

Insights Derived from Exploratory Data Analysis

Tool Used - Python

Dataset consists of 12043 rows and 23 columns.

On further Analysis, it was found out that.

Column Name	Number of Null Value
card_present_flag	4326
bpay_biller_code	11158
merchant_id	4326
merchant_code	11160
merchant_suburb	4326
merchant_state	4326
merchant_long_lat	4326

After cleaning the dataset,descriptive analysis was done.

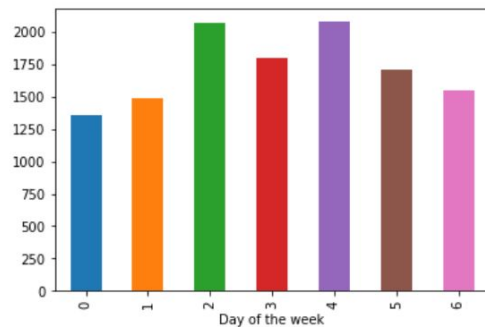
	card_present_flag	merchant_code	balance	age	amount
count	7717.000000	883.0	12043.000000	12043.000000	12043.000000
mean	0.802644	0.0	14704.195553	30.582330	187.933588
std	0.398029	0.0	31503.722652	10.046343	592.599934
min	0.000000	0.0	0.240000	18.000000	0.100000
25%	1.000000	0.0	3158.585000	22.000000	16.000000
50%	1.000000	0.0	6432.010000	28.000000	29.000000
75%	1.000000	0.0	12465.945000	38.000000	53.655000
max	1.000000	0.0	267128.520000	78.000000	8835.980000

The mean of transactional amount is **187.933588**.

After doing basic analysis and understanding the dataset. Date, Month, Year is separated from Datetime column to further analyse the data set.

```
In [21]: df_new.groupby('Day of the week')['transaction_id'].count().plot.bar()
```

```
Out[21]: <matplotlib.axes._subplots.AxesSubplot at 0x23e0c71b668>
```

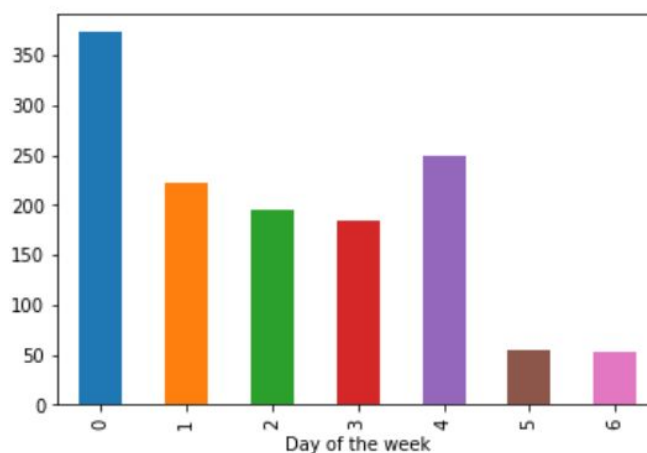


It was interesting to find out that the maximum number of transactions by volume took place on Wednesday and Friday. While, Maximum amount involved in transactions was seen on Monday.

```
In [18]: df_new["Day of the week"] = df_new['date'].dt.dayofweek
```

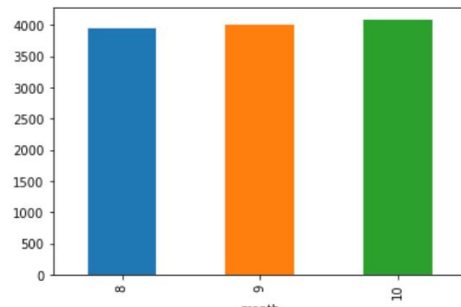
```
In [19]: df_new.groupby('Day of the week')['amount'].mean().plot.bar()
```

```
Out[19]: <matplotlib.axes._subplots.AxesSubplot at 0x23e0c2b3b00>
```



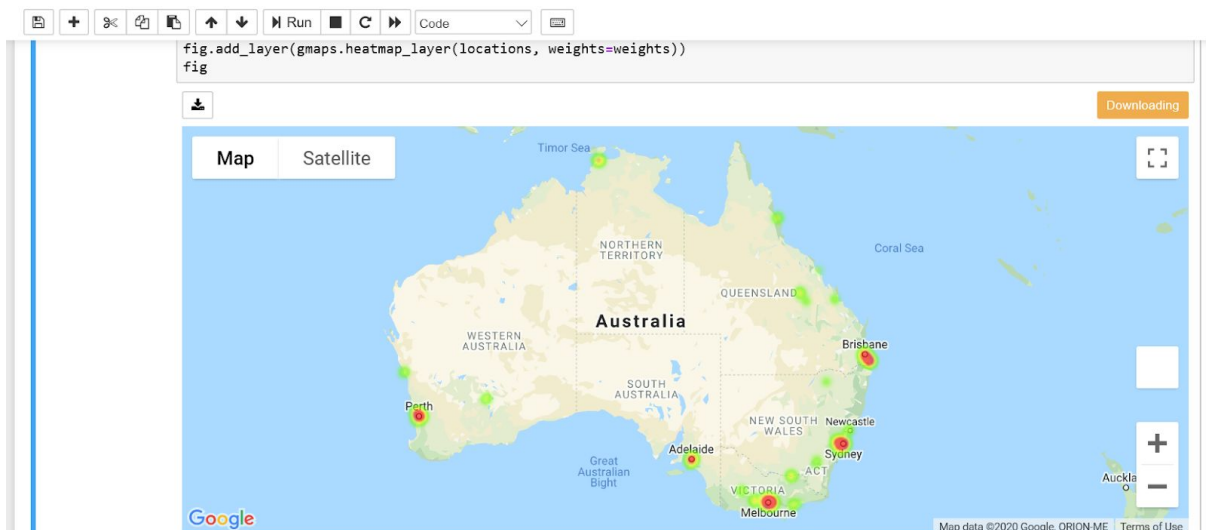
The volume of transactions showed an upward trend with increasing amounts of transactions each month.

```
In [20]: df_new.groupby('month')['transaction_id'].count().plot.bar()
Out[20]: <matplotlib.axes._subplots.AxesSubplot at 0x23e0c6fa0b8>
```



Another important and interesting finding was through the location of each transaction. When Latitude and Longitude coordinates were separated and were plotted in a heat map, it was found out that most of the transaction occurred in the following cities :

- 1) Perth
- 2) Adelaide
- 3) Melbourne
- 4) Sydney
- 5) Brisbane



With highest volume of transaction being 578 from a particular place.