**Section 1 (SQL)**

[Answer 1](http://sqlfiddle.com/?fbclid=IwAR1oUpQpWVqCkIj2BU6UoazP9i9Dj7CphxH-nXFXXfbMbXzla_2e6rxfy6U#!9/d32fd7/163)

Answer SQL (Q1, Q2)

[Answer 2](http://sqlfiddle.com/?fbclid=IwAR2wnT4uK8b75nmJ9Z_PgYAiVtqo496yD2mP4kYGeKhsNGTQN6-ywcWQFeg#!9/b60a85/360)

Answer SQL (Q1, Q2)

**Section 2 (Data Analysis)**

This is the customer level data from one telecom company. Each row represents a customer, each column contains customer’s attributes described on the column Metadata.

The raw data contains 7043 rows (customers) and 21 columns (features).

**Dataset Schema**

**customerID**

Customer ID

**gender**

Customer gender (female, male)

**SeniorCitizen**

Whether the customer is a senior citizen or not (1, 0)

**Partner**

Whether the customer has a partner or not (Yes, No)

**Dependents**

Whether the customer has dependents or not (Yes, No)

**tenure**

Number of months the customer has stayed with the company

**PhoneService**

Whether the customer has a phone service or not (Yes, No)

**MultipleLines**

Whether the customer has multiple lines or not (Yes, No, No phone service)

**InternetService**

Customer’s internet service provider (DSL, Fiber optic, No)

**OnlineSecurity**

Whether the customer has online security or not (Yes, No, No internet service)

**OnlineBackup**

Whether the customer has online backup or not (Yes, No, No internet service)

**DeviceProtection**

Whether the customer has device protection or not (Yes, No, No internet service)

**TechSupport**

Whether the customer has tech support or not (Yes, No, No internet service)

**StreamingTV**

Whether the customer has streaming TV or not (Yes, No, No internet service)

**StreamingMovies**

Whether the customer has streaming movies or not (Yes, No, No internet service)

**Contract**

The contract term of the customer (Month-to-month, One year, Two year)

**PaperlessBilling**

Whether the customer has paperless billing or not (Yes, No)

**PaymentMethod**

The customer’s payment method (Electronic check, Mailed check, Bank transfer (automatic), Credit card (automatic))

**MonthlyCharges**

The amount charged to the customer monthly in USD

**TotalCharges**

The total amount charged to the customer in USD

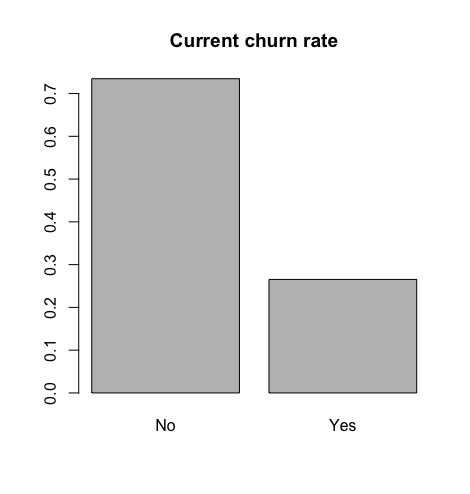
**Churn**

Whether the customer churned or not (Yes or No)

**Question**

1. What is current churn rate of this Telco company?

Ans = 26.54%

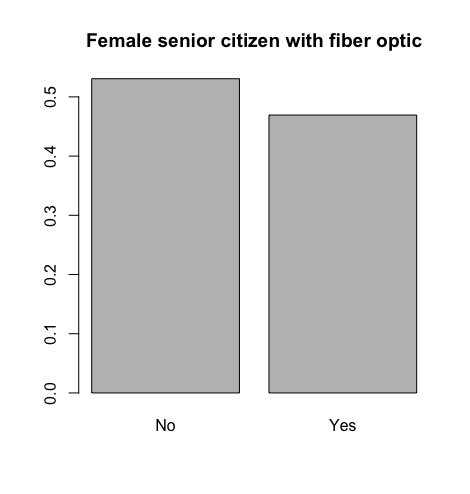
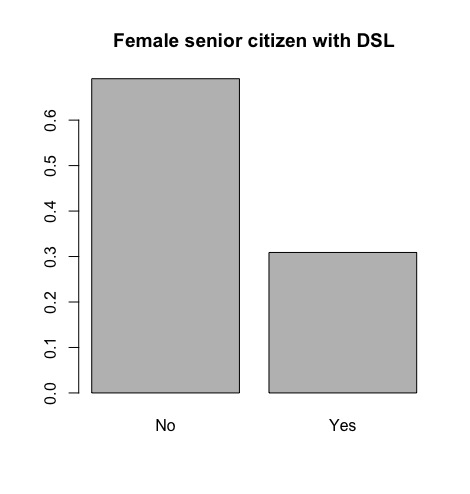
****

1. What is median of tenure of the customers?

Ans = Median is 29.

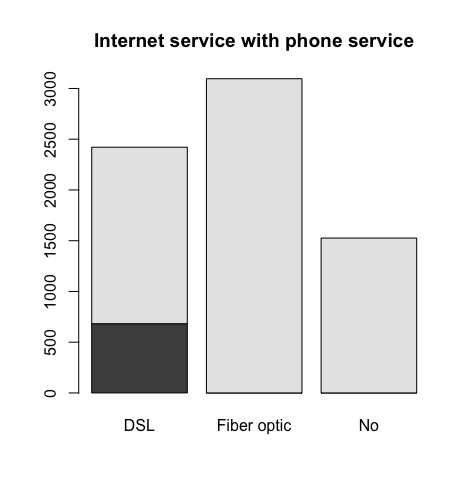
1. What is the difference in churn rate of female senior citizens who use Fibre Optic and DSL?

Ans = Different is 16.04% (fiber optic more churn rate than DSL.)



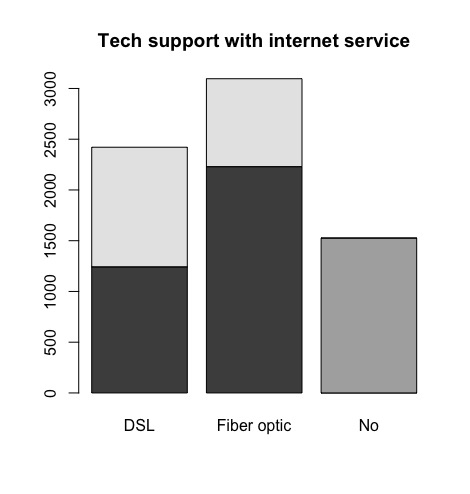
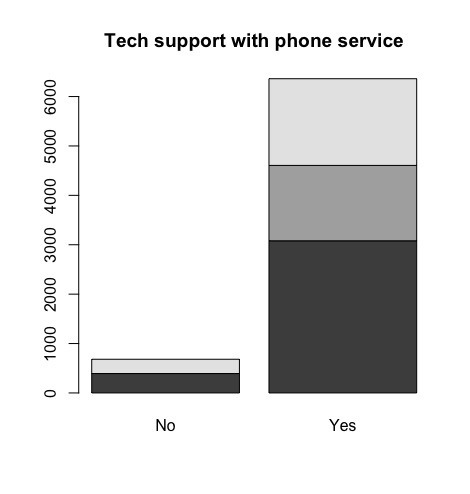
1. Can Fiber optic customers opt out for phone service?

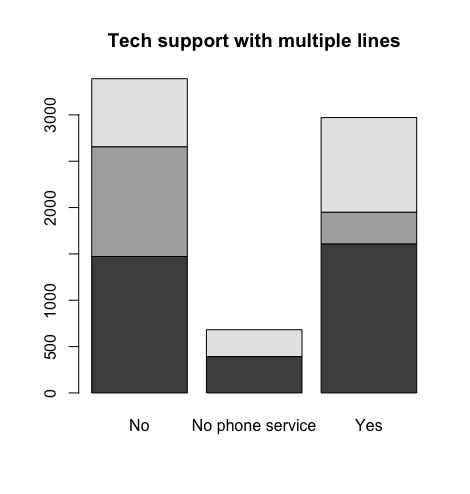
Ans = No because fiber optic still available with phone service.



1. Which services do not have tech support?

Ans = No internet services are no tech support.(At picture 3 in No)



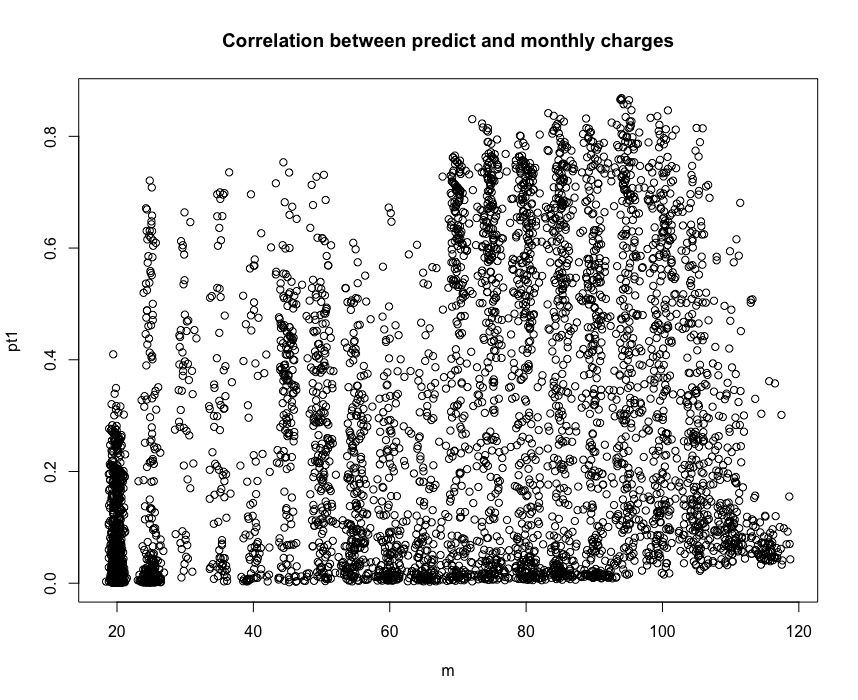


6) Is this statement true? And why do you think so?

*“The churn rate increases as the customers pay higher monthly charges”*

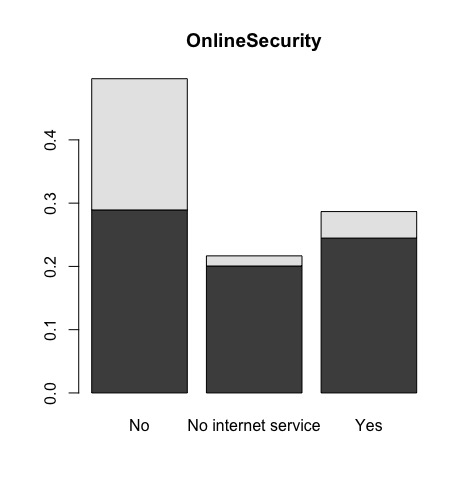
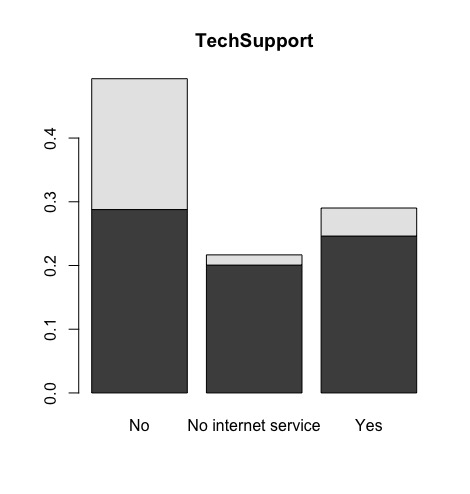
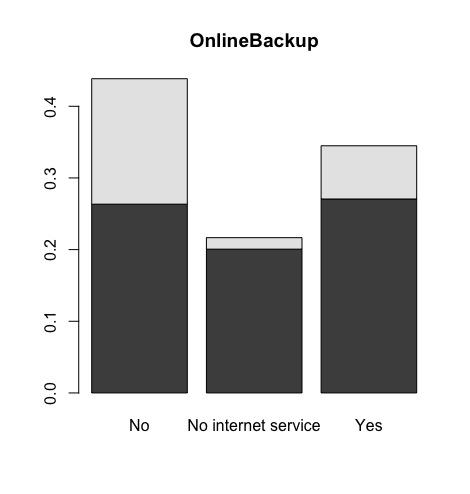
Ans = It not true because correlation between prediction and monthly charge around 0.36

It quite low correlation.

**

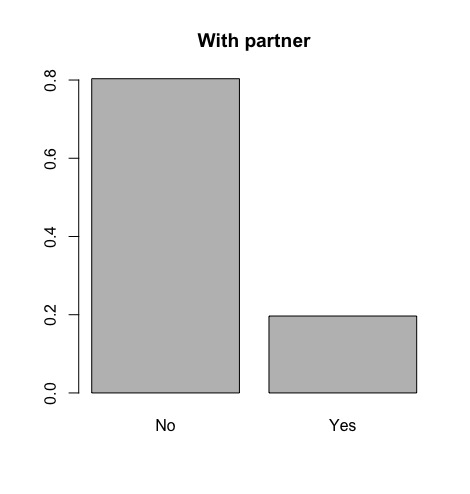
7) Do you think additional services such as Online backup, Online Security, and tech support are important from customers’ perspective? Why do you think so?

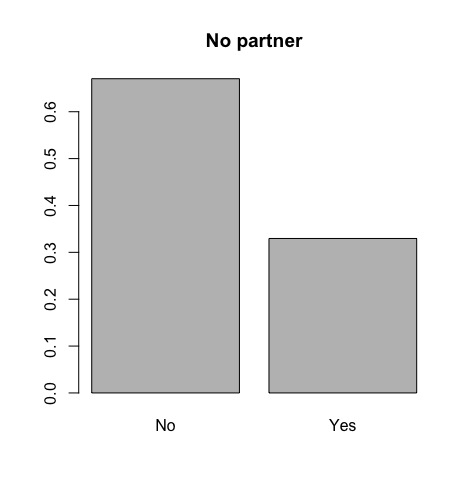
Ans = From data we can see churn rate of no additional services more than churn rate of have additional services.



8) Would you recommend the Telco company to target customers who are singles? Why do you think so?

Ans = No because no partner are more churn rate proportion than have partner





9) Which service you think has serious problems and would recommend the executive of this Telco company to look into?

Ans = From logistic regression model multiple service are significant with churn rate.

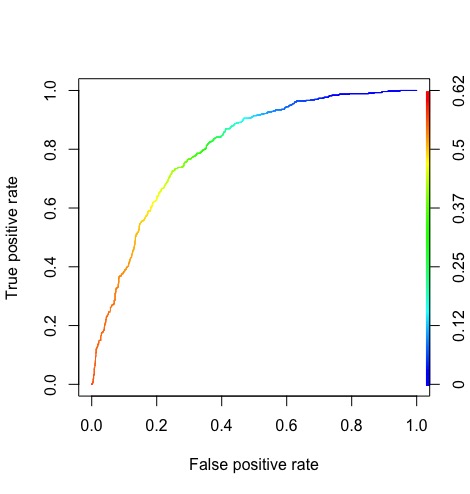
10) What features/attributes you think are the most predictive to churn behaviour? List the top 3 features with back up analysis.

Ans = From logistc regression model most features predictive is tenure,Contract,TotalChanges

**Bonus Section (Modeling)**

Build a churn prediction model (any technique is fine) and show the performance of the model in forms of confusion metrics / ROC / Precision vs Recall curve.

Ans = AUC is 0.8

****