Amazon-orders Project

```
AIM:
1.How much I spent on Amazon during this period.
2.What my highest, lowest, and average order totals were.
3.How much tax I paid, and the effective sales tax rate I paid.
4.How my spending fluctuated over time.
5.On which days I spent the most money.
BY: RUDRA PRATAP PADHI
```

In [1]:
```
import pandas as pd
%matplotlib inline
```

In [2]:
```
df = pd.read_csv(r'C:\Users\Rudra Pratap\Desktop\CSV\amazon-orders.csv')
```

In [3]:
```
df.head()
```

Out[3]:

	Order Date	Order ID	Payment Instrrument Type	Website	Purchase Order Number	Ordering CustoOdishar Email	Shipment Date	Shipping Address Name	Shipping Address Street 1	Shipping Address Street 2	...	Order Status	C T I
0	04-01-2019	112-0000000-0000001	Mastercard - 0001	Amazon.com	NaN	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	NaN	...	Shipped	OSP
1	04-01-2019	112-0000000-0000002	Mastercard - 0001	Amazon.com	NaN	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	NaN	...	Shipped	OSP
2	04-01-2019	112-0000000-0000003	Mastercard - 0001	Amazon.com	NaN	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	NaN	...	Shipped	OSP
3	04-04-2019	112-0000000-0000004	Mastercard - 0001	Amazon.com	NaN	rud**********@gmail.com	04-05-2019	Cutm,BBSR	Cutm,Jatni,BBSR	NaN	...	Shipped	OSP
4	04-05-2019	112-0000000-0000005	Mastercard - 0001	Amazon.com	NaN	rud**********@gmail.com	04-07-2019	Cutm,BBSR	Cutm,Jatni,BBSR	NaN	...	Shipped	OSP

5 rows × 23 columns

In [4]:
```
df.shape
```

Out[4]:
```
(59, 23)
```

In [5]:
```
df = df.fillna(0)
df.head()
```

Out[5]:

	Order Date	Order ID	Payment Instrrument Type	Website	Purchase Order Number	Ordering CustoOdishar Email	Shipment Date	Shipping Address Name	Shipping Address Street 1	Shipping Address Street 2	...	Order Status	C T I
0	04-01-2019	112-0000000-0000001	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
1	04-01-2019	112-0000000-0000002	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
2	04-01-2019	112-0000000-0000003	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
3	04-04-2019	112-0000000-0000004	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-05-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
4	04-05-2019	112-0000000-0000005	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-07-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP

5 rows × 23 columns

In [6]:
```
df["Total Charged in Rupees"] = df["Total Charged in Rupees"].astype(float)
df.head()
```

Out[6]:

	Order Date	Order ID	Payment Instrrument Type	Website	Purchase Order Number	Ordering CustoOdishar Email	Shipment Date	Shipping Address Name	Shipping Address Street 1	Shipping Address Street 2	...	Order Status	C T I
0	04-01-2019	112-0000000-0000001	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
1	04-01-2019	112-0000000-0000002	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
2	04-01-2019	112-0000000-0000003	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
3	04-04-2019	112-0000000-0000004	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-05-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
4	04-05-2019	112-0000000-0000005	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-07-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP

5 rows × 23 columns

In [7]:
```
df["Total Charged in Rupees"].sum()
```

Out[7]:
```
134313.0
```

In [8]:
```
df["Total Charged in Rupees"].mean()
```

Out[8]:
```
2276.491525423729
```

In [9]:
```
df["Total Charged in Rupees"].median()
```

Out[9]:
```
1124.25
```

In [10]:
```
df["Total Charged in Rupees"].max()
```

Out[10]:
```
15824.25
```

In [11]:
```
df["Total Charged in Rupees"].min()
```

Out[11]:
```
39.0
```

In [12]:
```
df["Tax Charged in Rupees"].sum()
```

Out[12]:
```
3945.75
```

In [13]:
```
df["Tax Charged in Rupees"].sum() / df["Total Charged in Rupees"].sum()
```

Out[13]:
```
0.02937727546849523
```

In [14]:
```
df['Order Date'] = pd.to_datetime(df['Order Date'])
df.head()
```

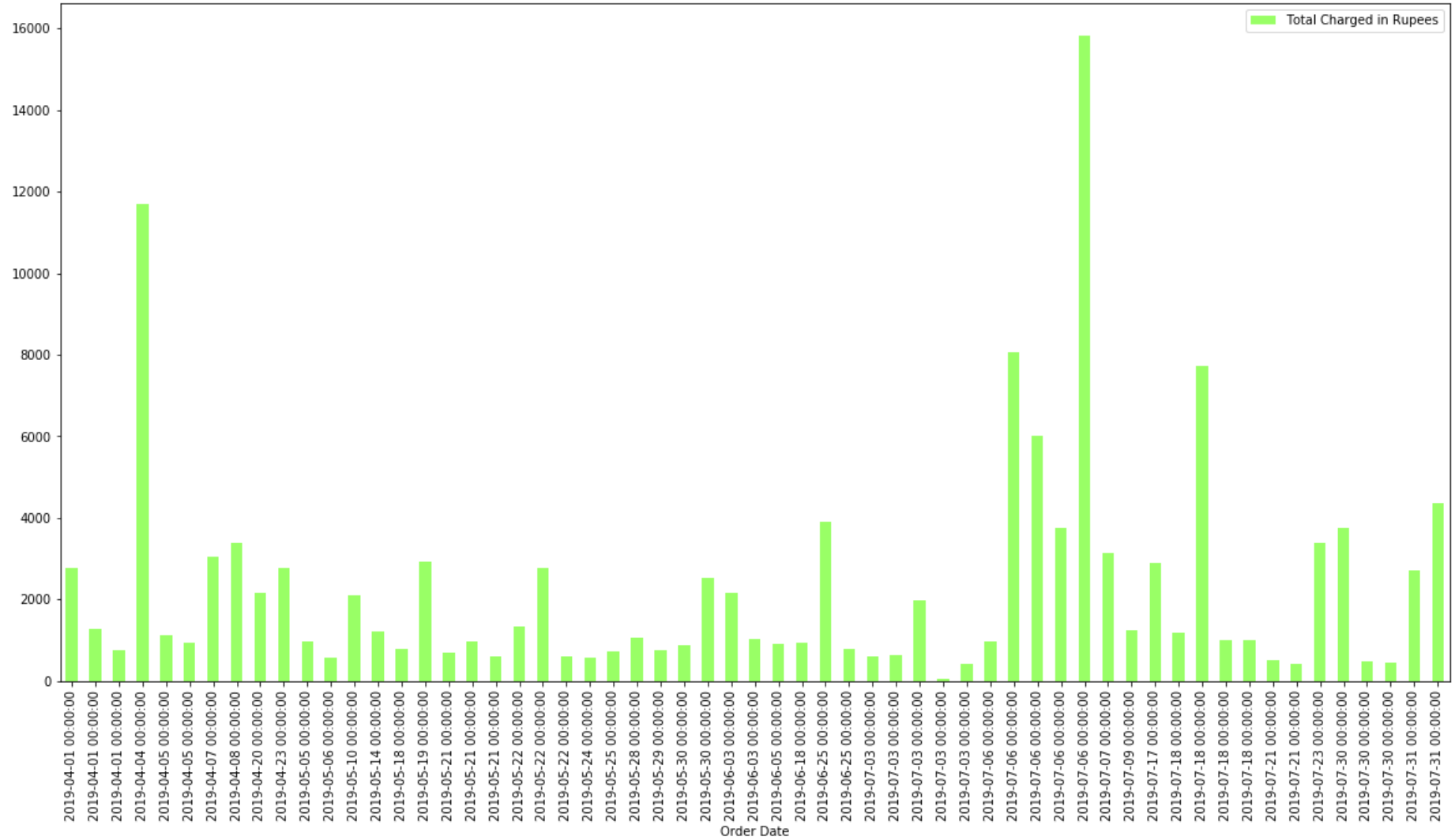
Out[14]:

	Order Date	Order ID	Payment Instrrument Type	Website	Purchase Order Number	Ordering CustoOdishar Email	Shipment Date	Shipping Address Name	Shipping Address Street 1	Shipping Address Street 2	...	Order Status	C T I
0	2019-04-01	112-0000000-0000001	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
1	2019-04-01	112-0000000-0000002	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
2	2019-04-01	112-0000000-0000003	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-01-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
3	2019-04-04	112-0000000-0000004	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-05-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP
4	2019-04-05	112-0000000-0000005	Mastercard - 0001	Amazon.com	0.0	rud**********@gmail.com	04-07-2019	Cutm,BBSR	Cutm,Jatni,BBSR	0.0	...	Shipped	OSP

5 rows × 23 columns

In [15]:
```
df.plot.bar(x='Order Date',y='Total Charged in Rupees',rot=90,figsize=(20,10),color='#99ff66')
```

Out[15]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x2557207a048>
```



In [16]:
```
daily_orders = df.groupby('Order Date').sum()["Total Charged in Rupees"]
daily_orders.head(56)
```

Out[16]:
```
Order Date
2019-04-01     4793.25
2019-04-04    11709.00
2019-04-05     2073.00
2019-04-07     3047.25
2019-04-08     3374.25
2019-04-20     2160.00
2019-04-23     2781.00
2019-05-05      974.25
2019-05-06      570.00
2019-05-10     2099.25
2019-05-14     1212.00
2019-05-18      787.50
2019-05-19     2916.75
2019-05-21     2281.50
2019-05-22     4705.50
2019-05-24      575.25
2019-05-25      735.00
2019-05-28     1046.25
2019-05-29      749.25
2019-05-30     3403.50
2019-06-03     3215.25
2019-06-05      919.50
2019-06-18      948.00
2019-06-25     4693.50
2019-07-03     3651.00
2019-07-06    34602.00
2019-07-07     3147.75
2019-07-09     1234.50
2019-07-17     2887.50
2019-07-18    10941.00
2019-07-21      945.75
2019-07-23     3396.75
2019-07-30     4674.75
2019-07-31     7062.00
Name: Total Charged in Rupees, dtype: float64
```

In [17]:
```
daily_orders.plot.bar(figsize=(20, 10), color='#ff99ff')
```

Out[17]: <matplotlib.axes._subplots.AxesSubplot at 0x255730a1248>