**Data-Driven Management Assignment**

Name: Keshav Sanwar Sharma

Dataset: parcel\_48 & driver\_48

Word count: 2030

# **EXECUTIVE SUMMARY**

The Board of Directors has ordered data analysis to understand the online delivery system's gaps in an effort to boost sales value by reducing lost and return orders that never made it to customers.

These significant discoveries were made:

* Only £0.2 million (7.5% of the total order value of £3.8 million) is actually being converted into sales.
* Out of 45,530 order parcels, only 41% parcels get delivered to customers (includes parcel return to warehouse by customer and lost during returned by customer), rest are either lost during delivery transit or returned to warehouse without reaching to customers.
* 89% of the orders are marked as priority and the average delivery time is slightly more than orders with non-priority.

The main suggestion to the Board is to reorganize and redesign the delivery system, considering the specific information of the customer's available day and time for receiving the package as well as alternate options to drop the package at a safe location (drop box, neighbourhood) while placing the order, better tracking of order delivery status, and timely updating of the customer on the progress. In order to streamline the priority delivery system and improve customer experience through quick and on-time delivery, the company should take advantage of the need for priority delivery and set up premium priority options on a pay-per-use basis.

# **BACKGROUND**

This management report's objective is to develop a plan to enhance the business's parcel delivery system using the most recent parcel database. The biggest obstacle to the success of home delivery operations is the delivery issue when the customer is not at home. By being aware of the gap between a package being sent out for delivery and it being delivered, lost, or returned to the warehouse. We must comprehend the preferences and openness of our customers to receive packages. In order to make decisions to increase revenue and sales, it is also necessary to understand the company's market area (urban/rural) and market share.

It is crucial for an online retail business to deliver products to customers on time and carve out a niche for itself in the industry. Decisions in organizations are now driven by customer needs, so it is critical for businesses to know their customers and focus on them. Customer loyalty would ultimately increase if businesses catered specifically to their needs. Businesses can only develop services and offers that are valuable to customers when they try to get to know them and when they pay attention to their feedback. Our company can build customer trust by paying a little more attention to delivery issues, which will ultimately increase customer retention rates.

The methodology involved performing data analysis on the supplied primary data, which also included data visualization. Two sources are included in the dataset:

* the company's parcel database data, and.
* the company's provided driver database information.

1. **DATA PREPARATION AND DATA RESTRICTION USING R**

To produce a complete data set with 45530 entries, data cleaning was done prior to beginning the data analysis in order to remove inconsistent and invalid data. From the original data set of 50065 entries, a total of 4535 were eliminated, resulting in a general loss of about 9.05% of data.

541 and 87 NULL values were removed from priority delivery and parcel payment columns respectively. 600 invalid values were removed from parcel returned, 349 from parcel status, and 2,958 from time of delivery. (See appendix table 1 for further variable definition)

To analyse driver’s data, the driver’s dataset was joined with the parcel dataset. The newly formed dataset comprised of a total of 45530 entries.

**Table 1: Variable Definition**

| Variable | Metric |
| --- | --- |
| parcel value | Price of order/parcel in £. |
| parcel status | Categorical variable with the values “delivered”, “lost” and “return to warehouse” |
| parcel returned | Categorical variable with the values 0 and 1.  1 indicating parcel was return to warehouse and 0 means parcel not returned to warehouse |
| parcel arrived | Indicates the scheduled parcel destination arrival (same-day, or  between 1 to 5 days late). |
| priority delivery | Categorical variable with values “yes” and “no”, indicating parcel was ordered with priority delivery option in the last three months |
| delivery location | Categorical variable with values “rural” and “urban”, indicating parcel delivery location |
| time of delivery | Categorical variable with values “morning”, “afternoon” and “evening”, indicating parcel delivery time of the day |
| promotion purchase | Categorical variable with values “yes” and “no”, indicating parcel was ordered by applying promo code in the last three months |
| parcel payment | Categorical variable with values “voucher”, “store card”, “credit card” and “debit card”, indicating parcel payment was completed by any one option |
| Parcel return case | Categorical variable with values “Returned by Customer”, Not Returned by Customer”,” Lost during Delivery”, “Lost during Return” |
| values | Categorical variable comprising of price grouped as range:  "0-10£", "10-20£", "20-30£", "30-40£", "40-50£", "50-60£", "60-70£", "70-80£", "80-90£", "90-100£", "100-110£", "110-120£", "120-130£", "130-140£", "140-150£", "150-160£", "160£-above". |
| Parcel arrived day | Categorical variable comprising of parcel arrival day grouped as range:  "Same Day", "First Day", "Second Day", "Third Day", "Fourth Day", "Fifth Day" |

See appendix table 2 for further variable definition.

# **DATA ANALYSIS AND VISUALISATION USING R**

Fig 1: Parcel Value Distribution

Chart, histogram

Description automatically generated

Fig 1, illustrates, the average parcel price is £83, with a minimum of £1 and a maximum of £163. Most packages value between £70-100.

Fig 2. Percentage-wise Distribution of Parcel Status

Chart, pie chart

Description automatically generated

From fig 2, we can conclude that, 76.3% of the company's packages, worth £2.9 million, are evidently returned to the warehouse, parcel lost during transit accounts for value £6,14,021 (16.2%). The least are delivered packages that result in company sales is 7.5% of all packages with a value of £0.2 million.

It is obvious that the return to warehouse status of the parcel predominates. Let's examine the impact of additional factors, such as delivery location, delivery time, or priority orders, on the status of the parcel.

Let us see the market share of parcels according to delivery location.

Figure 3: Location-wise Split of Parcels

Chart, bar chart

Description automatically generated

The distribution of packages according to delivery locations is shown in Fig 3. Only 25% (11408) of total orders come from urban areas, with 75% (34122) of all parcels going to rural areas. Also parcel status distribution is equal in both the locations.

Chart, bar chart

Description automatically generated

Fig 4: Price distribution by Parcel Status

Fig 4 shows the parcel status according to price segments, starting from £10-£160, all 3-parcel status has equal distribution, so we can rule out the possibility that price affects parcel status.

**Priority Status**

| Priority | Yes | No |
| --- | --- | --- |
| 40,955 | 4,575 |

Table 2: Count of Priority Status of Total Orders

Fig 5: Average Delivery Time for Parcels According to Location and Priority Status

Chart, bar chart

Description automatically generated

Table 2 shows the split of priority status, 89% of orders are marked as priority, which makes it difficult to prioritize order delivery time. Furthermore, from Fig 5, we can see the average delivery time for priority orders is more than non-priority orders.

**Time of Delivery**

**Chart, bar chart

Description automatically generated**

Fig 6: Parcel Status by Time of Delivery

Since most of the orders are out for delivery during morning and evening time, most return and lost orders also occur same time. With more vacant homes during the day and typical delivery times between 8am and 5pm, the challenge is obvious. Clearly, you need to better understand when your customers want to deliver and meet those needs.

**Parcel Returned to Warehouse**

To understand the parcel returned column, let us derive new terminology. Below is the detail explanation.

| Parcel Status | Parcel Returned | New Terms |
| --- | --- | --- |
| Returned to Warehouse | 1 | Parcel available in warehouse after returned by Customer |
| Returned to Warehouse | 0 | Parcel available in warehouse but failed to reach to customer/ came back during delivery |
| Lost | 1 | Lost during return (initiated by customer) |
| Lost | 0 | Lost during delivery (to customer) |

Table 3: New Terms derived to understand parcel arrived coulmn

Chart, bar chart, waterfall chart

Description automatically generated

Fig 7: Distribution of parcel returned to warehouse (New Term)

From fig 7, it is clear that parcel return to warehouse without being delivered/failed to deliver is definitely higher than parcel returned by customers, in fact, it is 23% higher. Same goes with lost parcels, loss during delivery is 37% higher than lost during returns.

28% of the returns are initiated by customers. The possible reason could be they didn’t like the product or received damage product. Home deliveries that are flawed or damaged result in high costs for supply chain partners. Additionally, this results in serious client inconvenience in having to make arrangements for both the delivery of replacement goods and the return of the original goods.

Fig 8: Parcel Returned to Warehouse by Parcel Arrival Day

Timeline

Description automatically generated

The majority of the parcels arrive, as shown in Fig. 8, the day after the order is placed, followed by the second, third, fourth, and fifth days. Fewest orders arrive on the same day. Since the first-day order delivery rate is high, both the customer-returned and non-customer-returned return rates are also high.

Fig 9: Payment Method Used for Parcel Payment differentiated by Delivery Location

Chart, bar chart, treemap chart

Description automatically generated

From Fig. 9, it is clear that vouchers are the most widely used form of payment, accounting for 38% of all transactions. Store cards are second, with a share of 21%, and debit and credit cards are close behind with a share of almost equivalent.

Table 4: Display promotion used at time of order.

| Promotion Used | No | Yes |
| --- | --- | --- |
| 42257 | 3273 |

Since 92% of orders do not use a promotion code when placing an order, either the company does not offer many promotion codes, or customer awareness of promotions discounts is very low.

Fig 10: Driver's delivery count according to experience and work pattern

Graphical user interface

Description automatically generated with low confidence

# Part-time drivers deliver more than other drivers with different work schedules, and this is also because they have twice the strength of other drivers. Additionally, it has been found that drivers on zero-hour contracts who’s experience ranges from 4 to 8 years delivers more than drivers who work full time.

# **RECOMMENDATIONS**

| **OBSERVATIONS** | **RECOMMENDATIONS** |
| --- | --- |
| 76.3% of all packages are returned to the warehouse, of which 63% are returned without even reaching the customers. A parcel that is returned to the warehouse prior to delivery states that the delivery driver either was unable to locate the location or the customer could not receive the package at delivery address. | The customer should be given options to choose one of the most recent available dates/days and delivery times when placing their order, in order to make sure they are aware of their availability to receive the package. Customers should be advised of safe delivery options, such as drop boxes like those offered by Amazon or their neighbourhoods. Additionally, the company can collaborate with a few major gas stations and convenience stores that are open 24 hours a day, have numerous locations across the nation, and allow customers to pick up their purchases from these establishments whenever it is most convenient for them.  Customers should receive regular updates on the status of their packages, and on the day of delivery, they should be given information about the van and a live tracking ID to make sure they are prepared. |
| 89% of customers choose priority delivery (Table 2) however priority and non-priority delivery time is almost same (Fig 5). | In order to introduce premium priority service on a pay-per-use basis based on the earliest time demanded at the time of placing the order, the company should take advantage of the fact that the majority of customers choose priority service. Additionally, only open slots should be shown to accommodate and provide real-time delivery and invest in delivery partners to meet this requirement. In turn, brand loyalty will increase. |
| 28% of the returned to warehouse parcels are returned by customers. | Reviews and ratings from customers should be gathered in order to comprehend the needs, trends, and degree of product standard of the market. In addition, more improved or safety measure should be used in packaging orders which will lead to less damage to products. |
| 16% of the orders are lost during delivery or during returns. | For the company to be able to claim the product price from the insurance company and avoid losses, all products should be insured against loss and theft. |
| Since our market share is 75% in rural area, it is a good opportunity to enhance and expand more in terms of sales | In rural areas, we can advertise store cards that function like credit cards, encouraging customers to make larger purchases. |
| In urban area, company’s market presence is 25% less than rural area. | New marketing strategies should be implemented by the company by offering more promotions and discounts to draw in urban customers in order to increase sales and popularity in the urban area. |
| Part-time and zero-hour contract drivers delivers more no. of delivery than full time drivers | More drivers should be hired for Part-time and zero-hour contract for normal operations and full-time drives should be utilized more for premium priority delivery service as they are available full time for the company. |

# **Further Investigations and Limitations**

# There are some restrictions in the datasets used for this report, which may have an impact on its conclusions. They are listed as follows:

# The likelihood of the product being purchased by the customer or any product reviews from customers were not disclosed.

# There is no explanation provided for products that are lost during delivery, returned to the warehouse, or lost during customer returns.

# With the exception of the product's price, no information was offered to help consumers comprehend the delivery, loss, and return patterns of packages.

# For the purpose of analysing consumer behaviour and purchasing patterns, no customer information is provided.

# **DASHBOARD**

Chart, bar chart, treemap chart

Description automatically generated

# **APPENDICES**

**Table 1: Further Cleaned Data Explained**

Below mentioned data were removed from the columns.

| Coulmn | Value | Entry count |
| --- | --- | --- |
| Parcel returned, | -122 | 600 |
| Parcel status | Pill | 349 |
| Time of delivery | Post | 2958 |

**Table 2: Further Variable Definition**

Variables define for creating Fig 2 and Fig 10

| Variable | Metric |
| --- | --- |
| parcel status count | Count of parcel according to parcel status |
| Parcel sum value | Sum of Count of parcel value according to parcel status in £ |
| Perc | To calculate percentage of each parcel status |
| Labels | To display % in each parcel status |
| Driver delivery count | Driver’s deliver count according to experience and work pattern |

Chart, bar chart

Description automatically generated

Chart, bar chart

Description automatically generated