**Freshco Hypermarket Capstone**

Freshco Hypermarket, situated in HSR, Bangalore, has established itself as a prominent supermarket in the region, catering to a wide range of customers. In response to evolving customer needs and to enhance convenience, Freshco introduced a home delivery service in the year 2021. To ensure seamless operations and optimize customer satisfaction, the store diligently maintained a comprehensive transaction data sheet, containing detailed information at the order level.

**Task to do -**

**Order level Analysis:**

1.    Identify order distribution at slot and delivery area level.  
2.    Identify the areas having highest increase in monthly orders (from Jan to Sep) in absolute orders.  
3.    Calculate delivery charges as a percentage of product amount at slot and month level.  
4.    Calculate discount as a percentage of product amount at slot and month level.  
5.    Calculate discount as a percentage of product amount at drop area and slot level.

**Completion Rate Analysis:**

6.    Identify Completion rate at slot vs day of the week (**Sunday to Saturday**) level. Can you spot some pattern in the data?  
7.    Calculate completion rate at drop area level.  
8.    Completion rate at number of products ordered level. For this first you need to create a column having number of products against every order.  
9.    Give you analysis on the any pattern you observe in the completion rate.

**Customer Level Analysis:**

10.    Identify Completion rate at source level.  
11.    Calculate LTV for every customer.  
12.    Calculate aggregated LTV at customer acquisition source level. Refer to aggregated LTV example.  
13.    Calculate aggregated LTV at acquisition month level. Refer to aggregated LTV example.  
14.    What is the average Revenue (Product amount after discount) per order at different customer acquisition source level?  
15.    What is the average Revenue (Product amount after discount) per order at acquisition month level?  
16.    Is there any pattern in order rating across slots, number of items placed, delivery charges, discount. For example, there might be an insight from the data that orders placed during late night are generally rated high. While orders placed in early morning are not rated high. OR orders having more than 5 items are generally rated high.

**Delivery Analysis:**

17.    Calculate average overall delivery time at month and delivery area level.  
18.    Calculate average overall delivery time at month and weekday/weekend level. You might need to create a column which will tag every date to either weekday or weekend.  
19.    Calculate average overall delivery time at slot level. Refer to the definition of slot.  
20.    Do you see any pattern in delivery charges with slot or delivery area.  
21.    Do you see any pattern in delivery time and delivery area. If yes then find out logical reason.

1. **Order Level Analysis**

**Objective:** To analyse the trends and characteristics of orders

**Metrics used:** No. of orders, Slot, Delivery area, Delivery charges, Product amount, Month/Date, Drop area, Discount.

**Analysis:**

1. Upon plotting a Clustered bar between number of orders at slot and delivery area level, we observe that

**High volume areas:** HSR layout has highest number of orders of approximately 4085 orders coming in the afternoon slot then followed by morning, night, evening and late night. Next closest area for maximum orders is ITI layout and Harlur with 1039 in afternoon and 382 in morning respectively.

**Moderate volume areas:** Kudlu and Bommanahalli- Mico layout have 518 and 551 number of orders mostly afternoon and morning times.

**Low volume areas:** Akshaya Nagar, BTM stage 1 and 2 and Sarjapur road have a smaller number of orders altogether as compared to other areas.

1. We can observe that high volume areas usually have delivery slots in afternoon and morning. And lowest number of orders across all areas are seen in the late-night slot.

**Inferences:**

1. HSR layout and ITI layout have high volume of orders possibly because they are more populated than other areas and people might prefer home deliveries as compared to shopping offline.
2. The slots that have most to least number of orders are afternoon, morning, night, evening and late night.

**Analysis:**

1. As we compare number of orders across all delivery areas for the months January to September, we observe that that is a steady increase in the number of orders and months across densely populated areas like HSR layout, ITI layout and Harlur.
2. Areas like Frazer Town, Brookefield, Mahadevapura have least number of orders across all months.

**Inferences:**

1. Areas such as HSR layout, ITI layout and Harlur have steady growth in orders. This could be due to certain factors like increasing population or inclination towards home deliveries because of traffic congestion.
2. Areas like Frazer town, Brookefield and Mahadevapura consistently have less number of orders across all months. This could be due to less population density, less preference of home deliveries or more competition from local stores, lesser price range for various products than supermarkets.
3. There is a steady increase in orders from January to September suggesting holiday seasons, changing weather conditions or an upward trend in consumer behaviour.

**Analysis:**

1. Upon plotting a combo graph of clustered column chart and line chart between delivery charges in percentage and product amount at slot and month level. Here we observe that September has highest product amount and less delivery charge percentage. And February has low product amount as compared to others but highest delivery charge in percentage.
2. Also, it can be observed that delivery charges are higher in late night slot as compared to other slots across all months except for May.

**Inferences:**

1. Product amount is higher in orders that are placed in afternoon and morning slots across all months. Product amount is lower in late night slots indicating that people order in bulk quantities during day times and less quantities during night times.
2. Also, it can be observed that people order during day times because delivery charges are high at night as compared to day and probably because a minimum cart amount is fixed which gives access to certain coupons making the delivery charges less or free.
3. And the months of January, February, March, April have more delivery charges as compared to other months. This could be due to seasonal trends, weather conditions and no holiday promotions or discounts available.

**Analysis:**

1. Product amount is highest in September and August as compared to other months.
2. Even discount prices and percentages are higher in August and September as compared to other months.
3. Discounts are higher during night and afternoon slots as compared to other slots. The least discount is at late night slot across all months.

**Inferences:**

1. Morning and afternoon slots show high product amounts indicating high customer activity with good promotions and coupons possibly because of high discount prices provided.
2. Evening and night slots have moderate product amounts with good discount percentages indicating these slots are popular irrespective of discounts provided.
3. The months of August and September show a peak in both product amount and discount offered, indicating festival offers and good promotional efforts.

**Analysis:**

1. We observe that areas with highest product amount have lowest discounts across all slots.
2. Areas with lowest product amount have highest discounts.
3. HSR layout has highest product amount across all slots as compared to any other area. Followed by ITI layout and Harlur.

**Inferences:**

1. It can be inferred that areas with high product amounts and low discounts have customers that are less price-sensitive and are willing to purchase without any or minimal discounts.
2. These areas could also have residents that have a better income and are less influenced by discounts and more focused on product availability and quality.
3. Areas with lowest product amount and high discount suggests that customer demand is weak. It could also mean that the brand has not been established yet in those areas hence the discounts are provided to attract more customers to build brand awareness.
4. **Completion rate analysis**

**Objective:** Measure the rate at which orders are completed or successfully delivered.

**Metrics:** Slot, day of the week, completion rate, drop area, number of products

**Analysis:**

1. The average completion rate across all slots and days of the week is pretty consistent.
2. The lowest completion rate is 97% which is on Saturday in the night slot.
3. The highest completion rate is 100% for late night slot on Wednesday and evening slot on Sunday.

**Inferences:**

It can be seen that almost all orders have a completion rate from 97% to 100% across all slots and all days of the week. This suggests that the business operations are running smoothly with good inventory management, customer satisfaction, delivery executives fulfilling duties irrespective of time slots and any day in the week, reliable services as there are a good guarantee of delivery.

**Analysis:**

1. It is observed that most of the locations have a completion rate of 99.55%.
2. Few locations like Cox town and Whitefield have a 0% completion rate.
3. Some locations like Bellandur, Marathahalli, Domlur EGL, Viveka Nagar, BTM Stage 1 which have a varied completion rate ranging from 49.78% to 96.71%.

**Inferences:**

We can observe that most locations have an excellent completion rate of 99.55%. However certain locations have very less completion rates indicating poor accessibility in those areas, high traffic congestion, bad road conditions, insufficient staff, high demand in peak hours and external factors like weather conditions. Areas with 0% completion rate indicates the orders were cancelled.

**Analysis:**

1. We can observe that the higher number of products the completion rate is higher.
2. Deliveries having 17 to 25 number of items have a completion rate of 99.55%.
3. There is no trend as such that a smaller number of products means bad completion rate.

**Inferences:**

1. High completion rates indicate a robust delivery team, good inventory management, customer satisfaction etc.
2. Possible factors contributing to lower completion rates for certain products could include factors such as demand variability, supply chain issues, product availability, or customer preferences.
3. **Customer level analysis**

**Objective:** To understand customer behaviour and factors influencing this behaviour

**Metrics:** Slot, day of the week, completion rate, drop area, number of products, customer acquisition source, LTV, aggregated LTV, average revenue, acquisition month, delivery charges, discount, order rating.

**Analysis:**

It can be observed that the completion rate is highest in organic acquisition source with 99.18%. Next is Facebook, Snapchat, Google, Instagram and offline campaign with completion rates of 99.13%, 99.12%, 99.11%, 99.02% and 99.00% respectively.

**Inferences:**

1. Customers acquired through organic channels have the highest average completion rate at 99.18%. This indicates that customers who find the platform organically, without specific advertising or promotional efforts, tend to have a slightly higher completion rate compared to other acquisition sources.

2. Other sources like Facebook, Snapchat, Google and Instagram have consistency in completion rates across different social media platforms. Offline campaigning has lower completion rate which could be due to customer behaviour or effectiveness of offline marketing strategies.

**Analysis:**

1. User ID APQ2413449 has the highest Customer lifetime value as compared to other user IDs. The LTV for this ID is 60925.
2. XTN17102597, WUO244697, VUO1878360, VNG1950019, SOG14109293, NRN1749047, PIW1787111, NQM134638, MJK644709, HYI10103110, IIQ9109104, EXR16101706, EDX1418408, ESP1681726, ECW1496573, BXE108709, BMS1940428 have an LTV of 0.

**Inferences:**

1. The highest LTV for customer meaning total revenue generated by the customer was highest in the case of user ID APQ2413449. This indicates that this particular customer finds home delivery convenient, frequently purchases products indicating customer satisfaction and loyalty towards the services provided. Also, this customer may buy high-value products which contributes significantly towards total revenue generated by the business.
2. Low LTVs could be a sign of dissatisfaction towards the services or has found an alternative that might be to their liking. It could also be the customer has had a negative encounter such as poor-quality product or poor delivery service. Sometimes, customers often purchase solely if there are any promotions or discounts with no intent to return.

|  |
| --- |
| **Analysis:** |
|  |

1. The total LTV is highest for organic acquisition source and lowest LTV is for Instagram.
2. The total number of customers acquired by different sources are- Facebook (2618), Google (5348), Instagram (2784), Offline campaign (2862), Organic (6680) and Snapchat (2531) indicating maximum customers were acquired through organic method of acquisition like content marketing, blogging, emails, organic social media posts etc.
3. Aggregated LTV was high through sources such as Snapchat and Google with Rs 364 and Rs 363 respectively.

**Inferences:**

1. High LTV for that platform indicates that it is a very effective way at attracting customers who are willing to spend more with time.
2. Marketing strategies have been well – targeted across these platforms.
3. Low LTV through that source suggests that a customer does not trust the platform enough to make decent purchases. Marketing might not be great in such platforms which may not reach the relevant audience.

**Analysis:**

1. It can be seen that most of the customers get acquired during the month of January. That is most customers spend a lot of money during January.
2. The least LTV is generated in the month of September.

**Inferences:**

1. We know that during the month of January there are a lot of people who take advantage of promotions or sales which provide discounted price on products and hence end up buying more contributing to increase in LTV.
2. Also, most companies provide bonuses at the end of previous year which can influence customer behaviour.
3. Less LTVs from the months of June to September suggest that there is an ineffective marketing strategy or no promotional offers, budget constraints or economic shifts.

**Analysis:**

1. Average revenue generated through organic sources and google accounts for 29% and 24% of the total revenue.
2. Rest platforms contribute 11% to 12% of total revenue.
3. 41% of revenue is contributed solely in the month of January and then a gradual decrease within the next months.

**Inferences:**

1. 29% and 24% average revenue is generated through sources like organic methods and google. This includes methods of search engine optimization, emails, google, blogging, social media influencing, social media posts etc.
2. Almost 41% revenue is generated in January indicating strong promotional offers, discounts, sales, holiday season that makes the customer buy more in that particular month.
3. There are no budgetary constraints as it is a start to new year and most companies provide bonuses at the end of year.
4. A lot of target audience may not behave the same way at the end of the year.

**Analysis:**

1. The average rating at late night is 3 as compared to other slots which have a rating of 4.
2. This could be due to lesser availability of staff which can lead to longer wait period or less attentive service.
3. Shortage of supplies affecting the quality of products or service provided.
4. An average rating of 4 means that the customer is satisfied with the products, delivery service was prompt, delivery executives were friendly and polite. However, there is a scope for improvement.

**Analysis:**

1. The average rating across most of the products is 4.
2. Highest number of products has a rating of 0.
3. However, no trend that says a greater number of products will always lead to bad ratings. Because 23 products ordered by a customer has a rating of 4.

**Inferences:**

1. Offering a large number of products may lead to a compromise in product quality. As in there can be inconsistency in the quality of different products leading to a low customer satisfaction.
2. Managing large product orders can lead to operational inefficiencies, logistical issues and overall service delivery.
3. It could also be that many products may not have a strong market-fit leading to poor ratings.

**Analysis:**

1. There is no necessary trend as such that high discount leads to better ratings.
2. However, it is observed that discounts ranging from 170 to 210 have a rating of 5.

**Inferences:**

1. It can be observed that ratings do not depend on the discounts provided if customers think products are already fairly priced.
2. Customers may value quality of products rather than compromise on the quality for a mere discount.
3. Customer satisfaction depends on overall experience including service quality, product quality, ease of delivery, delivery speed etc.

**Analysis:**

1. There is no such trend that suggests that the higher the delivery charges the lesser the ratings.
2. We can see that ratings are randomly dispersed for a wide variety of delivery charges.

**Inferences:**

1. It can be observed that ratings do not depend on the delivery charges.
2. Customers may value quality of products which can outweigh a minimal delivery fee.
3. Customer satisfaction depends on overall experience including service quality, product quality, ease of delivery, delivery speed etc.
4. If the delivery is fast then customers are willing to pay for good quality products to be delivered on time and in good condition.
5. **Delivery analysis**

**Objective:** To understand customer behaviour and factors influencing this behaviour

**Metrics:** Slot, weekday/ weekend, average overall delivery time, month, delivery area, delivery charges, delivery time.

**Analysis:**

1. Mahadevapura has the highest average overall delivery time in the month of May with 2.44 hrs.
2. Average delivery time in the month of May across all areas is higher as compared to others. Average overall delivery time in May is 0.74 hrs.
3. Average overall delivery time in February is lowest with 0.31 hrs.

**Inferences:**

1. Delivery times may increase in holiday seasons or sales because of high demand of orders.
2. Because of some hindrance in certain areas due to road constructions, or blockages or traffic congestion, delivery times can go higher.
3. Certain issues with logistic partners and transportation network can delay deliveries.
4. Other conditions like adverse weather conditions or strikes or public holidays can hinder delivery schedules.

**Analysis:**

1. It can be observed that there is not much significant difference in average overall delivery times during weekends or weekdays across all months.
2. Only during the month of May there is a slight difference in average overall delivery times during weekends or weekdays.

**Inferences:**

1. As seen above there is not much significant difference in average overall delivery time. This can be due to a robust logistics and well-optimized delivery processes.
2. There is a good balance of staff during weekends as well as weekdays across all months so that there is consistency in average overall delivery times.
3. Proper inventory management and proper planning so that resources are not exhausted during weekends.
4. A consistent transportation network to deliver in all weather conditions and different days across all months.

**Analysis:**

1. It can be observed that average overall delivery time during late night and night slots is quite lower as compared other time slots.

**Inferences:**

1. Most important factor to be considered is lesser traffic at night times, allowing vehicles to travel faster.
2. Most traffic signals are switched off at night so there are fewer delays on the road.
3. Delivery routes can be optimized at night due to lesser traffic and delivery executives can go through a direct path rather than take detours or long cuts.
4. Usually there are less volume of products ordered at night allowing the order to be processed quickly.

**Analysis:**

1. It can be observed that average delivery charges are higher at late night slot as compared to other slots of time.

**Inferences:**

1. Mostly there are lesser number of employees working at night. And if they are, the wages are higher for night shifts. Hence, the wages are compensated with higher delivery prices at night.
2. Lesser delivery executives are available at night hence there is more demand for limited source.
3. Additional safety and security have to be provided for night-time deliveries which increase operational costs.
4. And since night time deliveries are typically faster, customers don’t mind paying more delivery charges.

**Analysis:**

1. The average delivery charges are higher in areas like Brookefield, CV Raman Nagar, Frazer town, Vimanapura etc.
2. The delivery charges are lower at HSR layout, ITI layout, Harlur, Cox town etc.

**Inferences:**

1. The average overall delivery charges across certain areas is high due the long distance from the warehouse which is present at HSR layout. All these areas span more than 13km which is a long distance to be covered, hence more delivery charges.
2. Areas like Harlur, ITI layout, Kudlu are closer to HSR layout and hence delivery time is less as well as delivery charges are less.
3. There could be a possibility that there are certain promotional offers which offers free delivery or minimum delivery within a 5km radius from the warehouse present at HSR layout.
4. Also, certain areas with high operational costs, like labor, fuel, real estate can increase delivery charges.

**Analysis:**

1. Average overall delivery time for Mahadevapura is high with 2.44 hrs.
2. Other locations like Vimanapura, Brookefield, Pattandur are also high in average overall delivery time as compared to other locations.
3. Locations like Harlur, Bellandur, ITI layout and HSR layout have lower average overall delivery time with 35 to 40 mins.

**Inferences:**

1. Delivery time is higher in location that covers a longer distance specially more than 10 kms.
2. It is obvious that these locations will take more time to deliver because they are located farther away from the warehouse and will have to encounter traffic, delays in signals, adverse weather conditions etc.
3. Also, areas like ITI layout, Harlur, Bellandur are situated closer to the warehouse and have better connectivity in terms of infrastructure as compared to other locations. These locations fall within a 6km radius and hance delivery times are shorter.
4. Certain areas like Cox town and Frazer town have less delivery times but are situated further away than 10km radius. This is done probably because the customers in those areas do not buy online and buy in local stores. To promote demand and online shopping in such areas, delivery charges are exempted.

**Overall optimization:**

1. **Improve delivery efficiency** by using latest technologies to optimize delivery routes, reducing fuel costs and delivery times. Also establish regional warehouses closer to major markets to reduce costs and delivery times.
2. **Same-day and next-day deliveries** can be offered to customers that is according to their convenience.
3. **Transparent pricing** can be provided for certain aspects like delivery charges, explaining why there is higher cost in certain areas or if the weather condition is bad there is going to be a surge in delivery charge.
4. **Customer feedback** is important as it will help the business by gathering valuable insights to make necessary improvements.
5. **Loyalty programs and incentives** can be given to existing customers to refer new ones by giving referrals which benefit all of them.
6. **Promote or give ads** on various social media platforms to attract customers.
7. **New product launches** can attract wide variety of audience as this will keep inventory fresh and attract more customers.
8. **Market expansion** by expanding into other locations where supply is less and demand is high and there is potential for your business to grow.
9. **Website optimization** by providing users with a simple interface which is user-friendly and can offer a seamless shopping experience.
10. **Use data analytics** to understand customer behaviour, preferences and buying patterns to stock up inventory accordingly.