```
In [7]:
          # 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'
              Extract all the male senior citizens whose Payment Method is Electronic check & store the result in 'senior_n
          #d. Extract all those customers whose tenure is greater than 70 months or their Monthly charges is more than 100$
          #e. Extract all the customers whose Contract is of two years, payment method is Mailed check & the value of Churr
          #f. Extract 333 random records from the customer_churndataframe& store the result in 'customer_333' #g. Get the count of different levels from the 'Churn' column
In [3]:
          # ANSWER BELLOW :
In [3]:
          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 [5]:
          customer = pd.read csv("customer churn.csv")
In [6]:
          customer
Out[6]:
               customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceP
                    7590-
                                                                                            No phone
            0
                                                                                                               DSL
                           Female
                                             0
                                                   Yes
                                                               No
                                                                                    No
                                                                                                                              No
                   VHVEG
                                                                                              service
                    5575-
                             Male
                                             0
                                                   No
                                                                No
                                                                       34
                                                                                   Yes
                                                                                                 No
                                                                                                               DSL
                                                                                                                             Yes
                   GNVDE
                    3668-
            2
                             Male
                                             0
                                                   No
                                                               No
                                                                        2
                                                                                   Yes
                                                                                                 No
                                                                                                               DSL
                                                                                                                             Yes ...
                   QPYBK
                    7795-
                                                                                            No phone
            3
                                             0
                                                                                                               DSL
                             Male
                                                                       45
                                                                                    No
                                                                                                                              Yes ...
                                                   No
                                                               No
                  CFOCW
                                                                                              service
                    9237-
                           Female
                                             0
                                                                        2
                                                                                                                              No ...
            4
                                                   No
                                                               No
                                                                                   Yes
                                                                                                 No
                                                                                                          Fiber optic
                   HQITU
                    6840-
          7038
                                             0
                                                                       24
                                                                                                               DSL
                             Male
                                                   Yes
                                                               Yes
                                                                                   Yes
                                                                                                 Yes
                                                                                                                             Yes ...
                   RFSVB
                    2234-
          7039
                           Female
                                             0
                                                   Yes
                                                                       72
                                                                                                 Yes
                                                                                                          Fiber optic
                                                               Yes
                                                                                   Yes
                                                                                                                              No ...
                   XADUH
                    4801-
                                                                                            No phone
                                                                                                               DSL
                           Female
                                             0
                                                                                                                             Yes ...
          7040
                                                   Yes
                                                              Yes
                                                                       11
                                                                                    Nο
                    JZAZL
                    8361-
          7041
                             Male
                                                   Yes
                                                               No
                                                                        4
                                                                                   Yes
                                                                                                 Yes
                                                                                                          Fiber optic
                                                                                                                              No ...
                   LTMKD
          7042 3186-AJIEK
                             Male
                                             0
                                                   No
                                                                       66
                                                                                   Yes
                                                                                                 No
                                                                                                          Fiber optic
                                                               No
                                                                                                                             Yes ...
         7043 rows × 21 columns
In [7]:
          customer.head(5)
Out[7]:
            customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceProte
                 7590-
                                                                                         No phone
          0
                        Female
                                          0
                                                                                                            DSL
                                                Yes
                                                             No
                                                                     1
                                                                                 No
                                                                                                                           No ...
                VHVEG
                                                                                           service
                 5575-
                                          Λ
                                                                                                                           Yes ...
          1
                          Male
                                                 No
                                                             No
                                                                    34
                                                                                Yes
                                                                                              No
                                                                                                            DSL
                GNVDE
                 3668-
          2
                          Male
                                          0
                                                No
                                                            No
                                                                     2
                                                                                Yes
                                                                                              No
                                                                                                            DSI
                                                                                                                           Yes ..
                QPYBK
                  7795-
                                                                                         No phone
                                          0
                          Male
                                                 No
                                                             No
                                                                    45
                                                                                                            DSL
                                                                                                                           Yes ...
               CFOCW
                                                                                           service
                 9237-
                        Female
                                          0
                                                             No
                                                                     2
                                                                                 Yes
                                                                                                       Fiber optic
                                                                                                                           No ...
                                                 No
                                                                                              No
```

HQITU

5 rows × 21 columns In [11]: customer.iloc[:,5] 0 1 34 1 2 2 45 3 4 2 7038 24 7039 72 7040 11 7041 4 7042 66 Name: tenure, Length: 7043, dtype: int64 In [12]: customer.iloc[:,15] Out[12]: 0 Month-to-month One year 2 Month-to-month 3 One year 4 Month-to-month 7038 One year 7039 One year 7040 Month-to-month 7041 Month-to-month 7042 Two year Name: Contract, Length: 7043, dtype: object In [14]: seniorcitizen=customer[(customer["gender"]=="Male")&(customer["SeniorCitizen"]==1)&(customer["PaymentMethod"]=="E In [15]: seniorcitizen customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceP Out[15]: 8779-No phone 20 DSL Male 1 No No 1 No No ... QRDMV service 1658-55 Male No 18 Yes Yes No ... No Fiber optic **BYGOY** 5067-57 Male 1 Yes Yes 66 Yes Yes Fiber optic No ... **XJQFU** 0191-78 Male No No 30 Yes No DSL Yes ... **ZHSKZ** 2424-91 1 Male No No 1 Yes No Fiber optic No ... WVHPI 6229-No ... 6837 Male 1 No No 6 Yes No Fiber optic LSCKB 1400-6894 Male Yes No 3 Yes Fiber optic No MMYXY 7142-6914 Male 1 Yes No 43 Yes Yes Fiber optic No ... **HVGBG** 8739-6967 25 Yes Male No No Yes Fiber optic No WWKDU 6894-1 7032 Male No Yes Yes Fiber optic No ... No LFHLY 298 rows × 21 columns In [16]: seniorcitizen=customer[(customer["tenure"]>70)|(customer["MonthlyCharges"]<100)]</pre>

In [17]:

	senio	orcitizen																
ut[17]:		customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity		DeviceP					
	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							
	7037	2569- WGERO	Female	0	No	No	72	Yes	No	No	No internet service		N					
	7038	6840- RESVB	Male	0	Yes	Yes	24	Yes	Yes	DSL	Yes							
	7039	2234- XADUH	Female	0	Yes	Yes	72	Yes	Yes	Fiber optic	No							
	7040	4801- JZAZL	Female	0	Yes	Yes	11	No	No phone service	DSL	Yes							
	7041	8361- LTMKD	Male	1	Yes	No	4	Yes	Yes	Fiber optic	No							
	6310 ro	ows × 21 co	lumns															
	4												Þ					
n []:	# E)!	Evtract a	11 the	customers w	hose co	ntract is d	of two	vears naumen	t method is	mailed check	c & the value	0.1	churn					
	# L/L	LXTTACT A	cc the	customers wi	1036 601	illiact 13 C	or Ewo	years, paymen	t method 13	marted check	t & the value	. 01	CHUTH					
[18]:	senio	orcitizen_	_mailch	eck=seniorc	itizen[(seniorciti	izen['C	ontract']=='	Two year')	& (seniorciti	izen['Payment	Met	hod']==					
[20]:																		
	senio	orcitizen																
ut[20]:			gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity		DeviceP					
it[20]:			gender Female	SeniorCitizen 0	Partner Yes	Dependents No	tenure	PhoneService No	MultipleLines No phone service	InternetService DSL			DeviceP					
t[20]:		customerID 7590- VHVEG 5575-							No phone		No		DeviceP					
t[20]:	0	customerID 7590- VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No Yes		DeviceP					
it[20]:	0	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795-	Female Male	0	Yes	No No	1 34	No Yes	No phone service No No	DSL	No Yes Yes		DeviceP					
t[20]:	0 1 2	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW	Female Male Male	0 0	Yes No No	No No	1 34 2	No Yes Yes	No phone service No	DSL DSL	No Yes Yes		DeviceP					
t[20]:	0 1 2 3	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795-	Female Male Male Male	0 0 0	Yes No No	No No No	1 34 2 45	No Yes Yes	No phone service No No No phone service	DSL DSL DSL	No Yes Yes No		DeviceP					
it[20]:	0 1 2 3 4	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU	Female Male Male Male Female	0 0 0 0	Yes No No No	No No No No	1 34 2 45 2	No Yes Yes No Yes	No phone service No No No phone service No	DSL DSL DSL Fiber optic	No Yes Yes No No internet		DevicePi					
t[20]:	0 1 2 3 4	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU	Female Male Male Male Female	0 0 0 0	Yes No No No No	No No No No	1 34 2 45 2	No Yes Yes No Yes	No phone service No No No phone service No	DSL DSL DSL Fiber optic	No Yes Yes No No internet service							
ut[20]:	0 1 2 3 4 7037	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840-	Female Male Male Male Female Female	0 0 0 0	Yes No No No No No	No No No No	1 34 2 45 2 72	No Yes Yes No Yes Yes	No phone service No No No phone service No No	DSL DSL DSL Fiber optic No	No Yes Yes No No internet service Yes							
ut[20]:	0 1 2 3 4 7037	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH	Female Male Male Female Female	0 0 0 0	Yes No No No No Yes	No No No No No Yes	1 34 2 45 72 24	No Yes Yes No Yes Yes Yes	No phone service No No No phone service No No Yes	DSL DSL Fiber optic No	No Yes Yes No No internet service Yes							
t[20]:	0 1 2 3 4 7037 7038 7039	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH 4801- JZAZL	Female Male Male Male Female Male	0 0 0 0 0 0 0 0 0	Yes No No No No Yes Yes	No No No No No Yes Yes	1 34 2 45 2 72 24 72	No Yes Yes No Yes Yes Yes	No phone service No No No phone service No No Yes Yes No phone	DSL DSL DSL Fiber optic No DSL Fiber optic	No Yes Yes Yes No No internet service Yes No Yes							
	0 1 2 3 4 7037 7038 7039 7040	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH 4801- JZAZL 8361- LTMKD	Female Male Male Female Male Female Male Female Male	0 0 0 0 0 0 0 0	Yes No No No No Yes Yes	No No No No No Yes Yes	1 34 2 45 2 72 24 72 11	No Yes Yes No Yes Yes Yes No	No phone service No No No phone service No No Yes Yes No phone service	DSL DSL DSL Fiber optic No DSL Fiber optic DSL	No Yes Yes Yes No No internet service Yes No Yes							
	0 1 2 3 4 7037 7038 7039 7040	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH 4801- JZAZL	Female Male Male Female Male Female Male Female Male	0 0 0 0 0 0 0 0	Yes No No No No Yes Yes	No No No No No Yes Yes	1 34 2 45 2 72 24 72 11	No Yes Yes No Yes Yes Yes No	No phone service No No No phone service No No Yes Yes No phone service	DSL DSL DSL Fiber optic No DSL Fiber optic DSL	No Yes Yes Yes No No internet service Yes No Yes							
	0 1 2 3 4 7037 7038 7039 7040 7041 6310 ro	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH 4801- JZAZL 8361- LTMKD	Female Male Male Female Male Female Male Female Male	0 0 0 0 0 0 0 0	Yes No No No No Yes Yes	No No No No No Yes Yes	1 34 2 45 2 72 24 72 11	No Yes Yes No Yes Yes Yes No	No phone service No No No phone service No No Yes Yes No phone service	No DSL Yes No phone service DSL Yes No Fiber optic No No No intermet service No Yes DSL Yes No phone service DSL Yes Yes Fiber optic No Method is mailed check & the value of churn of churn O year') & (seniorcitizen['PaymentMethod']== Mo phone service DSL No No DSL Yes No phone service DSL Yes No Fiber optic No No Fiber optic No intermet service No intermet service No phone service DSL Yes No phone service DSL Yes No phone service DSL Yes Yes DSL Yes .								
	0 1 2 3 4 7037 7038 7039 7040 7041 6310 ro	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH 4801- JZAZL 8361- LTMKD	Female Male Male Male Female Female Male Female Male Itemale Itemale Male	0 0 0 0 0 0 1	Yes No No No No Yes Yes Yes	No No No No No No Yes Yes No	1 34 2 45 2 72 24 72 11 4	No Yes Yes No Yes Yes Yes Yes Yes Yes	No phone service No No No phone service No No Yes Yes No phone service Yes	DSL DSL DSL Fiber optic No DSL Fiber optic DSL Fiber optic	No Yes Yes No No internet service Yes No Yes							
	0 1 2 3 4 7037 7038 7039 7040 7041 6310 ro	7590- VHVEG 5575- GNVDE 3668- QPYBK 7795- CFOCW 9237- HQITU 2569- WGERO 6840- RESVB 2234- XADUH 4801- JZAZL 8361- LTMKD Dws × 21 col	Female Male Male Male Female Female Male Female Male Itemale Male Male	0 0 0 0 0 0 1	Yes No No No No Yes Yes Yes	No No No No No No Yes Yes No	1 34 2 45 2 72 24 72 11 4	No Yes Yes No Yes Yes Yes Yes Yes Yes	No phone service No No No phone service No No Yes Yes No phone service Yes	DSL DSL DSL Fiber optic No DSL Fiber optic DSL Fiber optic	No Yes Yes No No internet service Yes No Yes							

In [22]:

customer_333

```
4884-
          4590
                          Female
                                                                    57
                                           1
                                                  No
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
                                                             No
                                                                                                                        Yes ...
                   TVUQF
                    3892-
          4797
                            Male
                                           0
                                                 Yes
                                                            Yes
                                                                    72
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
                                                                                                                        Yes ...
                   NXAZG
                    9844-
          3117
                          Female
                                           1
                                                 Yes
                                                            Yes
                                                                    70
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
                                                                                                                         No ...
                    FELAJ
                    5781-
          2685
                          Female
                                                 Yes
                                                             No
                                                                    72
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
                                                                                                                        Yes ...
                   BKHOP
                    2523-
          6997
                          Female
                                                 Yes
                                                             No
                                                                    27
                                                                                Yes
                                                                                             No
                                                                                                     Fiber optic
                                                                                                                         No ...
                  EWWZL
          ---
                    6583-
          4741
                          Female
                                           1
                                                 Yes
                                                             No
                                                                    15
                                                                                Yes
                                                                                             No
                                                                                                     Fiber optic
                                                                                                                         No ...
                   KQJLK
                    9408-
           220
                                           0
                                                                     4
                          Female
                                                  No
                                                             No
                                                                                                     Fiber optic
                                                                                                                         No
                   SSNVZ
                    2391-
           647
                                           0
                                                                    50
                                                                                            Yes
                                                                                                          DSL
                            Male
                                                 Yes
                                                            Yes
                                                                                Yes
                                                                                                                         No ...
                    IPLOP
                    6302-
          6444
                                                                    31
                                                                                                          DSL
                            Male
                                                  No
                                                            Yes
                                                                                Yes
                                                                                            Yes
                                                                                                                         No
                   JGYRJ
                    2696-
                                           0
                                                                                                                        Yes ...
          6337
                          Female
                                                 Yes
                                                            Yes
                                                                    55
                                                                                Yes
                                                                                            Yes
                                                                                                     Fiber optic
         333 rows × 21 columns
 In [ ]:
           #g) Get the count of different levels from the 'churn' column
In [23]:
           customer['Churn'].value_counts()
Out[23]: No
                 5174
                  1869
          Name: Churn, dtype: int64
In [24]:
           # 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:
           #i. 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
           #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
In [31]:
           #i)set x-axis label to 'Categories of internet Service
In [26]:
           import matplotlib.pyplot as plt
In [32]:
           customer['InternetService'].value_counts()
Out[32]: Fiber optic
                          3096
                          2421
          DSL
                          1526
          Name: InternetService, dtype: int64
In [36]:
           x=customer['InternetService'].value counts()
In [37]:
           seniorcitizen
```

customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceP

Out[22]:

```
customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DevicePi
Out[37]:
                        7590-
                                                                                                       No phone
               0
                               Female
                                                   0
                                                                        No
                                                                                                                            DSL
                                                                                                                                             No
                      VHVEG
                                                                                                         service
                        5575-
                                                   0
                                                                                34
                                                                                                                            DSL
                                 Male
                                                          No
                                                                        No
                                                                                              Yes
                                                                                                             No
                                                                                                                                            Yes ...
                      GNVDE
                        3668-
               2
                                                   0
                                                                                 2
                                                                                                                            DSL
                                 Male
                                                          No
                                                                                              Yes
                                                                                                             No
                                                                                                                                            Yes ...
                                                                        No
                      QPYBK
                        7795-
                                                                                                       No phone
                                                   0
                                                                                45
                                                                                                                            DSL
               3
                                 Male
                                                          No
                                                                        No
                                                                                              No
                                                                                                                                            Yes
                      CFOCW
                                                                                                          service
                        9237-
               4
                               Female
                                                   0
                                                          No
                                                                        No
                                                                                 2
                                                                                              Yes
                                                                                                             No
                                                                                                                       Fiber optic
                                                                                                                                             No ...
                       HQITU
                        2569-
                                                                                                                                      No internet
            7037
                                                   0
                                                                                72
                                                          No
                               Female
                                                                       No
                                                                                              Yes
                                                                                                             No
                                                                                                                             No
                     WGERO
                                                                                                                                          service
                        6840-
            7038
                                                   0
                                                                                24
                                                                                                                            DSL
                                 Male
                                                          Yes
                                                                       Yes
                                                                                              Yes
                                                                                                            Yes
                                                                                                                                            Yes
                      RESVB
                        2234-
            7039
                               Female
                                                   0
                                                          Yes
                                                                       Yes
                                                                                72
                                                                                              Yes
                                                                                                            Yes
                                                                                                                       Fiber optic
                                                                                                                                             No
                      XADUH
                        4801-
                                                                                                       No phone
            7040
                               Female
                                                          Yes
                                                                       Yes
                                                                                11
                                                                                              No
                                                                                                                            DSL
                                                                                                                                            Yes
                       JZAZL
                                                                                                          service
                        8361-
            7041
                                 Male
                                                          Yes
                                                                        No
                                                                                 4
                                                                                              Yes
                                                                                                            Yes
                                                                                                                       Fiber optic
                                                                                                                                             No ...
                      LTMKD
           6310 rows × 21 columns
In [38]:
             y=customer['InternetService'].value_counts().tolist()
In [39]:
Out[39]: [3096, 2421, 1526]
In [43]:
             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 categor
              2000
              1500
              1000
               500
                 0
                       1600
                             1800
                                  2000 2200 2400
                                                      2600 2800
                                  categories of internet Service
In [44]:
```

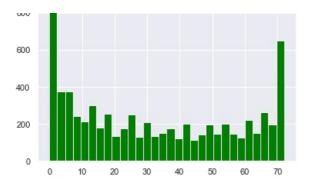
```
## i)set num of bins to be 30
## ii)set the color of bins to be 'green'
## iii)Assign the title 'Distribution of tenure'
In [45]:
plt.hist(customer['tenure'],bins=30,color='green')
```

plt.hist(customer['tenure'],bins=30,color='green')
plt.title("Distribution of tenure")
plt.show()

Distribution of tenure

b)Build a histogram for the 'tenure' column:

ดกล



```
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 [46]:
           import pandas as pd
In [47]:
           data=pd.read_csv("customer_churn.csv")
In [48]:
            data
Out[48]:
                customerID
                            gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DevicePi
                     7590-
                                                                                             No phone
                                                                                                                DSL
                            Female
                                                    Yes
                                                                No
                                                                                     No
                                                                                                                               No
                    VHVEG
                                                                                               service
                     5575-
                                              0
                                                                        34
                                                                                                                DSL
                              Male
                                                    No
                                                                No
                                                                                    Yes
                                                                                                  No
                                                                                                                              Yes
                    GNVDE
                     3668-
             2
                                              0
                                                                                                                DSL
                              Male
                                                                         2
                                                                                                                              Yes ...
                                                    No
                                                                No
                                                                                    Yes
                                                                                                  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
                                              0
                                                                        24
                                                                                                               DSI
                                                                                                                              Yes ...
                              Male
                                                    Yes
                                                                Yes
                                                                                    Yes
                                                                                                 Yes
                    RESVB
                     2234-
           7039
                            Female
                                              0
                                                    Yes
                                                                Yes
                                                                        72
                                                                                    Yes
                                                                                                  Yes
                                                                                                           Fiber optic
                                                                                                                               No
                    XADUH
                                                                                             No phone
                     4801-
           7040
                            Female
                                              0
                                                                Yes
                                                                        11
                                                                                     No
                                                                                                                DSL
                                                                                                                              Yes ...
                                                    Yes
                     JZAZL
                                                                                               service
                     8361-
           7041
                              Male
                                                    Yes
                                                                No
                                                                         4
                                                                                    Yes
                                                                                                  Yes
                                                                                                           Fiber optic
                                                                                                                               No
                    LTMKD
           7042 3186-AJIEK
                                              0
                                                                        66
                                                                                                                              Yes ...
                              Male
                                                    Nο
                                                                Nο
                                                                                    Yes
                                                                                                  Nο
                                                                                                           Fiber optic
          7043 rows × 21 columns
In [101...
           data.head(2)
```

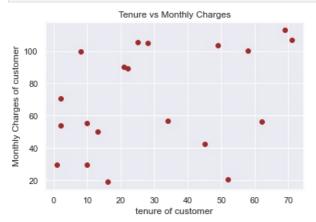
 custon	nerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	•••	DeviceProte
0 VH	7590- IVEG	Female	0	Yes	No	1	No	No phone service	DSL	No		
1 GN	5575- NVDE	Male	0	No	No	34	Yes	No	DSL	Yes		

2 rows × 21 columns

```
In [49]:
          import matplotlib.pyplot as plt
```

In [50]: $\verb|plt.scatter(x=data['tenure'].head(20),y=data['MonthlyCharges'].head(20),color='brown')|$ plt.xlabel('tenure of customer')

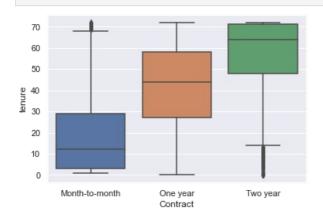
```
plt.ylabel('Monthly Charges of customer')
plt.title('Tenure vs Monthly Charges')
plt.show()
```



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

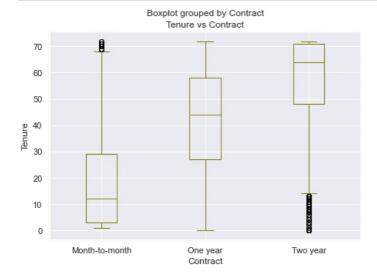
import seaborn as sns

In [52]:
 sns.boxplot(x=data['Contract'],y=data['tenure'])
 plt.show()



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

In [54]:
 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
           #i. Divide the dataset into train and test sets in 70:30 ratio.
           #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 [55]:
           from sklearn.model selection import train test split
           from sklearn.linear model import LinearRegression
In [56]:
           data.head(2)
Out[56]:
             customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DeviceProte
                  7590-
                                                                                      No phone
          0
                                                                                                                       No ...
                        Female
                                         0
                                               Yes
                                                           No
                                                                               No
                                                                                                         DSL
                VHVEG
                                                                                         service
                  5575-
                          Male
                                                                                            No
                                                                                                         DSL
                                                                                                                       Yes
                GNVDE
         2 rows × 21 columns
In [57]:
           x=pd.DataFrame(data['tenure']) #independent variable
In [58]:
           y=data['MonthlyCharges']
                                              # dependent variable
In [59]:
           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 [60]:
           data
                customerID gender SeniorCitizen Partner Dependents tenure PhoneService MultipleLines InternetService OnlineSecurity ... DevicePo
                     7590-
                                                                                         No phone
             0
                           Female
                                            0
                                                  Yes
                                                              No
                                                                      1
                                                                                  No
                                                                                                           DSL
                                                                                                                          No
                   VHVEG
                                                                                           service
                     5575-
                             Male
                                            0
                                                   No
                                                              No
                                                                     34
                                                                                 Yes
                                                                                              No
                                                                                                            DSI
                                                                                                                          Yes
                   GNVDE
                     3668-
             2
                                            0
                                                                      2
                                                                                                            DSL
                             Male
                                                  No
                                                              No
                                                                                 Yes
                                                                                              No
                                                                                                                          Yes ...
                   QPYBK
                     7795-
                                                                                         No phone
             3
                             Male
                                            0
                                                                     45
                                                                                  No
                                                                                                            DSL
                                                  No
                                                              No
                                                                                                                          Yes ...
                   CFOCW
                                                                                           service
                     9237-
             4
                                            0
                                                                      2
                                                                                                       Fiber optic
                                                                                                                          No ...
                           Female
                                                  No
                                                              No
                                                                                 Yes
                                                                                              No
                    HQITU
            ...
                     6840-
          7038
                             Male
                                            0
                                                                     24
                                                                                              Yes
                                                                                                            DSL
                                                                                                                          Yes ...
                                                             Yes
                   RESVB
                     2234-
          7039
                           Female
                                                  Yes
                                                             Yes
                                                                                 Yes
                                                                                              Yes
                                                                                                       Fiber optic
                                                                                                                          No
                   XADUH
```

	7040	4801- JZAZL	Female		0 Y	es `	Yes	11	No No pho		SL Yes		
	7041	8361- LTMKD	Male		1 Y	es	No	4	Yes Y	es Fiber opti	ic No		
	7042 318	6-AJIEK	Male		0	No	No	66	Yes	No Fiber opti	ic Yes		
	7043 rows	× 21 co	lumns										
	4												
In [122	data.he	ad(3)											
Out[122		7500	ender Se emale	niorCitizen 0	Partner Yes	Dependents No		PhoneService No	No phone	InternetService DSL	OnlineSecurity No	DeviceProte	
	1 5	575-	Male	0	No	No		Yes	service	DSL	Yes		
	2 3	VDE 8668- YBK	Male	0	No	No	2	Yes	No	DSL	Yes		
<u> </u>	3 rows × 2		ns										
	4											>	
In [61]:	lr=Line lr.fit(
Out[61]:	LinearRe	gressi	on()										
In [62]:	y_pred=	lr.pred	dict(x_	test)									
In [63]:	y_pred												
Out[63]:	array([6					1903979 ,	, 75	.62171426,					
	7	70.6336	3608, 6	5.6455579)])								
In [64]:	y_test.	values											
Out[64]:	array([58.2 ,	116.6	, 71.95,	,	109.95, 24	4.55,	81.6])					
In []:	<pre>#D)Logistic Regression # a)Built a simple logistic regressin model where dependent variable is 'churn' & independent variable is 'Month' #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¶</pre>												
In [65]:	<pre>from sklearn.metrics import mean_squared_error import numpy as np</pre>												
In [66]:	msc=mean_squared_error(y_test,y_pred)												
In [67]:	error=np.sqrt(msc)												
In [68]:	error												
Out[68]:	29.39458	3402727	3893										
In [69]:	from sk	learn.	linear_r	nodel imp	ort Log	isticRegre	ession						

```
In [70]: | x=pd.DataFrame(data['MonthlyCharges'])
In [71]:
          y=data['Churn']
In [72]:
          y pred
Out[72]: array([60.95089608, 72.98096699, 59.1903979 , ..., 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 [73]:
          y=data['Churn']
In [74]:
Out[74]: 0
                   No
                   No
         2
                  Yes
         3
                  No
                  Yes
         7038
                  No
         7039
                  No
          7040
                   No
         7041
                  Yes
         7042
                  No
         Name: Churn, Length: 7043, dtype: object
In [75]:
          mlr=LogisticRegression()
In [155...
          mlr
Out[155... LogisticRegression()
In [76]:
          y_pred
Out[76]: array([60.95089608, 72.98096699, 59.1903979 , ..., 75.62171426,
                70.63363608, 65.6455579 ])
In [77]:
          y_test.values
Out[77]: array([ 58.2 , 116.6 , 71.95, ..., 109.95, 24.55, 81.6 ])
In [79]:
          (934+156)/(934+156+212+107)
```

```
Out[79]: 0.7735982966643009
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 [80]:
          x=pd.DataFrame(data.loc[:,['tenure','MonthlyCharges']])
In [81]:
          y=data['Churn']
In [82]:
          x_{train}, x_{test}, y_{train}, y_{test} = train_{test}. Split(x, y, train_size=0.8, random_state=0)
In [83]:
          mlr=LogisticRegression()
In [84]:
          mlr.fit(x_train,y_train)
Out[84]: LogisticRegression()
In [85]:
          y_pred=mlr.predict(x_test)
In [86]:
          y_pred
Out[86]: array(['No', 'No', 'No', 'No', 'No', 'No'], dtype=object)
In [87]:
          y_test.values
Out[87]: array(['No', 'No', 'No', 'Yes', 'No', 'No'], dtype=object)
In [88]:
          from sklearn.metrics import confusion_matrix,accuracy_score
In [89]:
          print(confusion_matrix(y_pred,y_test))
         [[934 212]
          [107 156]]
In [90]:
          (934+156)/(934+156+212+107)
Out[90]: 0.7735982966643009
In [91]:
          accuracy_score(y_test,y_pred)
Out[91]: 0.7735982966643009
 In [ ]:
          # QUESTION
          # 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
          4627
          3225
                   Nο
          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
In [177...
           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')
            Accuracy Score: 0.7466288147622427
                                                   - 800
                    9.6e+02
                                                   700
```

- 600

- 500 - 400

- 300 - 200

87

Actual label

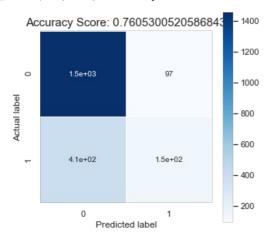
2.8e+02

```
In [92]:
          !pip install graphviz
         Requirement already satisfied: graphviz in /opt/anaconda3/lib/python3.8/site-packages (0.20)
In [93]:
          pip install pydotplus
         Requirement already satisfied: pydotplus in /opt/anaconda3/lib/python3.8/site-packages (2.0.2)
         Requirement already satisfied: pyparsing>=2.0.1 in /opt/anaconda3/lib/python3.8/site-packages (from pydotplus) (2
         .4.7)
         Note: you may need to restart the kernel to use updated packages.
In [101... from sklearn.tree import export_graphviz
          from sklearn.externals.six import StrinIO
          from IPython.display import pydotplus
          dot data = StringIO
          export_graphviz(clf, out_file=dot_data, filled=True, rounded=True, special_characters=True, feature_names = x_trai
          graph = pydotplus.graph from dot data(dot data.getvalue())
          graph.write png('diabetes.png')
          Image(graph.create_png())
         ModuleNotFoundError
                                                   Traceback (most recent call last)
         <ipython-input-101-d5db378b6599> in <module>
               1 from sklearn.tree import export_graphviz
         ----> 2 from sklearn.externals.six import StrinIO
               3 from IPython.display import pydotplus
               5 dot data = StringIO
         ModuleNotFoundError: No module named 'sklearn.externals.six'
 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 and predict the values on test set
          #iii. Build the confusion matrix and calculate the accuracy
In [102...
          from sklearn.model selection import train test split
          X=data[['tenure']] # Features
          y=data['Churn'] # Labels
          # Split dataset into training set and test set
          X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.3)
In [107...
          #Import Random Forest Model
          from sklearn.ensemble import RandomForestClassifier
          #Create a Random Forest Classifier
          seniorcitizen=RandomForestClassifier(n estimators=10)
          #Train the model using the training sets y pred=clf.predict(X test)
          seniorcitizen.fit(X train,y train)
          y_pred=seniorcitizen.predict(X_test)
          #Import scikit-learn metrics module for accuracy calculation
          from sklearn import metrics
          # Model Accuracy, how often is the classifier correct?
          print("Accuracy:",metrics.accuracy_score(y_test, y_pred))
```

Accuracy: 0.7605300520586843

```
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(seniorcitizen.score(X_test, y_test))
plt.title(all_sample_title, size = 15)
```

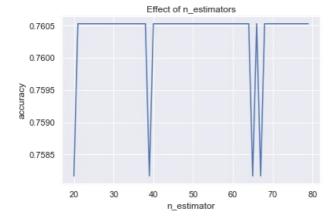
Out[109... Text(0.5, 1.0, 'Accuracy Score: 0.7605300520586843')



```
# Try different numbers of n_estimators
estimators = np.arange(20,80)
accuracy = []

for n in estimators:
    seniorcitizen.set_params(n_estimators=n)
    seniorcitizen.fit(X_train, y_train)
    y_pred=seniorcitizen.predict(X_test)
    accuracy.append(metrics.accuracy_score(y_test, y_pred))
plt.title("Effect of n_estimators")
plt.xlabel("n_estimator")
plt.ylabel("accuracy")
plt.plot(estimators, accuracy)
```

Out[116... [<matplotlib.lines.Line2D at 0x7fa40c950490>]



```
# 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
```

```
In [ ]:

In [ ]:
```

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js