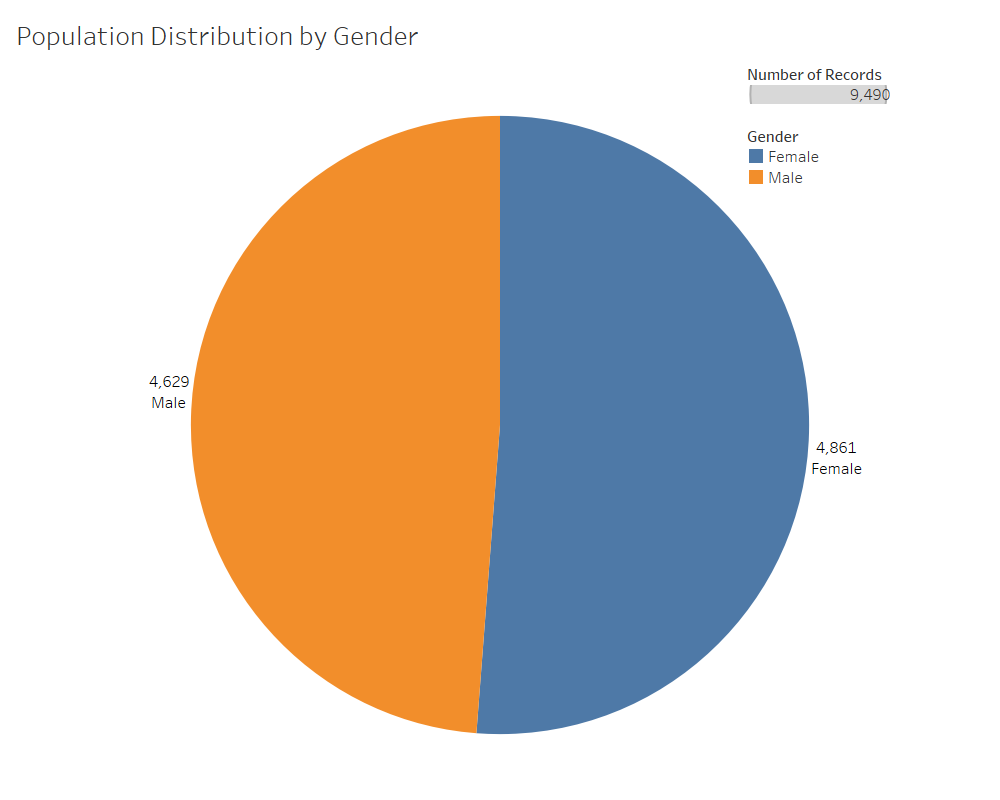
# Descriptive Analysis

## Column Definition

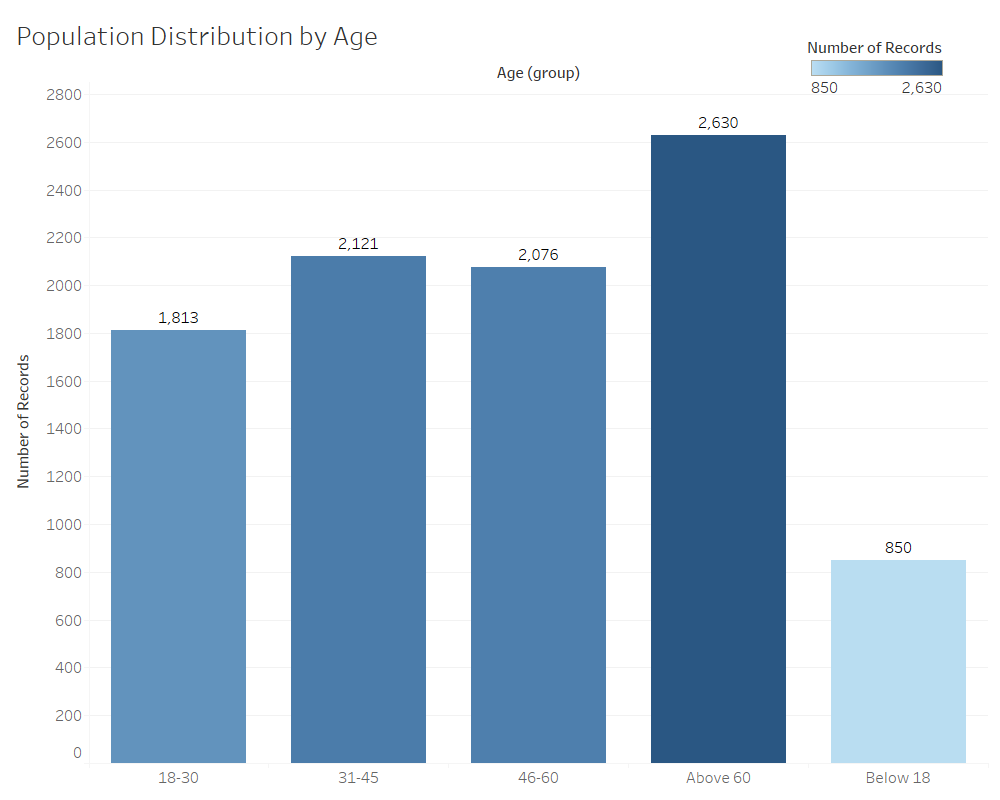
|  |  |  |
| --- | --- | --- |
| **Column Name** | **Description** | **Type** |
| CustomerId | Unique customer identification | Numeric |
| Age | Customer age | Numeric |
| Gender | Customer gender | String |
| isCustomerSuspended | Whether customer suspended earlier | String |
| CallDropRate | Call drop rate - customer experience data | Numeric |
| NumberOfComplaints | Customer complaints | Numeric |
| MonthlyBilledAmount - Rs | Avg. monthly bill | Numeric |
| UnpaidBalance - Rs | Total unpaid balance | Numeric |
| NumberOfMonthUnpaid | Total unpaid months | Numeric |
| TotalMinsUsedInLastMonth | Last month voice usage in minutes | Numeric |
| TotalCallDuration | Total voice usage in minutes | Numeric |
| AvgCallDuration | Average voice usage in minutes | Numeric |
| PercentageCallOutsideNetwork | Offnet voice usage % from total usage | Numeric |
| isChurned | Is customer disconnected? | String |

The selected dataset contains the telecom customer churn data with 9490 total records. It contains 14 columns with telecom user’s demographic data, behavioral data and transactional data including customer ID and customer churn status. A descriptive analysis was done for explore the selected telco churn dataset using telecom user’s demographic data variables such as gender and age. Further the dependent variable of dataset, churn status was analyzed to identify its distribution.

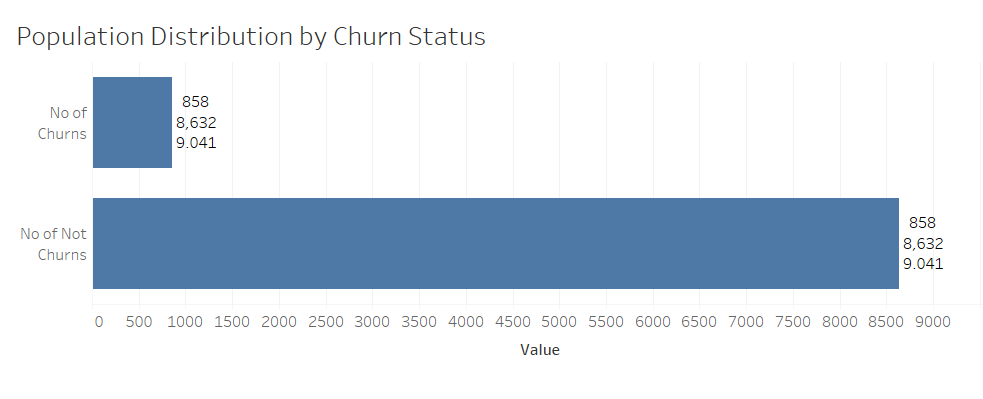
**Gender analysis:** The gender of majority users in selected telco dataset are females. However, the following figure shows that gender wise distribution of population is similar to each other.



**Customer Age analysis:** Most of the users in selected dataset are above 60 years old and there are few users less than 18 years of age. However, the following figure shows that age wise distribution of population is similar to each other.



**Churn distribution:** According to the following figure, there are 858 churned users and 8,632 active (not churned) users from the total of 9,490 users in selected telco dataset. The percentage of churn is 9.041%.



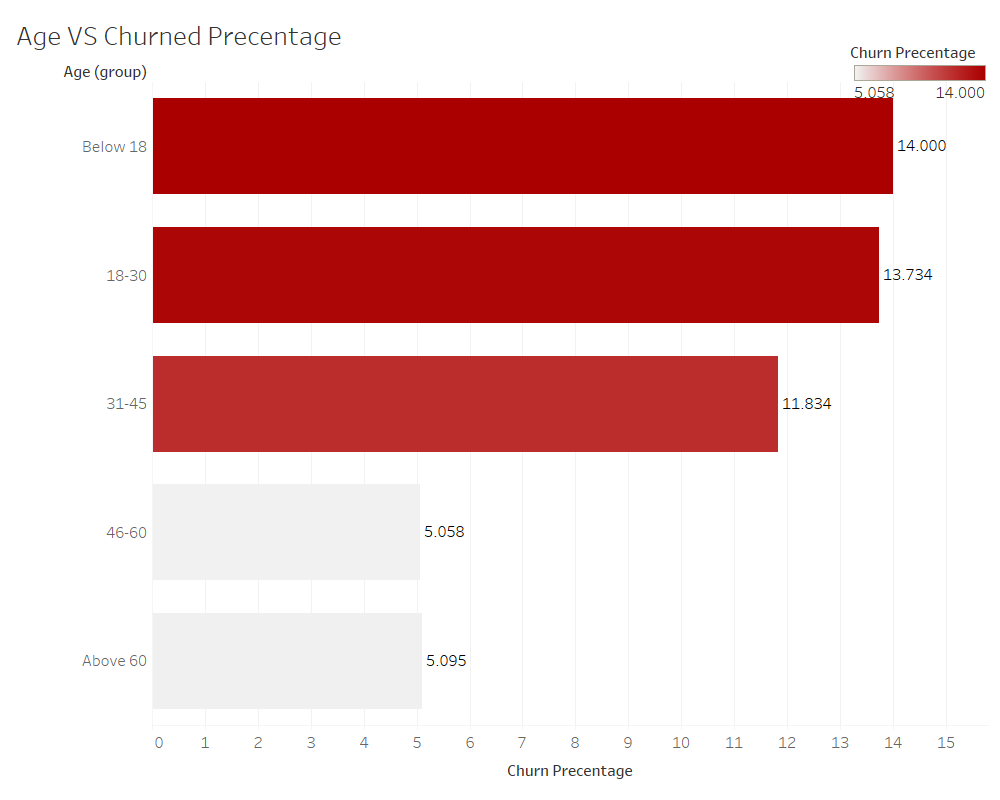
# Diagnostic Analysis

A diagnostic analysis was done for identify the relationship between customer churn and independent variables such as telco user’s demographic data, behavioral data and transactional data. These variables contain customer age, gender, customer was suspended earlier, call drop rate, number of complaints, monthly bill amount, number of unpaid months, unpaid bill value, average call duration, off-net call proportion and last month total usage. The analysis results are summarized below.

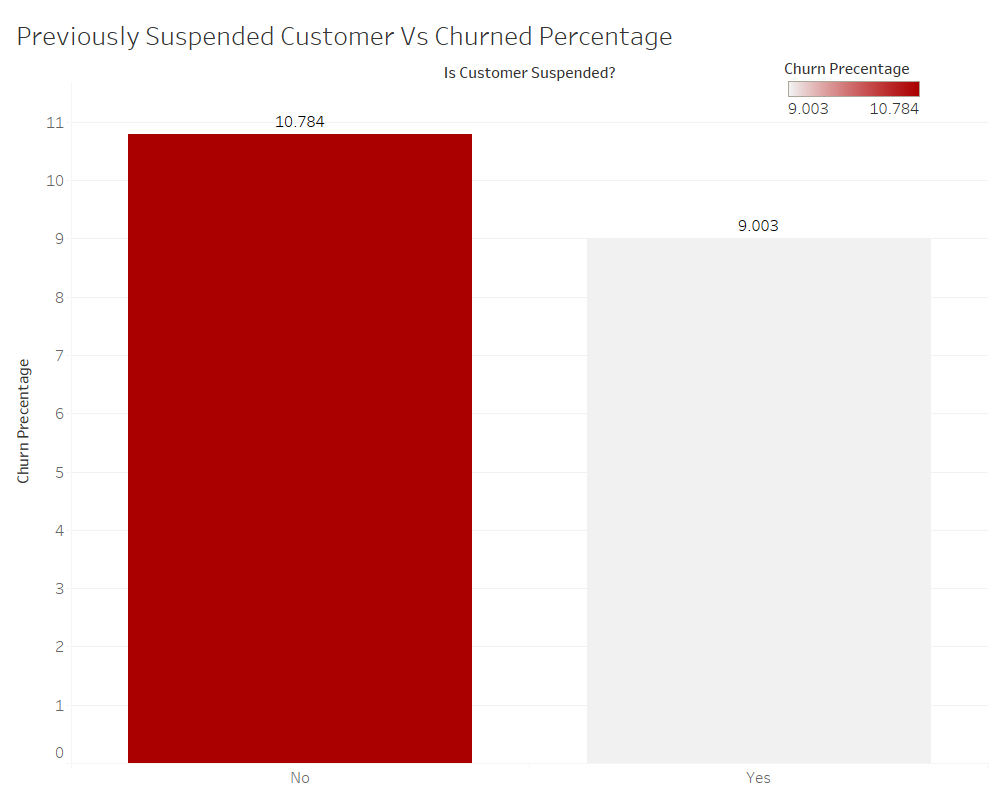
**Relationship between gender and customer churn:** According to the following figure, 8.7% has churned from total male population of telco users whereas 9.2% has churned from total female users. Therefore, churn percentage of females is 0.5% higher than churn percentage of males in the selected telco dataset.



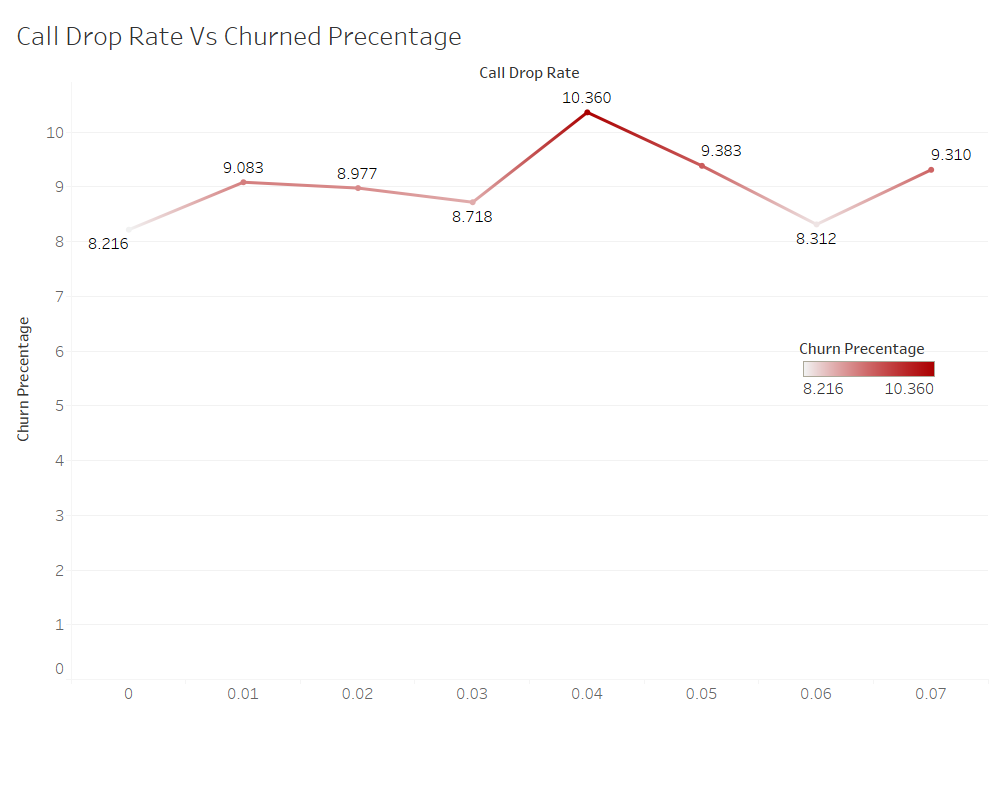
**Relationship between age and customer churn:** According to the following figure, 14% has churned from users age below 18, 13.7% has churned from users age between 18 - 30, 11.8% has churned from users age between 31 - 45, 5% has churned from users age between 46 - 60, and 5% has churned from users age above 60. Therefore, highest churn percentages are recorded from age groups below 45 in the selected telco dataset.



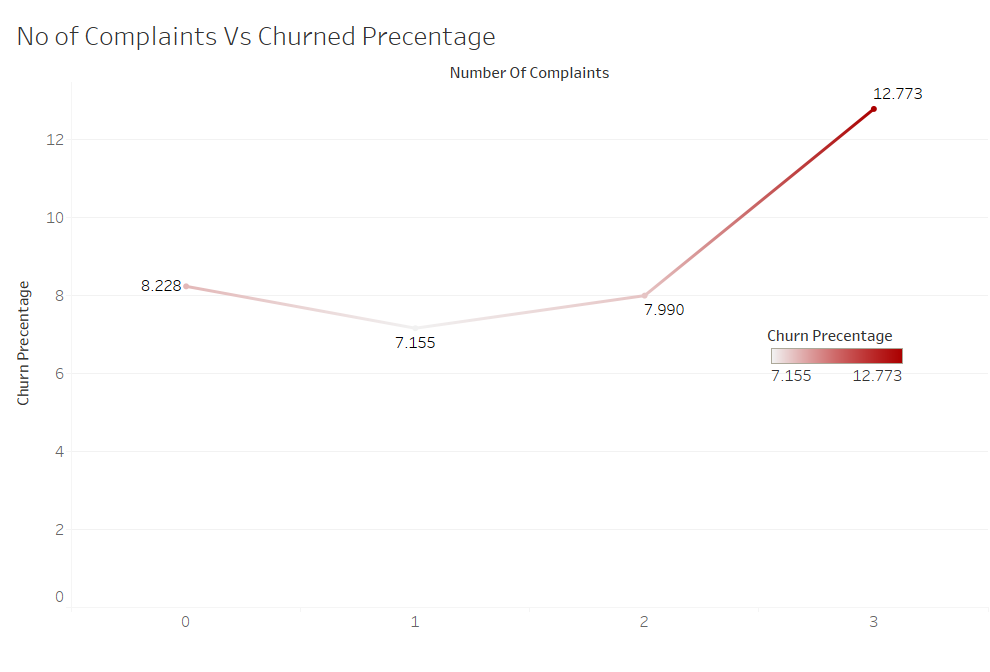
**Relationship between previously suspended customer and churn:** According to the following figure, 10.7% has churned from total users who have not suspended previously whereas 9.0% has churned from total users who have suspended at least one time previously. Therefore, churn percentage of never suspended users is 1.7% higher than at least one time suspended users in the selected telco dataset.



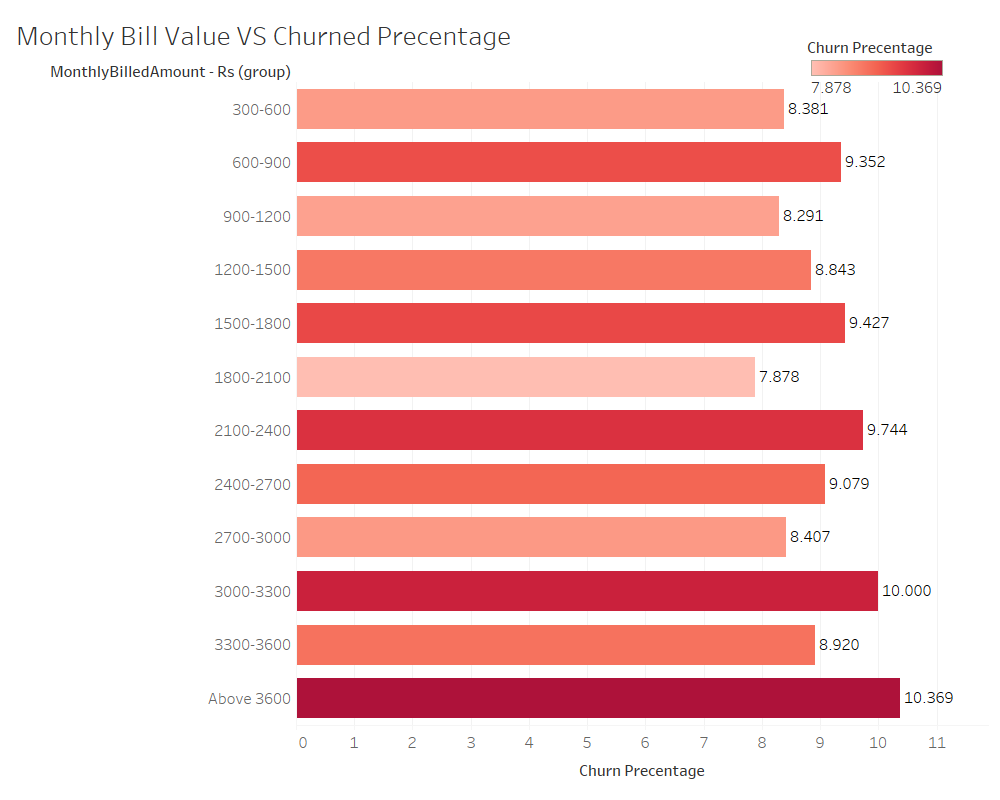
**Relationship between call drop rate and customer churn:** According to the following figure, customer churn percentage has increased with the call drop rate in the selected telco dataset.



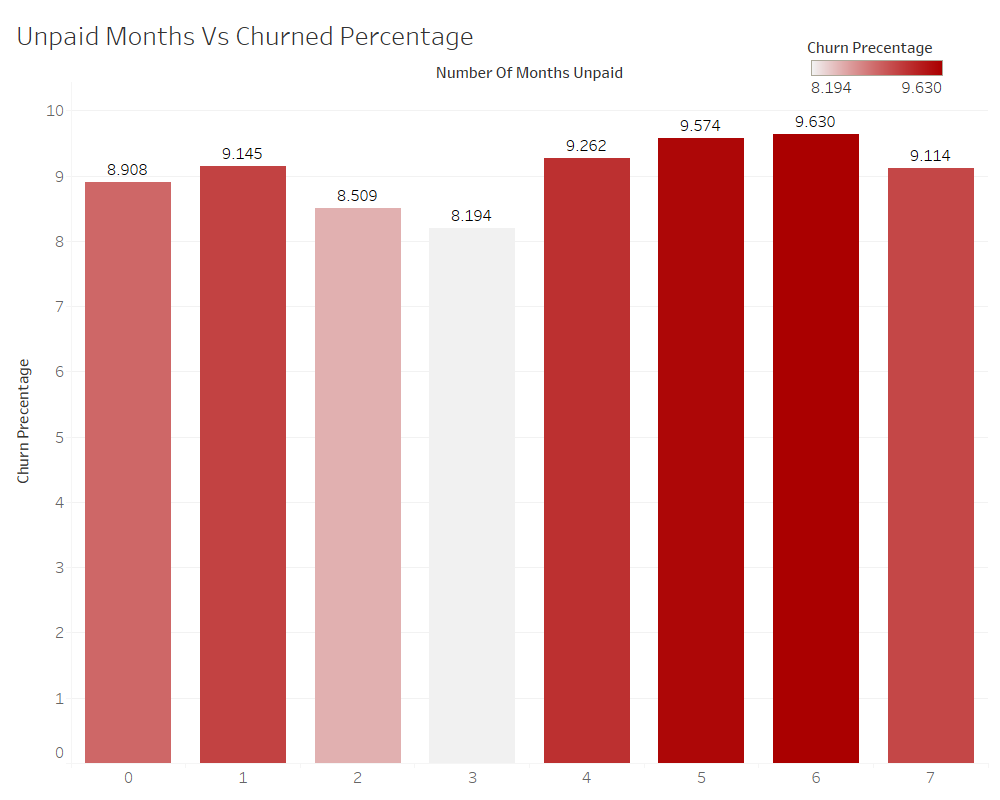
**Relationship between customer complaints and customer churn:** According to the following figure, customer churn percentage has increased with the no of customer complaints in the selected telco dataset.



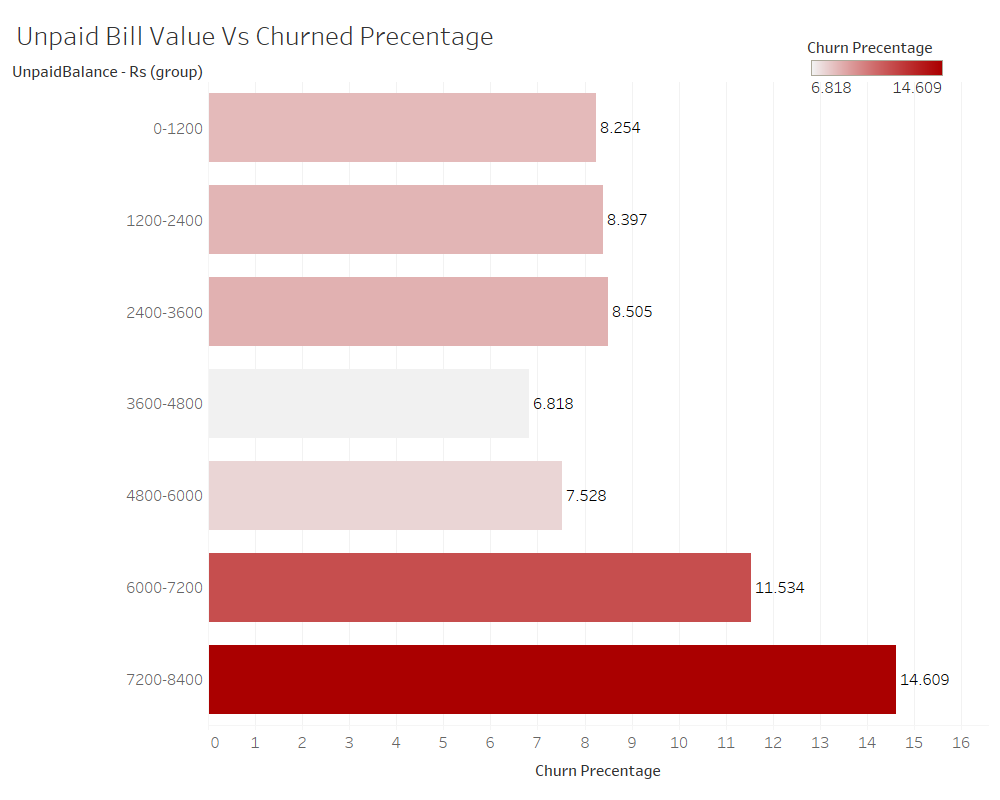
**Relationship between monthly bill amount and customer churn:** According to the following figure, customer churn percentage has increased with the monthly bill amount in the selected telco dataset.



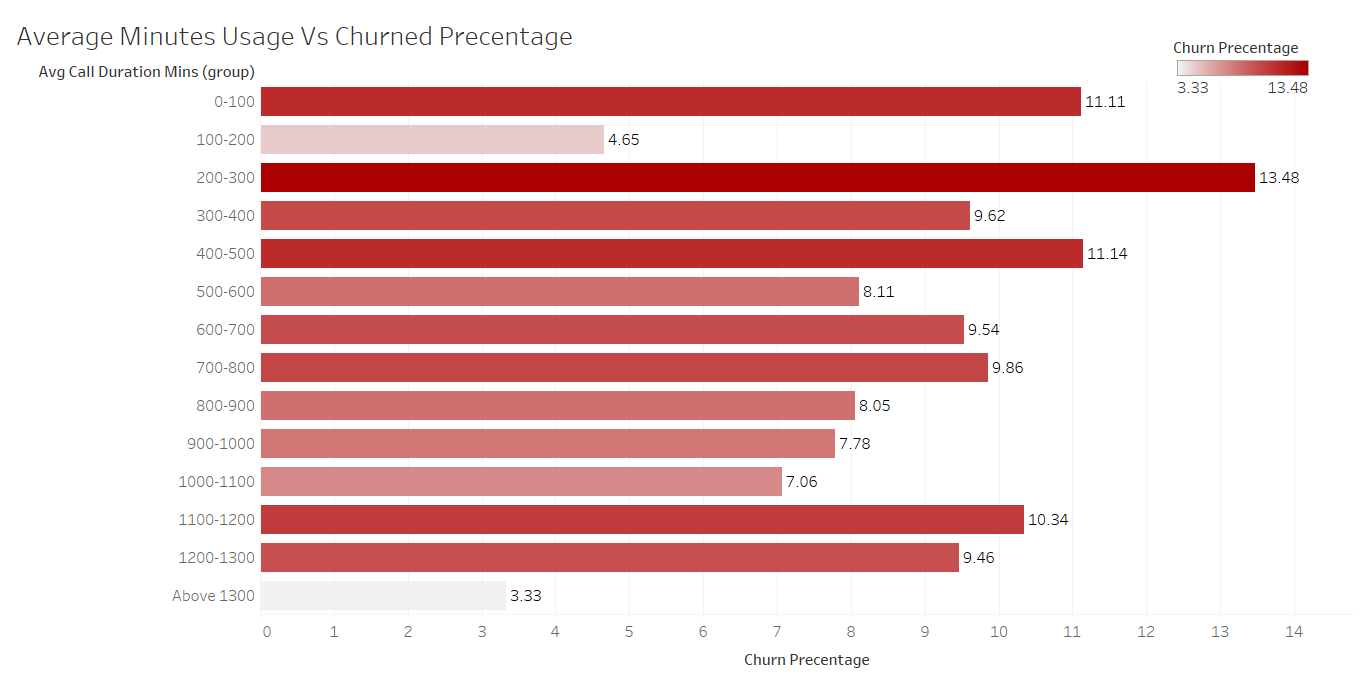
**Relationship between no of unpaid months and customer churn:** According to the following figure, customer churn percentage has slightly increased with the no of unpaid months in the selected telco dataset.



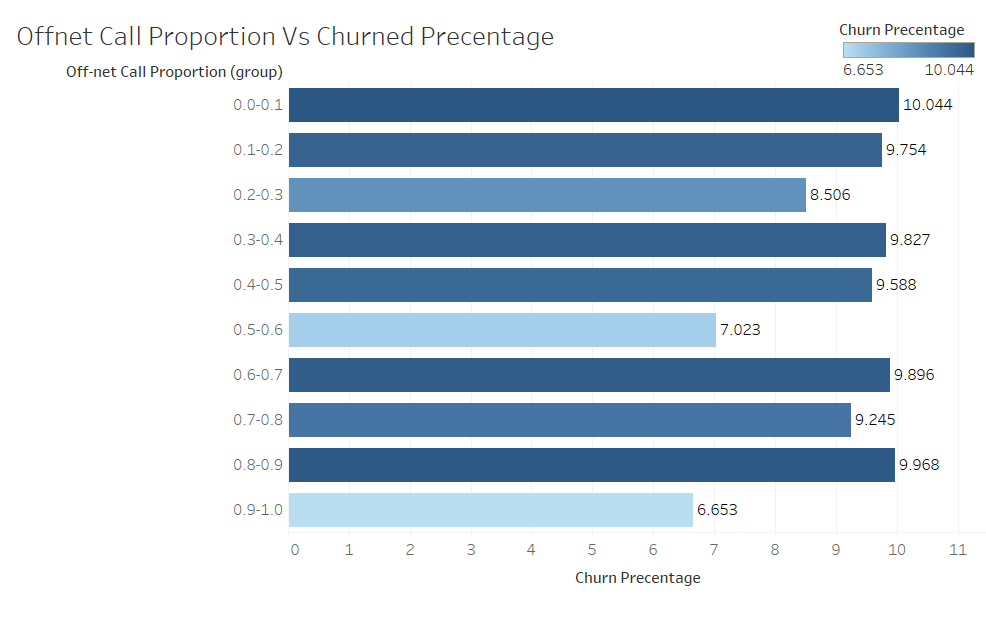
**Relationship between no of unpaid bill value and customer churn:** According to the following figure, customer churn percentage has increased with the unpaid bill value in the selected telco dataset.



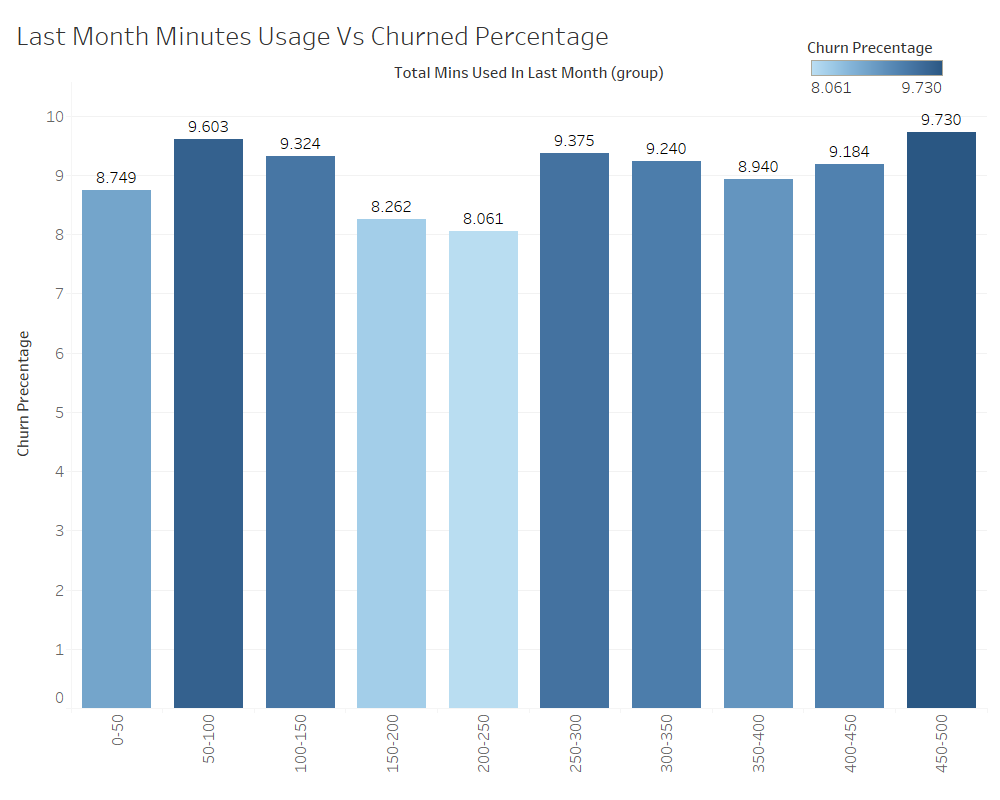
**Relationship between average call duration - minutes and customer churn:** According to the following figure, customer churn percentage has decreased with the average call duration in the selected telco dataset.



**Relationship between off-net call proportion and customer churn:** According to the following figure, customer churn percentage has not changed with the proportion of calls to other networks in the selected telco dataset. Therefore, there is no significant relationship between off-net calls and churn in selected dataset.



**Relationship between last month total usage – minutes and customer churn:** According to the following figure, customer churn percentage has not changed with the last month usage in the selected telco dataset. Therefore, there is no significant relationship between last month usage and churn in selected dataset.



Summary: According to the diagnostic analysis, it can be concluded that independent variables such as age, gender, customer was suspended earlier, call drop rate, number of complaints, monthly bill amount, number of unpaid months, unpaid bill value and average call duration have significant relationship to telco customer churn. However, there is a weak relationship between off-net call proportion and last month total usage with customer churn. Therefore, only the variables with significant relationship with churn are considered for further analysis.