

```
In [1]: import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline
```

- PowerCurve.csv is a worksheet having the device IDs of daily active users of each day of a month.
- Active\_Users.csv is a worksheet having the weekly active users' (WAU) that contains all device IDs that were active in a particular week.

```
In [2]: '''
excel1 = pd.read_excel("Active Users (1) (2).xlsx")
excel1.to_csv("Active_Users.csv", encoding='utf-8', index=False)

excel2 = pd.read_excel("PowerCurve-DAU_Device_ids (2) (1).xlsx")
excel2.to_csv("PowerCurve.csv", encoding='utf-8', index=False)
'''
```

```
Out[2]: '\nexcel1 = pd.read_excel("Active Users (1) (2).xlsx")\nexcel1.to_csv("Active_Users.csv", encoding=
\'utf-8\', index=False)\n\nexcel2 = pd.read_excel("PowerCurve-DAU_Device_ids (2) (1).xlsx")\nexcel2.to_
o_csv("PowerCurve.csv", encoding=\'utf-8\', index=False)\n'
```

```
In [3]: df_activeU = pd.read_csv("Active_Users.csv")
df_activeU.head()
```

```
df_activeU = pd.read_csv("Active_Users.csv")
df_activeU.head()
```

Out[3]:

	w1	w2	w3	w4	w5	w6	w7
0	fd7c28f9fd8045f2	fd7c28f9fd8045f2	fd7c28f9fd8045f2	fd7c28f9fd8045f2	fd7c28f9fd8045f2	231d91be38352d7a	306243851b716bf6
1	54910d2b363221e1	520443b0b8128202	a4bce0d054266d68	a4bce0d054266d68	c0bb01dbe2b2de0f	530104d139ed9d29f	d1afc6d7c4661d7e
2	520443b0b8128202	a4bce0d054266d68	7b042fcc54a45882	d98da6eaa4bb452f	7b042fcc54a45882	a4bce0d054266d68	3792a1c9395e3e2a
3	a4bce0d054266d68	d1afc6d7c4661d7e	aed9597fc6984d64	7b042fcc54a45882	c885d9690e13074	d1afc6d7c4661d7e	7b042fcc54a45882
4	3792a1c9395e3e2a	7b042fcc54a45882	407d67f50877e6f9	aed9597fc6984d64	aed9597fc6984d64	3792a1c9395e3e2a	a455b3d389d7d6a3b

In [4]:

```
df_activeU.shape
```

Out[4]:

```
(4014, 56)
```

In [5]:

```
df_activeU.isnull().sum()
```

Out[5]:

w1	2255
w2	2360
w3	2282
w4	1898
w5	1821
w6	1857
w7	1463
w8	1139
w9	1219
w10	1193
w11	1268
w12	1108
w13	934
w14	963
w15	1061
w16	948
w17	969
w18	915
w19	946
w20	912
w21	1023
w22	767
w23	820
w24	905
w25	913
w26	827
w27	829
w28	939
w29	1169
w30	1329
w31	1362
w32	1457
w33	1485
w34	1080
w35	1127
w36	833
w37	915
w38	613
w39	277
w40	268
w41	0
w42	566
w43	865
w44	796
w45	923
w46	421
w47	407
w48	192
w49	151
w50	260
w51	213
w52	246
w53	273
w54	105

```
In [6]: df_PCurve = pd.read_csv("PowerCurve.csv")
df_PCurve.head()
```

```
In [6]: df_PCurve = pd.read_csv("PowerCurve.csv")
df_PCurve.head()
```

```
Out[6]:
```

	2017-10-01	2017-10-02	2017-10-03	2017-10-04	2017-10-05	2017-10-06	2017-10-07
0	285020521074059	285020521074059	285020521074059	285020521074059	285020521074059	285020521074059	285020521074059
1	80964841145060060878848	001ee3d4fd0d1393	00a283e65135076e	001ee3d4fd0d1393	001ee3d4fd0d1393	00183d9b218571be	00183d9b218571be
2	001ee3d4fd0d1393	003e88d939e8e08c	00c7e013b900e087	002537a9182f4af0	0058339e87a02771	001ee3d4fd0d1393	001ee3d4fd0d1393
3	003e88d939e8e08c	0058339e87a02771	00d176adc3530e7a	007068741f1bb611	005f81093b87079e	005f81093b87079e	009197e411270447
4	00d176adc3530e7a	005a39201eca94a6	00ee8ca904293a0b	009197e411270447	00c5c32a8b4cf470	0083197a98ac8d3b	00c7e013b900e087

```
In [7]: df_PCurve.shape
```

```
Out[7]: (9551, 31)
```

```
In [8]: df_PCurve.isnull().sum()
```

```
In [8]:
```

## Power User Curve

```
In [9]: PCurve_col = df_PCurve.columns
```

```
In [10]: PCurve_col_val = []
size = df_PCurve.shape[0]

for _ in PCurve_col:
    no_notNull = size - df_PCurve[_].isnull().sum()
    PCurve_col_val.append(no_notNull)
```

```
In [11]: growthR_val = []
growth = []

for i in range(len(PCurve_col_val)):
    if i == (len(PCurve_col_val)-1):
        break
    else:
        diff = PCurve_col_val[i+1]-PCurve_col_val[i]
        percent = (diff/PCurve_col_val[i])
        growth.append(diff)
        growthR_val.append(percent*100)

print("Average growth rate: ",round(np.mean(growthR_val),2),"%")
```

Average growth rate: 3.07 %

```
In [12]: uniqueU = []
l = []

for _ in PCurve_col:
    l.append(list(df_PCurve[_]))
    for i in l:
        for k in i:
            uniqueU.append(k)

uniqueU = set(uniqueU)
print("Number of Unique users during the period of 31 days: ",len(uniqueU))

Number of Unique users during the period of 31 days: 57302
```

```
In [13]: print("Average number of active users during the period of 31 days: ",
            round(np.mean(PCurve_col_val),0))
```

Average number of active users during the period of 31 days: 6879.0

```
In [28]: new_growth = [0,992,-162, 200, 190, -368, 730, 94, -52, 21, -177, 419, -416, 853, 179, 387,
                    51, 113, 406,-1271, 1629, -183, -59, 377, -433, -1099, 44, 1669, 475, 397, 181]

new_users = [0,992,0, 200, 190, 0, 730, 94, 0, 21, 0, 419, 0, 853, 179, 387,
             51, 113, 406,0, 1629, 0, 0, 377, 0, 0, 44, 1669, 475, 397, 181]

churned = [0,0,-162, 0, 0, -368, 0, 0, -52, 0, -177, 0, -416, 0, 0, 0,
           0, 0, 0,-1271, 0, -183, -59, 0, -433, -1099, 0, 0, 0, 0]
```

retentionR = np.array(PCurve\_col\_val) + np.array(churned) - np.array(new\_users)

fig = plt.figure(figsize = (14, 8))

plt.bar(PCurve\_col, PCurve\_col\_val, label="Resurrected Users")

plt.bar(PCurve\_col,new\_users, label="New Users" )

plt.bar(PCurve\_col,churned, label="Churned Users")

plt.plot(retentionR, label="Retention rate",color="r",linewidth=1.5)

plt.xlabel("Days")

plt.xticks(ticks=activeU\_col)

plt.xticks(rotation=80)

plt.ylabel("Number of active users")

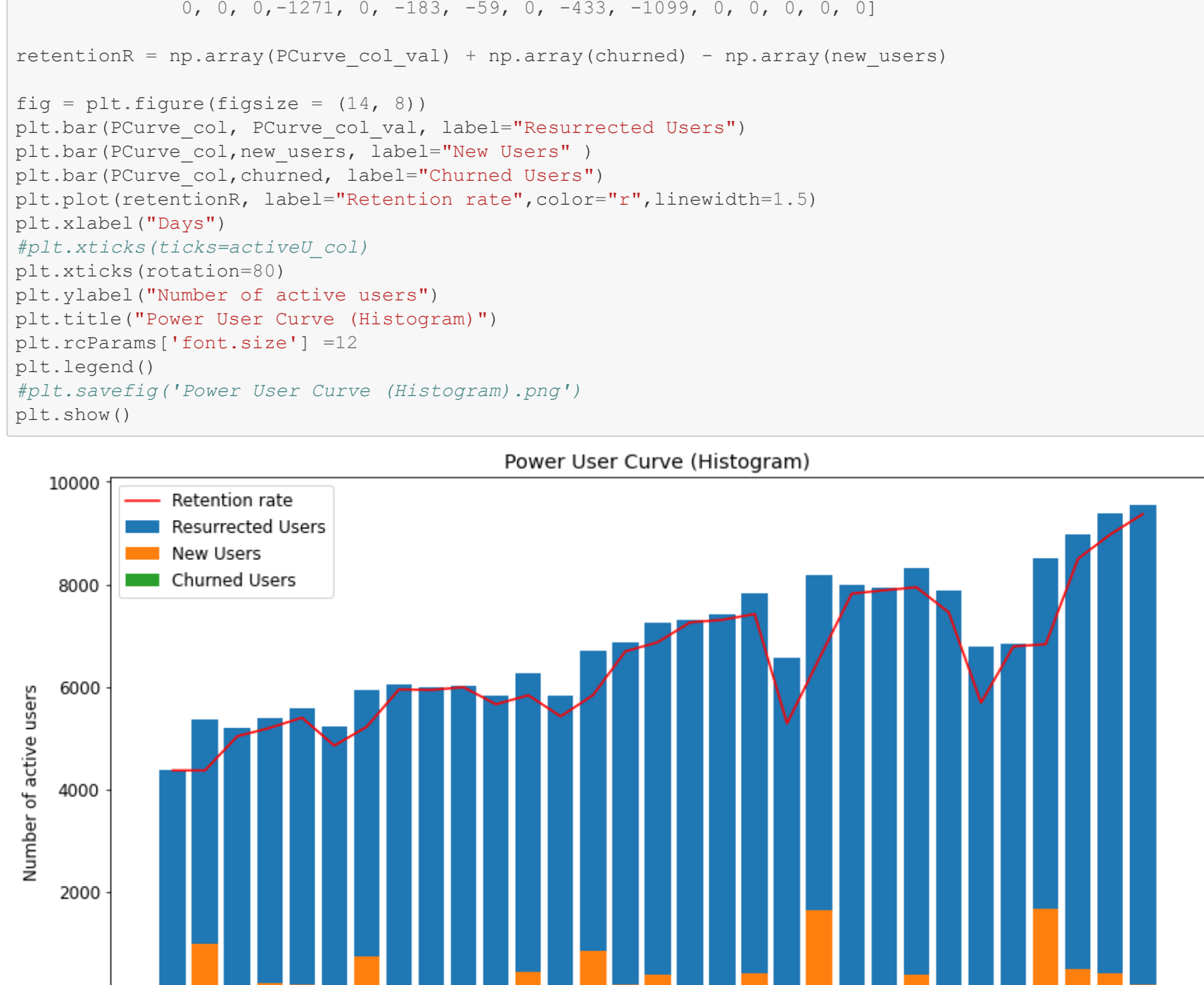
plt.title("Power User Curve (Histogram)")

plt.rcParams['font.size'] =12

plt.legend()

plt.savefig('Power User Curve (Histogram).png')

plt.show()



```
In [15]: print("Average number of new users: ", round((np.mean(new_users)),0))
print("Average number of churned users: ", round((np.mean(churned)),0))
```

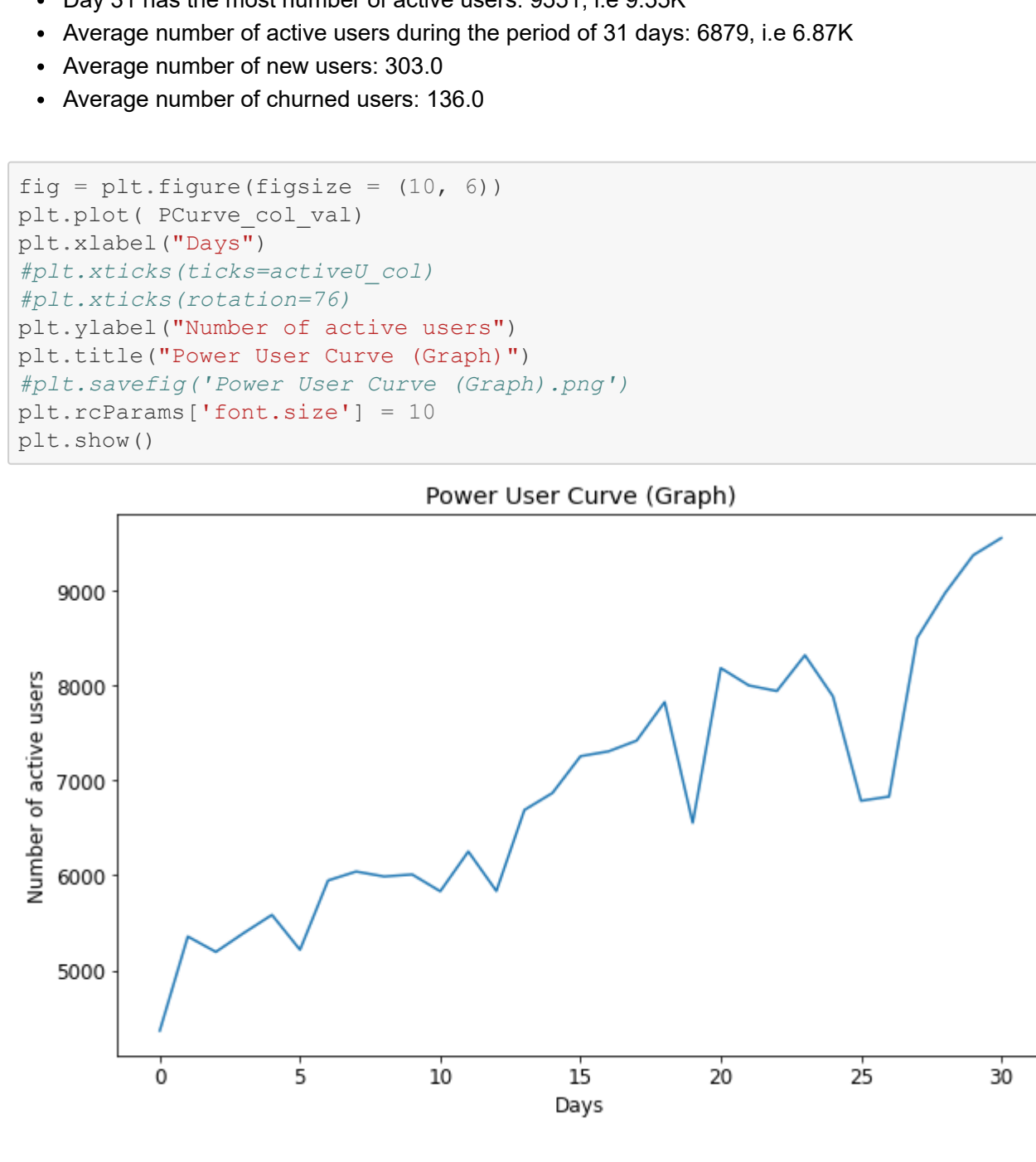
Average number of new users: 303.0

Average number of churned users: -136.0

## Insights from the above plot

- Total number of Unique users during the period of 31 days: 57302, i.e 57.30K
- From day 1 to day 31, number of active users grow almost continuously
- Average growth rate: 3.07 %
- Day 1 has the least number of active users: 4364, i.e 4.36K
- Day 31 has the most number of active users: 9551, i.e 9.55K
- Average number of active users during the period of 31 days: 6879, i.e 6.87K
- Average number of new users: 303.0
- Average number of churned users: 136.0

```
In [50]: fig = plt.figure(figsize = (10, 6))
plt.plot(PCurve_col_val)
plt.xlabel("Days")
plt.xticks(ticks=activeU_col)
plt.xticks(rotation=76)
plt.ylabel("Number of active users")
plt.title("Power User Curve (Graph)")
plt.savefig('Power User Curve (Graph).png')
plt.rcParams['font.size'] = 10
plt.show()
```



## Active Users distribution among Day of Week

```
In [17]: from datetime import date
import calendar
```

```
In [31]: df_timestamp = pd.DataFrame()
```

```
In [32]: df_timestamp['Timestamp'] = pd.to_datetime(df_PCurve.columns)
```

```
In [33]: df_timestamp['Day of Week'] = df_timestamp['Timestamp'].dt.day_name()
df_timestamp['Active Users'] = PCurve_col_val
```

```
In [34]: days = ['Tuesday','Monday','Sunday','Wednesday','Saturday','Thursday','Friday']
activeU_on_a_day = []

for _ in days:
    df = df_timestamp[df_timestamp['Day of Week'] == _]
    total = df['Active Users'].sum(axis=0)
    activeU_on_a_day.append(total)
```

```
In [35]: fig = plt.figure(figsize = (0.1,0.02,0.02,0.02,0.02,0.08))
```

explode=activeU\_on\_a\_day, labels = days, autopct='%1.1f%%', explode=explode, shadow=True )

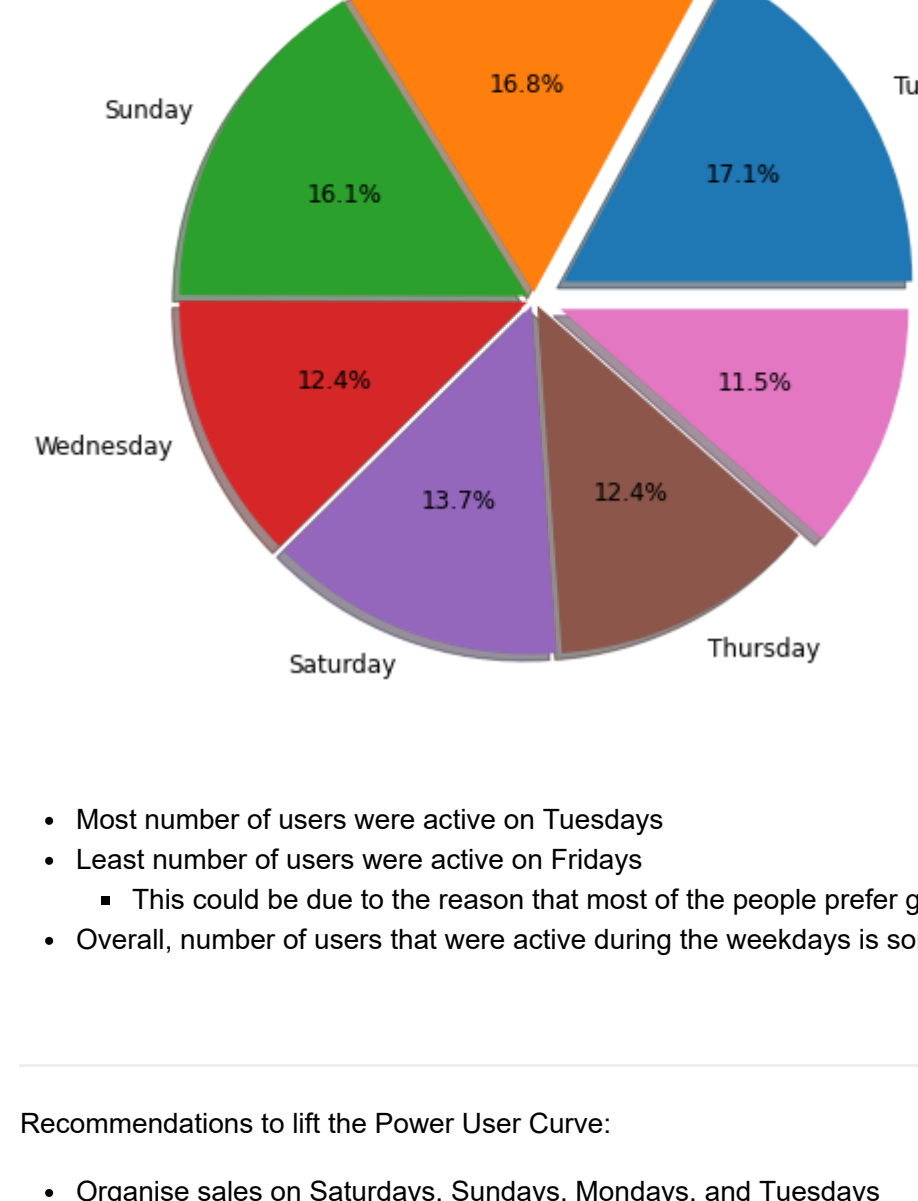
plt.title("Active Users distribution among Day of Week")

plt.savefig('Active Users distribution among Day of Week.png')

plt.rcParams['font.size'] = 16

plt.show()

## Active Users distribution among Day of Week



- Most number of users were active on Tuesdays
- Least number of users were active on Fridays
  - This could be due to the reason that most of the people prefer going out for parties and fun on friday nights
- Overall, number of users that were active during the weekdays is somewhat similar

## Recommendations to lift the Power User Curve:

- Organise sales on Saturdays, Sundays, Mondays, and Tuesdays
  - Because from Saturday to Tuesday number of active users grow consistently
- Put on hoarding advertisements during the weekends when people prefer going out
  - So that they could be attracted to use the Mylo app

```
In [35]:
```

## Weekly Growth Accounting

```
In [36]: activeU_col = df_activeU.columns
```

```
In [37]: activeU_col_val = []
size = df_activeU.shape[0]

for _ in activeU_col:
    no_notNull = size - df_activeU[_].isnull().sum()
    activeU_col_val.append(no_notNull)
```

```
In [37]:
```

```
In [38]: uniqueU = []
l = []

for _ in activeU_col:
    l.append(list(df_activeU[_]))
    for i in l:
        for k in i:
            uniqueU.append(k)

uniqueU = set(uniqueU)
print("Total number of Unique users during the period of 51 weeks: ",len(uniqueU))
```

Total number of Unique users during the period of 51 weeks: 31825

```
In [39]: print("Average number of active users during the period of 56 weeks: ",round(np.mean(activeU_col_val),0
))
```

Average number of active users during the period of 56 weeks: 3078.0

```
In [40]: growthR_val = []
growth = []

for i in range(len(activeU_col_val)):
    if i == (len(activeU_col_val)-1):
        break
    else:
        diff = activeU_col_val[i+1]-activeU_col_val[i]
        percent = (diff/activeU_col_val[i])
        growth.append(diff)
        growthR_val.append(percent*100)

print("Average growth rate: ",round(np.mean(growthR_val),2),"%")
```

Average growth rate: 1.58 %

```
In [49]: new_growth = [0,-105, 78, 384, 77, -36, 394, 324, -80, 26, -75, 160, 174, -29, -98, 113, -21, 54, -31,
                    34, -111, 256, -53, -85, -8, 86, -2, -110, -230, -160, -33, -95, -28, 405, -47, 294, -82,
                    302, 336, 9, 268, 0, 0, 69, 0, 492, 24, 215, 41, -109, 47, -33, -27, 168, -103,
                    -110]

new_users = [0, 0, 78, 384, 77, 0, 394, 324, 0, 26, 0, 160, 174, 0, 0, 113, 0, 54, 0,
             34, 0, 256, 0, 0, 0, 86, 0, 0, 0, 0, 0, 0, 405, 0, 294, 0,
             302, 336, 9, 268, 0, 0, 69, 0, 492, 24, 215, 41, 0, 47, 0, 0, 168, 0,
             0]
```

churned = [0,-105, 0, 0, 0, -36, 0, 0, -80, 0, -75, 0, 0, 0, -29, -98, 0, -21, 0, -31,
 0, -111, 0, -53, -85, -8, 0, -2, -110, -230, -160, -33, -95, -28, 0, -47, 0, -82,
 0, 0, 0, 0, -566, -299, 0, -127, 0, 0, 0, 0, -109, 0, -33, -27, 0, -103,
 -110]

retentionR = np.array(activeU\_col\_val) + np.array(churned) - np.array(new\_users)

fig = plt.figure(figsize = (18, 12))

plt.bar(activeU\_col, activeU\_col\_val, label="Resurrected Users")

plt.bar(activeU\_col, new\_users, label="New Users")

plt.bar(activeU\_col, churned, label="Churned Users")

plt.plot(retentionR, label="Retention rate",color="r",linewidth=1.5 )

plt.xlabel("Weeks")

plt.xticks(ticks=activeU\_col)

plt.xticks(rotation=60)

plt.ylabel("Number of active users")

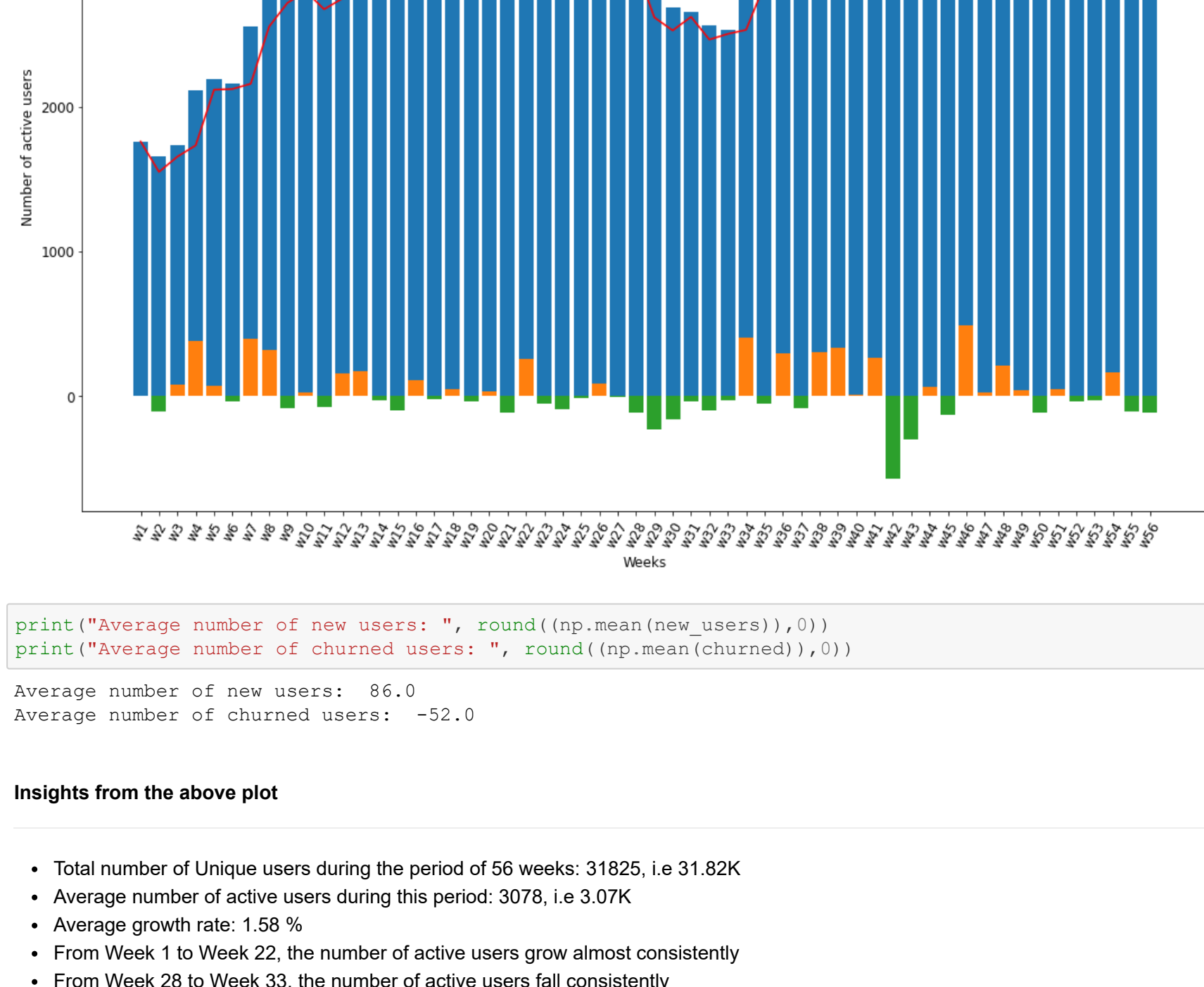
plt.title("Weekly Growth Accounting Chart (Histogram)")

plt.rcParams['font.size'] = 12

plt.legend()

plt.savefig('Weekly Growth Accounting Chart (Histogram).png')

plt.show()



```
In [42]: print("Average number of new users: ", round(np.mean(new_users),0))
print("Average number of churned users: ", round(np.mean(churned),0))
```

Average number of new users: 86.0

Average number of churned users: -52.0

## Insights from the above plot

- Total number of Unique users during the period of 56 weeks: 31825, i.e 31.82K
- Average number of active users during this period: 3078, i.e 3.07K
- Average growth rate: 1.58 %
- From Week 1 to Week 22, the number of active users grow almost consistently
- From Week 28 to Week 33, the number of active users fall consistently
- Week 2 has the least number of active users: 1654, i.e 1.65K
- Week 41 has the most number of active users: 4014, i.e 4.01K
- Average number of new users: 86.0
- Average number of churned users: 52.0

```
In [51]: fig = plt.figure(figsize = (10, 6))
plt.plot(activeU_col_val)
plt.xlabel("Weeks")
plt.xticks(ticks=activeU_col)
plt.xticks(rotation=60)
plt.ylabel("Number of active users")
plt.title("Weekly Growth Accounting Chart (Graph)")
plt.savefig('Weekly Growth Accounting Chart (Graph).png')
plt.rcParams['font.size'] = 10
plt.show()
```

