

Data Science Lab:Week 2

Find the below data set and perform the following operations:-

Dataset name: -Telco Customer Churn

```
import pandas as pd

df = pd.read_csv('WA_Fn-UseC_-Telco-Customer-Churn.csv')
df.head(5)
```

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	Mult
0	7590-VHVEG	Female	0	Yes	No	1	No	
1	5575-GNVDE	Male	0	No	No	34	Yes	
2	3668-QPYBK	Male	0	No	No	2	Yes	
3	7795-CFOCW	Male	0	No	No	45	No	
4	9237-HQITU	Female	0	No	No	2	Yes	

5 rows × 21 columns

1. Find the no. of duplicate records in the churn dataframe based on the customerID column.

```
duplicate_count = df.duplicated(subset=['customerID']).sum()

print("Number of duplicate records based on customerID column:", duplicate_count)
```

Number of duplicate records based on customerID column: 0

2. In the churn dataframe, what are the total no. of missing values for the variable TotalCharges?

```
missing_values_count = df['TotalCharges'].isnull().sum()
print("Total number of missing values for 'TotalCharges' variable:", missing_values_count)
```

Total number of missing values for 'TotalCharges' variable: 0

3. From the churn dataframe, what is the average monthly charge paid by a customer for the services he/she has signed up for?

```
import pandas as pd

average_monthly_charge = df['MonthlyCharges'].mean()
print("Average monthly charge paid by a customer for the services :", average_monthly_charge)
```

Average monthly charge paid by a customer for the services : 64.76169246059918

▼ **4. In the churn dataframe, under the variable Dependents how many records have "1@#" ?**

```
count = (df['Dependents'] == "1@#").sum()  
print("Number of records with 'Dependents' equal to '1@#':", count)
```

```
Number of records with 'Dependents' equal to '1@#': 0
```

▼ **5. Find the data type of the variable tenure from the churn dataframe.**

```
tenure_data_type = df['tenure'].dtype  
print("Data type of 'tenure' variable:", tenure_data_type)
```

```
Data type of 'tenure' variable: int64
```