Sergio Gutierrez

Telecom Customer Churn Prediction

Background:

Telco is a telecommunications company that offers a myriad of base and bonus packages surrounding their phone and Internet services. For their phone service, they offer the option to purchase multiple phone lines. For their Internet service, they offer add-ons like streaming TV, streaming movies, online security and online backup. Some other standalone add-ons are device protection and tech support. Telco's services are that of a traditional telecom giant, except for not offering TV as a base service.

Problem Statement:

With the increasing competitive offerings from other mobile, digital and streaming service providers contributing to the growing industry cord-cutting and customer churn, Telco is growing more concerned with understanding churn: that is, when the customer ends their contract and thus their relationship.

Telco understands churn is a normal part or the customer lifecycle but they are more interested in better serving its loyal customers and retaining the at risk ones in an increasingly competitive and fragmented market. Therefore, they would like to predict churn to proactively prevent and reduce customer losses.

Using their customer account information, demographic profiles and services purchased, the telecom provider expects to predict customer churn and thus proactively work to put in place customer programs and packages that will better retain and up-sale more current customers amidst the growing competition.

Dataset Description:

The original dataset was provided by Telco and it was published by IBM on their data science developer site. It eventually was posted and made only available on Kaggle. The dataset contains a little over 7,000 customer accounts, with information on their demographics, services purchased, tenure (or number of months as an ongoing customer), payment methods and whether they have churned or not. The maximum customer tenure provided in this dataset is 72 months (or six years).

Data Wrangling:

Upon loading the dataset from Kaggle, the dataset was inspected for its contents, containing 7,043 rows (or observations) and 21 columns (or variables). Surprisingly, given the large set of data points, the dataset did not have any missing or null values; so none of these data points had to be cleaned.

Of the 21 columns, only four columns were numeric, which were Senior Citizen, Tenure, Monthly Charges and Total Charges. However, Total Charges was loaded as a string like the other non-numeric columns. To correct this, Total Charges was converted into a float like Monthly Charges.

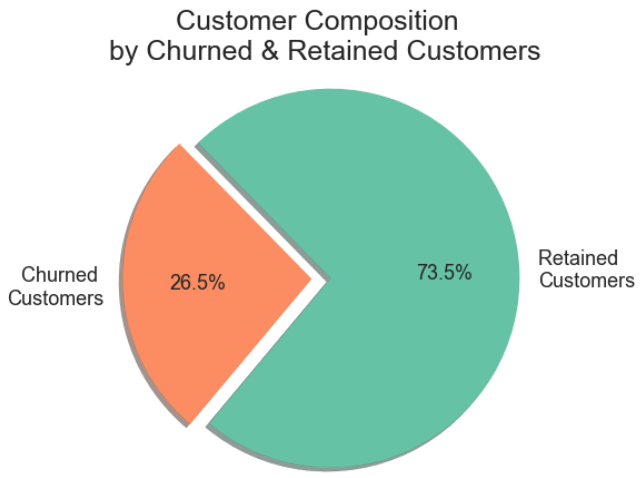
As for the other 17 non-numeric columns, each of these category variables contained string values to describe the customer demographic profiles, account information and contract details and purchased services. In order to make these string values into usable ones, the pandas get_dummies function was used to convert each category variable (and each of their values) into binary dummy variables.

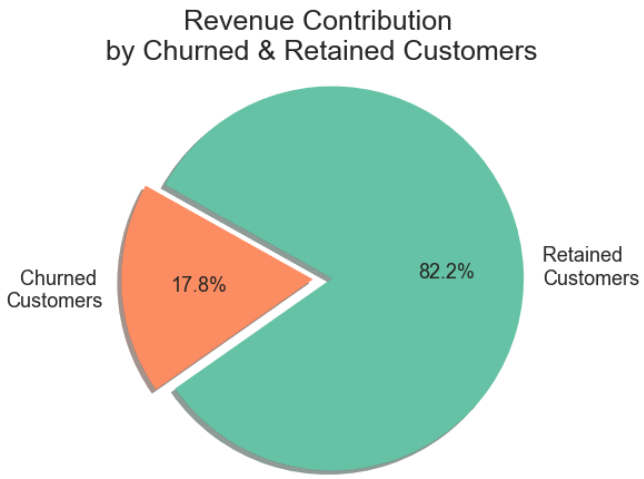
Lastly, upon converting categorical variables into dummy variables, some of the resulting dummy variable column names had blank spaces. To improve the ease of referencing these columns, blank spaces were replaced with underscores (e.g., Contract One_year to Contract_One_year).

Key Findings:

Customer Composition and Revenue Contribution

Based on this month's snapshot of its customer composition (illustrated below on the left), Telco has lost a little over a quarter of its customer base, which amounted to 26.5% (or close to 2,000 customers) of its total customers. This appears to be a substantial lost but we also need to assess the importance of these customers as it relates to their revenue contribution and tenure.



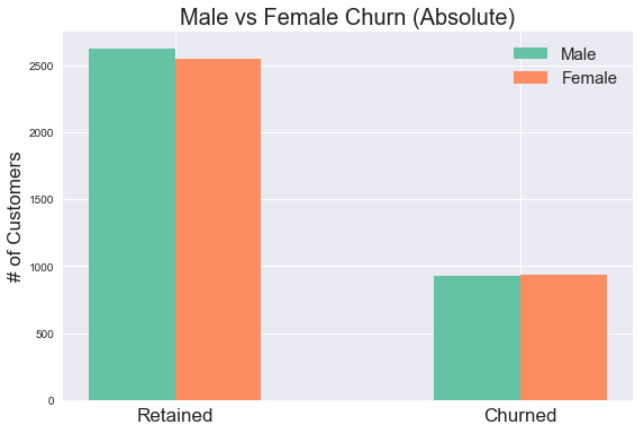


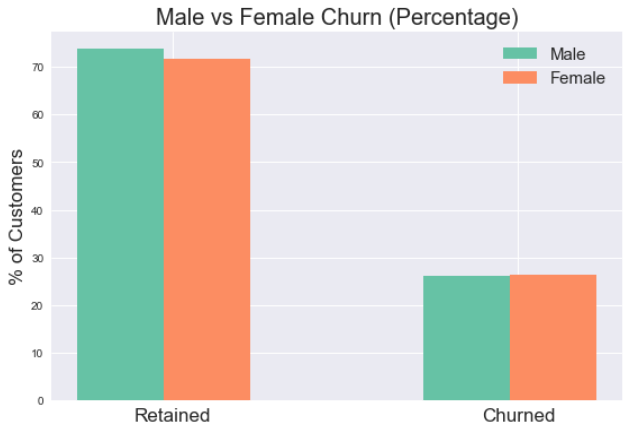
As we look to evaluate the revenue contribution from churned and retained customers (illustrated above on the right), the breakdown shows its churned customers makeup 18% (or close to $2.9MM) of its total revenue, while its retained customers makeup the remaining 82% (or $13.2MM). While Telco's churned customers comprised of 26.5% of customers, the provider's revenue losses from churned customers are not equally impacted since their revenue contribution percentage is less.

Additionally, the average revenue contribution from churned customers ($1,5312) was almost 60% lower than the average revenue contribution from retained customers ($2,555). While Telco is concerned with focusing on churn, the provider should acknowledge the value and importance of its retained customers as they are majority revenue contributors. Thus, any changes or improvements to policies and programs should not affect or harm its retained and most loyal customer base.

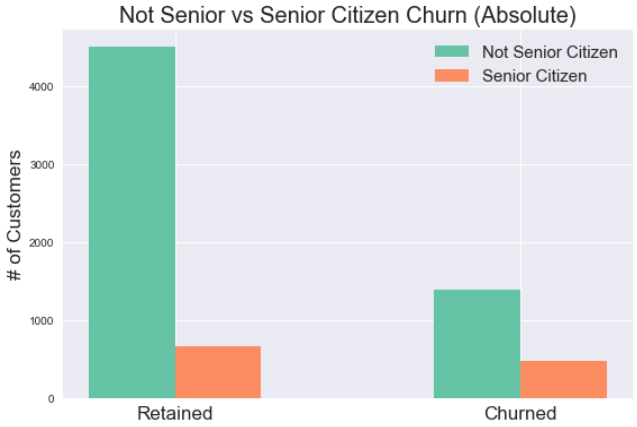
Customer Demographics and Households

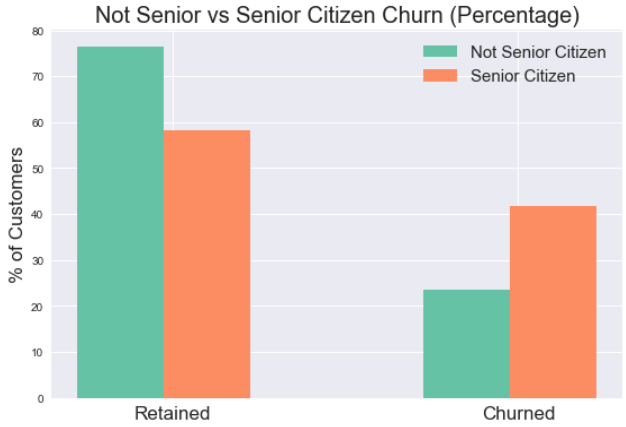
As we look to define and compare retained and churned customer profiles, several demographic and household profile insights can be inferred from the only four demographic variables from the data. Firstly, churned customers were equally male and female, with no particular skew toward a gender (pictured below on the left in absolute numbers and on the right in percentages). This comes as no surprise because such utility-like services for communication (and home entertainment) are essential for both genders.



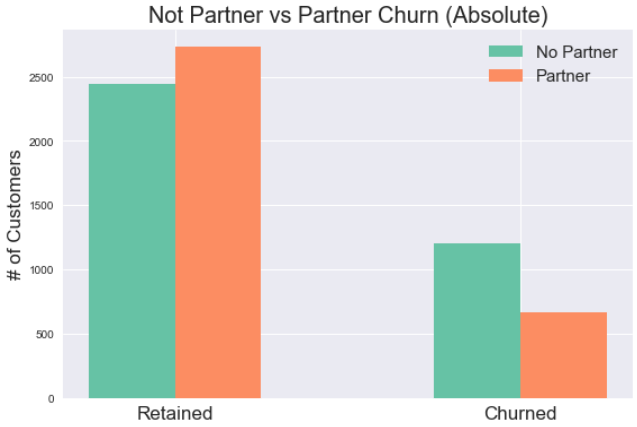


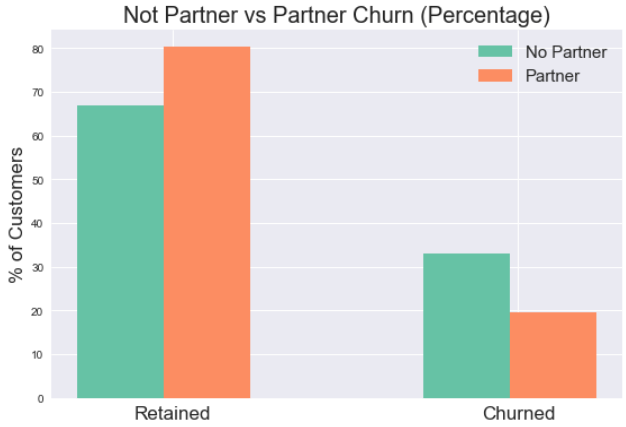
Moving to the next bar graphs comparing senior to non-senior citizens (below), senior citizens were more likely to churn than were non-senior citizens, as displayed in the percentage chart on the right. This is likely due to senior citizens being less likely to purchase and use Internet and streaming services. It could also be simply explained by natural causes in age that senior citizens are more likely to be lost customers.





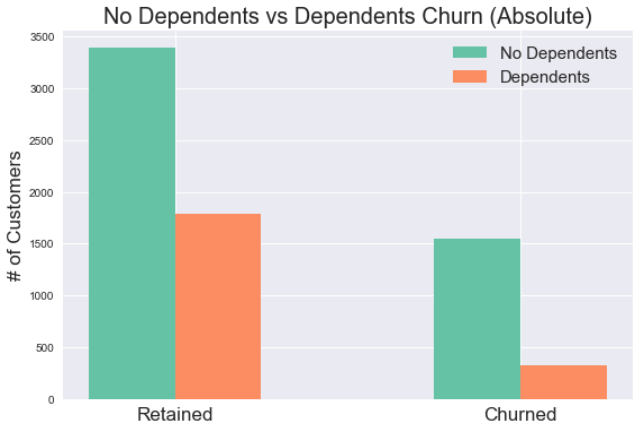
Continuing on to the partner versus no partner churned customer comparison, it seems that customers without partners are more likely to churn (pictured below on both the left and right bar graphs). This is likely driven by customers without partners having less financial support & security and thus are more likely to leave their service provider. It could also be that customers with partners are more likely to be families that heavily rely on these communication and entertainment services.

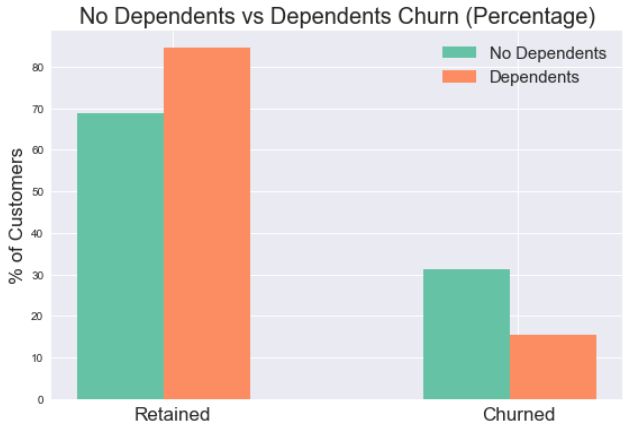




Lastly, comparing customers without dependents to those with, it appears that customers without dependents are likelier to churn than customers with. This is likely tied with partners comparison, as

customers with dependents are likely couple and/or family households with partners that can provide financial support and/or require such services for both adults and children.



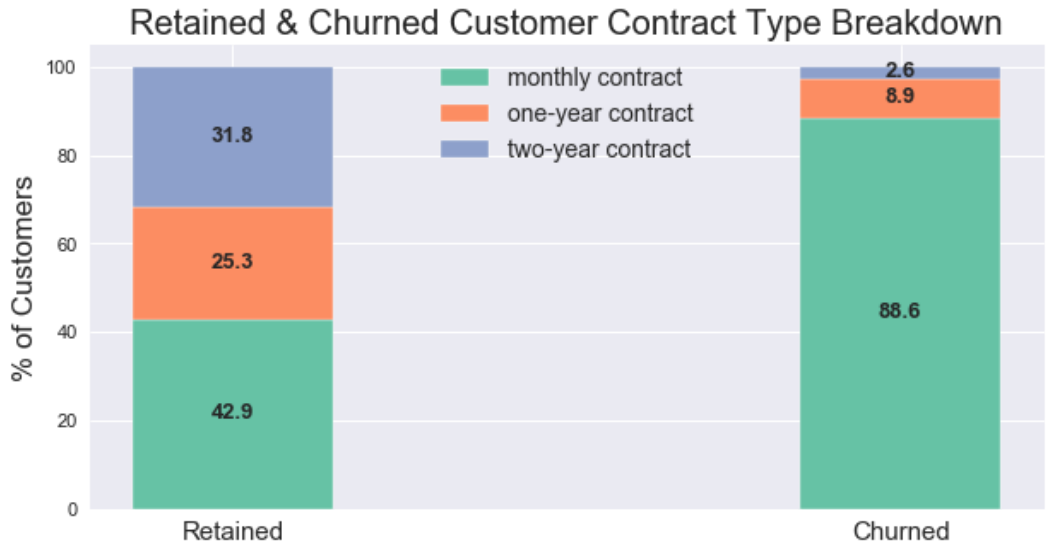


Customer Contract Timelines and Services Purchased

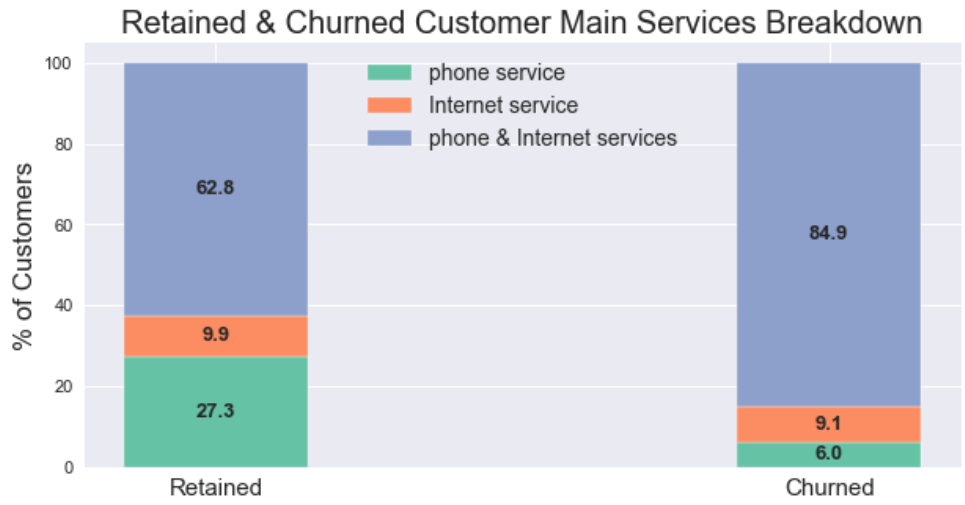
The next topic of interest is understanding the contracts from both retained and churned customers. What type of contract(s) was lead to the most churn? What combination of services lead to the same result?

Based on the stacked bar chart (shown below), customers on month-to-month contracts were the most likely to discontinue their service(s) with Telco (at 88.6% of churned customers), as these customers were neither obligated to stay with nor remain loyal to the provider (unlike those in one- or two-year contracts).

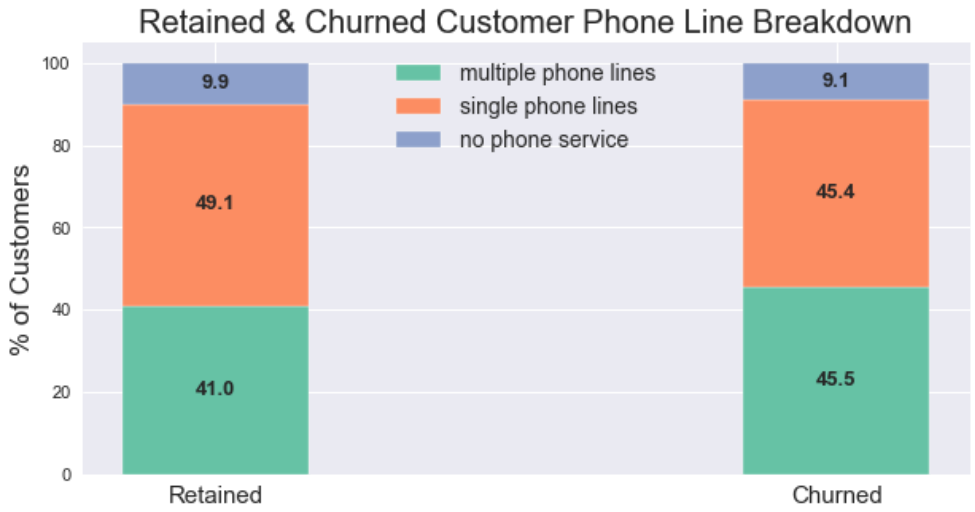
As it pertains to its retained customers, customers on month-to-month contracts comprised of over 40% of Telco's customer base, which was the highest amongst retained customer contracts. Unsurprisingly, customers on two-year contracts were more likely to be retained than those on one-year contracts.



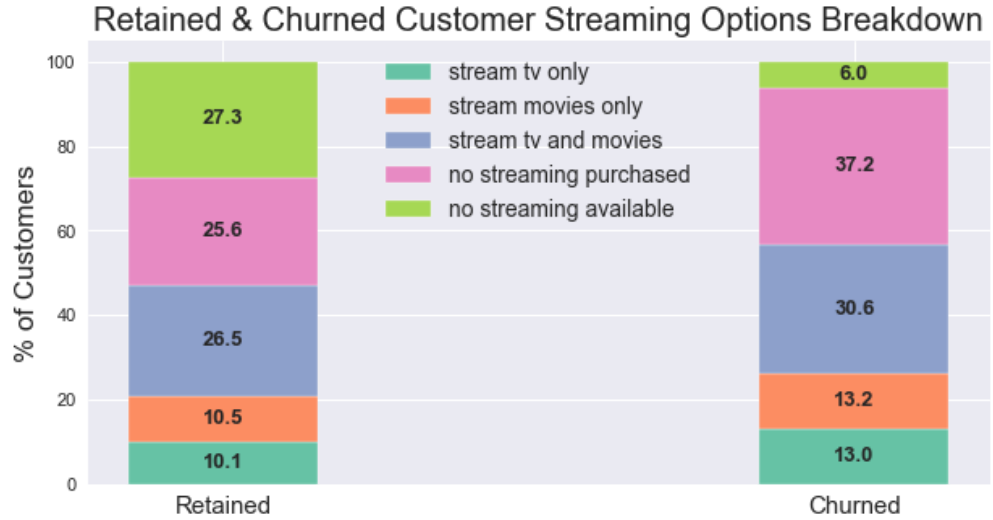
Looking into the main services purchasable in these contracts, over 60% of retained customers purchased both phone and Internet services, and over 25% of retained customers purchased phone service exclusively (illustrated below). As for churned customers, an overwhelming 85% of them purchased both phone and Internet services in tandem. Tying these results with contract types, many month-to-month customers purchased the phone and Internet services in tandem, which is surprising since a contract with more offerings should retain customers better but it does bring to question the weakness of that tandem.



Diving into the phone services purchased, the multiple phone line add-on was about equally popular with both groups and thus was likely not a major contributor to retaining or churning customers (shown below).

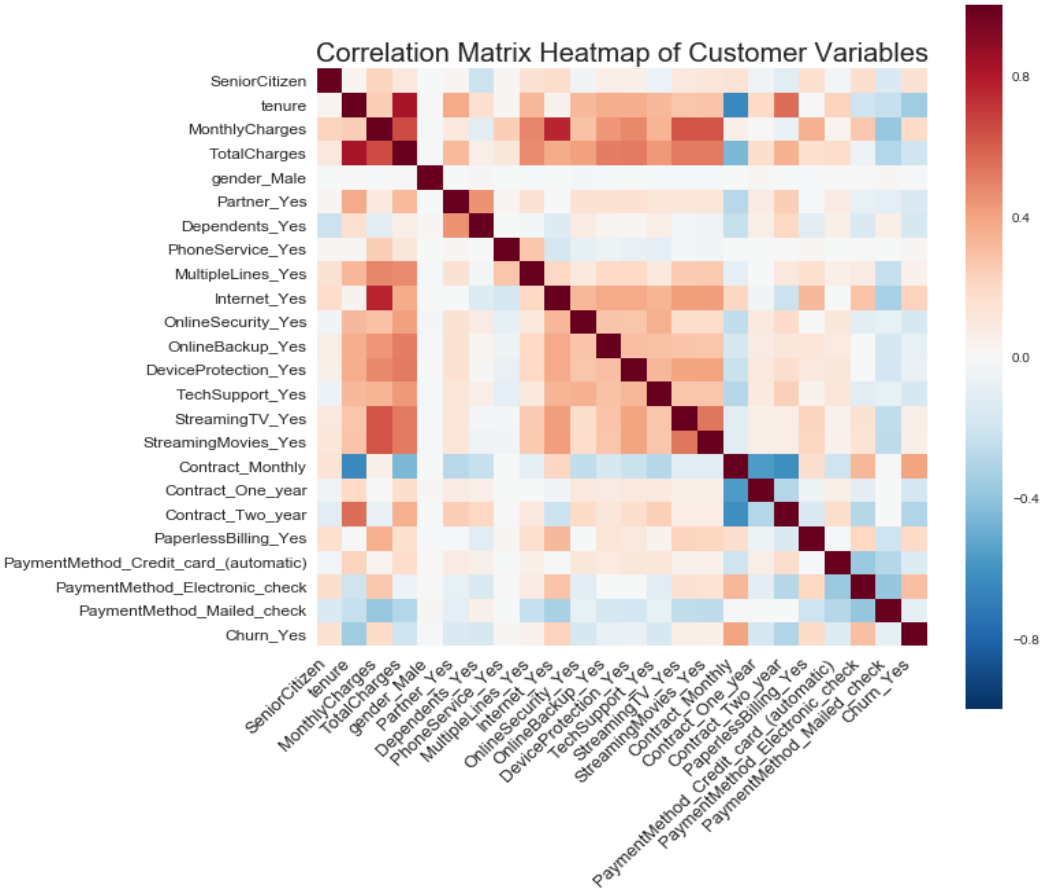


Lastly, diving into the Internet services purchased, the streaming add-on (for TV and/or movies) actually resulted in higher churn of customers than it did retain (shown below). Moreover, over 50% of retained customers were more likely to not purchase the add-on with their Internet or simply not purchase Internet. This brings to light that something may be wrong with the streaming service itself and/or Internet package.



Correlation Heatmap to Determine Key Variable Relationships

To narrow down which variables are highly correlated with each other and relate them to tenure or churn, the heatmap illustrated below summarizes these correlations. The strongest (or most positive) relationship was among tenure, monthly charges and total charges. This means that customers with higher monthly and total charges were more likely to have longer tenures with the provider and thus likelier to be retained.

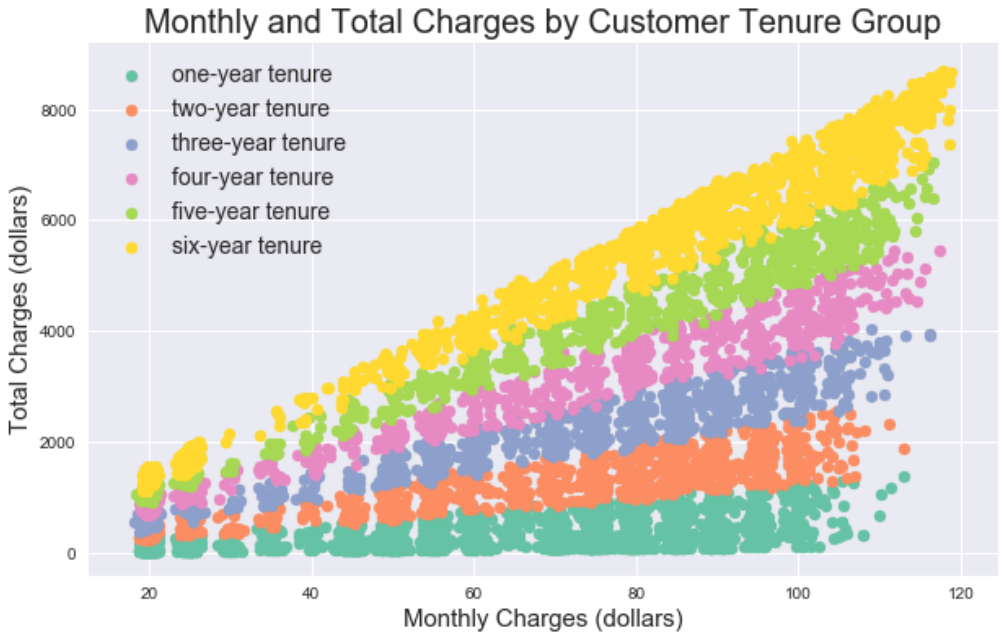


Another strong correlation is monthly charges and the combination of phone, Internet and add-on services (e.g., streaming TV and movies), which means the more services and add-ons purchased the higher monthly charges customers pay. Total charges also shares this relationship but to a slightly smaller extent.

Lastly, another standout correlation is the contract types (monthly, one year and two year) and tenure (months with the provider). As previously explored in the stacked bar graphs above, the higher number of months customers are contracted with the provider leads to longer time customers stay with the provider.

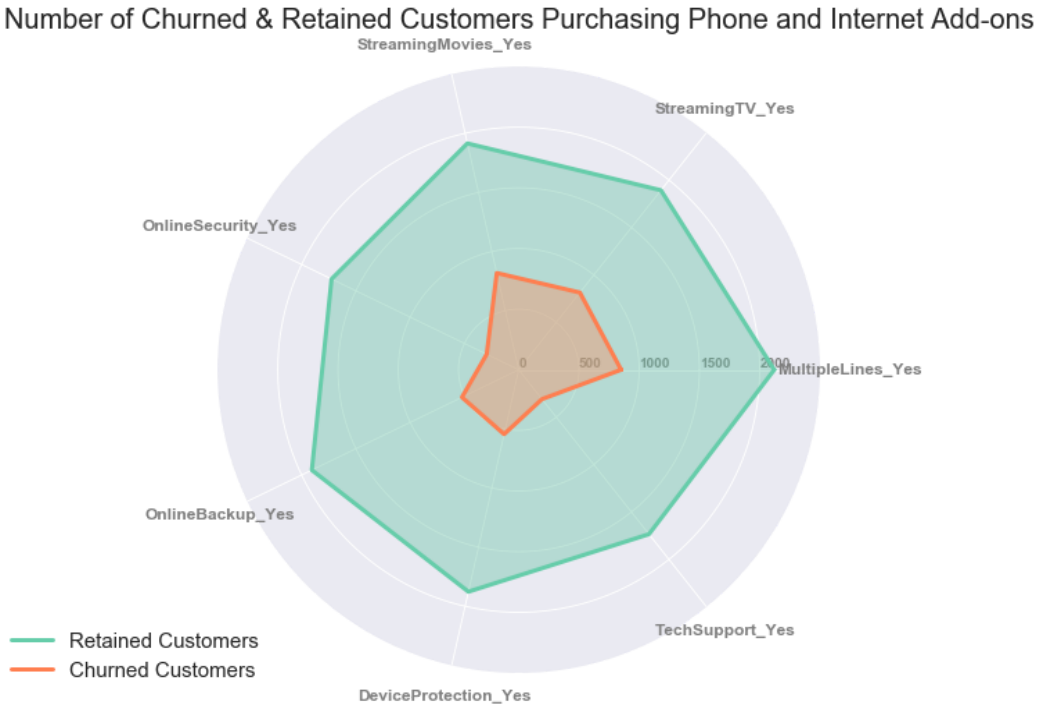
Customer Tenure Relationship with Monthly & Total Charges

Inspired by the heatmap above, the scatterplot below illustrates that customers have longer tenures when they have higher monthly charges or pay for more services and/or add-ons. This means customers that treated Telco as a one-stop shop telecom provider and purchased most or all of their services & add-ons were the most likely to continue their contracts for years and stay with the provider.



Customer Tenure Relationship with Phone and Internet Add-ons

As has been the trend, retained customers were more likely to purchase the add-ons (compared to churned customers) and in a well-rounded manner as shown below. As for churned customers, the add-ons they purchased more compared to others were streaming TV & movies and multiple phone lines.



Summary:

To review, Telco's primary issues with churn stem from customers on month-to-month contracts and from customers that do not bundle phone & Internet with the most add-on services. Additionally, Telco is losing more customers with demographic profiles of being single (no partner) and having no spouses and/or children (no dependents). With this foundational knowledge of the issues, we can move forward with creating a churn prediction model to empower Telco with the info to identify and retain at-risk customers.