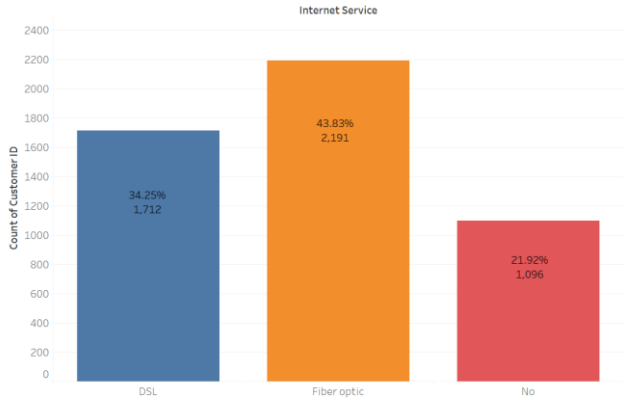
Digital Services Case Study

Distribution of Customers based on the **type of internet service**

Distiribution

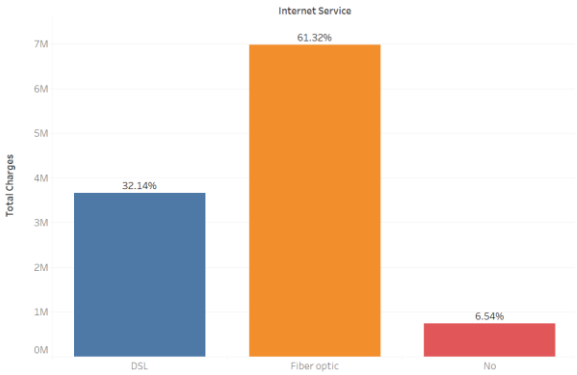


There are about **22%** of customers who do not avail Internet service.

Net Spend

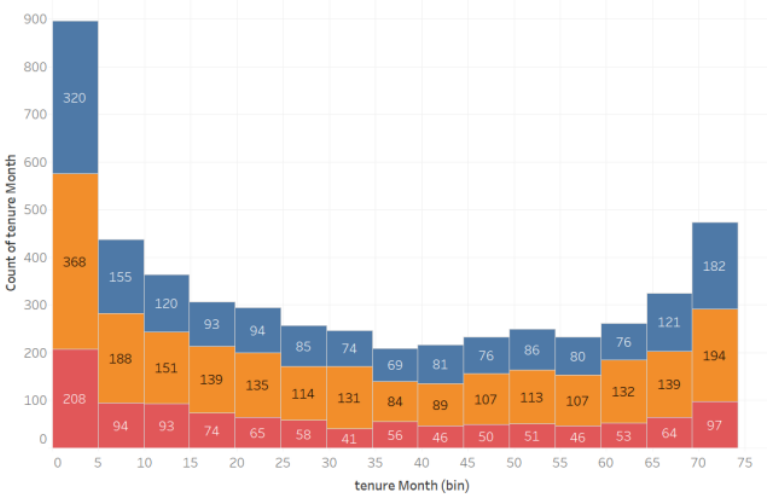
Internet ..	Avg. Total Charges	Avg. Monthly Charges
Fiber optic	3,183	92
DSL	2,136	59
No	679	21

Contritbution to Revenue



They contribute to **6.54%** of the total revenues .

Impact of the type of internet Services



The tenure of services availed is spread evenly across the three types of internet services, indicating that the duration of contract with the **company does not depend** on the type of service.

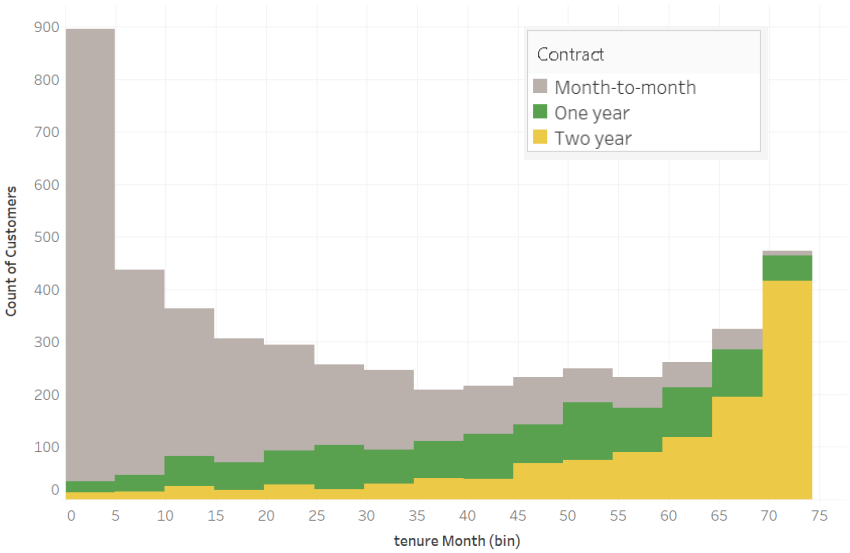
Churn rate for Internet Service Types

Churn	Internet Service DSL	Fiber optic	No
No	81.25% 1,391	58.19% 1,275	93.07% 1,020
Yes	18.75% 321	41.81% 916	6.93% 76

Indicates the Churn rate for **Fiber optic to be the highest.**

And the Churn rate for People who **do not avail internet services to be relatively negligible.**

Distribution of Customers based on the **type of Contract**



As observed in the graph.

Customers who tend to stay with the company for a longer time are usually on a one year or two year contract.

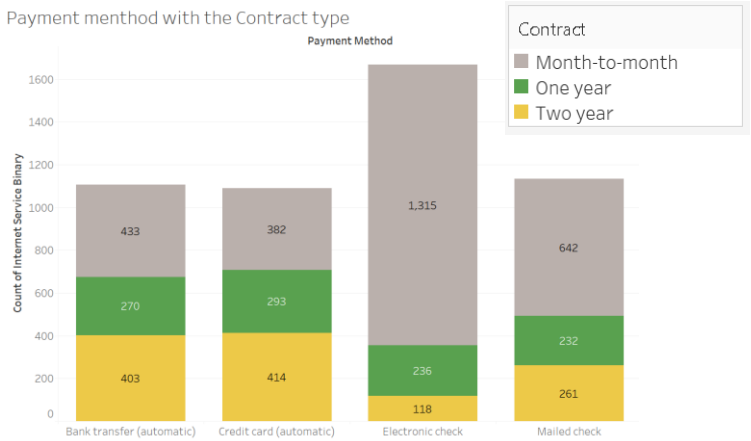
Churn Rate by Contract Type

Churn	Contract: Month-to-month	One year	Two year
No	57.72%	89.04%	97.66%
	1,600	918	1,168
Yes	42.28%	10.96%	2.34%
	1,172	113	28

Indicates the Churn rate for **Contract on Month basis to be higher**

Churn rate for customers who avail a long term contract is lesser.

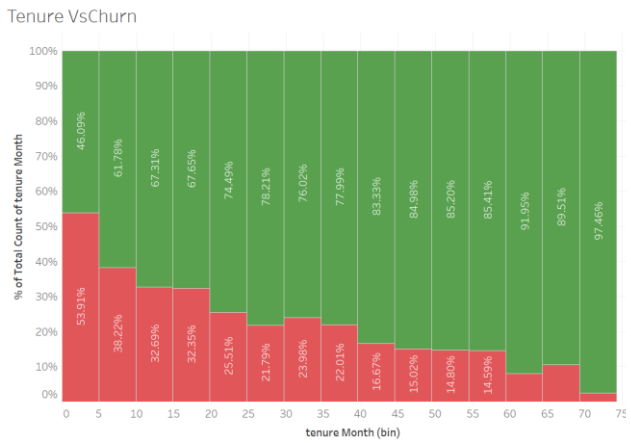
Distribution of Customers based on the **type of Payment Method**

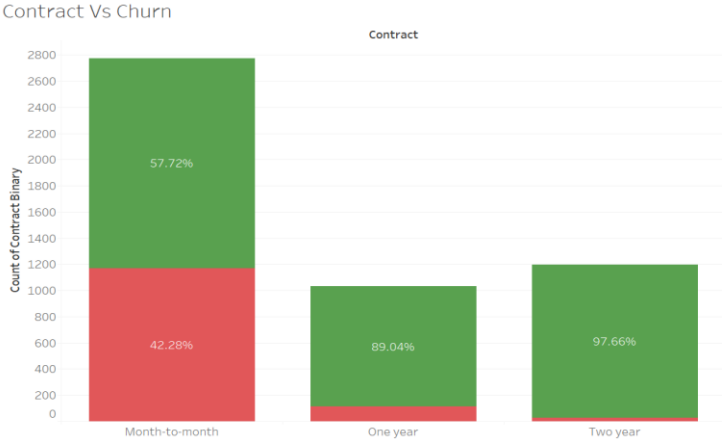


Shows the distribution of type of payments.

Here most of the electronic payments are done by monthly customers and this also impacts t the churn rate.

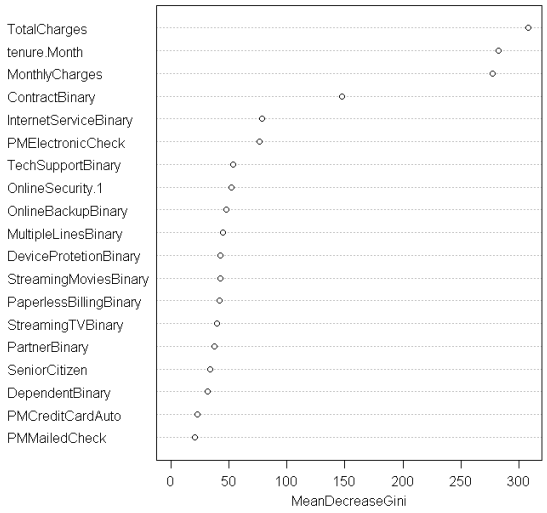
The Key Drivers for Churn Rate here are





Further Analysis using **Statistical techniques** indicate a similar story

Using Random Forest and checking the importance of variables and its impact on the churn rate, below ranks the effect of variables on the churn.




```
Call:
 randomForest(formula = Churn ~ . - GenderBinary - PhoneServiceBinary,      data = DSRf, importance = T
E, proximity = TRUE)
               Type of random forest: classification
                     Number of trees: 500
No. of variables tried at each split: 4

        OOB estimate of  error rate: 20.16%
Confusion matrix:
      No Yes class.error
No  3283 403   0.1093326
Yes  605 708   0.4607768
```

Churn rate prediction using Random forest is 80%.

Using Logistic Regression.: Importance of variables

```
Call:
glm(formula = Training$ChurnBinary ~ . - PMCreditCardAuto - PMMailedCheck -
    DeviceProtetionBinary - OnlineBackupBinary - PhoneServiceBinary -
    DependentBinary - PartnerBinary - SeniorCitizen - GenderBinary,
    family = binomial, data = Training)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.8039  -0.6770  -0.2928   0.7454   3.3616

Coefficients:
                         Estimate Std. Error z value Pr(>|z|)
(Intercept)            -1.024e+00  1.884e-01  -5.437 5.42e-08 ***
tenure.Month           -5.684e-02  8.300e-03  -6.848 7.48e-12 ***
MultipleLinesBinary     3.250e-01  1.185e-01   2.744 0.006075 **
InternetServiceBinary   1.635e+00  1.902e-01   8.598  < 2e-16 ***
OnlineSecurity.1       -3.227e-01  1.069e-01  -3.019 0.002540 **
TechSupportBinary      -4.825e-01  1.130e-01  -4.270 1.95e-05 ***
StreamingTVBinary       4.406e-01  1.256e-01   3.508 0.000452 ***
StreamingMoviesBinary   5.779e-01  1.243e-01   4.649 3.34e-06 ***
ContractBinary         -6.313e-01  1.093e-01  -5.775 7.67e-09 ***
PaperlessBillingBinary  4.457e-01  1.051e-01   4.242 2.22e-05 ***
PMElectronicCheck       3.966e-01  9.742e-02   4.071 4.68e-05 ***
MonthlyCharges         -3.700e-02  7.841e-03  -4.718 2.38e-06 ***
TotalCharges            3.498e-04  9.484e-05   3.688 0.000226 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 4047.3  on 3507  degrees of freedom
Residual deviance: 2927.7  on 3495  degrees of freedom
AIC: 2953.7

Number of Fisher Scoring iterations: 6
```

Churn rate prediction using Logistic regression is 81.1%

```
> table(ChurnValuesPredict,Testing$ChurnBinary)

ChurnValuesPredict   0   1
                 0 996 174
                 1 107 214
```