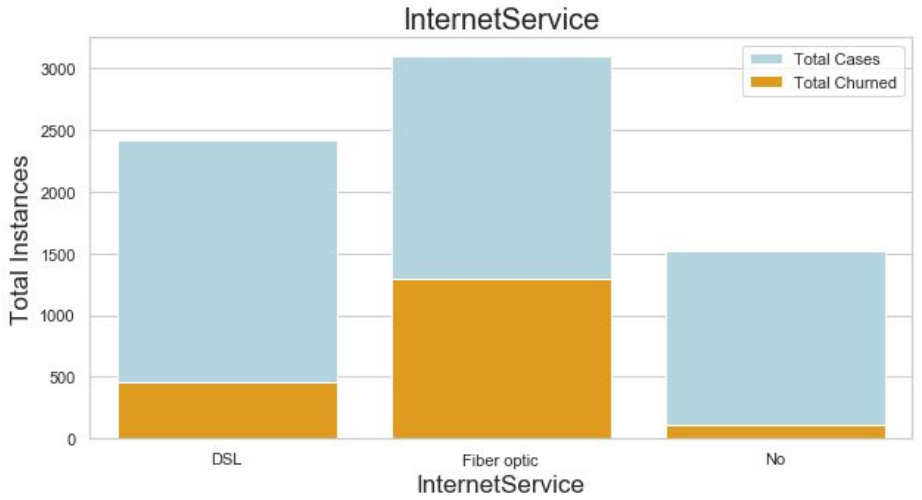
Don’t Give Customers the Option to Leave



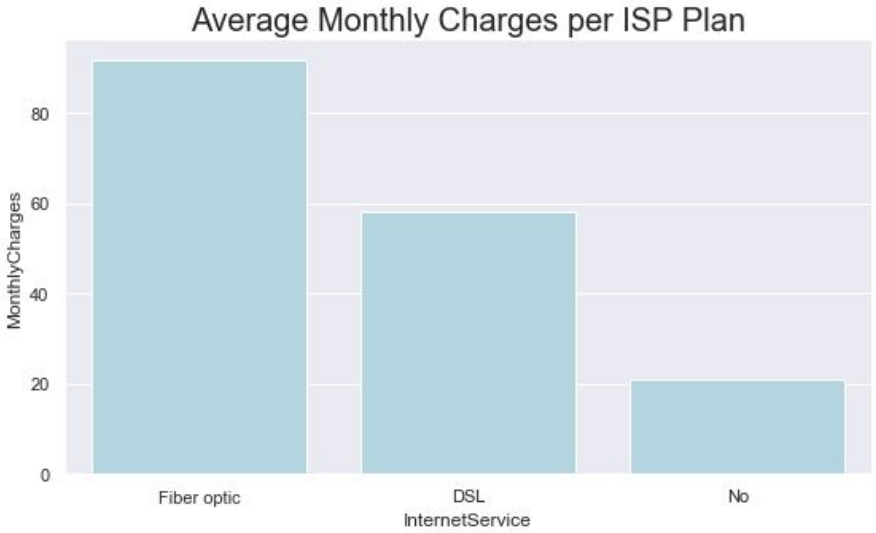


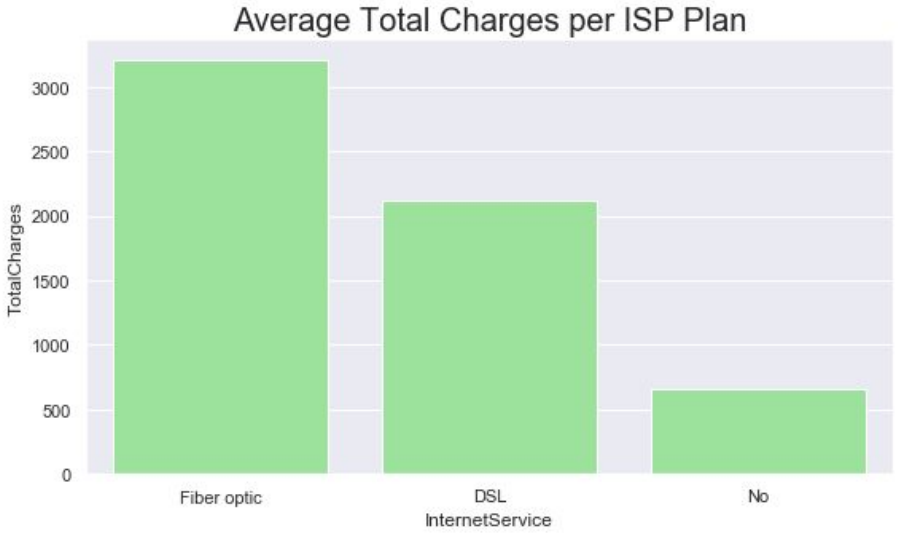
Internet Service Plan
DSL or Fiber Optic?



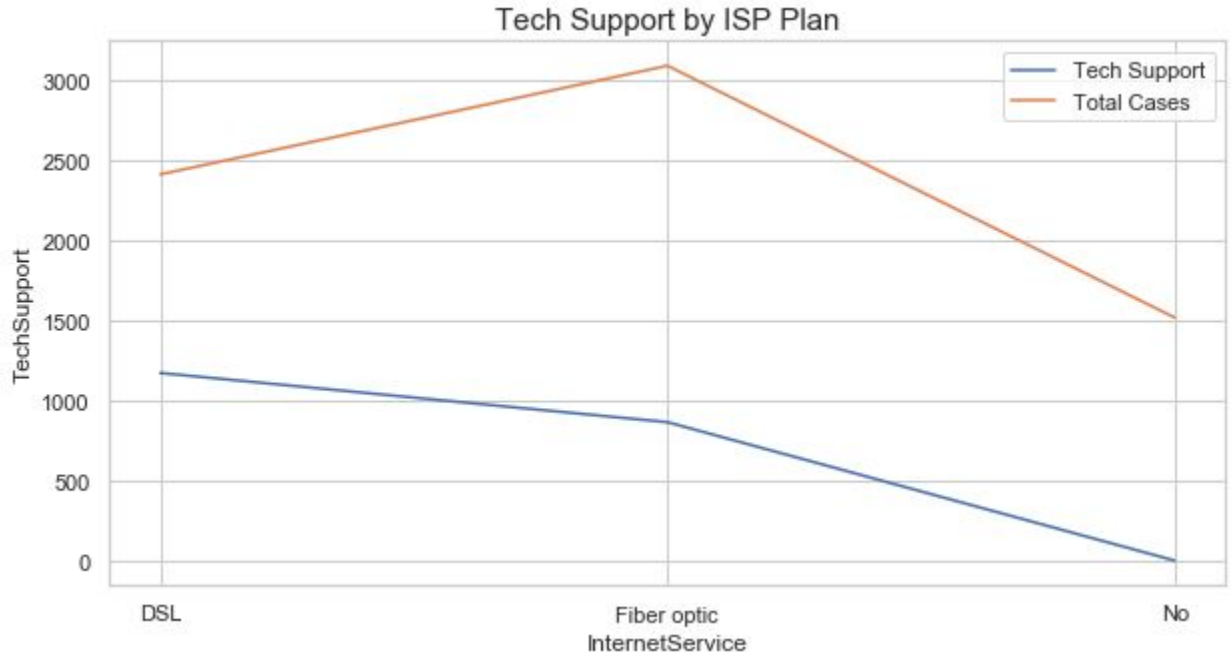


Fiber Optic Internet is a More Expensive Option

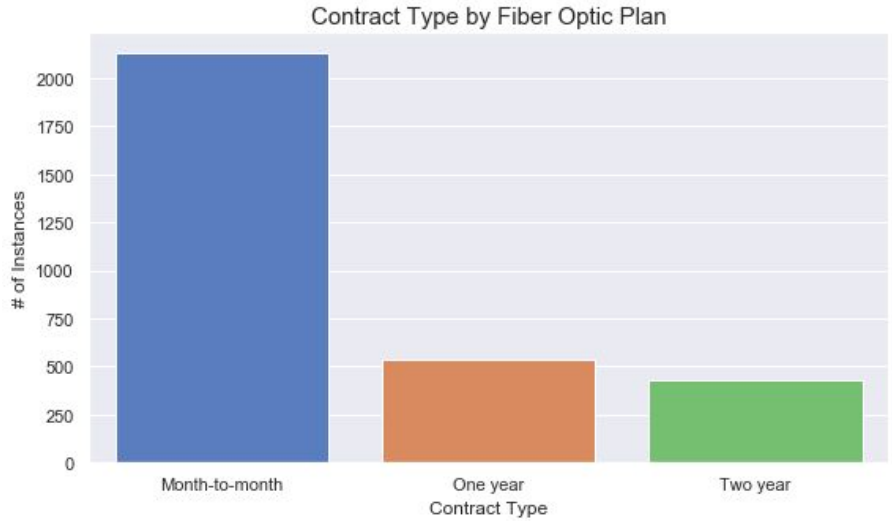


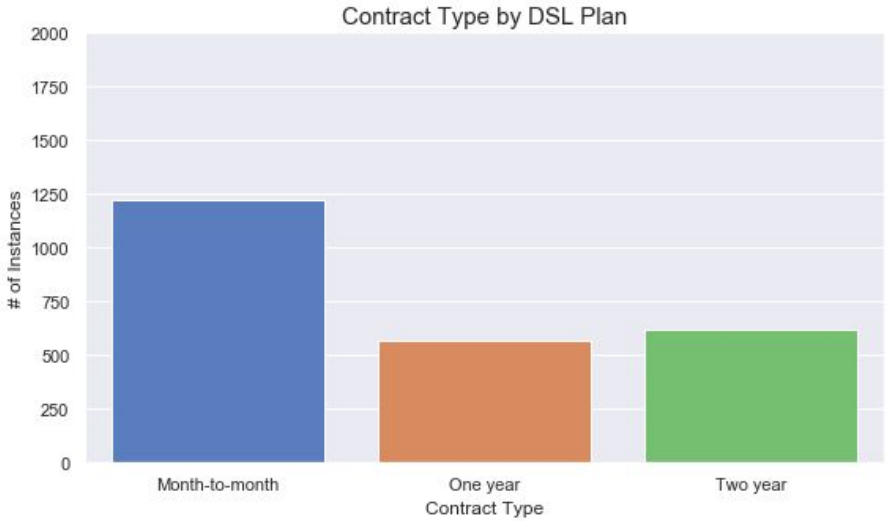


Less Tech Support Plans for Fiber Internet Subscribers



More Month to Month Contracts for Fiber Optic Subscribers





Future Work and Sources Cited

- ➔ Follow-up with the company after business recommendations have been followed
- ➔ Gather additional data
- ➔ Address further churn concerns after another round of modeling
- ➔ Use data gained from customer survey to further model and address concerns

Sources

- Churn Reduction in the Telecom Industry, Database Marketing Institute
  http://www.dbmarketing.com/telecom/churnreduction.html
- Customer Churn: 12 Ways to Stop Churn Immediately, Super Office
  https://www.superoffice.com/blog/reduce-customer-churn/
- DSL vs Cable vs Fiber: Comparing Internet Options, BroadbandNow
  https://broadbandnow.com/guides/dsl-vs-cable-vs-fiber

Thank you and I look forward to working together!