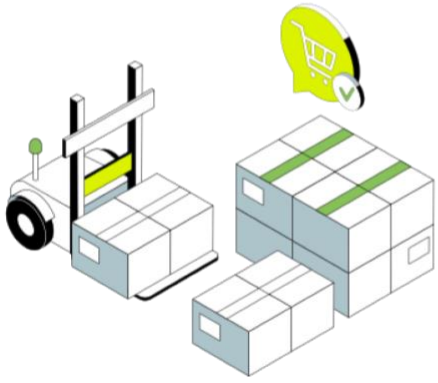
DISTRIBUTION CENTER 79

TRUCK CYCLE TIME ANALYSIS

PUNKT DC 79



Uliana Rybina
Business Planning
rybinaulianaa@gmail.com
6274953

Sander Lardinois
Process Control
sanderlardinois03@gmail.com
6285027

AJ Grant
Business Informatics
aj.grant6@icloud.com
6337349

Disclaimer (not a part of the report):

1. PUNKT is the branding name of the company to which distribution centre 79 belongs.
2. We assume that the management board is aware of how the business is running and knows crucial KPIs and metrics, like weight cycle time. Still, definitions are provided as a reminder.
3. As this report would be also used by the warehousing team and logistics supply and demand planners, we still use some technical terms, like accuracy and linear regression. These metrics are explained when used; moreover, such graphs as correlation matrices are assumed to be read correctly (an explanation is provided). All other data-related technical documentation is present in Appendix.
4. Green is the corporate colour, therefore the whole report is created keeping this colour in mind.
5. Here you can find the versions of the models, and KNIME docs: Case 3 Data - Versions.

EMAIL:

To: managementboard@gmail.com

From: dc79_analytics@gmail.com

Theme: DC79 Cycle Time Analysis - Final Report

Dear Management Board,

Hereby, we send you the final report of Distribution Center 79, where truck cycle time was analyzed (see the file attached). The results are also available at SharePoint in Teams.

The key findings are:

1. The manufacturer from which products are transported has the highest correlation with the predicted weighted cycle (manufacturer 5 has the lowest average weighted cycle time and manufacturer 13 has the highest weighted cycle time).
2. Average temperature, weight, fuel prices, the number of blocks in the truck and the number of SKUs[1] also correlate with the weighted cycle time. Weekdays, holidays and Covid-19 infection numbers do not correlate with weighted cycle time.
3. The trucks that belong to the cluster with the lowest average weighted cycle time take an average of 2.1 minutes for the pallet they unload. For the trucks in the cluster with the higher average weighted cycle time, it takes 5.2 minutes per pallet to unload. For the remaining trucks that belong to the medium cluster, the average weighted cycle time is 3.3 minutes per unloaded pallet. The most important factors which correlate with which cluster the truck belongs to are weight and SKUs.

[1] Number of different items in the truck.

It is recommended to continue gathering the data materials to increase the accuracy of the model. This mainly includes keeping track of the arrival and departure date of each truck, as well as their manufacturer, transporter, and cargo.

In case you have further questions, please email AJ Grant (aj.grant6@icloud.com), DC79 Business Informatics Specialist.

Kind regards,
DC79 Analytics Team

Table of Contents

Time Cycle Analysis

Goal & Results

There are two goals of the research:

1. Predict the time a truck spends at the distribution center 79 (DC79).
2. Determine the factors which influence the time a truck spends at DC79.

These findings would be used by the DC79 warehousing team to determine how many trucks should be at DC79 at each moment of the day and the number of employees being scheduled for a specific date. By adapting personnel shifts and regulating the truck movement, the PUNKT could save costs and increase efficiency.

The prediction scheme:

First, trucks are divided into three clusters with high, medium, and low weighted cycle times.[2] The model attempts to predict which cluster each truck belongs in; after that, it attempts to predict the total cycle time of that truck. This flow of data is summarised in Fig 1.

The following formulas represent cycle time and weighted cycle time.

- $\boldsymbol{Cycle\ Time\ =\ truck\ enters\ gate\ DC79\ -\ truck\ leaves\ gate\ \ DC79}$
- $Weighted\ Cycle\ Time\ = \frac{Cycle\ Time}{Amount\ of\ pallets\ in\ the\ truck}$; Weighted Cycle time is used as a score of how quickly a truck unloads each pallet on average. *The lower the number, the better.*

[2] The weighted cycle time is a score of how efficiently trucks can drop off their goods at the distribution centre. A truck in the high-weighted cycle time cluster is less efficient than the ones in the medium and lower-weighted cycle time clusters.

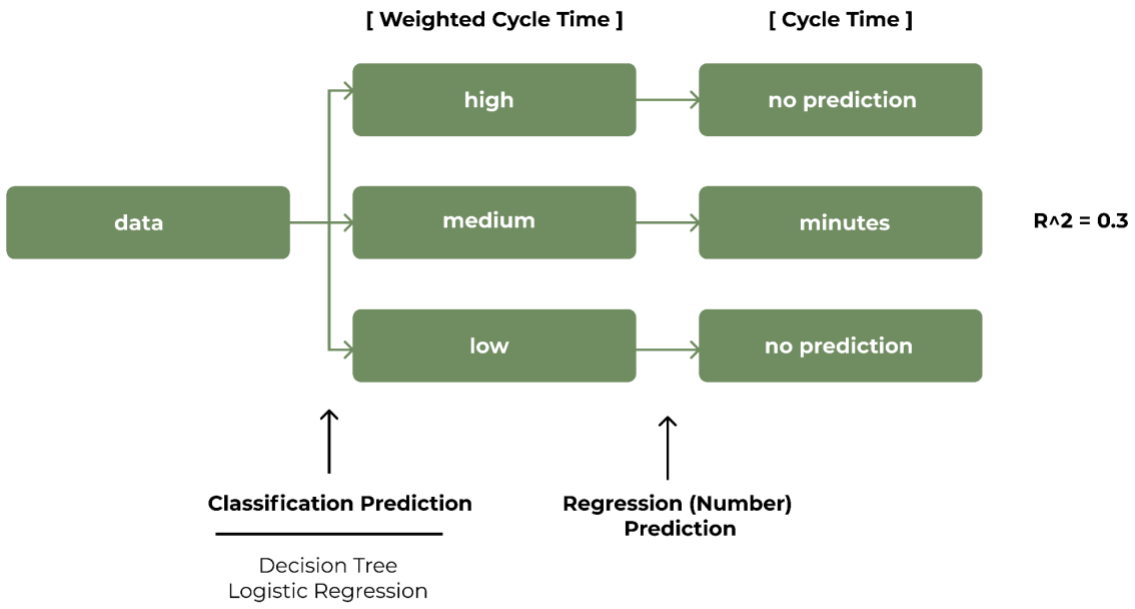


Fig 1. The scheme of the data analytics process used for time prediction and cluster analysis.

Weighted Cycle Time Group: Prediction

It is difficult to predict the weighted cycle time of each truck due to the small sample size and the number of missing values the dataset has. The highest certainty of the data (accuracy) was obtained by the random forest model - ~54%. Although the accuracy is low and thus insufficient to draw solid conclusions, the factors which correlate with weighted cycle time can be observed via the decision tree model.

1. The manufacturer from which products are transported has the highest correlation with the predicted weighted cycle time. Here, manufacturer 5 has the lowest average weighted cycle time and manufacturer 13 has the highest weighted cycle time.
2. Average temperature, weight, fuel prices, the number of blocks in the truck and the number of SKUs also correlate with weighted cycle time. Weekdays and Covid-19 infection rates do not correlate with weighted cycle time.
3. The KNIME models (see appendix) can be implemented into a tool that the logistics department can utilize with new data in order to predict the cycle time of any amount of trucks.

4. The trucks that belong to the cluster with the lowest average weighted cycle time take an average of 2.1 minutes for the pallet they unload. For the trucks in the cluster with the higher average weighted cycle time, it takes 5.2 minutes per pallet to unload. For the remaining trucks that belong to the middle cluster, the average weighted cycle time is 3.3 minutes per unloaded pallet. The factors which influence which cluster the truck belongs to are weight and SKUs [3].

It is recommended to continue gathering the data materials to increase the accuracy of the model. This mainly includes keeping track of the arrival and departure date of each truck, as well as their manufacturer, transporter, and cargo.

The presented internal and external data was cleaned and saved as a new database, which was used to determine factors which correlate with the weighed cycle time.

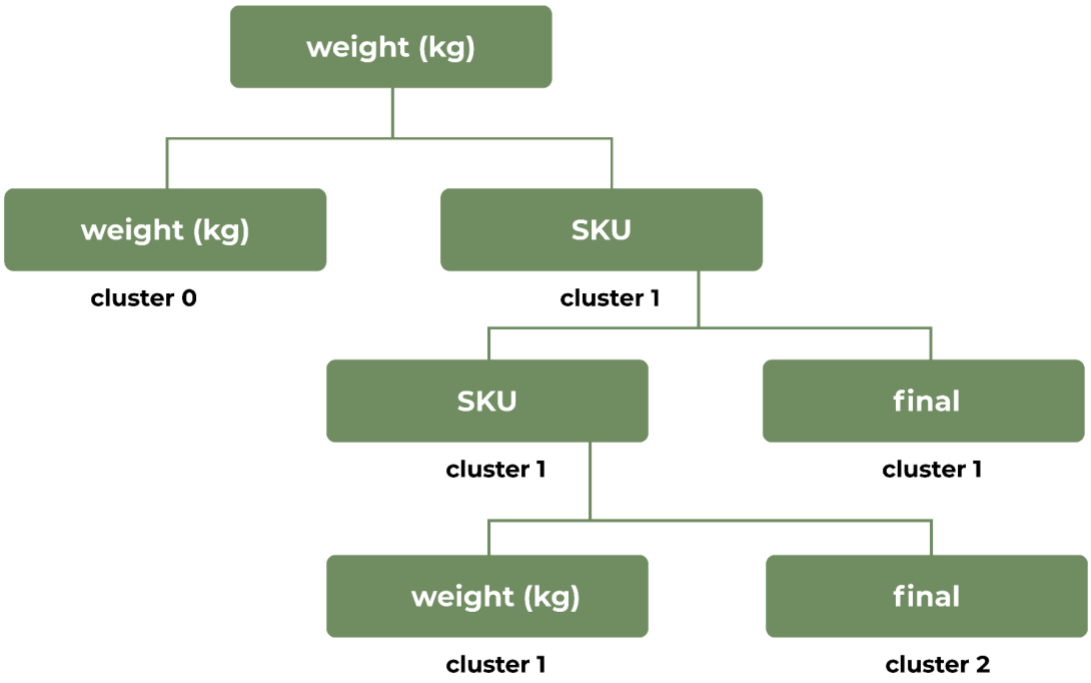


Fig 2. Factors which correlate with the weighted cycle time

[3] Number of different items in the truck.

Using the simple decision tree, it was found that the main factors which correlate with the prediction of the weighted cycle time clusters are SKUs (stock-keeping units) and weight.

Cycle Time: Factors

The factors which correlate with the cycle time were evaluated using KNIME decision trees and the linear correlation matrix. For the methodology, please review the Technical Documentation Section.

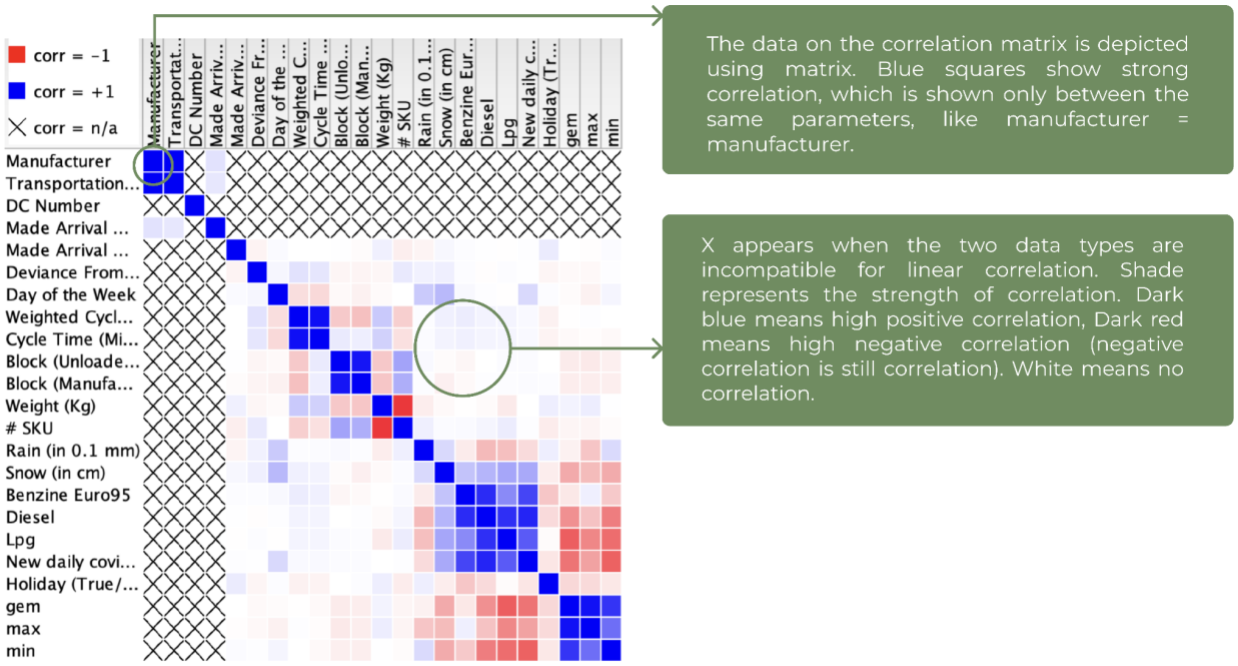


Fig 3. Correlation matrix explanation.

The linear regression analysis showed that there is not enough data to accurately predict the cycle time of a truck to be used by PUNKT (the parameter R^2 used in linear regression is less than 0.5; to confidently draw conclusions, it must be at least higher than 0.5). Additionally, by analyzing the linear correlation matrix between the cycle time and various variables from the dataset, we can see that none of the variables has a remarkable correlation to the cycle time. Therefore, we stress the importance of further data gathering to increase the certainty of the model.

Although the R^2 value for our predictions was low, the R^2 value for cycle time predictions of trucks in the medium-weighted cycle time cluster was significantly

higher than in other clusters (0.367). Although this is not high enough to exceed the 0.5 standards, it does suggest that trucks with a medium efficiency are much more consistent in the time they take than other trucks.

The team analyzed external and internal factors which could potentially correlate with the cycle time. These external variables include public holidays and weekdays, temperature and precipitation, fuel prices and the daily number of new Covid-19 cases. Factors such as the manufacturer from which products are transported, average temperature, weight, fuel prices, and the number of blocks seemed to also correlate with cycle time.

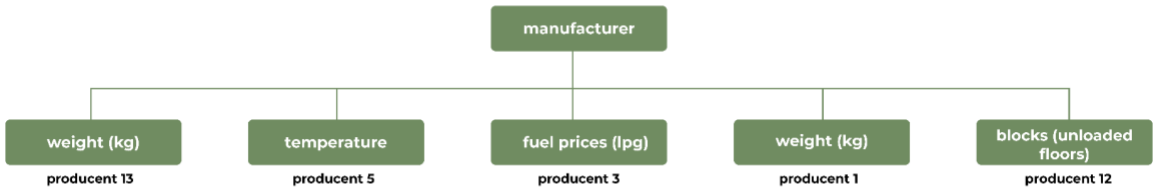


Fig 4. The decision tree showing the factors which correlate with cycle time.

Factors (sorted descendingly)	Does it correlate with cycle time or weighted cycle time?	Further tracking?
Manufacturer	Yes	Yes
Weight (kg)	Yes	Yes
Temperature and Precipitation	Yes	Yes
Fuel prices	Yes	Yes
Public Holidays and Weekdays	No	Yes
Covid-19 Data	No	No
SKUs	Yes	Yes

Manufacturer

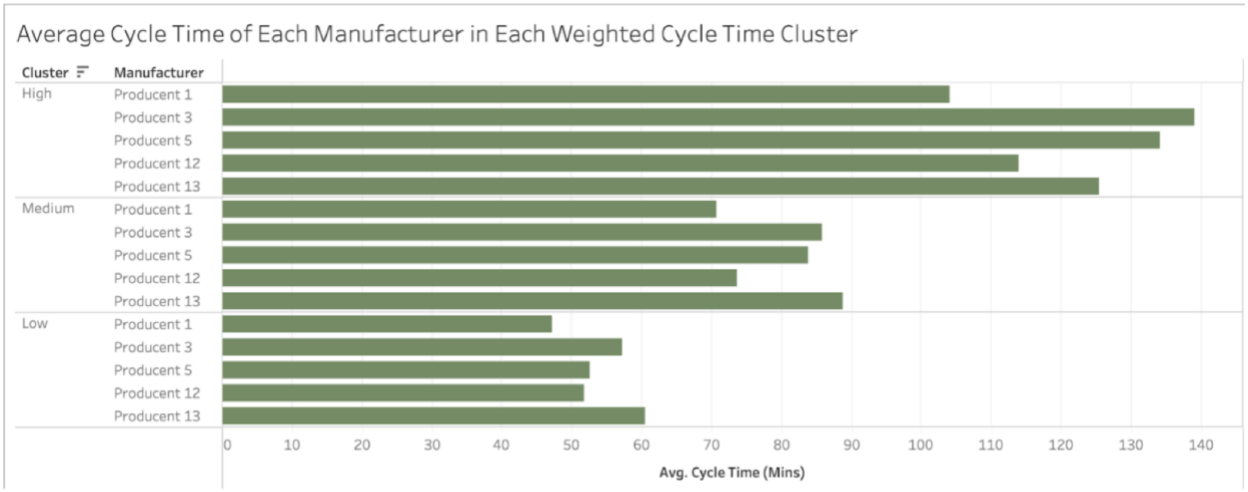


Fig 5. Average cycle time of each manufacturer in each weighted cycle time cluster

The manufacturer from which products are transported has the highest correlation with the predicted cycle time. This could be explained by the following reasons:

1. The manufacturer has different agreements with the company regarding what products it produces; therefore, some producents have products with a higher number of pallets, which can affect the cycle time.
2. The manufacturers may produce different products of different sizes and the products of some manufacturers might be more difficult to unload than those of others (due to weight, size or fragileness), this may result in a higher cycle time.
3. Some manufacturers put more different products (SKUs) in a truck than others. Due to the high variety of cargo that some manufacturers have, it might be more complex to unload the truck and thus the cycle time can be higher.

The average cycle time for each of the main manufacturers is presented in Figure 5. Based on this graph, it is possible to conclude that producent 5 on average has the lowest cycle time and producent 13 has the highest cycle time.

Public Holidays and Weekdays

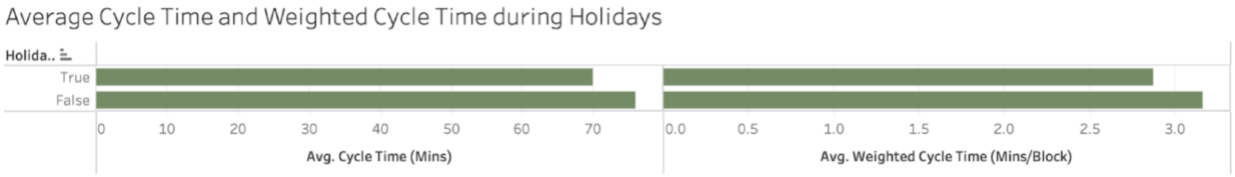


Fig 6. Average cycle time and weighted cycle time during holidays

Public Holidays and Weekdays did not show any significant correlation with the cycle time parameter. Still, it is necessary to continue monitoring this parameter due to several reasons:

1. On public holidays and weekends, there is often a change in traffic patterns, which could affect cycle time. Increased leisure activities and reduced work-related travel can lead to altered traffic conditions. This could cause spikes in the number of trucks at the same time at the distribution centre, thereby causing the cycle time to increase.
2. During public holidays and weekends, the demand for certain goods and services may vary. This may influence the number of trucks at the distribution centre and thereby change the cycle time.
3. Public holidays and weekends may impact the availability of labour resources. The capacity of a distribution centre may therefore be lower on a public holiday, causing the cycle time to increase when the delivery schedule is not adapted to this lower amount of workers or when delivery peaks occur.

Temperature and Precipitation

From the linear correlation matrix (fig 3), we can see that temperature and precipitation influence the cycle time. It is necessary to track this parameter on a weekly basis and take into account meteorological data in supply and demand planning due to the following reasons:

1. Extreme temperatures and precipitation can impact road conditions. In cases of heavy rain, snow, or icy conditions, the road surfaces may become slippery or unsafe for driving. This can result in reduced traction, slower speeds, and

potentially longer travel times for trucks approaching or leaving the distribution centre.

2. Inclement weather conditions, such as heavy rain or snowstorms, can lead to traffic congestion and delays. Reduced visibility, slower driving speeds, and potential accidents on the road may result in increased traffic volume or bottlenecks. As a consequence, trucks arriving at or departing from the distribution centre may encounter longer wait times or face delays caused by congested roads that could increase the cycle time.

Fuel Prices (LPG)

Fuel Prices showed some correlation with cycle time, as seen by the linear correlation matrix. This correlation can be explained in two following ways:

1. Higher fuel prices might cause transport companies to drive less often and make them wait until a truck is completely filled with goods before it is sent out. Therefore, trucks might drive less often, causing the number of trucks a distribution centre has to handle to decrease. This might decrease average cycle time as waiting times for trucks could decrease.
2. Trucks might only still drive when they are completely filled. This may increase the difficulty of unloading and loading the truck as personnel have to make sure that all the goods fit in perfectly and are handled carefully. This could increase cycle time.

Therefore, a monthly or quarterly report from the Sales team on fuel and gas prices would contribute to the higher efficiency of the DC79.

Covid-19 Epidemiologic Data

COVID-19 did not show any correlation with the cycle time. Especially after the WHO statement on the end of the COVID-19 pandemic, it is recommended not to track the COVID-19 data, as it has little or no effect on cycle time. Still, the HR data about illnesses would be beneficial for the planning teams.

Conclusion

The current findings are:

1. The manufacturer from which products are transported has the highest influence on the predicted cycle time and weighted cycle time. Here, manufacturer 5 has the lowest average cycle time and weighted cycle time and manufacturer 13 has the highest cycle time and weighted cycle time.
2. Average temperature, weight, fuel prices, and the number of blocks in the truck also affect the cycle time. Weekdays and Covid-19 do not affect weighted cycle time.
3. The KNIME models (see appendix) can be implemented into a tool that the logistics department can utilize with new data in order to predict the cycle time of any amount of trucks.

It is highly recommended to continue gathering the data, as the linear regression analysis showed that there is not enough data for accurately predicting the weighted cycle time of a truck to be used by PUNKT. Still, the model can be useful for predicting the cycle time of each truck, as its accuracy exceeds random chance. Therefore, we strongly believe that this model can be used for time prediction with additional in-depth research.

Technical Documentation

Hereby we present the methodology and technical details of the prediction time model and factors which influence the weighted cycle time.

Summary

After the data was cleaned and prepared, the entire remaining dataset (735 instances) was classified into 3 groups based on the weighted cycle time parameter. Using these clusters, a linear regression analysis was conducted on both the individual clusters and the entire remaining dataset in order to determine the accuracy of the classifier predicting the cycle time of a truck correctly. Besides the linear regression, a cluster prediction was conducted using basic decision trees and the random forest method.

Classification Model	Accuracy
Logistic Regression	50.679%
Decision Tree (Gini Index)	50.679%
Decision Tree (Gain Ratio)	49.321%
Random Forest	54.751%

Linear Regression

Cluster	Adjusted R^2 of Linear Regression Prediction of Cycle Time
Lower Weighted Cycle Time	-0.069
Medium Weighted Cycle Time	0.367
Higher Weighted Cycle Time	0.049

Clusters

Cluster	Average Weighted Cycle Time
Lower Weighted Cycle Time	2.102
Medium Weighted Cycle Time	3.332
Higher Weighted Cycle Time	5.244

Overall, accuracy values are between 49.321% and 54.751% when assigning unseen data to clusters. Thus it could be concluded that our models are not capable of accurately predicting the clusters of each truck.

Decision Trees

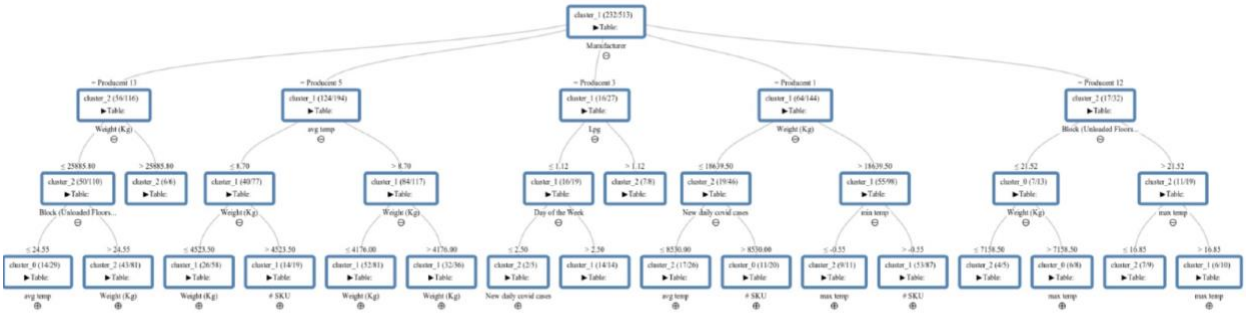


Fig 7. Decision Tree

The decision tree above is used to predict in which cluster an unseen variable belongs. The decision tree is constructed using the Gini index. The manufacturer has the strongest correlation with which cluster the model assigned to a truck; thus, the manufacturer is the crucial factor in predicting the weighted cycle time. Then, such factors as weight, average temperature, LPG (fuel/gas) prices and number of blocks in the truck also have a correlation with weighted cycle time.

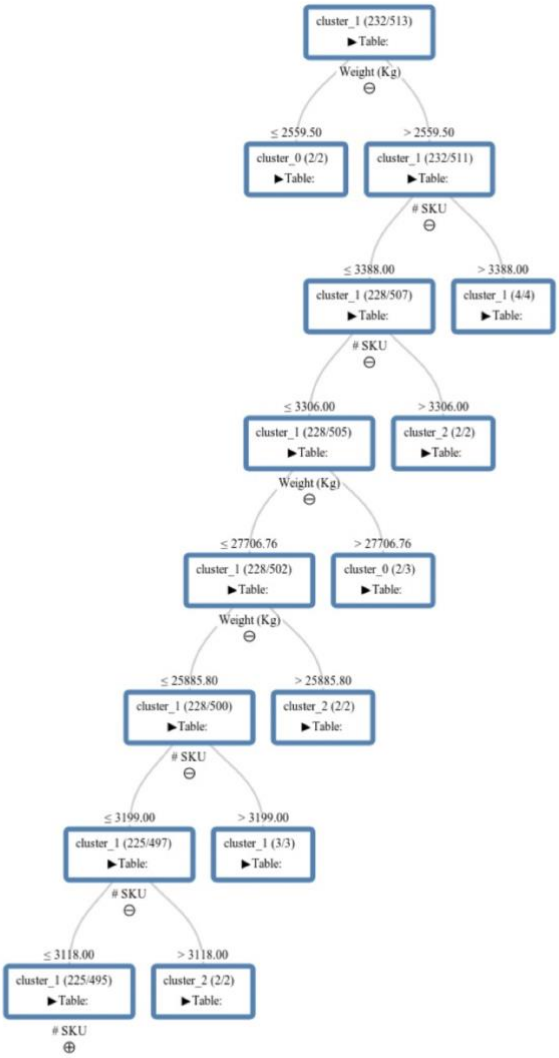


Fig 8. Decision Tree for time prediction variables

The second decision tree uses the gain ratio in order to predict which of the three clusters the truck should be assigned. It can be seen that both the weight of the truck and the number of SKUs in the truck are important factors that correlate with the weighted cycle time. The model suggests that a heavier-loaded truck has a higher weighted cycle time and that a truck that carries many different products also has a higher weighted cycle time. This seems logical, as a heavier truck might take special equipment or more personnel per crate to unload, making the unloading process go slower. Additionally, many different products of different sizes and weights might also

increase the difficulty of unloading and therefore also increase the weighted cycle time.

Data Model

Knime Model can be found in Case 3 Data - Versions. This is the current overview of the KNIME model.

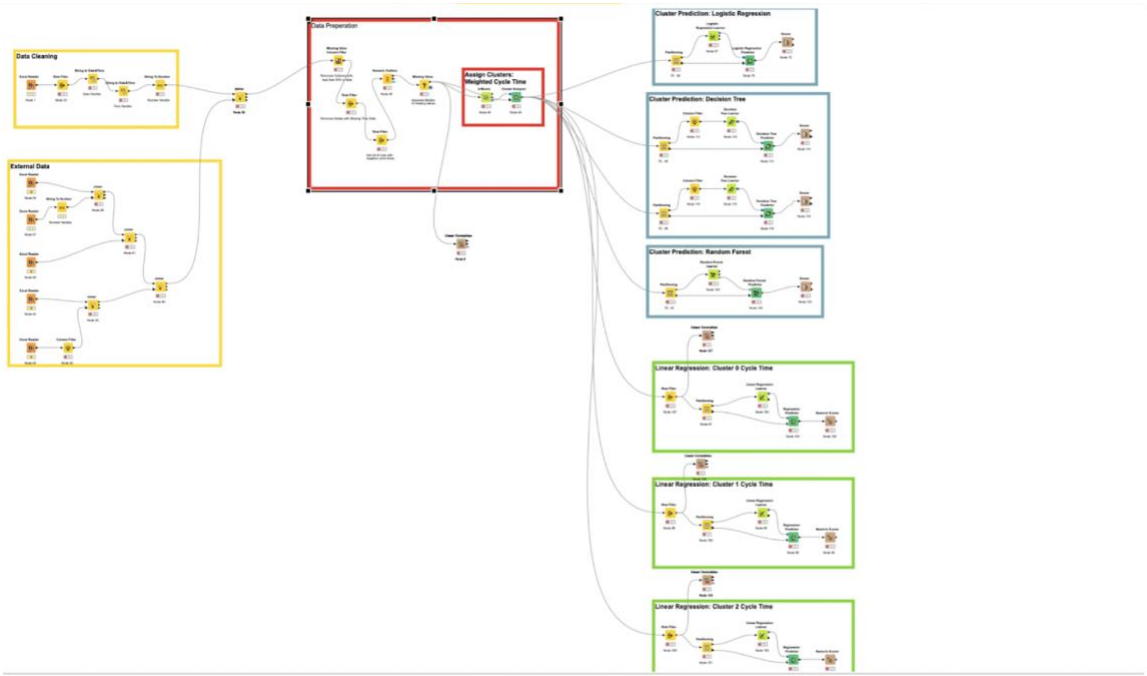


Fig 9. KNIME Model

KNIME Workspace

Case 3 Data - Versions.

Structure of KNIME workspace:

1. Data Cleaning
    a. The CAROZ DC 79 Excel file is cleaned.
2. External Data
    a. External Data is merged.
3. Data Preparation
    a. Handles outliers

        i. Anything outside of 1.5 * IQR is removed as a standard procedure.
    b. Removes Columns with missing data.
        i. Columns with less than half of their data are released.
    c. Removes Rows with missing data.
        i. Rows with no time data are removed, as it is impossible to get cycle time without the time data.
    d. Three clusters are created based on weighted cycle time.
4. Cluster Prediction
    a. Clusters are predicted.
5. Regression
    a. Splits data by cluster
        i. The cycle time is predicted from split data

External Data Sources & Bibliography

Weergegevens De Bilt Mei 2022 - Weerstatistieken KNMI. (n.d.). https://weerstatistieken.nl/de-bilt/2022/mei

KNMI - Dagwaarden neerslagstations. (n.d). https://www.knmi.nl/nederland-nu/klimatologie/monv/reeksen

Netherlands: daily new coronavirus cases 2023 | Statista. (2023, January 18). Statista.https://www.statista.com/statistics/1101300/coronavirus-cases-in-netherlands/

CBS Statline. (n.d.). https://opendata.cbs.nl/#/CBS/nl/dataset/80416ned/table?ts=1684759394992