

Name: Ahmed Yasser Mohammed Mohammed

Track: Data Analysis (Smart Village Branch)

CHURN ANALYSIS

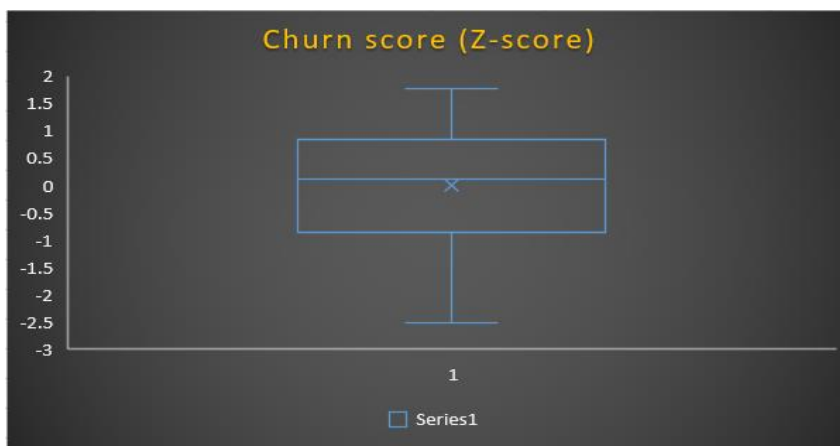
Data Validation and cleaning Report

1. There were 3 duplicate rows in (CUSTOMER STATUS) sheet and They were handled by being removed.
 2. Null values were Found within the churn label column and have been handled by filling it with the help of the Customer status column.
 3. The reason for the blank cells in churn category and churn reason is that those customers haven't left the company yet so they were filled with (not churned customer).
 4. Filling Null values in **Under30** and **senior citizen** columns by using age column.
 5. Filling the null values of **dependent** column by using the **number of dependents** column.
 6. Filling null values in **referred a friend** column by using **Number of referrals** column.
 7. Correcting the under 30 and senior citizen columns
-

Insights

1. Distribution of churn score:

No outliers found in the churn score



2. Investigating how the newly added measure **churn rate** changes between the 3 quarters:

Row Labels	Churn Rate
Q1	100%
Q2	100%
Q3	27%
Grand Total	27%

After investigating our data we found that the number of customers in Q1 and Q2 were very small and all of them were churned customers so the churn rate as shown is 100% in both Q1 and Q2, however, Q3 shows different results as the churn rate is 27%.



3. Investigating Churn reasons and Churn Category:

Most Frequent Churn reasons are (Competitor had better devices

- Competitor made better offer) and the reason with the highest average churn score is Service dissatisfaction

Row Labels	Count of Churn Reason	Average of Churn Score
Service dissatisfaction	63	83.6984127
Lack of affordable download/upload speed	30	82.8
Competitor made better offer	311	82.68167203
Lack of self-service on Website	29	82.51724138
Competitor had better devices	313	82.47923323
Long distance charges	64	82.1875
Don't know	130	82.06153846
Product dissatisfaction	77	82.03896104
Attitude of support person	220	81.79090909
Price too high	78	81.71794872
Limited range of services	37	81.7027027
Attitude of service provider	94	81.15957447
Extra data charges	39	81.15384615
Competitor offered more data	117	81.04273504
Competitor offered higher download speeds	100	81
Poor expertise of online support	31	80.32258065
Network reliability	72	78.91666667
Deceased	6	78.66666667
Poor expertise of phone support	12	77.83333333
Moved	46	77.67391304
Not Churned customers	5174	50.09818322
Grand Total	7043	58.50504047

However, on the level of Churn category we found that **Competitor** has the highest frequency and the highest average churn score

Row Labels	Count of Churn Reason	Average of Churn Score
Competitor	841	82.1783591
Price	211	81.90995261
Attitude	314	81.60191083
Dissatisfaction	303	81.42574257
Other	200	80.765
Not Churned customers	5174	50.09818322
Grand Total	7043	58.50504047



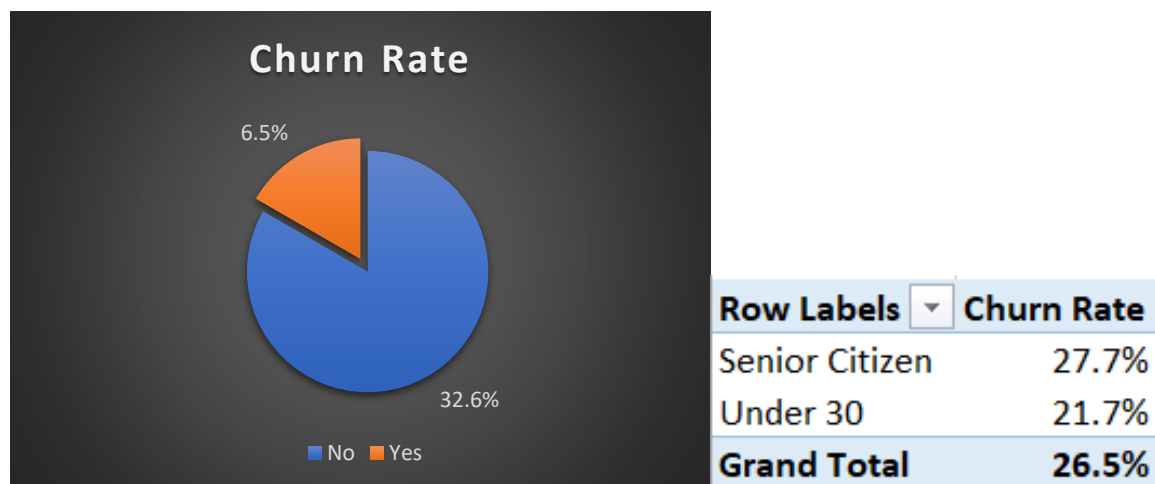
4. Investigating how demographics perform with churn score and Rate:

Gender→

Row Labels	Average of Churn Score	Churn Rate
Female	58.49139908	26.9%
Male	58.51842475	26.2%
Grand Total	58.50504047	26.5%

We can conclude that males and females have almost the same Average of Churn score but Female's churn rate is slightly higher than the Male's churn rate

Age→



Churn Rate for citizen whose age is greater than 30 is higher than the citizen who is under 30

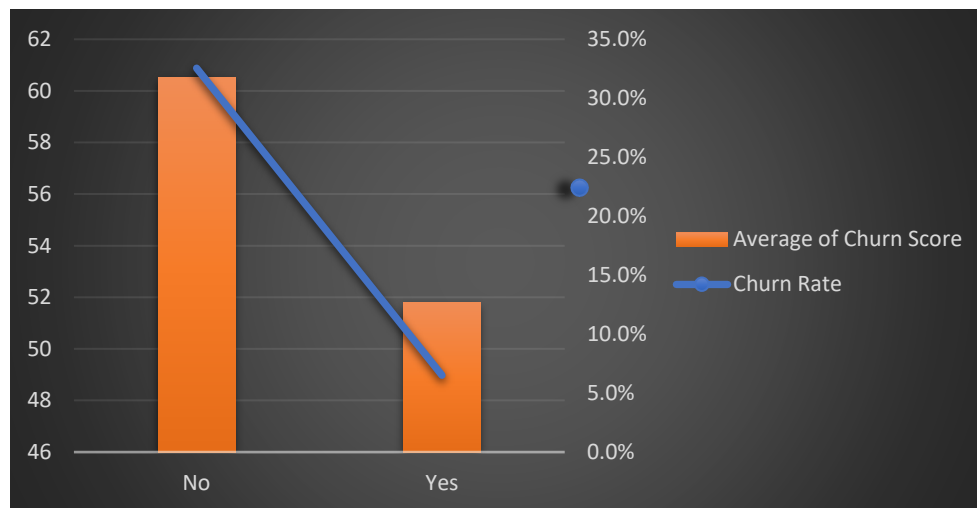
5. Marital Status:

Row Labels	Churn Rate	Average of Churn Score
No	33.0%	60.79071684
Yes	19.7%	56.05878895
Grand Total	26.5%	58.50504047

It seems that married people have less churn rate and average of churn score than unmarried people.

6. Having dependents using company's services:

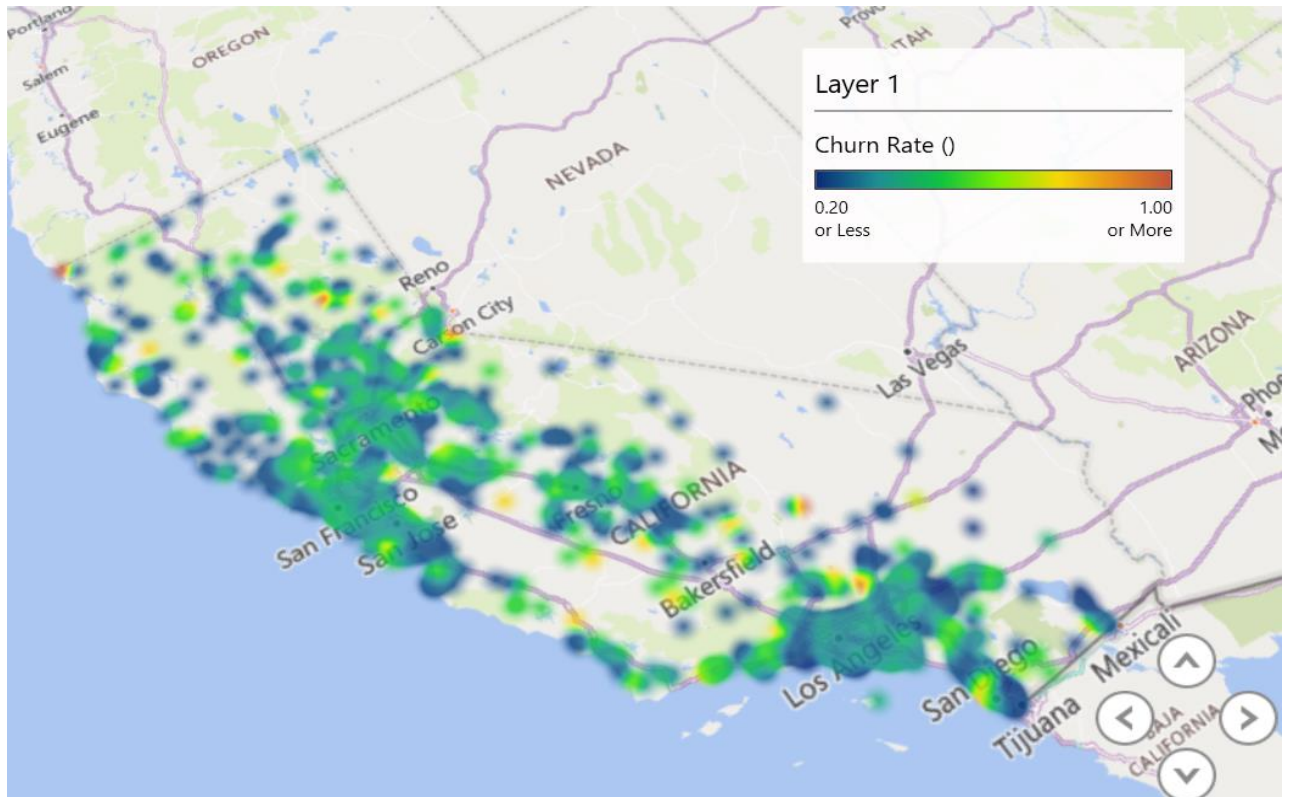
People that have dependents have less churn rate and average of churn score



7. Location:

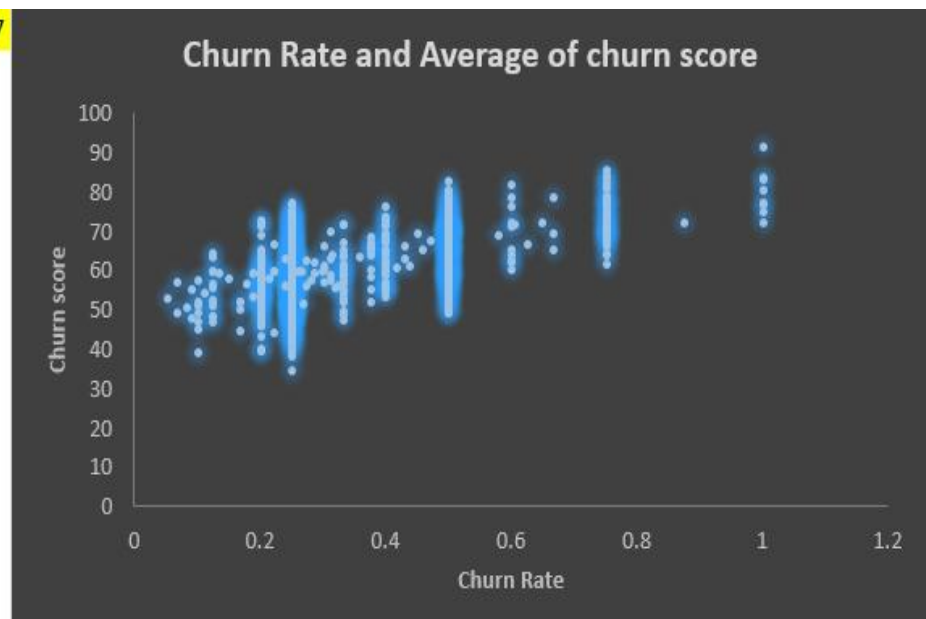
The churn rate is almost the same for most of the cities

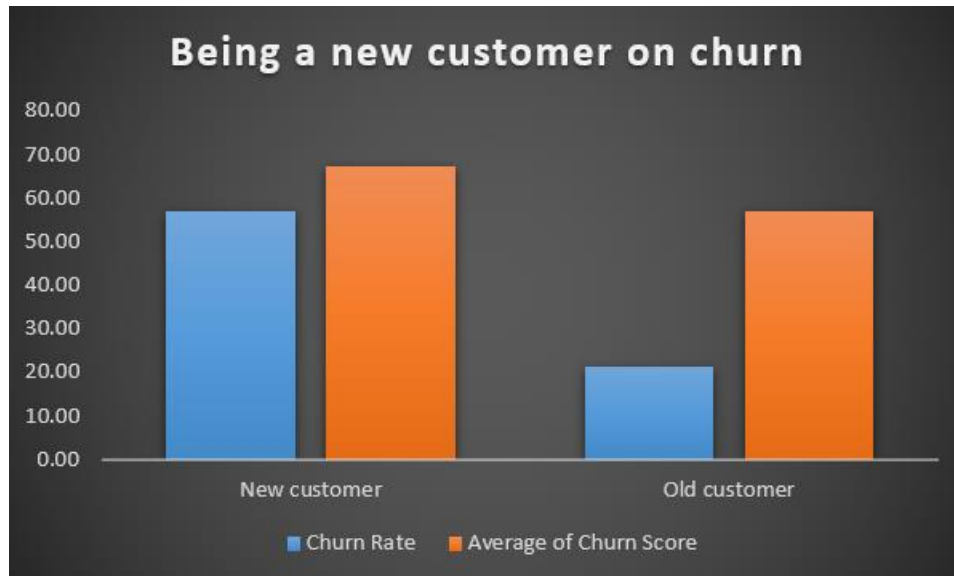
As shown:



We can also conclude that there is a moderate positive correlation between churn rate and average of churn score across cities:

Correlation coefficient 0.59917





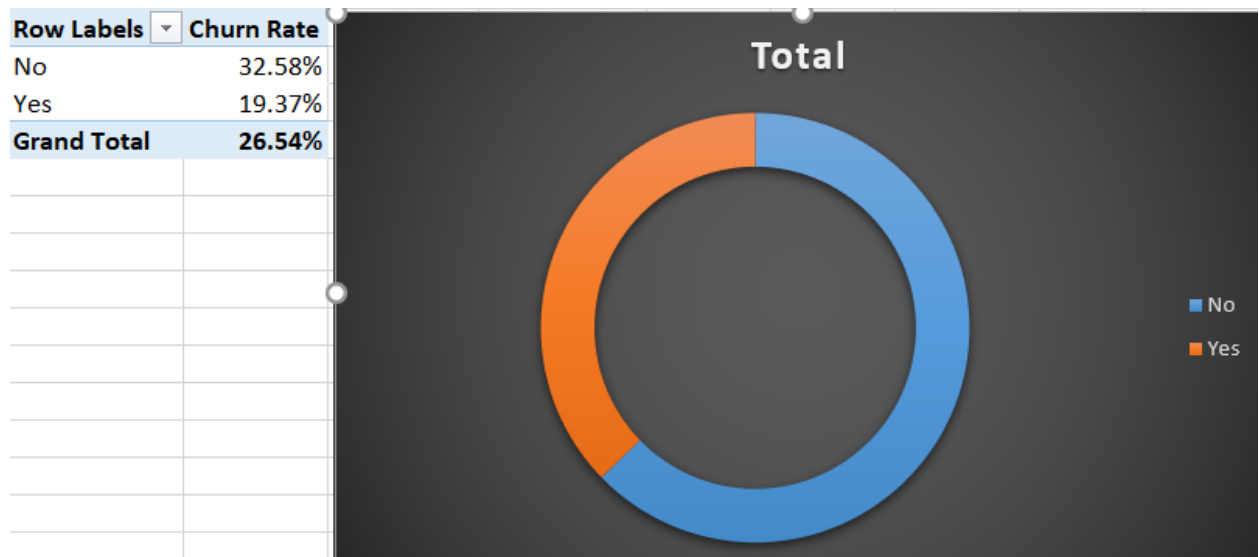
8.

It seems that new customers have more tendency to churn from the company

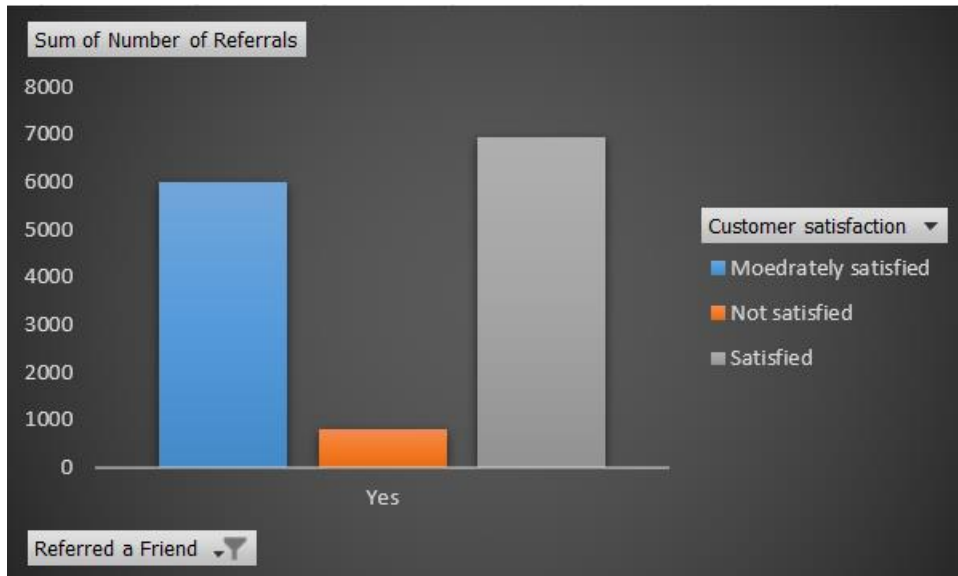


9. If the customer is satisfied that will encourage him to refer the company to a friend, and who is more likely to churn the one who referred a friend or not?

People who referred a friend less likely to churn from the company



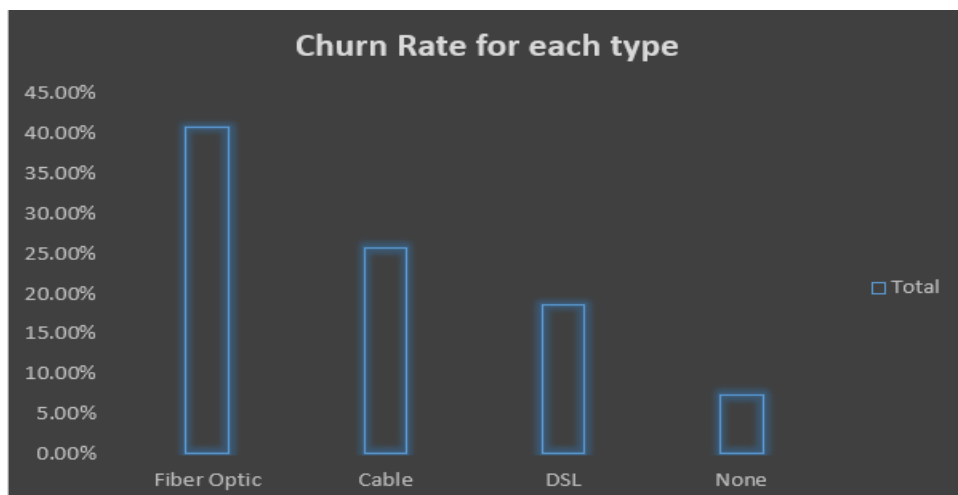
Satisfied customers more likely to refer a friend or more.



10. Churn rate is slightly higher for customers who have multiple lines than who don't:

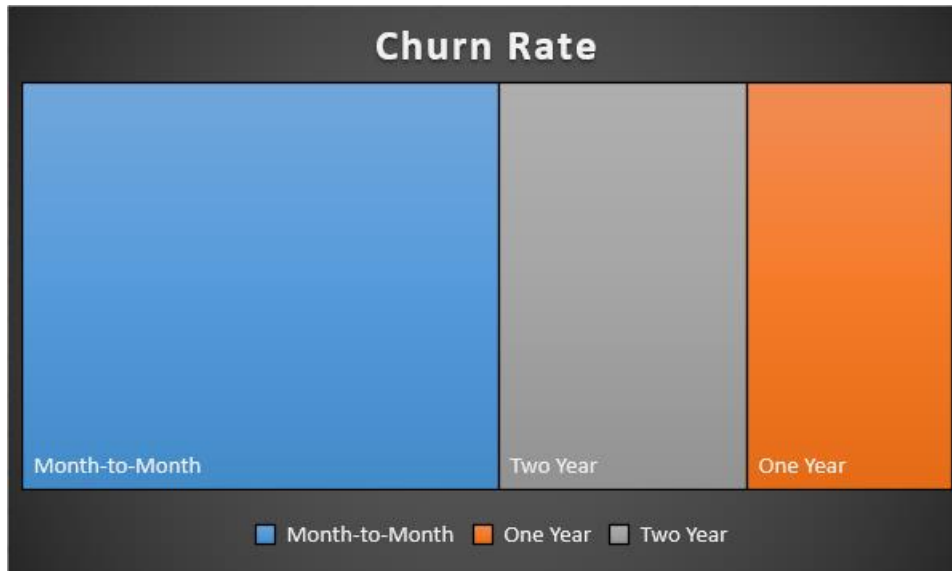
Row Labels	Churn Rate
No	0.25024558
Yes	0.286098957
Grand Total	0.265369871

11. What is the most frequently used internet type and which one has the highest churn rate?



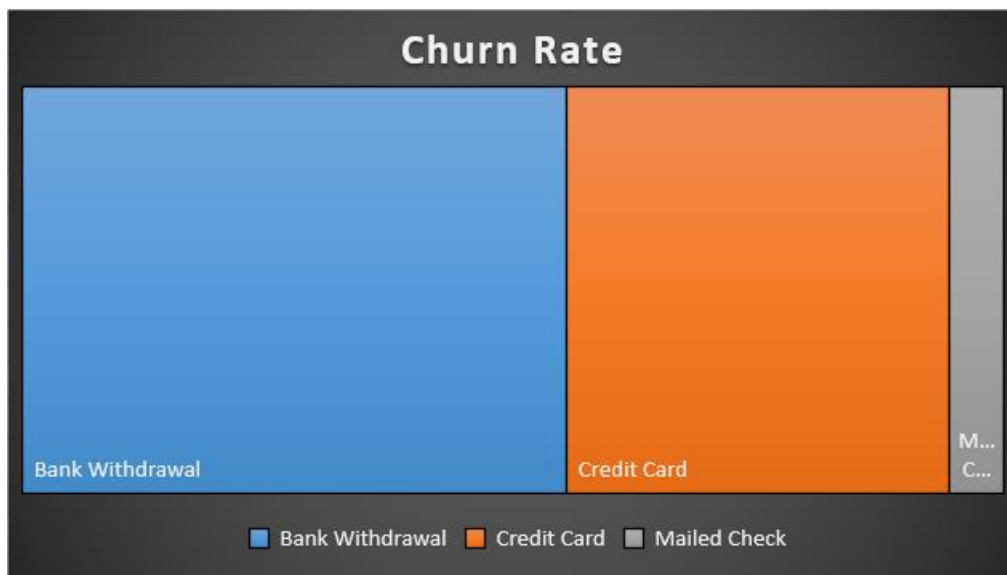
Fiber optic users have the highest churn rate and Fiber optic is also the most frequently used type.

12. Which contract type shows the highest churn?



Monthly contracts have the highest churn rate

13. Which payment method shows highest churn?



Bank withdrawals show highest churn rate

Factors affecting churn score

- Satisfaction Score: It is affecting the customer score, the lower the satisfaction score the higher the churn score.
- Senior level: Customers who are seniors are more likely to churn
- Dependents: Customers who have 0 dependents are more likely to churn
- Contract Type: Customers with Monthly contracts are more likely to churn
- Offers: Customers are more likely to churn with offer E
- Tenure: Customers who just at their 1 year in the company are more likely to churn, after one year has passed the customer will be more likely to stay in the company.

Applying Regression analysis

Analysis for categorical factors:

Regression Statistics								
Multiple R	0.3473026							
R Square	0.1206191							
Adjusted R Square	0.11861652							
Standard Error	19.8748538							
Observations	7043							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	16	380676	23792.2	60.2319889	3.899E-182			
Residual	7026	2775339	395.01					
Total	7042	3156015						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	46.8125083	1.11006	42.171	0	44.6364476	48.9886	44.6364476	48.9885689
Encoded Referred a Friend	0.66922808	1.5237	0.43921	0.66052017	-2.31767506	3.65613	-2.31767506	3.65613122
Encoded Male Gender	0.05296932	0.47414	0.11172	0.91105102	-0.87648532	0.98242	-0.87648532	0.98242397
Encoded Offer A	-1.81484857	1.00735	-1.8016	0.07164984	-3.7895563	0.15986	-3.7895563	0.15985916
Encoded Offer B	-1.18220186	0.78256	-1.5107	0.13091548	-2.71626294	0.35186	-2.71626294	0.35185921
Encoded Offer C	-0.64487913	1.03045	-0.6258	0.53145055	-2.6648674	1.37511	-2.6648674	1.37510914
Encoded Offer D	-2.35436832	0.88128	-2.6715	0.00756829	-4.08195192	-0.6268	-4.08195192	-0.62678471
Encoded Offer E	3.88023314	0.80638	4.81195	1.5259E-06	2.29949459	5.46097	2.29949459	5.46097168
Encoded Phone Service	1.41629369	0.86578	1.63585	0.10191479	-0.28090038	3.11349	-0.28090038	3.11348776
Encoded Multiple Lines	2.64587614	0.54421	4.86183	1.1883E-06	1.57905201	3.7127	1.57905201	3.71270028
Encoded Internet Service	5.84772726	0.74419	7.85789	4.4903E-15	4.38889854	7.30656	4.38889854	7.30655598
Encoded Month-to-Month Contract	10.2547131	0.6862	14.9443	9.7974E-50	8.90956364	11.5999	8.90956364	11.5998626
Encoded One Year Contract	1.57860323	0.70992	2.22364	0.02620438	0.18694843	2.97026	0.18694843	2.97025802
Encoded Paperless Billing	2.19615625	0.51818	4.23821	2.2819E-05	1.18036617	3.21195	1.18036617	3.21194632
Encoded Online Security	-3.79721087	0.59822	-6.3476	2.3237E-10	-4.96989453	-2.6245	-4.96989453	-2.6245272
Encoded Online Backup	-1.45929907	0.57157	-2.5531	0.01069642	-2.57974679	-0.3389	-2.57974679	-0.33885135
Encoded Married	-1.78189198	1.52487	-1.1686	0.24262304	-4.77109571	1.20731	-4.77109571	1.20731176

We can conclude from the previous analysis that the factors that impact the **Churn score** are (offer D - offer E - Multiple lines - Internet Service - Encoded Month-to-Month Contract - Encoded One Year Contract - Encoded Paperless Billing Encoded Online Security - Encoded Online Backup)

➔ And Monthly contract is the variable that impact the Churn score the Most

Analysis for numerical Data:

Regression Statistics								
Multiple R	0.523948107							
R Square	0.274521618							
Adjusted R Square	0.273283248							
Standard Error	18.0469608							
Observations	7043							
ANOVA								
	df	SS	MS	F	Significance F			
Regression	12	866394.2	72199.51896	221.6798168	0			
Residual	7030	2289620	325.6927943					
Total	7042	3156015						
	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	79.93769215	1.568057	50.97881275	0	76.86382737	83.01155694	76.86382737	83.01155694
Age	0.02843309	0.014888	1.90981367	0.056197864	-0.000751657	0.057617837	-0.000751657	0.057617837
Satisfaction Score	-7.646731197	0.193918	-39.43271381	0	-8.026869846	-7.266592547	-8.026869846	-7.266592547
CLTV	0.000170889	0.000198	0.861670256	0.388898388	-0.000217883	0.000559661	-0.000217883	0.000559661
Number of Referrals	-0.55535591	0.076628	-7.247436489	4.70022E-13	-0.705569716	-0.405142104	-0.705569716	-0.405142104
Tenure in Months	-0.026461023	0.024133	-1.096455376	0.2729171	-0.073769454	0.020847408	-0.073769454	0.020847408
Avg Monthly Long Distance Charges	0.011404107	0.023508	0.485110644	0.627612978	-0.03467917	0.057487385	-0.03467917	0.057487385
Avg Monthly GB Download	-0.029454313	0.013199	-2.231603662	0.025672568	-0.055327785	-0.003580842	-0.055327785	-0.003580842
Monthly Charge	0.100733944	0.014267	7.060413033	1.81907E-12	0.072765481	0.128702406	0.072765481	0.128702406
Total Charges	-0.001558023	0.000685	-2.274268411	0.022979855	-0.002900958	-0.000215088	-0.002900958	-0.000215088
Total Refunds	0.002952222	0.027337	0.107993952	0.914003572	-0.050636387	0.056540832	-0.050636387	0.056540832
Total Extra Data Charges	0.005893593	0.008689	0.678313274	0.497595411	-0.011138683	0.022925868	-0.011138683	0.022925868
Total Revenue	0.000188639	0.000577	0.3267461	0.743869657	-0.000943093	0.001320371	-0.000943093	0.001320371

Satisfaction score is the one with the highest impact on the churn score, as mentioned whenever the satisfaction score is increased by 1 this will decrease the churn score by 7.64.