Data_is_future

Improving the Average Order Value for Milk Basket

3rd August 2019

OVERVIEW

We aim to suggest methods and share insights into the order placing trends at Milk Basket and hope to improve the Average Order Value By increasing Revenue and Decreasing the Number of orders.

GOALS

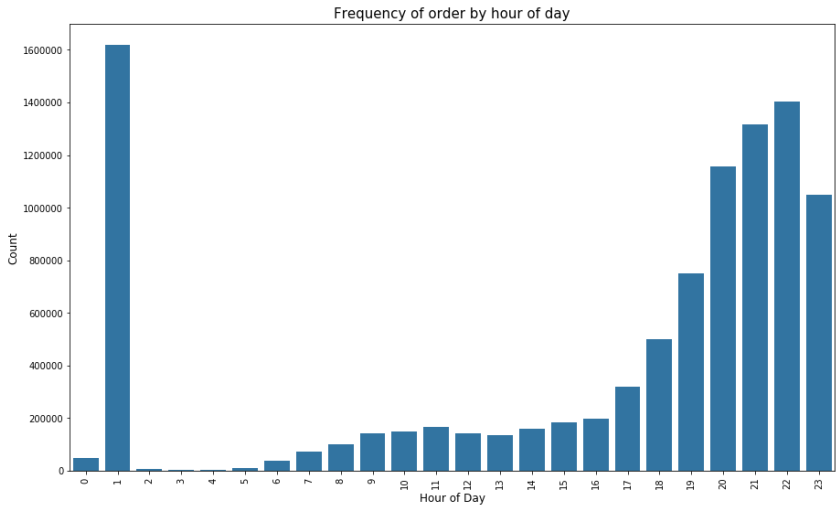
1. Improve the average order value
2. Recommend customers about related products

MILESTONES

Get basic and broad insights from the dataset

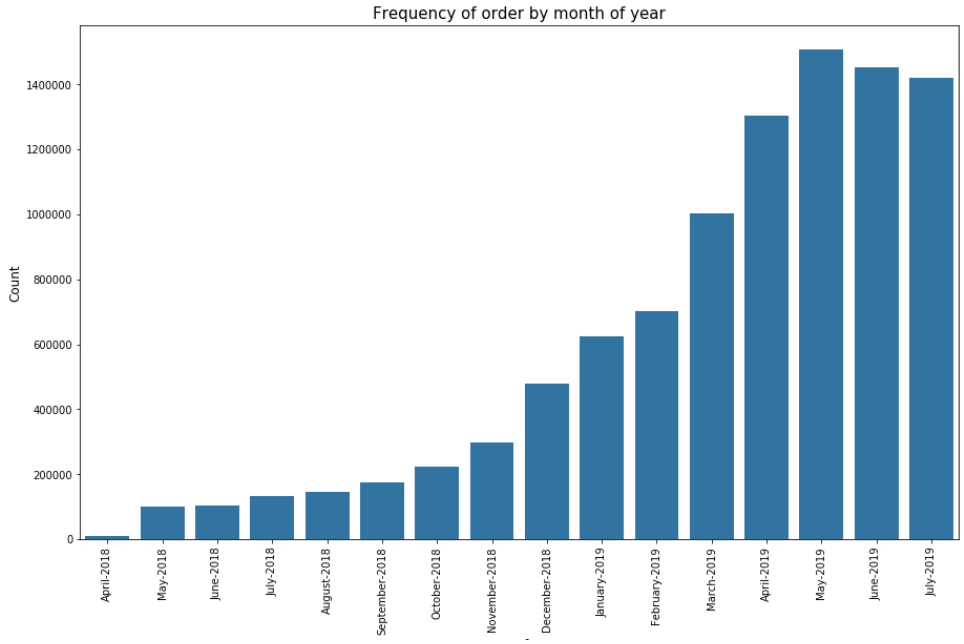
Using data analysis figure out basic trends such as

1. Peak hours during the day.



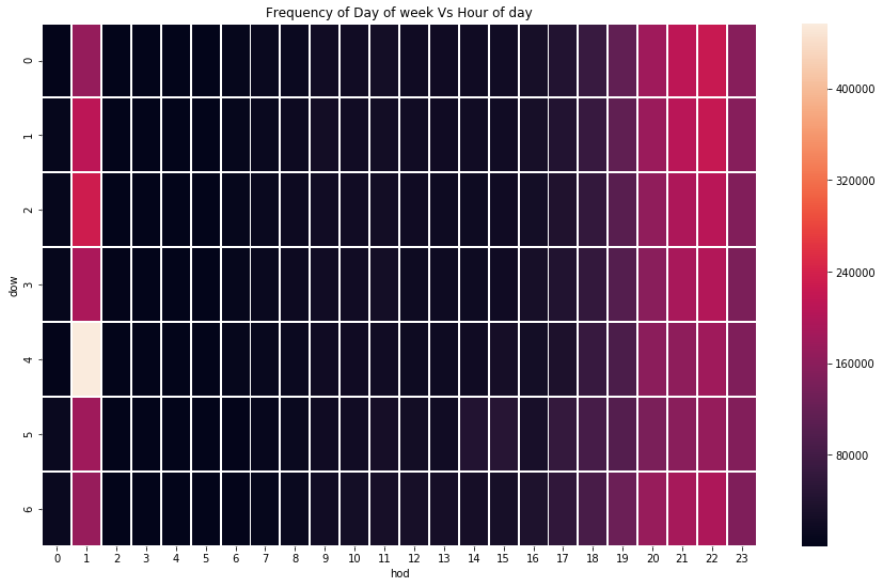
Inference: Peaks hours in a day are 1 am (Regular subscriptions) and 8-10 pm (More number of users free)

2. Number of orders placed per month.



Inference: The number of orders exponentially increased since the start, but slightly decreased in Jun-July 2019.

3. Peak days during the week



Inference: Friday is the peak day during the week

4. Customer segmentation based on purchase data.
5. Average Order Value distributions.

Study and Influence User Behavior

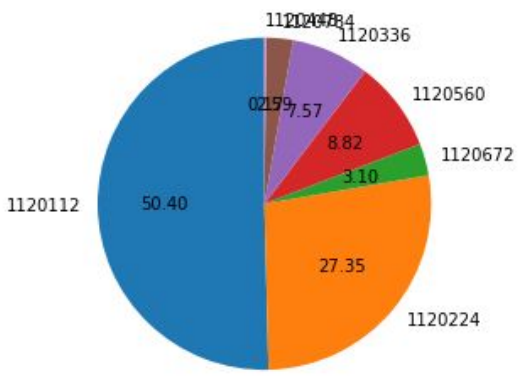
Identify Tier-1 customers (defined by certain purchase parameters, CLV etc). Cluster them meaningfully, and study their behaviour relative to other users in these clusters.

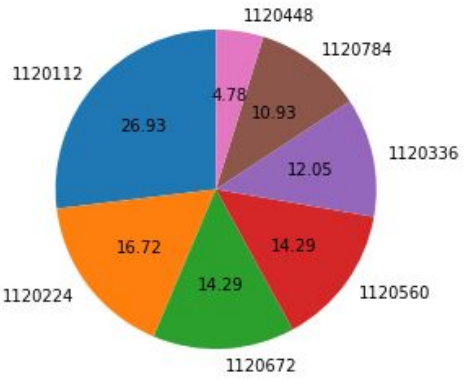
Try to predict the likelihood of a T-n customer becoming a T-1 customer given current data.

Shape the behavior of the T-n customer so that he becomes like a T-1 customer in the same cluster. Use insights from the most similar T-1 customers.

	UID	no_of_orders	mean_order_value	std_deviation	max_order	min_order	total_order_value
0	20932016	81	52.692469	25.650458	145.030	0.0000	4268.09000
1	11425904	50	56.211294	55.022871	265.200	4.0000	2810.56471
2	18601856	56	44.976607	27.907496	148.740	9.5000	2518.69000
3	14554736	38	79.396842	66.221772	327.020	0.0000	3017.08000
4	28784448	39	98.577949	49.434894	276.000	25.0000	3844.54000
5	14724528	16	322.198437	292.140210	813.155	18.0000	5155.17500
6	29194144	54	68.268148	50.994404	210.000	8.0000	3686.48000

City Specific Considerations





Total Transaction Volume

Days Since Launch

Significant skew in data was observed. Given this we concluded that it would be best to study user behavior specific to each city

Study Purchasing Trends Using FPGrowth Algorithm

Items Commonly Purchased together + Confidence score for the pair (based on training data)

```
(2361744,) ((1576512,), 1.0)
(1120112, 3954496) ((1120224,), 0.9666666666666667)
(1120224, 3954496) ((1120112,), 0.8529411764705882)
(2373952,) ((1386672,), 1.0)
(1120112, 3921008) ((1120224,), 0.9583333333333334)
(1120224, 3921008) ((1120112,), 1.0)
```

Influence Average Order Value by:

Use recommendations to:

1. **Increasing Order Size**

Recommend commonly grouped purchases to

2. **Decreasing number of orders by clubbing them**

Use FPgrowth growth groups to identify cases where customers forget to order something on one day and place on next.

Tech Stack

Python,Jupyter, Numpy, Pandas, Scikit learn,pickle and similar python libraries