DS_Mod1_Data_Exploration

In this session, will learn how to read a .csv file and also how to access individual elements from the "customer_churn" dataset.

---

read.csv()->This function helps you to load a .csv file into r-studio.

```
customer_churn<-read.csv("C:/Users/INTELLIPAAT/Desktop/customer_churn.csv")
```

---

class()-> This function helps you look at the class of the object

```
class(customer_churn)
```

```
## [1] "data.frame"
```

Here we are looking at the class of one individual column ("tenure") of the customer_churn dataframe

```
class(customer_churn$tenure)
```

```
## [1] "integer"
```

Here we are looking at the class of one individual column ("gender") of the customer_churn dataframe

```
class(customer_churn$gender)
```

```
## [1] "factor"
```

Now we will access the entire DeviceProtection column from the customer_churn dataframe.

So, we can access one individual column from a data-frame by using the "$" symbol.

```
customer_churn$DeviceProtection->c_device_protection


head(c_device_protection)
```

```
## [1] No  Yes No  Yes No  Yes
## Levels: No No internet service Yes
```

Now we will access the entire PaymentMethod column from the customer_churn dataframe. Then the head() function will help us to have a look at the first few rows of the dataset.

```
customer_churn$PaymentMethod->c_payment
head(c_payment,10)
```

```
##  [1] Electronic check          Mailed check
##  [3] Mailed check              Bank transfer (automatic)
##  [5] Electronic check          Electronic check
##  [7] Credit card (automatic)   Mailed check
##  [9] Electronic check          Bank transfer (automatic)
## 4 Levels: Bank transfer (automatic) ... Mailed check
```

Now we will take out the gender column separately by using the index. So, will give the column number-2 to extract this column.

```
customer_churn[,2]->c_gender
head(c_gender)
```

```
## [1] Female Male   Male   Male   Female Female
## Levels: Female Male
```

Now we will take out the customerID column separately by using the index. So, we will give the value to be 1 here.

Then use the head() function to have a glance at the first few records.

```
customer_churn[,1]->c_id
head(c_id)
```

```
## [1] 7590-VHVEG 5575-GNVDE 3668-QPYBK 7795-CFOCW 9237-HQITU 9305-CDSKC
## 7043 Levels: 0002-ORFBO 0003-MKNFE 0004-TLHLJ 0011-IGKFF ... 9995-HOTOH
```

Now we will take out the 2nd,6th and 7th column separately using the combine function c(), again leaving the row number empty as we need to see entire set of rows. Then use the head() function.

```
customer_churn[,c(2,6,7)]->c_267
head(c_267)
```

```
##   gender tenure PhoneService
## 1 Female      1           No
```

```
## 2   Male     34          Yes
## 3   Male      2          Yes
## 4   Male     45           No
## 5 Female      2          Yes
## 6 Female      8          Yes
```

Now same way we can use the column names to take out the column separately, leave the row number empty as we need to see entire set of rows. Then use the head() function

```
customer_churn[,"MonthlyCharges"]->c_month
head(c_month)

## [1] 29.85 56.95 53.85 42.30 70.70 99.65
```

Now we will take out a continuous set of columns by their respective column numbers. Then use the head() function.

```
customer_churn[,3:8]->c_38
head(c_38)

##   SeniorCitizen Partner Dependents tenure PhoneService    MultipleLines
## 1             0     Yes         No      1           No No phone service
## 2             0      No         No     34          Yes               No
## 3             0      No         No      2          Yes               No
## 4             0      No         No     45           No No phone service
## 5             0      No         No      2          Yes               No
## 6             0      No         No      8          Yes              Yes
```

Similarly access the rows with their respective row number. here 2nd row is shown.

```
customer_churn[2,]->c_2
head(c_2)

##   customerID gender SeniorCitizen Partner Dependents tenure PhoneService
## 2 5575-GNVDE   Male             0      No         No     34          Yes
##   MultipleLines InternetService OnlineSecurity OnlineBackup
## 2            No             DSL             Yes           No
##   DeviceProtection TechSupport StreamingTV StreamingMovies Contract
## 2              Yes          No          No              No One year
##   PaperlessBilling PaymentMethod MonthlyCharges TotalCharges Churn
## 2               No  Mailed check          56.95       1889.5    No
```

Similarly,we can access the 100th row.

```
customer_churn[100,]->c_100
head(c_100)
```

```
##     customerID gender SeniorCitizen Partner Dependents tenure PhoneService
## 100 4598-XLKNJ Female             1     Yes         No     25          Yes
##     MultipleLines InternetService OnlineSecurity OnlineBackup
## 100            No     Fiber optic             No          Yes
##     DeviceProtection TechSupport StreamingTV StreamingMovies
## 100              Yes          No         Yes             Yes
##          Contract PaperlessBilling    PaymentMethod MonthlyCharges
## 100 Month-to-month              Yes Electronic check           98.5
##     TotalCharges Churn
## 100       2514.5   Yes
```

Here we will use the combine function to access the 1st, 5th and 10th row. Then we will have a look at these using the head() function.

```
customer_churn[c(1,5,10),]->c_1_5_10
head(c_1_5_10)
```

```
##    customerID gender SeniorCitizen Partner Dependents tenure PhoneService
## 1  7590-VHVEG Female             0     Yes         No      1           No
## 5  9237-HQITU Female             0      No         No      2          Yes
## 10 6388-TABGU   Male             0      No        Yes     62          Yes
##       MultipleLines InternetService OnlineSecurity OnlineBackup
## 1  No phone service             DSL             No          Yes
## 5                No     Fiber optic             No           No
## 10               No             DSL            Yes          Yes
##    DeviceProtection TechSupport StreamingTV StreamingMovies       Contract
## 1                No          No          No              No Month-to-month
## 5                No          No          No              No Month-to-month
## 10               No          No          No              No       One year
##    PaperlessBilling            PaymentMethod MonthlyCharges TotalCharges
## 1               Yes         Electronic check          29.85        29.85
## 5               Yes         Electronic check          70.70       151.65
## 10               No Bank transfer (automatic)          56.15      3487.95
##    Churn
## 1     No
## 5    Yes
## 10    No
```

Again, we will take out a continuous set of rows by their respective row numbers. Here We have shown from row number 100 to 200. Then use the head() function.

```
customer_churn[100:200,]->c_100_200
head(c_100_200)
```

```
##     customerID gender SeniorCitizen Partner Dependents tenure PhoneService
## 100 4598-XLKNJ Female             1     Yes         No     25          Yes
## 101 6380-ARCEH   Male             0      No         No      1          Yes
## 102 3679-XASPY Female             0     Yes        Yes      1          Yes
## 103 7123-WQUHX   Male             0      No         No     38          Yes
## 104 5386-THSLQ Female             1     Yes         No     66           No
## 105 3192-NQECA   Male             0     Yes         No     68          Yes
##        MultipleLines InternetService      OnlineSecurity
## 100               No     Fiber optic                  No
## 101               No              No No internet service
## 102               No              No No internet service
## 103              Yes     Fiber optic                  No
## 104 No phone service             DSL                  No
## 105              Yes     Fiber optic                  No
##            OnlineBackup    DeviceProtection         TechSupport
## 100                 Yes                 Yes                  No
## 101 No internet service No internet service No internet service
## 102 No internet service No internet service No internet service
## 103                  No                 Yes                 Yes
## 104                 Yes                 Yes                  No
## 105                 Yes                 Yes                 Yes
##             StreamingTV     StreamingMovies       Contract
## 100                 Yes                 Yes Month-to-month
## 101 No internet service No internet service Month-to-month
## 102 No internet service No internet service Month-to-month
## 103                 Yes                  No       One year
## 104                 Yes                  No       One year
## 105                 Yes                 Yes       Two year
##     PaperlessBilling             PaymentMethod MonthlyCharges TotalCharges
## 100              Yes          Electronic check          98.50      2514.50
## 101               No              Mailed check          20.20        20.20
## 102               No          Electronic check          19.45        19.45
## 103               No Bank transfer (automatic)          95.00      3605.60
## 104               No Bank transfer (automatic)          45.55      3027.25
## 105              Yes Bank transfer (automatic)         110.00      7611.85
##     Churn
## 100   Yes
## 101    No
## 102    No
## 103    No
## 104    No
## 105   Yes
```

Similarly We have taken out a chunk of rows from row number 5000 to 7000. Then use the head() function.

```
customer_churn[5000:7000,]->c_50007000
head(c_50007000,3)
```

```
##      customerID gender SeniorCitizen Partner Dependents tenure
## 5000 1699-TLDLZ Female             0     Yes        Yes     16
## 5001 5600-PDUJF   Male             0      No         No      6
## 5002 8292-TYSPY   Male             0      No         No     19
##      PhoneService MultipleLines InternetService      OnlineSecurity
## 5000          Yes            No               No No internet service
## 5001          Yes            No              DSL                  No
## 5002          Yes            No              DSL                  No
##             OnlineBackup    DeviceProtection         TechSupport
## 5000 No internet service No internet service No internet service
## 5001                  No                  No                 Yes
## 5002                  No                 Yes                 Yes
##              StreamingTV     StreamingMovies       Contract
## 5000 No internet service No internet service       Two year
## 5001                  No                  No Month-to-month
## 5002                  No                  No Month-to-month
##      PaperlessBilling           PaymentMethod MonthlyCharges TotalCharges
## 5000               No            Mailed check           19.7       301.55
## 5001              Yes Credit card (automatic)           49.5       312.70
## 5002              Yes Credit card (automatic)           55.0      1046.50
##      Churn
## 5000    No
## 5001    No
## 5002   Yes
```

Now we will take the out rows from 50 to 60 but just the 2nd and 3rd column.

```
customer_churn[50:60,c(2,3)]->c_random1
head(c_random1)
```

```
##    gender SeniorCitizen
## 50 Female             0
## 51 Female             1
## 52 Female             0
## 53 Female             1
## 54 Female             1
## 55 Female             1
```

Here we will access the rows from 100 to 200 and 1000:2000. And we will also access columns from 5 to 8.

```
customer_churn[c(100:200,1000:2000),5:8]->c_random2
head(c_random2)

##     Dependents tenure PhoneService    MultipleLines
## 100         No     25          Yes               No
## 101         No      1          Yes               No
## 102        Yes      1          Yes               No
## 103         No     38          Yes              Yes
## 104         No     66           No No phone service
## 105         No     68          Yes              Yes
```