Data-Driven Management assignment

Name: Amusan Ibrahim Oyelola

Title: Data-Driven Management Individual Coursework.

Dataset: driver\_31 and parcel\_31

Word count: 2,135

# **Executive Summary**

This report aims to identify the main factors that influence the driver’s performance in terms of value generated from delivered parcels. The study focused on four variables: work pattern, van type, time of delivery, and gender. The analysis was conducted on a dataset of 48,032 observations, and 14 variables were used in the analysis.

The result of the study indicates that work pattern is a significant factor that influences a driver’s performance. Drivers who work part-time tend to generate higher values from delivered parcels compared to those who work full-time and zero-hours contract. The study also found out that the type of van plays a crucial role in the driver’s performance. Drivers who use large vans tend to generate more value from delivered parcels because of the ability to carry larger quantities of parcels.

Furthermore, the study found that the time of delivery also plays a significant role in the driver’s performance. Those drivers that delivered parcels in the morning and evening generated higher value from parcels delivered than those delivered in the afternoon. The study also found that gender showed a positive correlation with the total value generated from parcels delivered. Male drivers tend to generate more value from the total parcel delivered than females.

The report concludes with recommendations for the company to utilize the identified areas of strength, such as the best delivery windows, larger vans, lost parcel tracking, and training for beginner drivers, to improve their efficiency and profitability.

# **Background**

The key business question that will be answered in this report is: “What are the main factors that influence a driver’s performance in terms of value generated from delivered parcels?”. This question will be measured based on the information provided in the dataset as follows:

1. Is there a correlation between the driver’s experience and the value generated from the delivered parcel?
2. Is there a connection between the driver’s work pattern and the value generated from the delivered parcel?
3. Is there a correlation between the type of van a driver use for work and the value generated from the delivered parcel?
4. Is there a connection between the Time of delivery of the parcel and the value generated from the delivered parcel?
5. Is there a connection between the Gender of the driver and the value generated from the delivered parcel?

These questions will be used to analyze the main business question and clearly come to a conclusion on factors that have impacted the performance of the driver in terms of the value generated from the delivered parcel.

# **Data Preparation and Variable Definition**

This section will discuss various aspects of data preparation such as setting up the R-script by calling necessary libraries, reading the dataset, cleaning data, and performing data transformation to generate necessary variables for analysis.

* 1. **Data Cleaning.**

The R-script was set up by calling the necessary libraries, setting the working directory, and the datasets were read and assigned accordingly. The map function () was used to check the total number of NA values in each dataset across the entire column at once and it was discovered that priority delivery and parcel payment were the only variables with NA values while there were no NA values in the driver dataset. The str () was used to check the nature of observations in both variables to decide whether to drop the NA values or replaced them. The NA values in priority delivery were replaced with “no” because if its value is not “yes” that is the parcel was on priority delivery then it is not, and those NA values for parcel payment were dropped. Outlier was detected and removed from the parcel returned variable. Finally, the cleaned parcel dataset was joined with the driver dataset using full\_join () function and called data\_driver\_parcel\_joined.

Descriptive analysis was carried out on the joined dataset to check for outliers and no outlier was detected as shown in appendix 1. A further check was carried out using boxplot, especially on the parcel returned variable and it shows there was no outlier in it. The number of observations in the joined data was reduced to 48,771 from 50,065 after cleaning. From this joined dataset I created a sub-dataset called driverCategory\_data with 48,032 observations which is the final dataset for analysis as stated on the R-script.

* 1. **Data Transformation.**

This section will describe some important variables used in the analysis. The business question is focused on parcel value generated from delivered parcels, hence, for the purpose of this analysis the following variables were generated.

Experience level: The experience variable in the original dataset was used to categorize drivers into three categories. These categories are Beginner, Intermediate, and Expert.

Total parcel value: This is the sum of all the parcel values under consideration. This is important in order to understand the total value (in £) generated by the parcel delivered and the total sum of value from those lost and returned to the warehouse.

Percentage total parcel value: This is the percentage of the total parcel value by parcel status. This is important to understand the percentage contribution of each parcel status in the total parcel value generated. This will clearly show areas of strength and weakness in the analysis.

Percentage delivered: This is the percentage of the total parcel delivered by each category of the driver. This is important to understand the contribution of each category of driver.

Percentage lost: This is the percentage of the total parcel lost by each category of the driver. This is important to understand the contribution of each category of driver.

Percentage returned: This is the percentage of the total parcel returned to the warehouse by each category of the driver. This is important to understand the contribution of each category of driver.

Table 1: Variable definition

|  |  |
| --- | --- |
| Variable | Metric |
| Parcel value | Parcel value of each parcel, measured in £. |
| Gender | Categorical variable with the values “Male” and “Female”. |
| Time of Delivery | Categorical variable with the values “Morning”, “Afternoon” and “Evening”. |
| Driver id | Driver identifier number with values from 1 to 100. |
| Experience | Indicate the years of experience of the driver with values from 1 to 10 |
| Experience level | Categorical variable with the values “Beginner”, “Intermediate” and “Expert”. |
| Parcel status | Categorical variable with the values “Delivered”, “Lost” and “Returned to warehouse”. |
| Work pattern | Categorical variable with the values “Full time”, “Part time” and “Zero hours contract”. |
| Van type | Categorical variable with the values “Large van”, “Medium van” and “Small van”. |
| Total parcel value | The total sum of parcel value, measured in £ |
| Percentage total parcel value | The percentage of the total parcel value, measure in £ |
| Percentage delivered | The percentage of the total parcel delivered, measured in £ |
| Percentage lost | The percentage of the total parcel lost, measured in £ |
| Percentage returned | The percentage of the total parcel returned to warehouse, measured in £ |

# **Analysis**

**3.1 Distribution of the Parcel value.**

The main objective of an organization is to make a profit, which is why it is crucial to analyze the distribution of parcel values across parcel status to evaluate the performance of each category. Table 3.1 illustrates the percentage contribution of each parcel status, indicating that parcel delivery has the lowest contribution of 6%, while parcel return to the warehouse has the highest contribution of 84%. This suggests that the company's delivery operation has some shortcomings, and Figure 3.1 provides a graphical representation of this distribution.

Table 3.1: This table shows the percentage of the total parcel value for each category in parcel status.

Graphical user interface, application

Description automatically generated

Figure 3.1: This graph shows the percentage of the total parcel value for each category in parcel status.

Chart, bar chart

Description automatically generated

The success of a delivery operation depends on the coordination and effectiveness of various factors such as drivers, logistics team, and support staff involved in the operation. The study investigated if the driver is responsible for the lost and returned parcels using their experience level and it was discovered that the most experienced among the driver recorded most parcels lost and returned to the warehouse as shown in figures 3.2 and 3.3 respectively (The data table is in appendix 2 and 3 respectively).

Figure 3.2: This graph shows the relationship between Driver’s experience level and the Parcel value lost.

Chart, bar chart

Description automatically generated

Figure 3.3: This graph shows the relationship between Driver’s experience level and the Parcel value returned to the warehouse.

Chart, bar chart

Description automatically generated

The study concludes that the driver's experience level is not the reason for lost and returned parcels. The poor delivery performance of the company could be caused by factors such as incorrect information given to the driver or unavailability of the recipient during delivery.

Also, the study investigated the relationship between a driver's experience level and the total parcel value delivered. The findings revealed that there is a positive correlation between the driver's experience level and the total parcel value delivered, meaning that the higher the experience level the higher the total parcel values delivered. Table 3.2 below shows that the expert, intermediate, and beginner drivers contributed 40%, 32%, and 29% respectively. Figure 3.4 shows a clear picture of this contribution based on experience level.

Table 3.2: This table shows the percentage of total parcel value delivered for each category of Driver.

Graphical user interface, application

Description automatically generated

Figure 3.4: This graph shows the relationship between Driver’s experience level and the Parcel value delivered.

Chart, bar chart

Description automatically generated

The study further investigated factors that affect this impact of drivers based on their experience level to improve the total parcel value delivered and the organization's profitability. Factors such as the work pattern, type of van used for delivery, time of delivery, and gender of the driver were analyzed.

**3.2 Impact of Work Pattern on Total Parcel Delivered.**

The work pattern of the driver was investigated, and it was discovered that part-time drivers delivered more parcel value than other categories (full-time and zero-hours contract). For the beginners, intermediates, and experts, part-time drivers recorded 58%, 45%, and 47% respectively as shown in table 3.3 below. Hence, the graphical representation in Figure 3.5, clearly shows that the work pattern of the driver affected the total parcel delivered with part-time being the best pattern that works best for all the categories of the driver.

Table 3.3: This table shows the impact of work patterns on the total parcel value delivered.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 3.5: This graph shows the relationship between Driver’s work pattern and the Parcel value delivered.

Chart, bar chart

Description automatically generated

**3.3 Impact of Time of Delivery on Total Parcel Delivered.**

The impact of time of delivery on the total parcel value delivered was examined, and the analysis found that more parcels were delivered in the morning and evening across all driver categories. The study suggests that customer availability in the morning and evening affects delivery volume, indicating that the logistics team should plan driver journeys accordingly and avoid afternoon deliveries, which may result in returned or lost parcels. Table 3.4 and Figure 3.6 provide numerical and graphical representations of the findings.

Table 3.4: This table shows the impact of Time of delivery on the total parcel value delivered.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 3.6: This graph shows the relationship between the Driver’s time of delivery and the Parcel value delivered.

Chart, bar chart

Description automatically generated

**3.4 Impact of Van type on Total Parcel Delivered.**

The study also explored whether the type of van used by drivers had any correlation with the total value of parcels delivered. According to Table 3.5, drivers using large vans contributed over 60% of the total value of parcels delivered across all driver categories.

Table 3.5: This table shows the impact of the type of Van on the Total parcel value delivered.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 3.7 provided a more pronounced visualization of the relationship between the type of van used by drivers and the total value of parcels delivered. Additionally, the analysis revealed that drivers using medium vans contributed the lowest percentage of total parcel value delivered. The findings suggest that the company could focus on effectively utilizing its resources by prioritizing the use of large vans.

Figure 3.7: This graph shows the relationship between the Driver’s time of delivery and the Parcel value delivered.

Chart, bar chart

Description automatically generated

**3.5 Impact of Gender on Total Parcel Delivered.**

Table 3.6 and Figure 3.8 explore the correlation between drivers' gender and the total parcel value delivered. Male drivers across all driver categories delivered over 80% of the total value of parcels delivered, while female drivers delivered an average of 9.6% of the total value of parcels across all driver categories. Figure 3.8 provides a visual representation of the relationship between gender and the percentage of total parcel value delivered.

Table 3.6: This table shows the impact of Gender on the Total parcel value delivered.

A screenshot of a computer

Description automatically generated with medium confidence

Figure 3.8: This graph shows the relationship between the Driver’s gender and the Parcel value delivered.

Chart, bar chart

Description automatically generated

This report revealed that the total parcel value delivered was impacted by factors such as work pattern, van type, time of delivery, and gender. Therefore, it is crucial for the company to optimize its operations based on these findings to improve its delivery performance.

# Recommendations.

The following are recommendations based on the above analysis:

1. Delivery operation: The company should put in place an effective tracking and tracing mechanism and take appropriate steps to locate, and re-cover lost parcels.
2. Van Type: The company should consider investing more in large vans. This will help to increase the number of parcels that can be delivered per trip, reducing the number of trips required to deliver the same volume of parcels.
3. Work patterns: The company should further investigate why full-time drivers are not the highest-performing drivers. However, priority should be given to part-time drivers while the investigation is going on. This will allow the company to improve its profitability.
4. Time of delivery: The company should make use of the optimal delivery windows identified in this report (Morning and Evening) to increase the number of parcels delivered per driver.
5. Gender: This report revealed that male drivers performed better than female drivers. However, further research needs to be carried out to ensure the company is not gender biased.
6. Training: Training should be provided for beginner drivers to improve their effectiveness on the job. Figure 3.4 above (figure 2 on the dashboard), shows that if more training is provided for the low-experience driver (Beginner), this will improve their efficiency and increase the total parcel value delivered.

# Dashboard

Graphical user interface, application

Description automatically generated with medium confidence

# Appendices

Appendix 1: This table shows the descriptive statistics of the cleaned joined dataset.

Graphical user interface

Description automatically generated

Appendix 2: This table shows the percentage of total parcel value lost for each category of Driver.

Table

Description automatically generated with medium confidence

Appendix 3: This table shows the percentage of total parcel value returned to the warehouse for each category of Driver.

Graphical user interface, application

Description automatically generated