Data Science Lab:Week 2

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

Dataset name: -Telco Customer Churn

```python
import pandas as pd

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

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

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

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

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

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

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

In [6]:

```
import pandas as pd

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

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)
```

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)
```