Import Required Packages

In [1]:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [2]:

```
net_usage = pd.read_csv('internet_session.csv', parse_dates=['start_time'])
```

In [8]:

```
net_usage
```

Out[8]:

	name	start_time	usage_time	IP	MAC	upload	download	tot
0	user1	2022-05-10 02:59:32	00:00:36:28	10.55.14.222	48:E7:DA:58:22:E9	15861.76	333168.64	
1	user1	2022-05-10 18:53:27	00:01:49:56	10.55.2.253	48:E7:DA:58:22:E9	16957.44	212152.32	
2	user1	2022-05-10 21:20:44	00:01:35:00	10.55.2.253	48:E7:DA:58:22:E9	14080.0	195153.92	
3	user1	2022-05-11 00:37:42	00:00:26:00	10.55.2.253	48:E7:DA:58:22:E9	5242.88	40806.4	
4	user1	2022-05-11 02:59:38	00:00:11:52	10.55.2.253	48:E7:DA:58:22:E9	22067.2	10772.48	
...	...	...	...	...	...	...	...	
4707	user9	2022-11-04 01:11:34	00:06:54:32	10.55.4.189	DA:2F:97:0E:B7:D0	107960.32	2390753.28	
4708	user9	2022-11-04 10:26:09	00:00:23:49	10.55.4.59	DA:2F:97:0E:B7:D0	11407.36	209674.24	
4709	user9	2022-11-04 20:41:42	00:01:24:13	10.55.15.186	DA:2F:97:0E:B7:D0	18995.2	373657.6	
4710	user9	2022-11-05 00:21:06	00:08:49:43	10.55.4.159	DA:2F:97:0E:B7:D0	46602.24	593766.4	
4711	user9	2022-11-05 20:55:37	00:01:06:20	10.55.2.33	DA:2F:97:0E:B7:D0	21237.76	298536.96	

4712 rows × 9 columns

```
Q1:What is the most frequent internet activity time of the day ?
```

In [3]:

```
print('The maximum usage time per user:')
max_usage = net_usage.groupby('name').usage_time.max()
max_usage
```

```
The maximum usage time per user:
```

Out[3]:

```
name
user1    00:19:35:11
user2    00:20:39:52
user3    00:17:01:28
user4    01:00:21:07
user5    00:06:36:11
user6    00:19:35:11
user7    00:22:00:07
user8    00:17:24:26
user9    00:19:26:09
Name: usage_time, dtype: object
```

In [ ]:

```
# Obervation on Maximum Usage -->User7-->22 hrs 7 Seconds, User2-->20 hrs 39 Minutes and 52
```

```
Q2:How often the ip changes ?
```

In [9]:

```
net_usage.IP.value_counts()
```

Out[9]:

```
10.55.0.89      80
10.55.14.148    64
10.55.15.221    55
10.55.1.50      48
10.55.10.46     44
                ..
10.55.14.67      1
10.55.7.44       1
10.55.12.225     1
10.55.12.190     1
10.55.2.33       1
Name: IP, Length: 1302, dtype: int64
```

In [10]:

```
base_ip = '48:E7:DA:58:22:E9'
ip_count = 0
for i in range(1, net_usage.shape[0]):
    if net_usage.iloc[i]['IP'] != base_ip:
        ip_count +=1
        base_ip = net_usage.iloc[i]['IP']

print('The IP Adress changed ' + str(ip_count) + ' times')
```

```
The IP Adress changed 2308 times
```

Observation: The IP addressed Changes 2308 Times in the Usage.

Q3A:How often the device changed.

In [14]:

```
device = []
basename = 'device'
mac = net_usage['MAC'][0]
device_number = 1
for i in net_usage['MAC']:
    if i == mac:
        device.append(basename + str(device_number))
    else:
        device_number += 1
        device.append(basename + str(device_number))
        mac = i
net_usage['device'] = device
```

In [15]:

```
base_device = 'device1'
device_count = 0
for i in range(1, net_usage.shape[0]):
    if net_usage.iloc[i]['device'] != base_device:
        device_count +=1
        base_device = net_usage.iloc[i]['device']

print('The device changed ' + str(device_count) + ' times')
```

```
The device changed 1226 times
```

Observation: Device Got Changes 1226 Times

Q3B:What is the average usage per hour ?

In [16]:

```
net_usage.reset_index(inplace=True)
net_usage['hour'] = pd.to_datetime(net_usage['start_time']).dt.hour
net_usage['day'] = net_usage['start_time'].dt.day
net_usage['month'] = net_usage['start_time'].dt.month

hourly_average = net_usage.groupby('hour').total_transfer.mean()
print('The Average usage per hour is:\n ' +  str(round(hourly_average, 2)))
```

```
The Average usage per hour is:
 hour
0     464530.44
1     530880.86
2     431576.11
3     345303.34
4     359809.44
5     275960.91
6     468959.59
7     292886.83
8     366681.92
9     377480.64
10    393259.12
11    309492.45
12    310137.98
13    335270.58
14    472403.71
15    517005.11
16    403919.40
17    525423.69
18    665414.45
19    390839.43
20    355740.06
21    471461.40
22    449600.50
23    407785.08
Name: total_transfer, dtype: float64
```

Observation: The Maximum Usage is on 18th Hour(Actually 19th as Index starts from 0)

Q3C:What is the Average Usage per day ?

In [17]:

```
daily_average = net_usage.groupby('day').total_transfer.mean()
print('The Average usage per day is:\n ' +  str(round(daily_average, 2)))
```

```
The Average usage per day is:
 day
1     396705.04
2     494496.48
3     445865.63
4     676332.03
5     634564.05
6     396261.75
7     402259.89
8     301859.57
9     393521.97
10    350665.02
11    729857.65
12    346695.95
13    501906.70
14    352701.10
15    521520.51
16    426719.39
17    475795.71
18    337490.93
19    301941.32
20    365130.12
21    462211.69
22    486595.37
23    383153.93
24    320598.94
25    443689.47
26    463432.02
27    324318.12
28    494576.34
29    363645.61
30    361418.88
31    369118.01
Name: total_transfer, dtype: float64
```

```
The Observation is 11th Day the Maximum Usage Happend
```

```
What is the Average Usage and per month ?
```

In [18]:

```
monthly_average = net_usage.groupby('month').total_transfer.mean()
print('The Average usage per month is:\n ' +  str(round(monthly_average, 2)))
```

```
The Average usage per month is:
 month
5     311177.16
6     338418.08
7     418583.99
8     479042.44
9     482955.52
10    549467.63
11    399675.45
Name: total_transfer, dtype: float64
```

In [ ]:

```
# The Average Usage per Month Max is on 10th Month(Actually 11th as Index starts from 0)
```