Food & goods delivery analysis

https://sites.google.com/view/ajaikumar/food-goods-delivery-analysis?authuser=0

Introduction

The aim of this project is to conduct an analysis of the food & goods delivery data of a delivery business called "Delivery Center", to answer various business questions related to customer preferences, order volume, delivery time, cost, and payment methods. By analyzing these datasets, we can identify key insights that can help the business optimize its operations, improve customer satisfaction, and increase profitability.

Data description

The analysis is based on seven datasets:

- Channels ==> contains information about the delivery channels where the goods and food of retail partners are sold.
- Deliveries ==> contains information about deliveries made by the delivery drivers of the Delivery Center.
- Drivers ==> contains information about delivery drivers who operate from Delivery Center hubs and make deliveries to customers when an order is processed.
- Hubs ==> contains information about the hubs of Delivery Center, which are the centers of distribution for orders and from where deliveries are made.
- Orders ==> contains information about the orders processed through the Delivery Center platform during the year 2021.
- Payments ==> contains information about the payments made to Delivery Center.
- Stores ==> contains information about retail partners who use the Delivery Center platform to sell their goods and/or food on marketplaces.

Each dataset provides specific information related to the delivery process, including order and payment details, delivery drivers' information, and hub locations. By combining and analyzing these datasets, we can gain insights into various aspects of the delivery process, such as delivery times, delivery costs, payment methods, and order volume.

Link to the data source 👉 http://bit.ly/42Mxc2t

Methodology

The analysis involved several steps.

First, the datasets were cleaned and preprocessed to ensure data quality and consistency. Data cleaning involved handling duplicate records, null records, inconsistencies, format errors, misspellings, and data type conversions.

Then, performed exploratory data analysis to understand the distribution and trends in the data.

Next, conducted statistical analysis to answer the business questions posed. Finally, interpreted the results and provided insights based on the analysis.

The Python programming language and various data analysis libraries (Pandas, NumPy, Matplotlib, Seaborn, Datetime, etc.) were used throughout the analysis process.

Jupyter Notebook

Objectives

The following are the business questions to be answered through this analysis:

1. What are the most popular delivery channels among customers?
2. Which hubs have the highest and lowest order volume?
3. What are the most popular store segments among customers?
4. What is the average delivery time for different delivery channels?
5. How does the delivery distance affect the delivery cost and delivery time?
6. Which payment methods are most frequently used by customers?

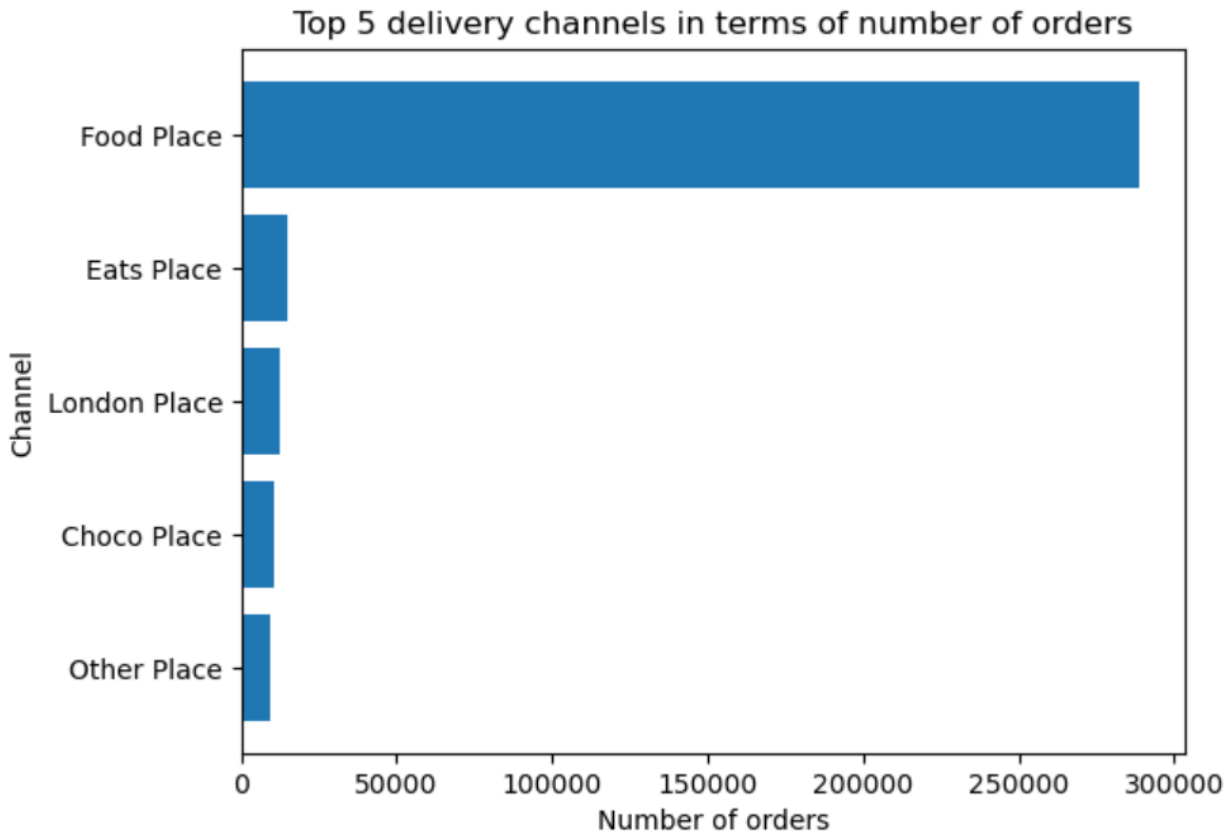
7. How do the delivery distance and delivery status affect the drivers' earnings?
8. How does the order amount affect the delivery fee and delivery cost?
9. How does the delivery cost vary with respect to the day of the week?
10. Is there a significant difference in the average order amount between different payment methods?

Results & Insights

Based on the analysis, the results and insights derived for each of the above business questions are as follows:

What are the most popular delivery channels among customers?

Calculating the number of orders per delivery channels

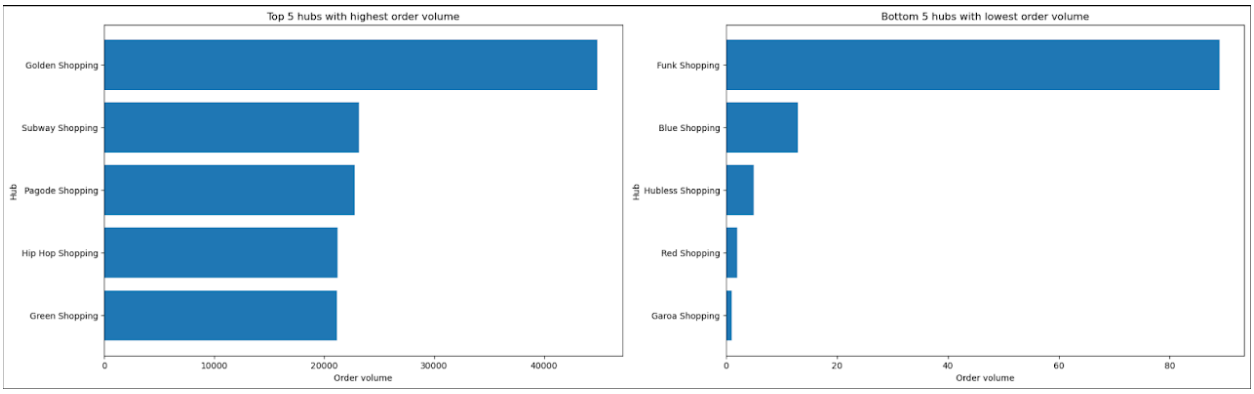


- "Food Place" has the highest number of orders, with a count of 2,88,723. This suggests that "Food Place" is the most popular delivery channel among customers.
- On the other hand, "Other Place" has the lowest number of orders, with a count of 9,204.
- This information can be used by the business to understand which delivery channels are most effective in reaching out to customers and where they should focus their marketing efforts.

- -

2. Which hubs have the highest and lowest order volume?

Finding out the hubs with the highest and the lowest order volume

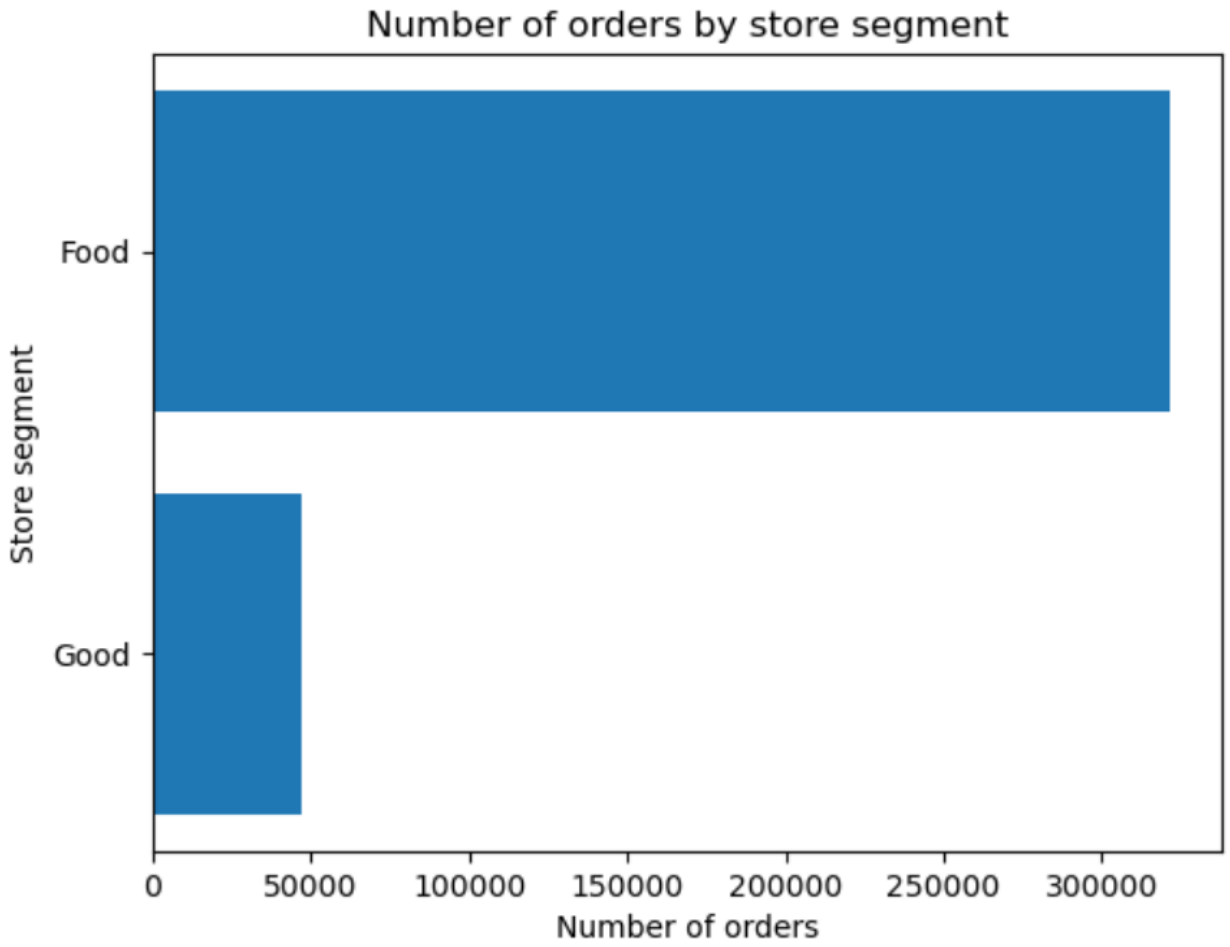


- The top five hubs with the highest order volume are "Green Shopping", "Hip Hop Shopping", "Pagode Shopping", "Subway Shopping", and "Golden Shopping".
- On the other hand, the hubs with the lowest order volume are "Garoa Shopping", "Red Shopping", "Hubless Shopping", "Blue Shopping", and "Funk Shopping".
- The high order volumes in certain hubs could be due to the location of the hub and the number of orders being processed in that particular area.
- The low order volumes could be due to various reasons such as the location of the hub, the demand in the area, or the availability of delivery partners in that location.

- -

3. What are the most popular store segments among customers?

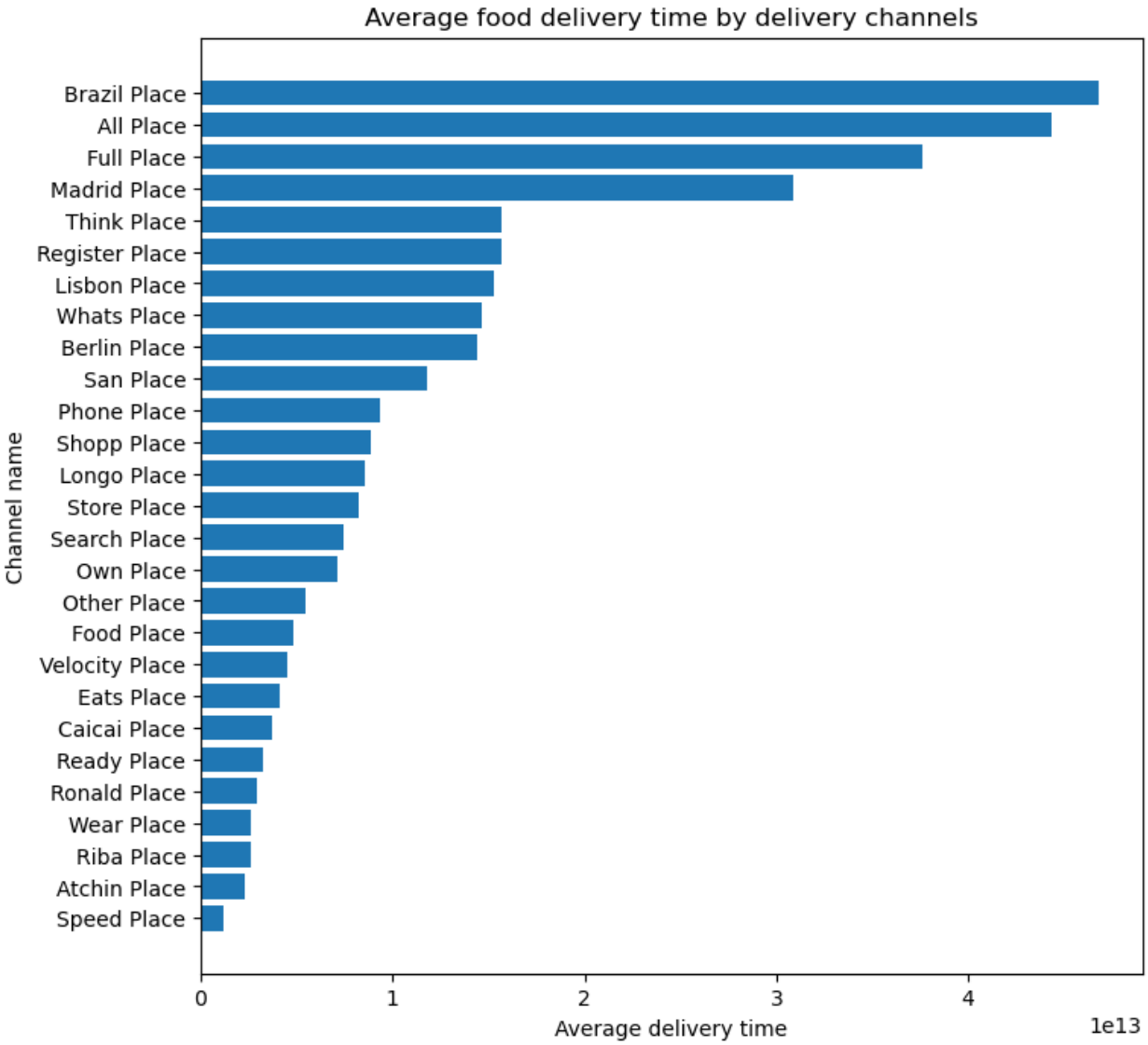
Calculating the number of orders for each of the store segments

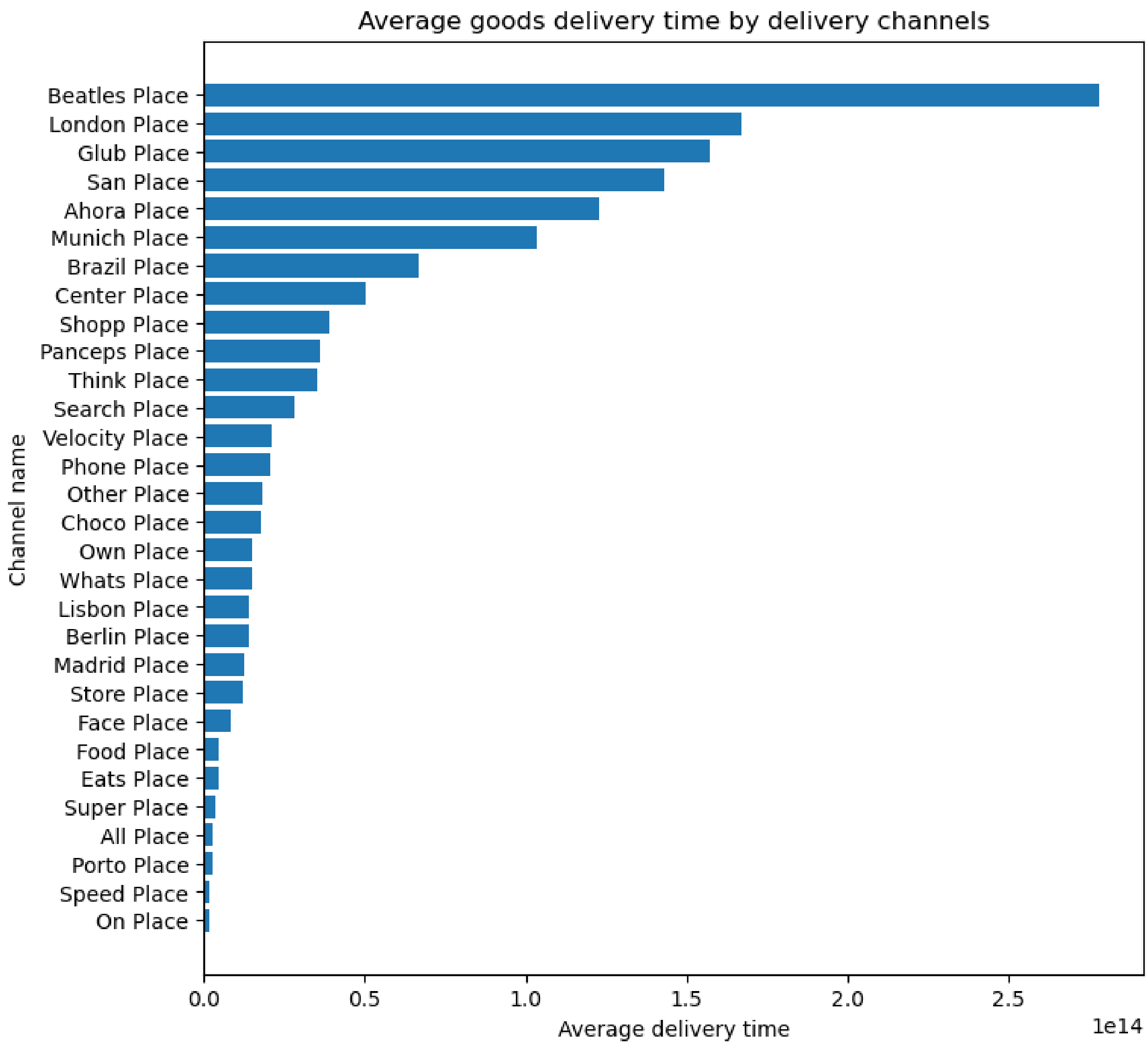


- It seems that the majority of the customers are more inclined towards food-related stores as compared to good-related stores, with the food segment having significantly more sales.
- This could be due to a higher demand for food items, especially with the rise of food delivery services during the pandemic.
- It could also indicate that the platform may have more food-related stores available for customers to choose from. As a result, the platform could focus on expanding its selection of stores in the good segment to attract more customers interested in purchasing non-food items.

- -

Calculating the average delivery time for each of the delivery channels


Average food delivery time by delivery channels
Brazil Place
All Place
Full Place
Madrid Place
Think Place
Register Place
Lisbon Place
Whats Place
Berlin Place
San Place
Phone Place
Shopp Place
Longo Place
Store Place
Search Place
Own Place
Other Place
Food Place
Velocity Place
Eats Place
Caicai Place
Ready Place
Ronald Place
Wear Place
Riba Place
Atchin Place
Speed Place
Channel name
0
1
2
3
4
Average delivery time
1e13

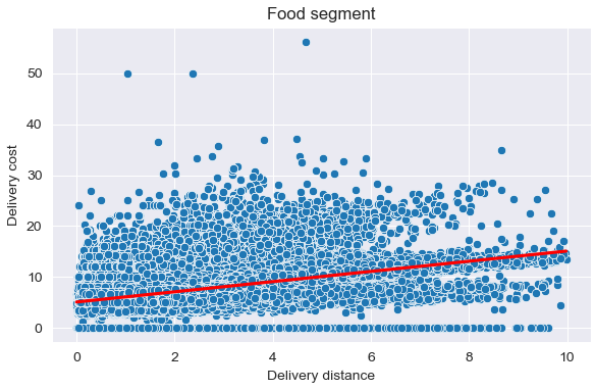


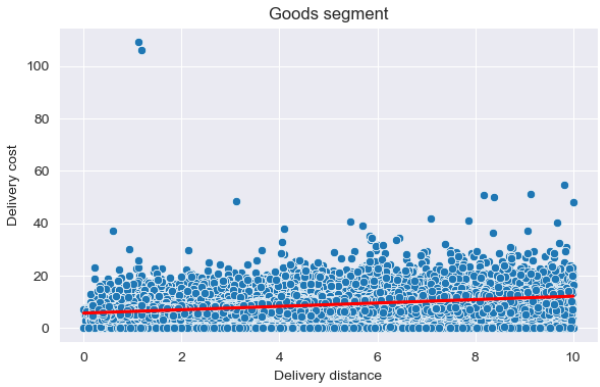
- The delivery times vary significantly across channels and store segments.
- For the food segment, the fastest delivery time is provided by the "Speed Place" channel, with an average delivery time of 00:20:23, and the slowest delivery time is provided by the "Brazil Place" channel, with an average delivery time of 12 hours and 59 minutes.
- For the goods segment, the fastest delivery time is provided by the "On Place" channel, with an average delivery time of 00:31:30, and the slowest delivery time is provided by the "Beatles Place" channel, with an average delivery time of 3 days, 5 hours, and 10 minutes.

- -

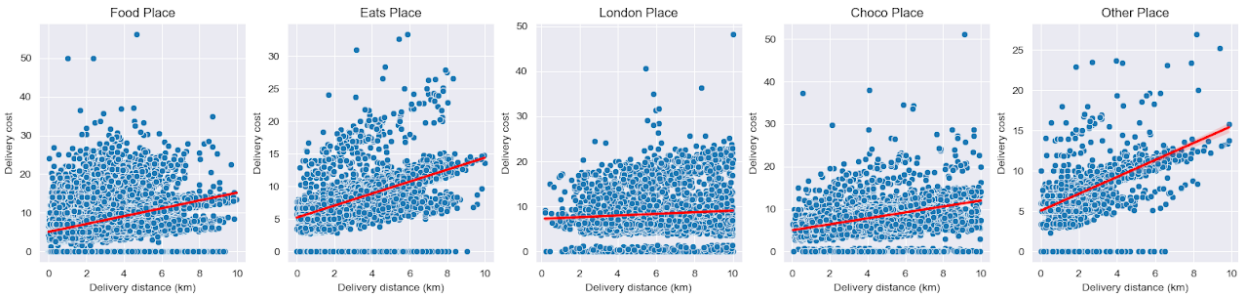
5. How does the delivery distance affect the delivery cost and delivery time?

Checking the relationship between the delivery distance and the delivery cost for the store segments as well as the delivery channels





Store segments



Delivery channels

Coefficients of correlation:

1. Delivery distance vs Delivery cost
    - Food place = 0.503
    - Eats place = 0.573
    - London place = 0.07
    - Choco place = 0.42
    - Other place = 0.521

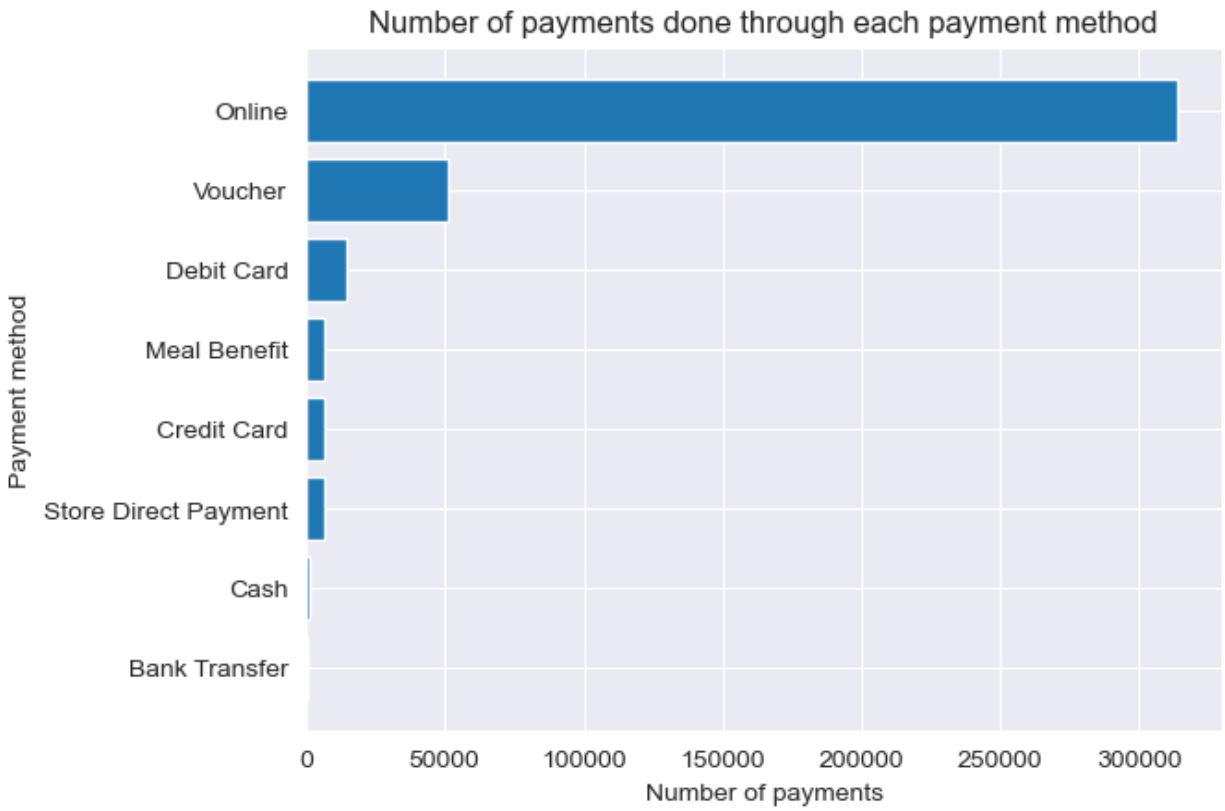
2. Delivery distance vs Delivery time

- Food place = 0.003
- Eats place = 0.012
- London place = -0.052
- Choco place = 0.034
- Other place = -0.015
- For the purpose of this project, I've included only the top 5 delivery channels for the analysis based on their order volume.
- Based on the coefficients of correlation, there seems to be a positive correlation between delivery distance and delivery cost for most delivery channels and segments, with the strongest correlation seen in "Eats Place" and "Food Place".
- This suggests that customers may have to pay more for their delivery as the distance between the restaurant and the delivery location increases.

- -

6. Which payment methods are most frequently used by customers?

Calculating the number of payments through each payment method

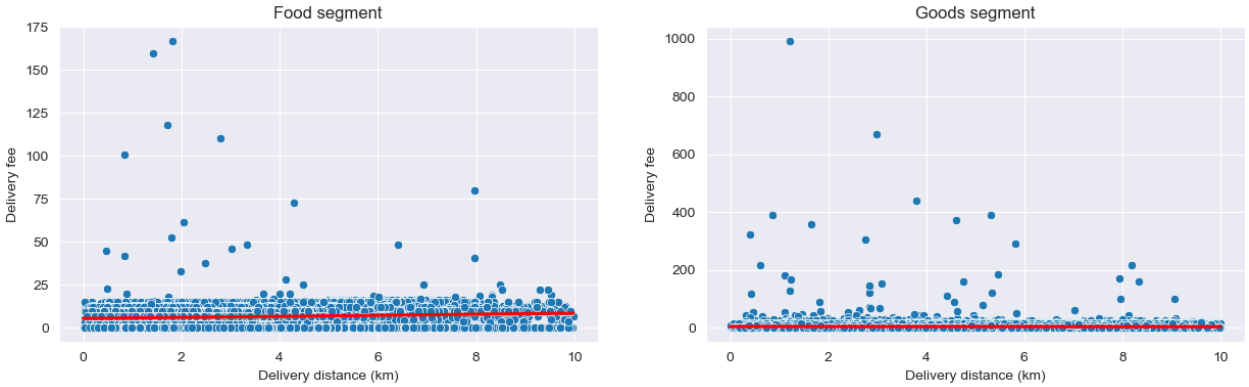


- The most frequently used payment method by customers is "Online" with 313,990 transactions.
- The second most popular payment method is "Voucher" with 51,065 transactions, followed by "Debit Card" with 14,344 transactions, and "Credit Card" with 6,684 transactions.
- "Meal Benefit" and "Store Direct Payment" methods have almost the same number of transactions with 6,687 and 6,574, respectively.
- "Cash" has 1,190 transactions, while "Bank Transfer" has only 300 transactions.

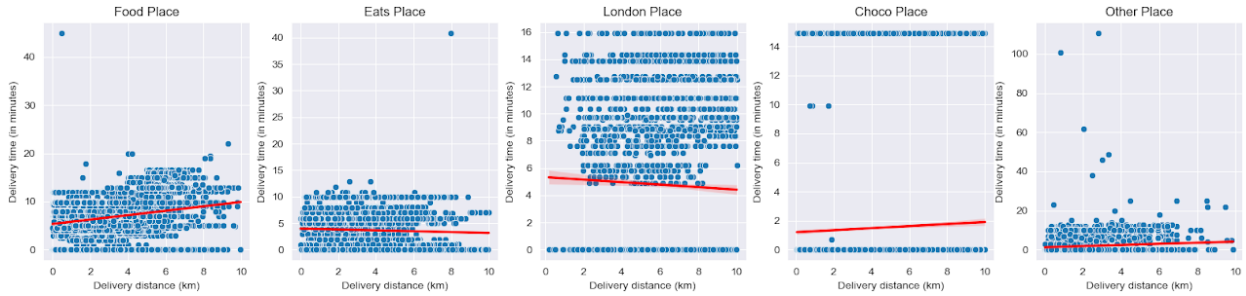
---

7. How do the delivery distance and delivery status affect the drivers' earnings?

Checking the relationship between the delivery distance and the delivery fee for the store segments as well as the delivery channels



Store segments

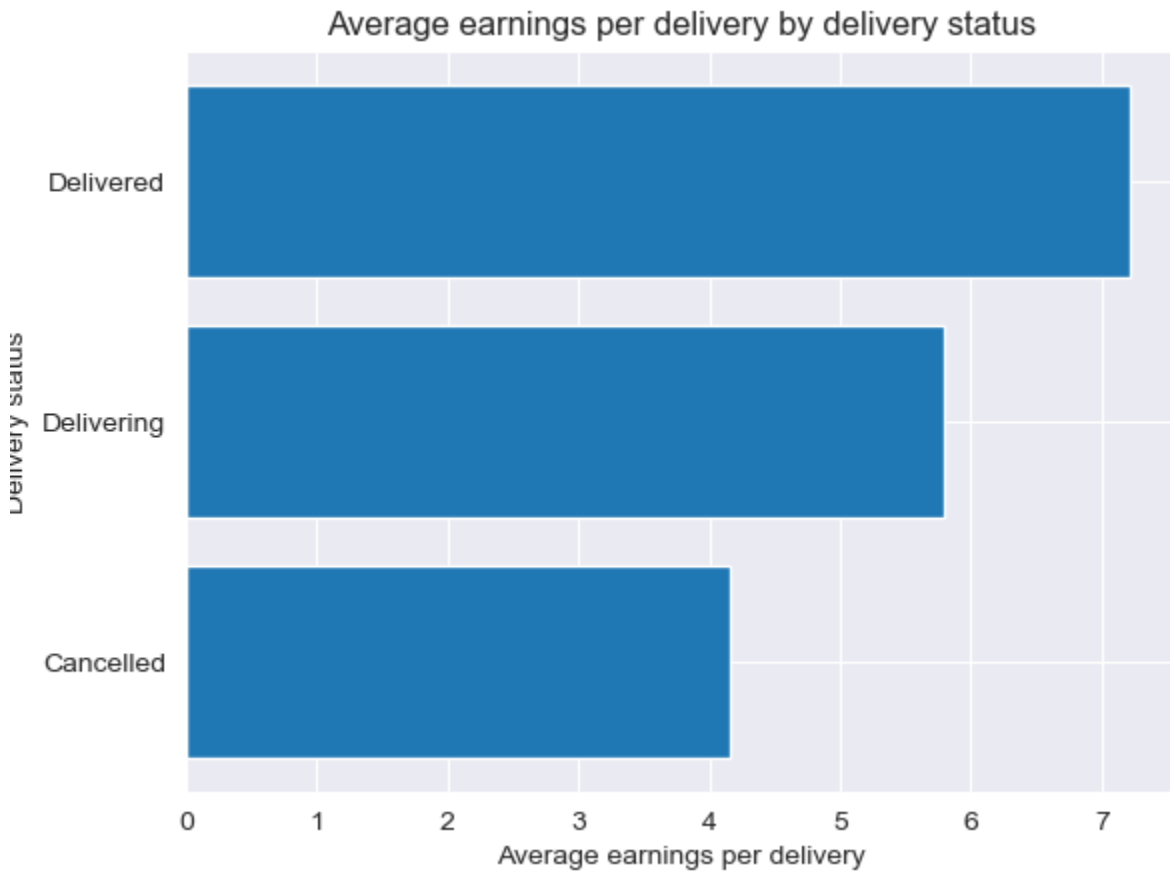


Delivery channels

Coefficients of correlation: Delivery distance vs Delivery fee

- Food place = 0.128
- Eats place = -0.049
- London place = -0.039
- Choco place = 0.037
- Other place = 0.095
- Based on the coefficients of correlation, it seems that there is a weak positive correlation between delivery distance and delivery fee for the "Food Place" and "Choco Place" delivery channels, and a weak negative correlation for the "Eats Place" and "London Place" channels.
- The "Other Place" channel has a weak positive correlation.
- However, it's important to note that correlation does not necessarily imply causation, and there could be other factors at play.

Calculating the average earnings per delivery based on delivery status and store segments



Average earnings per delivery for each store segments:

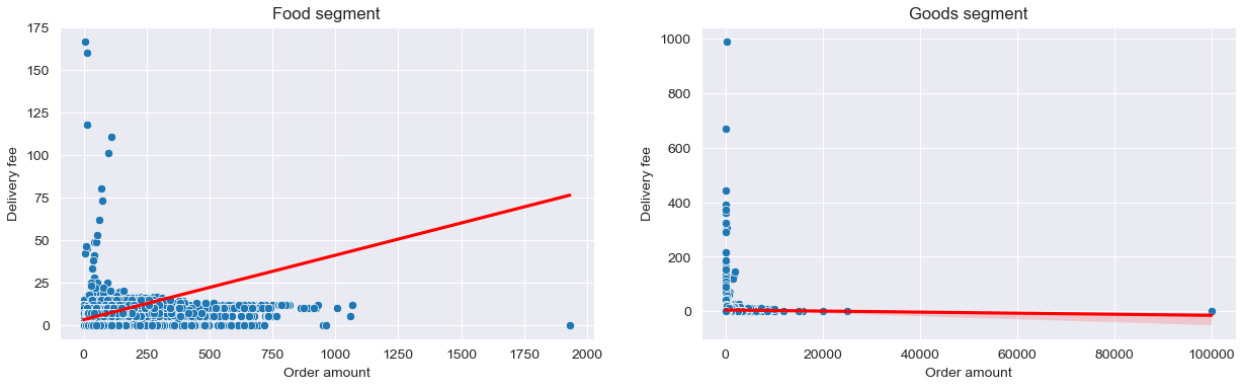
- Food segment = 7.46
- Goods segment = 4.90

Average earnings per delivery by delivery status:

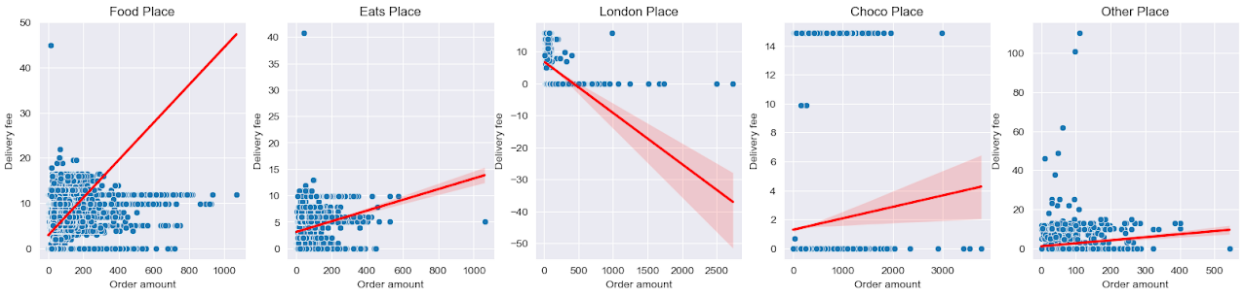
- Delivered = 7.22
- Delivering = 5.79
- Cancelled = 4.16
- In terms of the driver's earnings, it appears that the Good store segment has the lowest average earnings per delivery, while the Food segment has the highest.
- When it comes to delivery status, drivers earn the most on average for deliveries that have been successfully completed (Delivered), followed by deliveries that are currently being made (Delivering), and the least for deliveries that have been cancelled.

8. How does the order amount affect the delivery fee and delivery cost?

Checking the relationship between the order amount and the delivery fee for the store segments and the delivery channels



Store segments



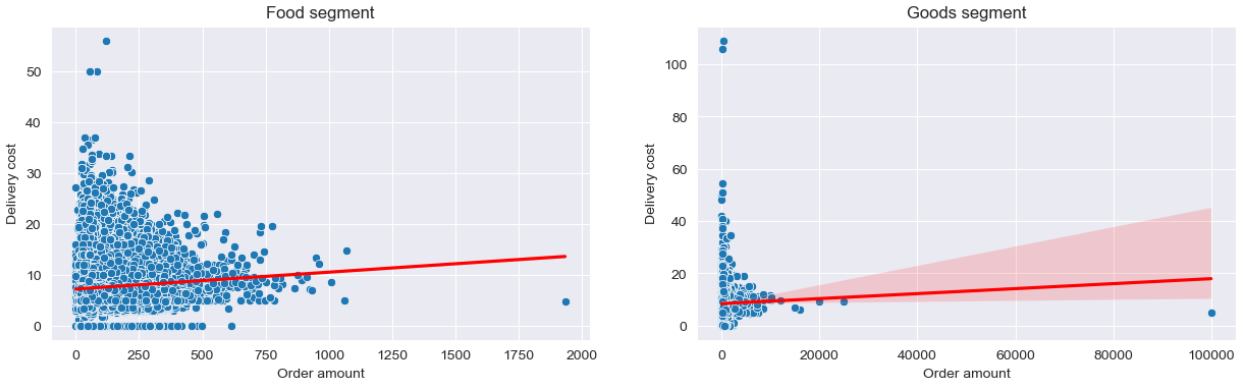
Delivery channels

Coefficients of correlation: Order amount vs Delivery fee

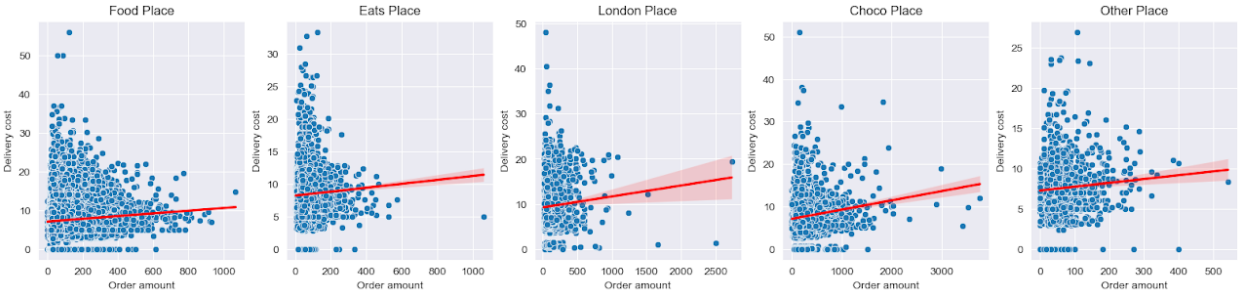
- Food place = 0.491
- Eats place = 0.128
- London place = -0.415
- Choco place = 0.036
- Other place = 0.144

- Based on the coefficients of correlation, it appears that there is a positive relationship between order amount and delivery fee for most of the delivery channels except for "London Place", where there is a negative relationship.
- However, the correlation coefficients are generally weak, indicating that the relationship is not very strong.

Checking the relationship between the order amount and the delivery cost for store segments and the delivery channels



Store segments



Delivery channels

Coefficients of correlation:

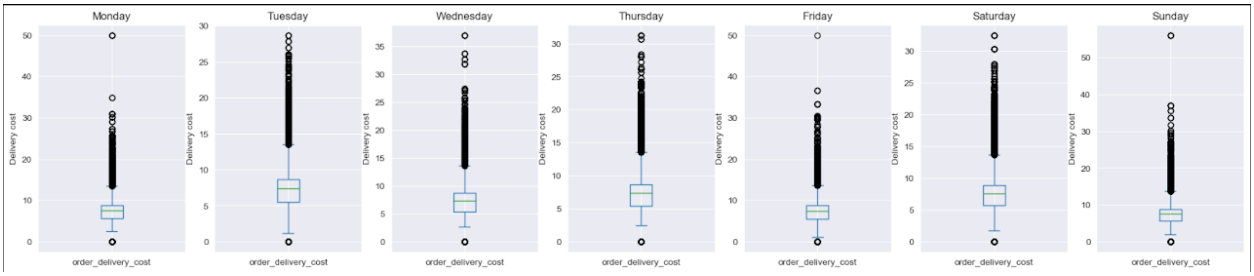
- Food place = 0.078
- Eats place = 0.043
- London place = 0.06
- Choco place = 0.114

- Other place = 0.07
- For the relationship between order amount and delivery cost, the correlation coefficients are also generally weak across all the delivery channels and segments.
- This suggests that there is not a strong relationship between order amount and delivery cost.

- -

9. How does the delivery cost vary with respect to the day of the week?

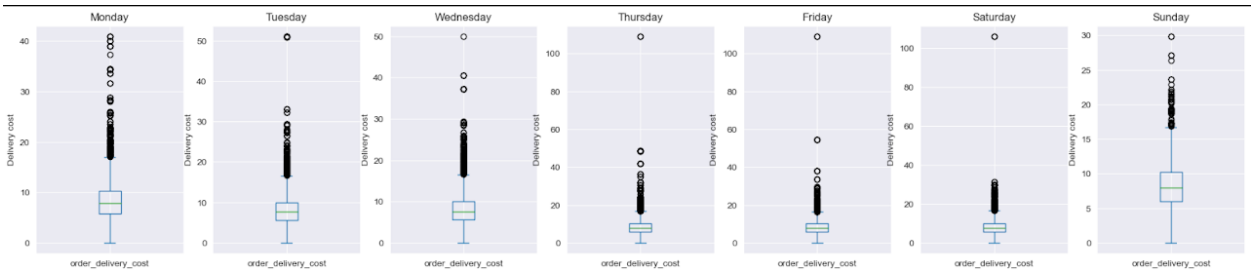
Checking the distribution of delivery cost throughout the weekdays for each of the store segments



Food segment

Average delivery cost for each day in the food segment:

- Friday = 7.44
- Monday = 7.46
- Saturday = 7.55
- Sunday = 7.56
- Thursday = 7.38
- Tuesday = 7.41
- Wednesday = 7.37



Goods segment

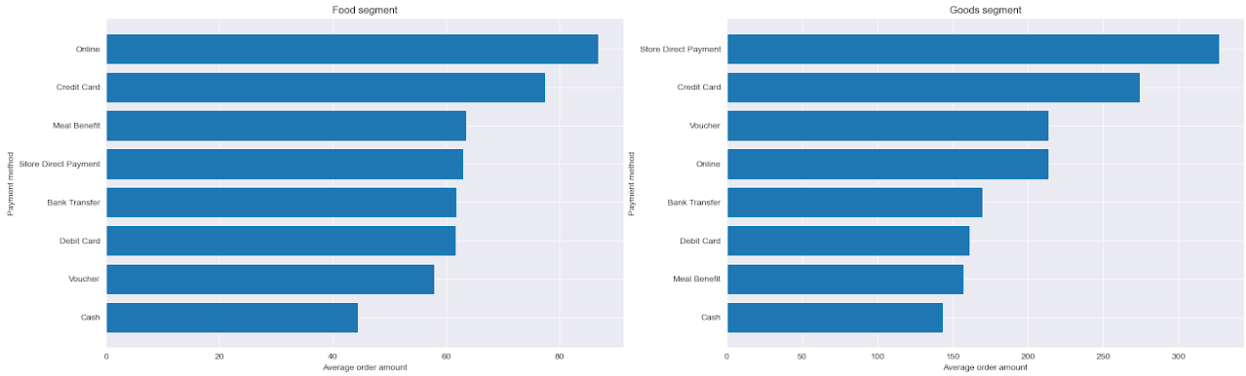
Average delivery cost for each day in the goods segment:

- Friday = 8.52
- Monday = 8.49
- Saturday = 8.33
- Sunday = 8.46
- Thursday = 8.60
- Tuesday = 8.33
- Wednesday = 8.28
- The average delivery cost doesn't vary significantly with respect to the day of the week, both in the food and goods segments.
- The average delivery costs incurred during each weekday are relatively similar across all days of the week, with only small fluctuations of a few cents between the highest and lowest values.
- This suggests that the delivery cost is not significantly affected by the day of the week.
- However, there are some orders where the delivery cost is high for certain days.

- -

10. Is there a significant difference in the average order amount between different payment methods?

Calculating the average order amount for each of the payment methods



Average order amount vs payment method for each store segment

- For the food segment, the average order amount for "Cash" is the lowest at 44.47, while the average order amount for "Online" is the highest at 86.91.
- For the goods segment, the average order amount for "Cash" is again the lowest at 143.70, while the average order amount for "Store Direct Payment" is the highest at 327.25.
- The differences between the average order amount across different payment methods appear quite substantial, suggesting that the choice of payment method may play a role in determining the order amount.

Conclusions

From the analysis, the following conclusions can be derived:

- The most popular product categories in the store are food and goods, with food being the most popular.
- The store's sales are highest on Fridays and Saturdays, indicating that weekends are the busiest time for the store.
- The average order amount in the store is higher for the goods segment compared to the food segment, which suggests that customers are willing to spend more on goods than on food.
- Online mode of payment is the most frequently used payment method by customers, followed by vouchers and debit cards.

- Customers who pay using online payment methods, such as credit card or online payment, tend to spend more on their orders than customers who pay using other payment methods.
- The delivery distance has a positive correlation with the delivery cost for most delivery channels, which means that longer delivery distances result in higher delivery costs.
- The average earnings per delivery are highest for the food segment, indicating that drivers earn more for delivering food orders compared to goods orders.
- The order amount has a positive correlation with the delivery fee for most delivery channels, which suggests that customers who place larger orders are charged higher delivery fees.
- The average delivery cost is highest on Fridays and lowest on Wednesdays for both the food and goods segments.

Suggestions

Based on the conclusions derived from the analysis the following are the possible suggestions for improving the business operations of Delivery Center:

- To increase the overall revenue, the store should focus on increasing the average order amount, as this has a strong positive correlation with the delivery fee and delivery cost. One way to achieve this is by offering bundle deals or discounts for larger orders.
- The store should also focus on promoting online payment methods, as they have the highest average order amount and are the most frequent payment method used by customers. The store can also introduce more payment options as well.
- The store should consider offering more delivery options, such as same-day delivery, to attract more customers and increase customer satisfaction.
- The store should also monitor the delivery distance and delivery status, as they have a significant impact on the driver's earnings and can affect the store's revenue in the long run.

- Since the delivery costs incurred during weekdays vary, the store should optimize its delivery schedules and routes to minimize the delivery costs.
- The store should also consider expanding its product offerings, especially in the goods segment, to attract a wider customer base and increase revenue.
- The store can consider partnering with other businesses, such as restaurants or grocery stores, to offer bundled deals or cross-promotions to increase customer engagement and revenue.
- Finally, the store should invest in marketing and advertising to increase brand awareness and attract new customers, especially in competitive markets.