```
In [2]: # QUESTION
           #Problem Statement:
           #You are the Data Scientist at a telecom company "Neo" whose customers are churning out to
           #its competitors. You have to analyse the data of your company and find insights and stop your
           #customers from churning out to other telecom companies.
           # Tasks to be done:
           A) Data Manipulation:
           a. Extract the 5th column & store it in 'customer 5'
           b. Extract the 15th column & store it in 'customer_15'
           c. Extract all the male senior citizens whose Payment Method is Electronic check &
           store the result in 'senior male electronic'
           d. Extract all those customers whose tenure is greater than 70 months or their
           Monthly charges is more than 100$ & store the result in 'customer_total_tenure'
           e. Extract all the customers whose Contract is of two years, payment method is Mailed
           check & the value of Churn is 'Yes' & store the result in 'two_mail_yes'
           f. Extract 333 random records from the customer_churndataframe& store the result in
           'customer 333'
           q. Get the count of different levels from the 'Churn' column
            File "<ipython-input-2-f41482e7624f>", line 8
              A) Data Manipulation:
          SyntaxError: unmatched ')'
 In [3]:
           # ANSWER BELLOW
In [11]:
           import pandas as pd
           import numpy as np
           import seaborn as sns
           import matplotlib.pyplot as plt
           import warnings
           warnings.filterwarnings('ignore')
           sns.set(style="darkgrid")
           %matplotlib inline
In [12]:
           customer = pd.read csv("customer churn.csv")
In [13]:
           customer
               customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DevicePi
                    7590-
                                                                                        No phone
             0
                          Female
                                           0
                                                 Yes
                                                                     1
                                                                                No
                                                                                                         DSL
                                                                                                                        No ...
                                                             No
                   VHVEG
                                                                                          service
                    5575-
                            Male
                                           0
                                                  No
                                                             No
                                                                    34
                                                                                Yes
                                                                                             No
                                                                                                          DSI
                                                                                                                       Yes ...
                   GNVDE
                    3668-
             2
                            Male
                                           0
                                                 No
                                                             No
                                                                     2
                                                                                Yes
                                                                                             No
                                                                                                         DSL
                                                                                                                       Yes ...
                   QPYBK
                    7795-
                                                                                        No phone
                            Male
                                           0
                                                  No
                                                             No
                                                                    45
                                                                                No
                                                                                                          DSL
                                                                                                                        Yes ...
                  CFOCW
                                                                                          service
                    9237-
             4
                                           0
                                                                     2
                          Female
                                                 No
                                                             No
                                                                                Yes
                                                                                             No
                                                                                                     Fiber optic
                                                                                                                        No ...
                   HQITU
                    6840-
          7038
                            Male
                                           0
                                                 Yes
                                                            Yes
                                                                    24
                                                                                Yes
                                                                                            Yes
                                                                                                         DSL
                                                                                                                       Yes ...
                   RESVB
                    2234-
          7039
                          Female
                                                 Yes
                                                            Yes
                                                                    72
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
                                                                                                                        No
                   XADUH
                    4801-
                                                                                        No phone
          7040
                                           0
                                                 Yes
                                                                                                         DSL
                                                                                                                       Yes ...
                          Female
                                                            Yes
                                                                    11
                                                                                No
                    JZAZL
                                                                                          service
                    8361-
          7041
                                                 Yes
                                                                     4
                            Male
                                                             No
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
                                                                                                                        No ...
                   LTMKD
          7042 3186-AJIFK
                                           Λ
                            Male
                                                 Nο
                                                             Nο
                                                                    66
                                                                                Yes
                                                                                             Nο
                                                                                                     Fiber optic
                                                                                                                       Yes ...
```

customer.head(5)

7043 rows × 21 columns

In [7]:

	0	7590- VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	
	1	5575- GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	
	2	3668- QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	
	3	7795- CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	
	4	9237- HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	
		× 21 colu	mns									
	4											<b>)</b>
In [14]:	<pre>customer.iloc[:,5]</pre>											
Out[14]:	1 2 3 4 7038 7039 7040 7041 7042	1 34 2 45 2  24 72 11 4 66 : tenure,	, Length:	7043, dtyp	e: int6	4						
In [15]:	cust	omer.ilo	oc[:,15]									
Out[15]:	0 1 2 3 4	Month	n-to-mont One yea n-to-mont One yea n-to-mont	r h r								
	7038		One yea	r								
	7039 7040 7041 7042	Month	One yea n-to-mont n-to-mont Two yea	r h h	ype: ob	ject						
In [18]:	7039 7040 7041 7042 Name:	Month	One yea n-to-mont n-to-mont Two yea ct, Lengt	r h h r h: 7043, dt			e")&(cu	ıstomer[" <mark>Sen</mark> :	iorCitizen"]	==1)&(custome	er["PaymentMe	thod"]=="E
In [18]:	7039 7040 7041 7042 Name:	Month	One yea n-to-mont n-to-mont Two yea ct, Lengt	r h h r h: 7043, dt			<u>•</u> ")&(cu	ustomer[" <mark>Sen</mark> :	iorCitizen"]	==1)&(custome	er["PaymentMe	thod"]=="E
	7039 7040 7041 7042 Name:	Month Contrac Orcitize Orcitize customer	One yean-to-mont n-to-mont Two yean ct, Lengt en=customen D gender	r h h r h: 7043, dt	r["gend	er"]=="Male			MultipleLines	==1)&(custome		
In [19]:	7039 7040 7041 7042 Name:	Month Contract Orcitize Orcitize Customerl 8778 QRDM	One yean-to-mont Two yean-to-mont Two yean-to, Lengt  en=custom  D gender  Male	r h h r h: 7043, dt	r["gend	er"]=="Male						DeviceP
In [19]:	7039 7040 7041 7042 Name:	Month Contract Orcitize Orcitize Customerl 8779 QRDM 1658 BYGO	One yea n-to-mont n-to-mont Two yea ct, Lengt en=custome  D gender V Male Male Male	r h h r h: 7043, dt er[(custome	r [ "gend Partner No	er"]=="Male	tenure	PhoneService	MultipleLines No phone	InternetService	OnlineSecurity	DeviceP
In [19]:	7039 7040 7041 7042 Name: seni	Month Contract Corcitize Corcitize CustomerI 8779 QRDM 1658 BYGO 5067 XJQF	One yean one yean one year one year one year one year one	r h h r h: 7043, dt er[(custome  SeniorCitizen	r [ "gend Partner No	er"]=="Male Dependents No	tenure	PhoneService No	MultipleLines  No phone service	InternetService DSL	OnlineSecurity No	DeviceP
In [19]:	7039 7040 7041 7042 Name: seni 20 55 57 78	Month Contract Corcitize Corcitize Customerl 8778 QRDM 1658 BYGO 5067 XJQF 0199 ZHSK	One yean one yean one yean one yean one year one year one custom one	r h h r h: 7043, dt er[(custome  SeniorCitizen  1 1 1	Partner No No Yes	Dependents  No  No  Yes	tenure  1 18 66 30	PhoneService  No Yes Yes Yes	MultipleLines  No phone service  Yes  Yes  No	InternetService  DSL  Fiber optic  Fiber optic  DSL	OnlineSecurity  No  No  No  Yes	DevicePi
In [19]:	7039 7040 7041 7042 Name:  seni  20  55  78  91	Month Contract Corcitize Corcitize Customerl 8773 QRDM 1658 BYGO 5067 XJQF 0197 ZHSK 2422 WVHP	One yean one yean one yean one yean one year one year one customer.  Denote one one one one one one one one one on	r h h r h: 7043, dt er[(custome  SeniorCitizen  1 1 1 1	Partner No No Yes No	Dependents  No  No  Yes  No	tenure  1 18 66 30 1	PhoneService  No Yes Yes Yes Yes	MultipleLines  No phone service  Yes  Yes  No	InternetService  DSL  Fiber optic  Fiber optic  DSL  Fiber optic	OnlineSecurity  No  No  No  Yes  No	DevicePi
In [19]:	7039 7040 7041 7042 Name: seni 20 55 57 78	Month Contract Corcitize Corcitize Customerl 8773 QRDM 1658 BYGO 5067 XJQF 0199 ZHSK 2422 WVHP	One yean one yean one yean one yean one yean one year one year one	r h h r h: 7043, dt er[(custome  SeniorCitizen  1 1 1	Partner No No Yes	Dependents  No  No  Yes	tenure  1 18 66 30	PhoneService  No Yes Yes Yes	MultipleLines  No phone service  Yes  Yes  No	InternetService  DSL  Fiber optic  Fiber optic  DSL	OnlineSecurity  No  No  No  Yes	DeviceP
In [19]:	7039 7040 7041 7042 Name:  seni  20 55 57 78 91	Month Contract Orcitize Orcitize Customerl 8778 QRDM 1658 BYGO 5063 XJQF 019° ZHSK 2422 WVHP	One yea not one year. Length of the year o	r h h r h: 7043, dt er[(custome  SeniorCitizen  1 1 1	Partner No No Yes No No No	Dependents  No  No  Yes  No  No  No  No  No  No  No  No  No  N	tenure  1 18 66 30 1	PhoneService  No Yes Yes Yes	MultipleLines  No phone service  Yes  Yes  No  No  No	InternetService  DSL Fiber optic  DSL Fiber optic  DSL Fiber optic	OnlineSecurity  No  No  No  Yes  No	DeviceP
In [19]:	7039 7040 7041 7042 Name:  seni  20  55  57  78  91  6837	Month Contract Corcitize Corcitize Customerl 8779 QRDM 1658 BYGO 5067 XJQF 0197 ZHSK 2422 WVHP 6229 LSCK	One yean to mont to mont Two years to Lengt to mont Two years to Lengt to mont Two years to month to m	r h h r h: 7043, dt  er[(custome  SeniorCitizen  1 1 1 1 1 1 1	Partner No No Yes No No Yes	Dependents  No  No  Yes  No  No  No	tenure  1 18 66 30 1 6	PhoneService  No Yes Yes Yes Yes Yes	MultipleLines  No phone service  Yes  Yes  No  No	InternetService  DSL Fiber optic  DSL Fiber optic  Fiber optic	OnlineSecurity  No No No No Yes No	DevicePi
In [19]:	7039 7040 7041 7042 Name:  seni  20  55  57  78  91  6837	Month Contract Corcitize C	One yean-to-mont Two yearst, Lengt Two yearst, Lengt Penson Two yearst,	r h h r h: 7043, dt er[(custome  SeniorCitizen 1 1 1 1 1 1 1 1 1 1	Partner No No Yes No No Yes Yes	Dependents  No  No  Yes  No  No  No  No	tenure  1  18  66  30  1   6  3	PhoneService  No Yes Yes Yes Yes Yes Yes Yes	MultipleLines  No phone service  Yes  Yes  No  No  No  Yes	InternetService  DSL  Fiber optic  DSL  Fiber optic   Fiber optic   Fiber optic  Fiber optic	OnlineSecurity  No No No Yes No No No	DeviceP

7032 6894- Male 1 No No 1 Yes Yes Fiber optic No

298 rows × 21 columns

In [16]: seniorcitizen=customer[(customer["tenure"]>70 )|(customer["MonthlyCharges"]<100)]</pre>

In [39]: seniorcitizen

Out[39]: customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity 8779-No phone 20 Male 1 No No No DSL No QRDMV service 1658-55 Male 1 No No 18 Yes Yes Fiber optic No BYGOY 5067-57 1 Male Yes Yes 66 Yes Yes Fiber optic No ... **XJQFU** 0191-78 1 30 DSL Male No No Yes No Yes **ZHSKZ** 2424-1 91 Male No No 1 Yes No Fiber optic No ... WVHPL 6229-6837 1 6 Male No No Yes No Fiber optic No ... LSCKB 1400-6894 3 Yes Male Yes Nο Yes Fiber optic No MMYXY 7142-6914 Male 1 Yes No 43 Yes Yes Fiber optic No ... HVGBG

298 rows × 21 columns

8739-

WWKDU 6894-

**LFHLY** 

Male

Male

In [ ]: # E)Extract all the customers whose contract is of two years, payment method is mailed check & the value of churn

25

1

Yes

No

No

Yes

Yes

Fiber optic

Fiber optic

No

No ...

In [41]: seniorcitizen mailcheck=seniorcitizen[(seniorcitizen['Contract']=='Two year') & (seniorcitizen['PaymentMethod']==

In [42]: seniorcitizen

6967

7032

customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceP Out[42]: 8779-No phone 20 Male No No 1 No DSL No QRDMV service 1658-55 Male No No 18 Yes Yes Fiber optic No **BYGOY** 5067-57 Male 1 Yes 66 Yes Fiber optic Yes Yes No ... XJQFU 0191-30 DSI 78 Male 1 Nο Nο Yes Nο Yes ... ZHSKZ 2424-91 Male 1 No No 1 Yes No Fiber optic No ... WVHPL 6229-6837 Male 6 Fiber optic 1 No No Yes No No ... **LSCKB** 1400-6894 3 Male Yes Yes Yes Fiber optic No ... Nο MMYXY 7142-6914 Male 1 Yes No 43 Yes Yes Fiber optic No ... **HVGBG** 8739-6967 Male No No 25 Fiber optic No WWKDU 6894-7032 Yes No No Yes Fiber optic No ... LFHLY

298 rows × 21 columns

```
In [ ]:
           ## F)Extract 333 random records from the customer_churn dataframe & store the result in 'customer_333'
In [50]:
           customer_333=customer.sample(n=333)
In [82]:
           customer_333
                customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceP
Out[82]:
                     8144-
           4808
                            Female
                                                                       54
                                                                                    Yes
                                                                                                               DSL
                                                                                                                             Yes
                    DGHXP
                     9948-
           1039
                              Male
                                                   Yes
                                                                       38
                                                                                    Yes
                                                                                                 No
                                                                                                          Fiber optic
                                                                                                                             Yes
                                                                No
                    YPTDG
                     4760-
           2783
                                             0
                            Female
                                                   Yes
                                                                       43
                                                                                   Yes
                                                                                                 Yes
                                                                                                          Fiber optic
                                                                                                                             Yes ...
                                                                No
                    THGOT
           6208 0909-SELIE
                                             0
                                                                                                              DSL
                                                                                                                             Yes ...
                              Male
                                                   Yes
                                                                No
                                                                       61
                                                                                    Yes
                                                                                                 Yes
                     7473-
                                                                                                                       No internet
                                                                                                                                           Ν
           2658
                            Female
                                             0
                                                   Yes
                                                               Yes
                                                                       14
                                                                                    Yes
                                                                                                 No
                                                                                                                No
                    ZBDSN
                                                                                                                          service
                     2040-
                                             0
                                                                                                               DSL
           218
                              Male
                                                   Yes
                                                               Yes
                                                                       65
                                                                                    Yes
                                                                                                Yes
                                                                                                                              No ...
                    LDIWQ
                     5696-
                                                                                                                        No internet
           3032
                              Male
                                             0
                                                    No
                                                                Nο
                                                                       17
                                                                                    Yes
                                                                                                 Nο
                                                                                                                Nο
                    EXCYS
                                                                                                                           service
                     4797-
           5678
                            Female
                                             0
                                                    No
                                                               Yes
                                                                        1
                                                                                    Yes
                                                                                                 No
                                                                                                               DSL
                                                                                                                             Yes ...
                    AXPXK
                     0336-
                                                                                            No phone
           5690
                              Male
                                                                                                               DSL
                                                                                                                             Yes
                    KXKFK
                                                                                              service
                     8229-
           3479
                            Female
                                             0
                                                    No
                                                                No
                                                                                    Yes
                                                                                                 No
                                                                                                               DSL
                                                                                                                              No ...
                    TNIQA
          333 rows × 21 columns
 In [ ]:
           #g) Get the count of different levels from the 'churn' column
In [83]:
           customer['Churn'].value counts()
Out[83]: No
                  5174
                  1869
           Yes
          Name: Churn, dtype: int64
 In [4]:
           # QUESTION
           # plot for the 'InternetService' column:
           i. Set x-axis label to 'Categories of Internet Service'
           ii. Set y-axis label to 'Count of Categories'
           iii. Set the title of plot to be 'Distribution of Internet Service'
           iv. Set the color of the bars to be 'orange'
           b. Build a histogram for the 'tenure' column:
              Set the number of bins to be 30
           ii. Set the color of the bins to be 'green'
           iii. Assign the title 'Distribution of tenure'
           c. Build a scatter-plot between 'MonthlyCharges' & 'tenure'. Map 'MonthlyCharges' to
           the y-axis & 'tenure' to the 'x-axis':
           i. Assign the points a color of 'brown'ii. Set the x-axis label to 'Tenure of customer'
           iii. Set the y-axis label to 'Monthly Charges of customer'
           iv. Set the title to 'Tenure vs Monthly Charges'
d. Build a box-plot between 'tenure' & 'Contract'. Map 'tenure' on the y-axis &
            'Contract' on the x-axis
             File "<ipython-input-4-4bf9dc5dc80a>", line 2
               i. Set x-axis label to 'Categories of Internet Service'
          SyntaxError: invalid syntax
```

In [53]:

#i)set x-axis label to 'Categories of internet Service

```
In [84]:
           import matplotlib.pyplot as plt
In [85]:
           customer['InternetService'].value_counts()
Out[85]: Fiber optic
                           3096
          DSL
                           2421
                           1526
          Name: InternetService, dtype: int64
In [86]:
           x=customer['InternetService'].value_counts()
In [87]:
Out[87]: Fiber optic
                            3096
                           2421
          DSL
                            1526
          Name: InternetService, dtype: int64
In [77]:
           y=customer['InternetService'].value counts().tolist()
In [78]:
Out[78]: [3096, 2421, 1526]
In [98]:
           import numpy as np
           import matplotlib.pyplot as plt
           plt.bar(x,y,color='red')
plt.xlabel("categories of internet Service")
plt.ylabel("Count of categories")
           plt.title("Distribution of internet Service")
           plt.show()
                              Distribution of internet Service
             3000
             2500
           Count of categories
             2000
             1500
            1000
             500
               0
                         1800 2000 2200 2400 2600 2800
                              categories of internet Service
 In [ ]:
           # b)Build a histogram for the 'tenure' column:
           ## i)set num of bins to be 30
           ## ii)set the color of bins to be 'green'
           ## iii)Assign the title 'Distribution of tenure'
In [94]:
           plt.hist(customer['tenure'],bins=30,color='green')
           plt.title("Distribution of tenure")
           plt.show()
                               Distribution of tenure
           800
```

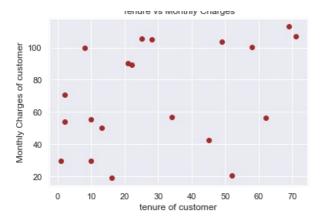
600

```
200
```

```
In [ ]:
           ### Built a scatter-plot between 'MonthlyCharges' & 'tenure'. Map 'MonthlyCharges' to the y-axis & 'tenure' to th
           #### i)Assign the points a color of 'brown'
           #### ii)Set the x-axis label to tenure of customer
           #### iii)Set the y-axis label to 'Monthly Charges of customer'
           #### iV)Set the title to the 'Tenure vs Monthly Charges'
 In [ ]:
           import pandas as pd
In [99]:
           data=pd.read_csv("customer_churn.csv")
In [100...
           data
                customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DevicePi
                     7590-
                                                                                             No phone
             0
                                              0
                                                                                                               DSL
                            Female
                                                                No
                                                                                                                               No
                    VHVEG
                                                                                              service
                     5575-
                              Male
                                             0
                                                    No
                                                                       34
                                                                                    Yes
                                                                                                               DSL
                                                                No
                                                                                                  No
                                                                                                                              Yes ...
                    GNVDE
                     3668-
             2
                                             0
                                                                        2
                                                                                                               DSL
                                                                                                                              Yes ...
                              Male
                                                    No
                                                                                    Yes
                                                                                                  No
                                                                No
                    QPYBK
                     7795-
                                                                                            No phone
             3
                              Male
                                             0
                                                    No
                                                                No
                                                                        45
                                                                                     No
                                                                                                               DSL
                                                                                                                              Yes
                   CFOCW
                                                                                               service
                     9237-
             4
                            Female
                                             0
                                                    No
                                                                No
                                                                        2
                                                                                    Yes
                                                                                                  No
                                                                                                          Fiber optic
                                                                                                                               No ...
                    HQITU
                     6840-
           7038
                                                                                                               DSL
                                             0
                              Male
                                                    Yes
                                                                       24
                                                                                                 Yes
                                                                                                                              Yes ...
                                                               Yes
                                                                                    Yes
                    RESVB
                     2234-
                                             0
           7039
                            Female
                                                    Yes
                                                               Yes
                                                                        72
                                                                                    Yes
                                                                                                 Yes
                                                                                                          Fiber optic
                                                                                                                               No
                    XADUH
                     4801-
                                                                                            No phone
           7040
                            Female
                                              0
                                                    Yes
                                                               Yes
                                                                        11
                                                                                     No
                                                                                                               DSL
                                                                                                                              Yes ...
                     JZAZL
                                                                                               service
                     8361-
           7041
                              Male
                                                    Yes
                                                                No
                                                                                    Yes
                                                                                                 Yes
                                                                                                          Fiber optic
                                                                                                                               No
                    LTMKD
           7042 3186-AJIEK
                                              0
                              Male
                                                    No
                                                                No
                                                                       66
                                                                                    Yes
                                                                                                  No
                                                                                                          Fiber optic
                                                                                                                              Yes ...
          7043 rows × 21 columns
In [101...
           data.head(2)
Out[101...
             customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceProte
                  7590-
                                                                                         No phone
           0
                         Female
                                           0
                                                 Yes
                                                             No
                                                                                  No
                                                                                                            DSL
                                                                                                                           No
                 VHVEG
                                                                                            service
                  5575
                                                                                                            DSL
                           Male
                                                                                                                           Yes ...
                 GNVDE
          2 rows × 21 columns
In [102...
           import matplotlib.pyplot as plt
In [105...
           plt.scatter(x=data['tenure'].head(20),y=data['MonthlyCharges'].head(20),color='brown')
           plt.xlabel('tenure of customer')
           plt.ylabel('Monthly Charges of customer')
           plt.title('Tenure vs Monthly Charges')
```

Tanura ve Monthly Charace

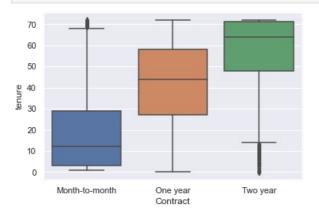
plt.show()



In [106. ## d)built a boxplot between tenure & contract. Map 'tenure' on y-axis and 'Contract' on the x-axis

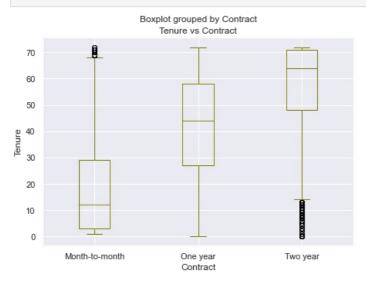
In [108... import seaborn as sns

sns.boxplot(x=data['Contract'],y=data['tenure'])
plt.show()



```
import seaborn as sns
import matplotlib.pyplot as plt
```

```
data.boxplot(by=['Contract'],column="tenure",figsize=(7,5),color='olive')
plt.xlabel("Contract")
plt.ylabel("Tenure")
plt.title("Tenure vs Contract")
plt.show()
```



```
In [ ]:
    # QUESETION
    # C) Linear Regression:
    # a. Build a simple linear model where dependent variable is 'MonthlyCharges' and independent variable is 'tenure'
```

```
#ii. Build the model on train set and predict the values on test set
           #iii. After predicting the values, find the root mean square error
           #iv. Find out the error in prediction & store the result in 'error'
           #v. Find the root mean square error
 In [ ]:
           # i)Divide the dataset into train and test sets is 70:30 ratio
In [112...
           from sklearn.model_selection import train_test_split
           from sklearn.linear_model import LinearRegression
In [113...
           data.head(2)
             customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceProte
Out[113...
                  7590-
                                                                                          No phone
           0
                                           0
                                                                                                             DSL
                         Female
                                                 Yes
                                                             No
                                                                                  No
                                                                                                                            No
                 VHVEG
                                                                                            service
                  5575-
                           Male
                                           0
                                                  No
                                                             No
                                                                     34
                                                                                 Yes
                                                                                               No
                                                                                                             DSL
                                                                                                                           Yes ...
          1
                 GNVDE
          2 rows × 21 columns
In [114...
           x=pd.DataFrame(data['tenure']) #independent variable
In [116...
           y=data['MonthlyCharges']
                                               # dependent variable
In [117...
           x\_train, x\_test, y\_train, y\_test=train\_test\_split(x,y,test\_size=0.3, random\_state=0)
           print(data.shape)
           print(x train.shape)
           print(x_test.shape)
           print(y_train.shape)
           print(y_test.shape)
           (7043, 21)
           (4930, 1)
           (2113, 1)
           (4930,)
           (2113,)
In [118...
           import numpy as np
           import pandas as pd
In [119...
           data=pd.read_csv("customer_churn.csv")
In [120...
           data
                customerID gender SeniorCitizen Partner
                                                        Dependents
                                                                           PhoneService
                                                                                         MultipleLines InternetService OnlineSecurity ... DeviceP
Out[120...
                                                                    tenure
                     7590-
                                                                                             No phone
              0
                            Female
                                              0
                                                    Yes
                                                                No
                                                                         1
                                                                                     No
                                                                                                                DSL
                                                                                                                               No
                    VHVEG
                                                                                               service
                     5575-
                                              0
                                                                        34
                                                                                                                DSL
                              Male
                                                    No
                                                                No
                                                                                    Yes
                                                                                                  No
                                                                                                                              Yes
                    GNVDE
                     3668-
                                                                                                               DSI
                                                                                                                              Yes ...
              2
                              Male
                                              0
                                                    Nο
                                                                Nο
                                                                         2
                                                                                    Yes
                                                                                                  Nο
                    QPYBK
                     7795-
                                                                                             No phone
              3
                              Male
                                              0
                                                    No
                                                                No
                                                                        45
                                                                                     No
                                                                                                                DSL
                                                                                                                              Yes ...
                   CFOCW
                                                                                               service
                     9237-
              4
                            Female
                                              0
                                                    No
                                                                No
                                                                         2
                                                                                    Yes
                                                                                                           Fiber optic
                                                                                                                               No ...
                    HOITU
            ...
                     6840-
           7038
                                                                                                               DSL
                                                                                                                              Yes ...
                              Male
                                              0
                                                    Yes
                                                               Yes
                                                                        24
                                                                                    Yes
                                                                                                 Yes
                    RESVB
                     2234-
                                                                                                                               No ...
           7039
                            Female
                                              0
                                                    Yes
                                                                Yes
                                                                        72
                                                                                    Yes
                                                                                                 Yes
                                                                                                           Fiber optic
                    XADUH
                     4801-
                                                                                             No phone
           7040
                                              0
                                                                                                                DSL
                            Female
                                                    Yes
                                                                Yes
                                                                        11
                                                                                     No
                                                                                                                              Yes ...
                     JZAZL
                                                                                               service
                     8361-
           7041
                              Male
                                                    Yes
                                                                No
                                                                                    Yes
                                                                                                 Yes
                                                                                                           Fiber optic
                                                                                                                               No ...
                    LTMKD
```

#i. Divide the dataset into train and test sets in 70:30 ratio.

```
Yes ...
          7042 3186-AJIEK
                            Male
                                                                                Yes
                                                                                             No
                                                                                                     Fiber optic
         7043 rows × 21 columns
In [122...
           data.head(3)
            customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceProte
Out[122...
                  7590-
                                                                                     No phone
          0
                                         0
                       Female
                                              Yes
                                                          No
                                                                              No
                                                                                                       DSL
                                                                                                                      No ...
                VHVEG
                                                                                       service
                  5575-
                          Male
                                         0
                                               No
                                                          No
                                                                             Yes
                                                                                          No
                                                                                                       DSL
                                                                                                                     Yes ...
          1
                GNVDE
                  3668-
          2
                                         0
                                                                  2
                                                                             Yes
                                                                                                       DSI
                                                                                                                     Yes ...
                          Male
                                               Nο
                                                          Nο
                                                                                          Nο
                QPYBK
         3 rows × 21 columns
In [123...
           lr=LinearRegression()
           lr.fit(x_train,y_train)
Out[123. LinearRegression()
In [124...
           y_pred=lr.predict(x_test)
In [125...
           y_pred
Out[125... array([60.95089608, 72.98096699, 59.1903979 , ..., 75.62171426,
                  70.63363608, 65.6455579 ])
In [126...
           y_test.values
Out[126... array([ 58.2 , 116.6 , 71.95, ..., 109.95, 24.55, 81.6 ])
 In [ ]:
           #D)Logistic Regression
           # a)Built a simple logistic regressin model where dependent variable is 'churn' & independent variable is 'Monthi
           #i)Divide the dataset into 65:35 ratio
           #ii)Build the model on train set and predict the values on test set
           # iV)Build the confusion matrix and get the accuracy score¶
In [127...
           from sklearn.metrics import mean_squared_error
           import numpy as np
In [128...
           msc=mean squared error(y test,y pred)
In [129...
           error=np.sqrt(msc)
In [130...
           error
Out[130... 29.394584027273893
In [131...
In [132...
           from sklearn.linear_model import LogisticRegression
In [133...
```

x=pd.DataFrame(data['MonthlyCharges'])

```
In [134...
           y=data['Churn']
In [139...
           y_pred
\texttt{Out} [\texttt{139}\_\texttt{array}([\texttt{60.95089608}, \ \texttt{72.98096699}, \ \texttt{59.1903979} \ , \ \dots, \ \texttt{75.62171426},
                  70.63363608, 65.6455579 ])
In [140...
           y_test.values
Out[140... array([ 58.2 , 116.6 , 71.95, ..., 109.95, 24.55, 81.6 ])
In [145...
            from sklearn.metrics import confusion matrix,accuracy_score
In [146...
            (1815+0)/(1815+651+0+0)
Out[146... 0.7360097323600974
In [148...
           x=pd.DataFrame(data.loc[:,['tenure','MonthlyCharges']])
In [151...
           y=data['Churn']
In [153...
Out[153... 0
                     No
                     No
           1
                    Yes
          3
                    Nο
                    Yes
          7038
                     No
           7039
                    No
          7040
                    No
          7041
                   Yes
           7042
                    No
          Name: Churn, Length: 7043, dtype: object
In [154...
           mlr=LogisticRegression()
In [155...
           mlr
Out[155... LogisticRegression()
In [159...
           y_pred
Out[159... array([60.95089608, 72.98096699, 59.1903979 , ..., 75.62171426,
                  70.63363608, 65.6455579 ])
In [160...
           y_test.values
Out[160... array([ 58.2 , 116.6 , 71.95, ..., 109.95, 24.55, 81.6 ])
In [162...
           (934+156)/(934+156+212+107)
```

```
In [164...
           # QUESTION
          # D) Logistic Regression:
          # a. Build a simple logistic regression modelwhere dependent variable is 'Churn' & independent variable is 'Month
          # i. Divide the dataset in 65:35 ratio
          # ii. Build the model on train set and predict the values on test set
          # iii. Build the confusion matrix and get the accuracy score
 In [ ]:
          # OUESTION
          # b. Build a multiple logistic regression model where dependent variable is 'Churn' & independent variables are
          # i. Divide the dataset in 80:20 ratio
          # ii. Build the model on train set and predict the values on test set
          # iii. Build the confusion matrix and get the accuracy score
In [165...
          x=pd.DataFrame(data.loc[:,['tenure']])
          y=data['Churn']
In [166...
          x\_train, x\_test, y\_train, y\_test=train\_test\_split(x, y, train\_size=0.8, random\_state=0)
In [167...
          from sklearn.tree import DecisionTreeClassifier
In [168...
          # Create Decision Tree classifer object
          clf = DecisionTreeClassifier()
          # Train Decision Tree Classifer
          clf = clf.fit(x_train,y_train)
          #Predict the response for test dataset
          y_pred = clf.predict(x_test)
In [169...
          y pred
Out[169... array(['No', 'No', 'No', 'No', 'No', 'Yes'], dtype=object)
In [170...
          y_test
Out[170... 2200
                   No
          3225
                   No
          2828
                   No
         3768
                   No
         2631
                 Yes
         5333
                 Yes
         6972
                 Yes
          4598
                   No
         3065
                   No
         Name: Churn, Length: 1409, dtype: object
In [172...
          #Import scikit-learn metrics module for accuracy calculation
          from sklearn.metrics import confusion_matrix,accuracy_score
          from sklearn import metrics
```

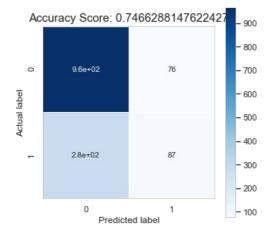
In [173... # Model Accuracy how often is the classifier correct?

```
print("Accuracy:", metrics.accuracy_score(y_test, y_pred))

Accuracy: 0.7466288147622427
```

```
cm = confusion_matrix(y_test, y_pred)
plt.figure(figsize=(5,5))
sns.heatmap(data=cm,linewidths=.5, annot=True,square = True, cmap = 'Blues')
plt.ylabel('Actual label')
plt.xlabel('Predicted label')
all_sample_title = 'Accuracy Score: {0}'.format(clf.score(x_test, y_test))
plt.title(all_sample_title, size = 15)
```

## Out[177... Text(0.5, 1.0, 'Accuracy Score: 0.7466288147622427')



In [ ]:

In [ ]:

In [ ]:

In [ ]:

```
In [178...
          !pip install graphviz
         Collecting graphviz
           Downloading graphviz-0.20-py3-none-any.whl (46 kB)
                                               | 46 kB 313 kB/s eta 0:00:01
         Installing collected packages: graphviz
         Successfully installed graphviz-0.20
In [179...
          !pip install pydotplus
         Collecting pydotplus
           Downloading pydotplus-2.0.2.tar.gz (278 kB)
                                        | 278 kB 503 kB/s eta 0:00:01
         Requirement already satisfied: pyparsing>=2.0.1 in /opt/anaconda3/lib/python3.8/site-packages (from pydotplus) (2
          .4.7)
         Building wheels for collected packages: pydotplus
           Building wheel for pydotplus (setup.py) ... done
           Created wheel for pydotplus: filename=pydotplus-2.0.2-py3-none-any.whl size=24566 sha256=12e67bb213ac49306445a6
         75f4026cbb2e3ddd3192bc720bba2ab611a8332b75
           Stored in directory: /Users/surajjamadar/Library/Caches/pip/wheels/fe/cd/78/a7e873cc049759194f8271f780640cf96b3
         5e5a48bef0e2f36
         Successfully built pydotplus
         Installing collected packages: pydotplus
         Successfully installed pydotplus-2.0.2
 In [ ]:
```

```
In [ ]:
    In [9]: # QUESTION
               E) Decision Tree:
               a. Build a decision tree model where dependent variable is 'Churn' & independent
               variable is 'tenure'
               i. Divide the dataset in 80:20 ratio
               ii. Build the model on train set and predict the values on test set
               iii. Build the confusion matrix and calculate the accuracy
                File "<ipython-input-9-f3425134eba2>", line 2
                   E) Decision Tree:
              SyntaxError: unmatched ')'
    In [ ]:
               # QUESTION
               F) Random Forest:
               a. Build a Random Forest model where dependent variable is 'Churn' & independent
               variables are 'tenure' and 'MonthlyCharges'
i. Divide the dataset in 70:30 ratio
               ii. Build the model on train set {\color{blue} \text{and}} predict the values on test set iii. Build the confusion matrix {\color{blue} \text{and}} calculate the accuracy
Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js
```