A Shipment Difficulty Model for Third-Party Logistics Resource Allocation

by

Lili Yao

Doctor of Veterinary Medicine, China Agricultural University, 2015 and

Master of Clinical Veterinary Medicine, China Agricultural University, 2017

and

Christine M. Mueller
Bachelor of Science in International Business, Northeastern University, 2017

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 2022

Signature of Author: ______________________________

Lili Yao
Department of Supply Chain Management
May 6, 2022

Signature of Author: ______________________________

Christine Mueller
Department of Supply Chain Management
May 6, 2022

Certified by: ______________________________

James Rice
Deputy Director, Center for Transportation and Logistics
Capstone Advisor

Accepted by: ______________________________

Prof. Yossi Sheffi
Director, Center for Transportation and Logistics
Elisha Gray II Professor of Engineering Systems
Professor, Civil and Environmental Engineering

A Shipment Difficulty Model for Third-Party Logistics Resource Allocation

by

Lili Yao

And

Christine M. Mueller

Submitted to the Program in Supply Chain Management
on May 6, 2022 in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

The market for obtaining truckload capacity is becoming more dynamic as demand for truckload freight capacity in the US increases. Freight brokers form a vital connection between shippers and the hundreds of thousands of truckload transportation providers in the US and are critical to unlocking all available freight capacity. Previous research has focused on network and load optimization for freight brokerage firms, but not on optimizing the internal resources dedicated to booking and managing shipments. This study investigates commonalities in features between shipments that require similar amounts of resources to manage using feature engineering to quantify various shipment characteristics and unsupervised machine learning to cluster features. The results of this study found that there is overlap between the shipping cost per mile, the number of carrier cancellations, and the lead times between shipment request, shipment booking, and pickup time. Understanding how these shipment features relate to one another and contribute to overall shipment difficulty will help freight brokerages and third-party logistics providers better anticipate which types of shipments will require more the allocation of more internal resources in order to more effectively manage internal operations.

Capstone Advisor: James Rice
Title: Deputy Director, Center for Transportation and Logistics

ACKNOWLEDGMENTS

I would like to thank my advisor, Mr. Jim Rice, for his guidance and support. He has always been an inspiring role model to us with his open-mindedness and profound knowledge of the field. I would like to thank my capstone partner, Christine. Thanks for your hard work and dedication. Thanks for your help, listening, and understanding during the whole year. I am especially grateful to my dear parents and husband, Yiding, for their continuous and unparalleled love, help, and support. I always knew that you believed in me and stood with me.

Lili Yao

Thank you to our advisor, Jim, for your guidance and for always providing an ear to listen to. Thank you to my partner, Lili, for your patience, coding abilities, and for always bringing candy to our meetings. Thank you to my SCM 2022 cohort, a group of incredible individuals I have had the privilege of getting to know and whom I learn something from every day. Finally, thank you to my family and friends for your unconditional and limitless love, support and patience.

Christine Mueller

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

1.1 Industry and Company Background

Every year, more than 17 billion tons of freight valued at a total of $17 trillion are moved in the United Sates (Bureau of Transportation Statistics, 2020). Trucking is the primary mode of commercial freight transportation in the United States, representing 80.4% of total freight spend and 72.5% of total domestic tonnage shipped in 2019 (American Trucking Associations, 2020). As of April 2020, there were 928,647 for-hire carriers on file with the Federal Motor Carrier Safety Administration (American Trucking Associations, 2020). Freight brokers serve as a vital connection between these carriers and the millions of businesses that have goods that need to be transported.

The sponsoring entity for this research project is a 3PL and freight brokerage serving North America, which we will refer to as Company A. As a non asset-based brokerage and third-party logistics (3PL) provider, Company A operates none of its own equipment, instead utilizing its relationships with carriers and a proprietary load matching system to match shippers with carriers based on available capacity, lane preferences, and other service requirements. Its network of carriers is capable of performing standard dry van shipments as well as refrigerated trucking, flatbed trucking, rail/intermodal, and drayage. Its capabilities include both full truckload (TL) and less-than-truckload (LTL) shipments, with a variety of service options including expediting critical loads and providing drivers to carriers with vehicular assets but no available drivers. Daily variabilities in capacity and demand as well as overarching industry issues can make it difficult for freight brokers like Company A to match shipments requiring pickup with the appropriate carrier; for example, the United States trucking industry is facing increasing fuel costs and a shortage of more than 60,000 drivers (CTL, 2019). For this reason, there is a critical need to allocate resources more effectively within the freight transportation industry.

1.2 Problem Statement

The sponsoring entity is interested in identifying which characteristics of a shipment most strongly impact its ability to find a carrier; specifically, which shipment characteristics draw the most internal resources and have the greatest impact on overall operating efficiency. In other words, will the amount of resources required to fulfill a given shipment affect the company's ability to fulfill other shipments?

Company A currently utilizes a proprietary Transportation Management System (TMS) to collect information about each carrier's preferences and abilities, such as available equipment, lane preferences, and any relevant historical data. When a shipment request is received, a company representative will utilize the TMS to find a suitable carrier for the shipment. Shipment management systems such as this company's TMS are commonly used across other freight brokerages and 3PLs, matching carriers to shipments using real-time data. The matching of carriers to shipments through such systems, however, is typically a reactive process that does not consider the effect that utilizing carrier resources for one shipment will have on the ability to fulfill future shipment requests.

To begin evaluating shipment requests more proactively, the sponsoring entity would like to create a "shipment difficulty score" model that assigns a categorical score to full truckload shipments based on the amount of internal resources a shipment is expected to require during the brokerage process. A dataset that includes a list of all shipments transacted by Company A over a particular time period, along with relevant shipment characteristics such as the commodity, customer information, and timing of various milestones in the brokerage process, will be provided to facilitate the calculation of this score. If a quantitative difficulty score can be assigned to each shipment, then Company A will be able to allocate its resources more effectively and provide higher service levels, ultimately improving the firm's competitiveness among other 3PLs within the market. The analysis will also serve as a model for how freight brokerages in general should approach finding the best carrier match for a shipment while optimizing their resources.

1.3 Methodology and Hypothesis

Research regarding network and load optimization in the context of third-party logistics has been done, but limited research has explored or attempted to quantify the concept of shipment difficulty from the standpoint of utilization of resources internal to a 3PL. The aim of this project is to establish a shipment difficulty rating method based on shipment information that would be available at the time of shipment request. The calculation of this rating method will be based on three years' worth of shipment data provided by the sponsoring entity.

The first step in the creation if this rating method was to identify key quantitative indicators of shipment difficulty that could be derived from the provided dataset to be used as the dependent variables. These quantitative indicators would form the basis of our shipment difficulty score. Second, we determined the relevant shipment characteristics that could be used as attributes or independent variables for our analysis. Third, we correlated these various shipment characteristics with the quantitative indicators of shipment difficulty to find relationships and identify attributes of significance. And finally, we used machine learning methods to develop a predictive model that would provide the quantitative measures and the means with which to score shipments on their degree of difficulty. Our hope is that this model can be used as a tool for 3PLs to quantitatively determine the amount of resources different types of shipments might require in order for them to manage, and to help drive overall better decision making when allocating resources within their firms.

2 Literature Review

To design our method for developing a predictive shipment difficulty model, we first assessed some of the current factors affecting brokers' ability to acquire trucking capacity when they need it, and considered which statistical approaches might be appropriate for the analysis of these factors. This review of relevant literature will explore the role of third-party logistics providers in the current state of goods movement in the United States, as well as some of the unique challenges brokerages face in navigating the motor freight industry. It will also explore the available methods for the evaluation of certain shipment characteristics as they relate to overall shipment difficulty, as well as the proposed methodology for the creation of a predictive shipment difficulty score.

2.1 Shippers, Brokers, and Carriers

For every physical good produced for sale, there must be a physical movement of product. Some firms choose to invest in and arrange for the transportation of their goods in-house through the purchase of capital equipment and hire of dedicated personnel whose responsibility it is to facilitate transportation. Due to the high cost of investment and high overhead, however, many firms choose to outsource transportation to a third party.

Third-party logistics is a term broadly used to refer to the outsourcing of a company's logistics process to a specialized logistics company (Council of Supply Chain Management Professionals, 2013). Services offered by a third-party logistics provider typically include warehousing, inventory management, and transportation, among other services. This review will focus on the transportation management aspect of third-party logistics with an emphasis on freight brokering. A broker is defined by the Federal Motor Carrier Safety Administration as "a person or an entity which arranges for the transportation of property by a motor carrier for compensation" (FMCSA, 2014). Rather than provide transportation to a shipper via proprietary trucks, a freight brokerage is a non-asset business that locates a third-party carrier to meet a particular shipper's requirements. Fees charged vary by brokerage, but are typically set such that the brokerage retains around 20% of the net profit to the carrier (Carroll, 2021).

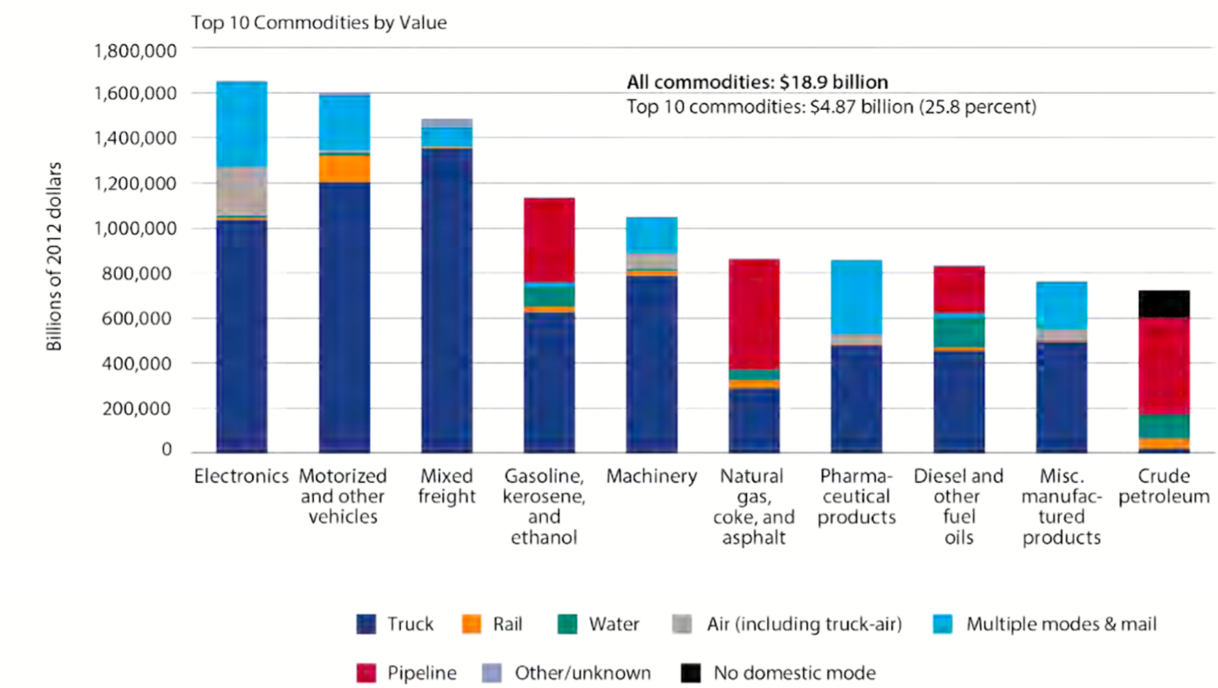
Without a broker, most shippers would be unable to connect with the vast majority of carriers in the U.S.; as of April 2020, there were 928,647 for-hire carriers registered, 91.3% of which operated six or fewer vehicles (ATA, 2021). Freight brokers bridge a gap that is crucial to the facilitation of much of the movement of goods in the United States. Properly leveraging available carrier capacity, however, can pose a challenge to many freight brokers. These challenges will be explored in later sections of the literature review.

2.2 Overview of Freight Trucking in the United States

According to the U.S. Department of Transportation's Bureau of Transportation Statistics, nearly $20 billion of freight are moved domestically every year, over half of which are moved via truck (US Department of Transportation, 2020). Figure 2.1 details the method of shipping used for all major commodities shipped in the US in 2020. Trucking was the primary mode of transportation used to ship nearly all of the top 10 commodities shipped in the US annually by value. Even when not utilized as the primary mode of transportation for a particular shipment, trucking is often required to carry out the final stretch or "last mile" of a delivery to a shipper or receiver’s premise. Despite heavy reliance on trucking across all industries, freight trucking is a relatively low-margin business. Costs of fuel, driver wages, and maintenance, as well as high initial costs of investment for trailers and other equipment, contribute to a margin of around 5% for most carriers (Aoaeh, 2016).

Figure 2.1

Value of Top 10 Commodities by Transportation Mode



Note. Adapted from US Department of Transportation, Bureau of Transportation Statistics, p. 97

While the term freight trucking is generally associated with full truckload, dry van shipments, many variations and sub-categories of this method of transportation can be used to serve different categories of goods. For example, freight can be transported as either a full truckload (TL) or less-than-truckload (LTL) shipment (Aoaeh, 2016). A TL shipment involves a movement of goods from a single shipper that requires a full trailer or container. In an LTL movement, a number of shipments from a number of different customers can be combined to form a full truckload, with costs allocated to each shipper accordingly.

Many equipment types can be used to move freight, depending on the unique characteristics and requirements of a shipment. General, dry van shipments are the most common and refer to freight moved in an enclosed trailer without any temperature controls or other particularized conditions. For freight with certain size or environmental requirements, specialized equipment can be used, including, but not limited to, flatbed (unenclosed) trucks, refrigerated trucks, and trucks with containers specially designed to transport certain hazardous or sensitive materials. For freight shipped across multiple modes of transportation, special intermodal containers designed to be used in truck, ocean, and rail shipments can be used to facilitate the seamless transfer of freight between different modes of transportation without requiring the removal of goods until the container has reached the end destination (Aoaeh, 2016). Because there are so many different ways in which to accommodate the unique requirements of a shipment, freight brokerages face the unique challenge of not only having to locate available carrier capacity, but also ensuring that all other requirements of transporting the freight can be met.

2.3 Challenges within the US Freight Trucking Industry

Rising consumer demand for goods is driving the demand for trucking capacity upwards, putting a strain on the trucking industry's already-limited capacity. Retail sales, including sales transacted in retail stores, restaurants, and online sellers, increased 13.9% in September of 2021 over September of 2020 (Harrison, 2021). An increase in demand for consumer goods cascades to an increase in demand for transportation of those goods, almost all of which will be moved via truck at some point.

There are a number of challenges to meeting this increase in demand for trucking capacity. At the forefront of the discussion are personnel and hiring issues. In a survey of industry stakeholders conducted by the American Transportation Research Institute in 2021, 30% of respondents ranked the national shortage of truck drivers as their number one concern, followed by 7% who ranked driver retention first, and an additional 7% who ranked driver compensation as their number one concern (McReynolds et al., 2021). While it is difficult to precisely quantify the extent of the driver shortage and its direct impact on the economy, the American Trucking Associations estimates the industry is short approximately 80,000 drivers of the number required to meet full transportation demand (ATA, 2021). Industry stakeholders believe this shortage to be caused primarily by the difficulties in recruiting young drivers. The federal minimum age required to obtain a Commercial Driver's License (CDL) is 21, three years later than the average high school graduation age of 18. It is hypothesized that by age 21, many potential candidates for commercial truck driving will have already begun to embark on a different career path for which the minimum age of entry is lower (McReynolds et al., 2021). Federal regulations restricting the number of hours a driver may spend driving also contribute to the shortage of drivers and pose a challenge to capacity. Hours-of-service regulations currently restrict the number of hours a driver spends on the road to 11 cumulative hours during a maximum 14-hour shift, after which the driver is mandated to take a minimum 10-hour rest period (Federal Motor Carrier Safety Administration, 2020). Timing of pickups, drop offs, and the amount of time spent waiting to be loaded and unloaded, which can be significant, further restrict the potential utilization of a driver (Correll, 2019). Another important note on this topic is that drivers are not paid for hours spent waiting to load or unload; when wait times are significant, drivers' paid hours will decrease, rendering a career in truck driving less lucrative and further exacerbating the driver shortage.

Supply shortages have also hampered trucking industry capacity. A global shortage of automotive parts, particularly semiconductors, has extended into the production of heavy-duty trucks, severely limiting production of new Class 8 trucks and leading to a significant backlog in new equipment orders (Smith, 2021). This shortage of equipment, combined with the shortage of driving labor, is leading to capacity constraints that create significant challenges not only for carriers, but also for freight brokerages looking for carrier capacity.

2.4 Transportation Management Systems

To streamline the process of matching a carrier with a shipper, many freight brokers have chosen to adopt Transportation Management Systems (TMS's), which are systems designed to centralize and organize shipper and carrier information in order to optimize operations (Coyote Logistics, 2021). The exact information collected, and functionalities of the system vary by brokerage, as the systems are often proprietary, but at minimum, a TMS will contain information pertaining to shipment request characteristics and carrier capabilities, allowing the freight brokerage to quickly match shipment requests with a carrier that is able to fulfill all requirements of the shipment.

The capacity challenges within the freight trucking industry outlined in section 1.2.1 necessitate freight brokerages to effectively leverage the capabilities of their carrier networks in order to maximize the service level to shippers. Because the fulfillment of certain shipments might require more resources than others, there is potential for some shipments to drain a higher proportion of the freight brokerage's resources (in the form of broker hours and available carrier capacity) than others. This ultimately impacts the freight brokerage's ability to fulfill future shipment requests, thereby lowering its realized service level. Some research has been done around the appropriate logic and algorithms to use when assigning a carrier to a shipment. However, limited research exists in regard to the development of predictive modeling that would actually help a freight brokerage to anticipate the impact of a shipment to the overall service level of the business.

2.5 Methodological Approaches to Difficulty Score Models

Several of the relevant studies that we reviewed used multiple quantitative models to offer improvements to efficiency in freight trucking. Procter and Sousa used clustering and regression algorithms to find that dispatchers were the root cause of poor driver performance and provided three potential solutions to overcome the challenges in the freight trucking industry (Procter & Sousa, 2021). Shen and his colleagues developed six models, including a regression model and a time series model, to forecast demand for freight transportation in Great Britain (Shen et al., 2009).

Although there is currently no model being used to identify and evaluate the difficulty of an entire freight truck shipment, the difficulty score model in other industries and other analytical models in the freight trucking industry provide some areas in which to benchmark. The difficulty score is most commonly used in the medical field to evaluate the feasibility and safety of specific treatment methods. Lee and his colleagues validated the difficulty score of Laparoscopic Liver Resection(LLR) based on

Ban's research (Lee et al., 2019 and Ban et al., 2014). They combined various related parameters, such as tumor location, the extent of liver resection, tumor size, liver function, and proximity of major vessels, to detect LLR's difficulty score. The research objects and data processing methods here are quite different from those in the freight trucking industry; however, we used the method of breaking down the difficulty score into multiple factors as a framework for our research, once we were able to evaluate which parameters were most relevant to freight trucking for third-party logistics companies.

Apostolides's research outlined the main factors that affected truck transportation's multifactor productivity or MFP (Apostolide, 2009). Table 2.1 shows the detailed factors that can affect the MFP of trucks. For purposes of our research, we roughly divided MFP into three categories: 1) driver income of each haul; 2) equipment condition, which includes the fuel efficiency of the vehicle as well as the use of other advanced technologies; and 3) the constraints of the actual route. These three factors formed the basis on which we established our shipment difficulty score. Because these are factors that only consider the costs to the carrier, we will also integrate the concerns of third-party logistics companies, such as booking lead time and profit to the 3PL.

Table 2.1

Factors Affecting Truck MFP in Three Time Periods

	Time period		
	1987-1995	1995-2001	2001-2003
MFP movement	+	_	+
Factors affecting MFP			
1. Increase in capital per worker; improved quality of capital.	+	_	0
2. Increased use of:			
a. computers	+	+	+
b. software	+	+	_
3. Efficiency of using intermediate inputs.	_	_	+
4. Improved fuel efficiency.			
a. Single-unit trucks	+	+	_
b. Combinaton trucks	+	_	_
5. Average length of haul.	+	+	NA
6. Containerization	+	_	+
7. Interstate deregulation	+	+	+
8. Intrastate deregulation		_	+
9. Mergers/acquisitions			+
10. Recesssion\9-11-2001		_	

Note. Adapt from Apostolides, 2009

With regards to data processing and model development, several studies in the field of supply chain management provide us with some ideas. Sireethorn Benjatanont and Dylan Francisco Tantuico successfully predicted truck dwell time by using linear model, random forest, and gradient boosting methods (Benjatanont and Tantuico, 2020). In this study, more than 19 million data records about the entire truck transportation process were collected from a sponsor company. The data points were categorized into three groups: Customer, Load, and Driver. Three data processing methods were then used to engineer variables that could be compared and analyzed together. First, time-related variables were extracted as integers to present date and time information. Second, aggregated variables were able to be created by using aggregation functions, such as sum, average and unique. Third, one-hot encoding and numerical encoding methods were used to adapt categorical variables to linear and tree-based models. A comparison of the parameters of the six models shows that the random forest classification model with one-hour bins is more suitable for the real freight trucking business, as shown in Table 2.2. The results of this comparison allowed the authors to conclude that characteristics of shipper facilities have the greatest impact on truck dwell time.

Table 2.2

Evaluation Metrics Comparison Summary

Models	Regression			Classification		
Evaluation	Ridge Regress	Random Forest	Gradient Boost	Logistic Regress	Random Forest	Gradient Boost
RMSE	1.040	1.032	1.031	1.336	1.216	1.263
Mean Error	0.0005	0.0006	0.0004	0.638	0.513	0.487
Accuracy	0.370	0.372	0.380	0.430	0.470	0.467
F1 score	0.309	0.307	0.327	0.357	0.401	0.424
Error by bin	0.807	0.800	0.794	0.840	0.743	0.773

Note. Adapted from Application of linear models, random forest, and gradient boosting methods to identify key factors and predict truck dwell time for a global 3PL company, Sireethorn Benjatanont and Dylan Francisco Tantuico, May 2020, p.52

Davis and Figliozzi built and combined four models to evaluate the competitiveness of electric delivery trucks based on truck characteristics and logistical constraints (Davis and Figliozzi, 2013). The models are categorized as follows: 1) ownership cost; 2) power consumption; 3) practical routing constraints; and 4) additional needs for real-world business. The authors derived an innovative formula that shows the relationship between energy costs and the operation of electric trucks as they are operated in everyday practice, including factors such as "distance traveled, speed, route/vehicle characteristics, and key logistical planning parameters." (Davis and Figliozzi 2013, p.22). While the data pool for our research will consist primarily of data collected from gasoline-powered trucks, the methodology used in this study will also help provide a framework for the evaluation of certain operational factors of trucking.

The above studies will be able to provide valuable frameworks and methods for developing a difficulty score for trucking shipments arranged by freight brokerages. We can benchmark from quantification methods used in other industries, as well as data processing methods and some model formulas from other logistics research. However, none of the studies that we reviewed provided comprehensive analysis of factors that affect trucking shipments from the standpoint of a freight brokerage. We cover this gap by focusing on the complete freight trucking process, from the initiation of a shipment request to the completion of the shipment and establishing a difficulty score model to provide third-party logistics companies with suggestions for effective resource allocation.

3 Data and Methodology

3.1 Description & Origin of Data

The data used in this analysis was provided by the sponsoring entity. The dataset provided contains load characteristics for truckload shipments brokered over three years, 2,867, 961 line items in total. The information contained in the dataset was recorded by carrier and customer representatives at the time of arrangement of a shipment. For each shipment line item, 56 associated shipment characteristics were provided. From this dataset, we identified a number of critical shipment characteristics that would form the basis of our analysis, which are outlined in Figure 3.1.

Table 3.1

Data Dictionary of Key Variables Used for Analysis

Variable	**Definition**
CreateDate	Create date/time of shipment
ActiveDate	Date/time shipment became active in TMS
BookedDate	Book date/time of carrier for a shipment
PickupDate	Pickup date/time of shipment
ProductCategory	Category of product shipped
Miles	Shipment distance
CarrierFallOffs	Number of carrier cancellations for shipment
CustomerRate	Rate charged to customer
CarrierRate	Rate charged by carrier
CargoWeight	Shipment weight
StopCount	Number of stops during shipment

We used Python to import the dataset from Excel and conduct subsequent analyses.

3.2 Data Cleaning and Pre-processing

3.2.1 Data Cleaning

We began the process of preparing the data analysis by removing rows for which any of the key shipment characteristics mentioned above were blank. The columns in the dataset that did not contain any unique identifiers or characteristics required for our analysis were also removed to reduce dataset size, which would improve the ease of data processing and development of the model. In order to filter out any blank or dummy entries, the dataset was filtered to include only shipments in which the distance traveled was greater than 0 miles and the CarrierRate charged was greater than 0 dollars. After data cleaning was complete, the final dataset contained 42,291 entries.

3.2.2 Data Pre-processing

Before data analysis and model building, some additional pre-processing of the data was also required. This was to ensure that the target features would be able to be converted into numerical values and then fed into the model for subsequent analysis. We divided the data preprocessing into three parts: first, converting time-related variables into numerical values; second, converting categorical variables into numerical values; and third, scaling all of the numerical values to allow them to be compared within the same model.

3.2.2.1 Processing of Time-Related Variables

In our study, four time-dependent variables were used. Each represents a different stage of the booking process from the 3PL perspective:

a) The “CreateDate” represents the date and time on which the shipper contacted the 3PL and shipment info was entered into the TMS.
b) The “ActiveDate” represents the date and time on which the 3PL began working on booking the shipment and locating a carrier.
c) The “BookedDate” represents the date and time on which a carrier was located and booked. In the case of carrier cancellation, this time would be updated to reflect the time at which the final carrier was booked.
d) The “PickupDate” represents the date and time on which the shipment was picked up by the carrier.

These time-related variables were converted to DateTime Format in Python in order to allow for ease of analysis and creation of new time-related variables later on.

3.2.2.2 Processing of Categorical Variables

Categorical variables cannot be analyzed using regression or any other quantitative method of analysis. In order to incorporate these variables into the model, we chose to apply one-hot encoding to convert the categorical variables into binary variables (Al-Shehari & Alsowail, 2021). Before using one-hot encoding, we had to determine how to manually divide specific categorical features into different groups, and then convert the nominal variables to 0 or 1 through one-hot encoding to represent a certain group. There were two features processed through one-hot encoding:

a) ProductCategory: category of the commodity being shipped, divided into 8 categories according to product attributes and business characteristics. The original dataset included over 50 different categories. We worked with representatives from the sponsoring entity to group these categories into clusters with similar characteristics and shipping requirements. The final categories were: Electronics, Clothing, Liquid, Fresh, Chemicals, Metal, Plastics, and Medicine. One-hot encoding was applied to these groupings.

2) PickupTime: the time of day the shipment was picked up, categorized into one of the following groups: Early morning, Morning, Afternoon, and Night. We felt that understanding the time of day of a pickup would give us important insights, but that this information didn't necessarily need to be detailed to the hour or minute. Carrier availability at 9 AM and at 10 AM would be very similar, while availability at 3 AM would be very different. Dividing the day into four blocks of time and one-hot encoding the categories would allow us to capture insights into the time of day of the pickup, while still keeping the number of characteristics contained in the model to a feasible level.

3.2.3 New Variables

We calculated three new time-related variables and one new monetary variable on which to conduct our analysis:

a) The “PrebookingTime” is calculated as the “BookedDate” minus the “CreateDate” and represents how long it took the 3PL to locate and confirm a carrier for a shipment

b) The “BookingTime” is calculated as the PickupDate minus the BookedDate and represents how long prior to the desired pickup date that a carrier was able to be found.

c) The “LeadTime” is calculated as the PickupDate minus the ActiveDate and represents how long prior to the desired pickup date the customer chose to contact the 3PL. These newly calculated variables were combined with the previously identified key variables to form a single dataframe.

d) “CostPerMile” is calculated as the CarrierRate (cost) divided by the Miles (shipment distance) and allows us to compare costs between shipments of differing distances.

3.2.3.1 Processing of All Numerical Variables and Scaling

For our analysis, which was conducted in python, we determined that there were two numerical datatypes we would be able to use: ‘int’ and ‘float’. Both data types would be able to be incorporated seamlessly and interchangeably. However, each of the different numerical variables had different units and different ranges; for example, “CarrierFallOffs” contained whole integer numbers ranging mostly from 1-3 with a maximum of 5, while “Miles” contained floats with an average if 585.55 and a maximum of 6069. In order to fit a model capable of comparing between features, we converted each numerical variable to the same range of values by using the MinMaxScaler method. For a single feature, MinMaxScaler subtracts the minimum value of the feature from each value and dividing by the range of the feature, which can be expressed as in equation (1).

$$\text{MinMaxScaler:x-minmax-min} \quad (1)$$

3.3 Analysis and Investigation

To create a difficulty score model to help 3PLs better allocate their resources, we used several analytical methods in this study, including creating a heatmap to show correlation between features, and using K-means clustering to create different unsupervised classification models of features.

3.3.1 Correlation Heatmap

Understanding the relationship between variables was the first step in the analysis of our data. A correlation heatmap between numerical variables would be able to tell us about the relationship between numerical variables very intuitively. The correlation heatmap assigns different colors according to the correlation value between variables, so as to clearly show the cause-and-effect relationship between variables (*Seaborn.Heatmap — Seaborn 0.11.2 Documentation*, n.d.). We created a heatmap of all of the above mentioned critical values, with the exception of the one-hot encoded variables, which were analyzed using regression analysis.

3.3.2 K-Means

K-Means is an unsupervised learning method for clustering unlabeled datasets, where "K" refers to the number of clusters, and "means" refers to find the center point of the cluster. The algorithm receives the parameter K, and then divides the n data which is inputted in advance into K clusters so that the obtained clusters are satisfied. Objects in the same cluster have high similarity of characteristics, while objects in different clusters have low similarity. K-means performs clustering with K number of points in the space as centers and classifies the objects closest to them. Through an iterative method, the value of each cluster center is gradually updated until the best clustering result is obtained. Because our data is an unlabeled dataset, and we expect to find certain patterns from the data, we decided K-Means would be the best method by which to cluster our variables.

In our study, we selected different features and tried to give different weights to these features to derive a model from K-Means clusters. Further detail on the model will be introduced in the results chapter.

4 Results

The processed dataset contains 390,785 lines and 72 related features. We started exploring the dataset by finding the Pearson correlation coefficients between our identified features of importance to better understand the relationships between these variables. After this exploratory phase, we used unsupervised machine learning to cluster the data and create groups of similar features, which would create our final difficulty score model.

4.1 Correlation between each numerical feature

After standardization, the data was ready for input into the Pearson correlation coefficient model. This model helped us quantify the magnitude of the correlation between each of our variables. The correlation matrix in Table 4.1 shows the exact correlation coefficients. Figure 4.1 depicts a heat map of the correlation coefficients. It is consistent with the correlation matrix in Table 4.1 and more intuitively expresses the relationship between these variables.

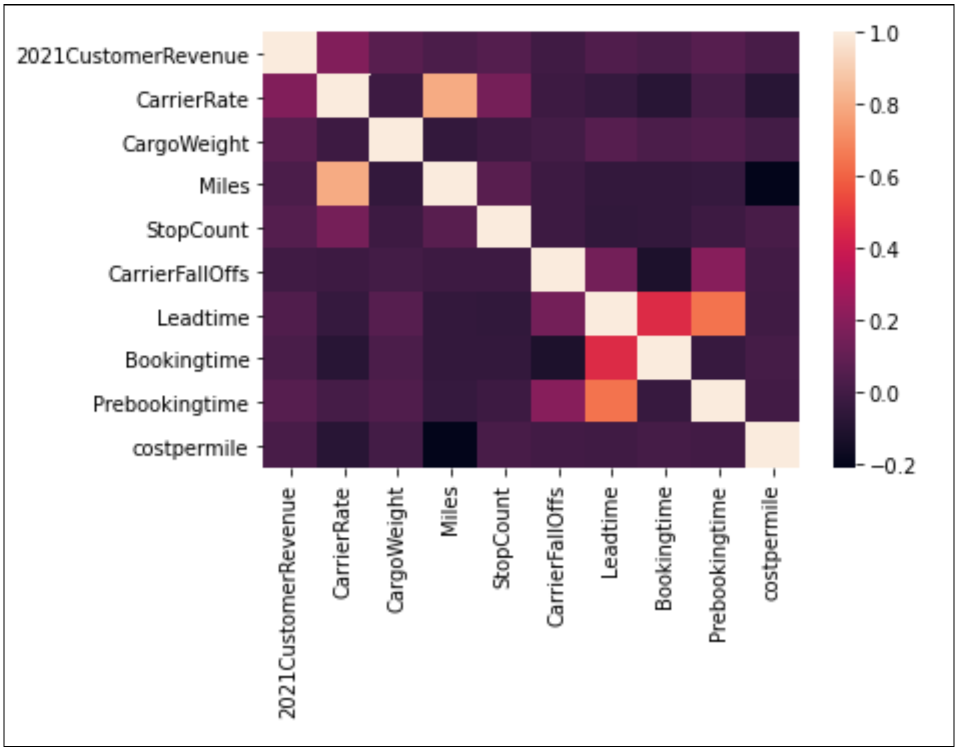
Table 4.1

Correlation Coefficient Between Numerical Features

	2021 Customer Revenue	Carrier Rate	Cargo Weight	Miles	Stop Count	Carrier FallOffs	Leadtime	Booking time	Pre Booking time	Cost per mile
2021CustomerRevenue	1.0000	0.1884	0.0671	0.0300	0.0595	-0.0024	0.0465	0.0265	0.0649	0.0202
Carrier Rate	0.1884	1.0000	-0.0126	0.8016	0.1532	-0.0131	-0.0325	-0.0786	0.0110	-0.0780
Cargo Weight	0.0671	-0.0126	1.0000	-0.0426	-0.0125	0.0077	0.0653	0.0335	0.0444	0.0061
Miles	0.0300	0.8016	-0.0426	1.0000	0.0703	-0.0118	-0.0456	-0.0448	-0.0329	-0.2067
Stop Count	0.0595	0.1532	-0.0125	0.0703	1.0000	-0.0134	-0.0495	-0.0419	-0.0134	0.0224
Carrier FallOffs	-0.0024	-0.0131	0.0077	-0.0118	-0.0134	1.0000	0.1510	-0.1141	0.2044	0.0047
Leadtime	0.0465	-0.0325	0.0653	-0.0456	-0.0495	0.1510	1.0000	0.4618	0.6437	-0.0023
Booking time	0.0265	-0.0786	0.0335	-0.0448	-0.0419	-0.1141	0.4618	1.0000	-0.0315	0.0146
Pre Booking time	0.0649	0.0110	0.0444	-0.0329	-0.0134	0.2044	0.6437	-0.0315	1.0000	0.0023
Cost Per mile	0.0202	-0.0780	0.0061	-0.2067	0.0224	0.0047	-0.0023	0.0146	0.0023	1.0000

Figure 4.1

Heatmap of Numerical Features



4.2 K-Means clustering of features

After exploring the data and identifying variables that might be related to one another, we used unsupervised machine learning in the form of K-Means clustering to divide data into groups with similar features. For each model, we determined which features of the original data should be used in the model, then used the "elbow method" to determine how many clusters should be set. The elbow method is done by graphing the sum of the squared distance errors between the mass points of each cluster and the sample points in the cluster, also known as the degree of distortion, for each possible number of clusters. For data with more defined clusters, the degree of distortion will decrease significantly as the number of clusters increases from zero toward a critical point, at which point it will decrease more slowly. The ideal number of clusters is found at the critical point, which forms an "elbow" along the graph.

In our work, we considered both the elbow method and practical business operations to determine the optimal number of clusters for our model. In the following sections, we will discuss how we used Python to complete several K-Means models with differing features, feature weights, and different numbers of clusters for comparison.

4.2.1 Model 1

In Model 1, we considered CarrierFallOffs, Leadtime, and costpermile and gave them the same weight. Based on the results of the elbow method plot, shown in Figure 4.2, we determined the appropriate number of clusters should be three or four. In Model 1, we chose to use 3 clusters. The three data clusters' spot plot and sizes after conducting K-Means clustering are shown in Figure 4.3 and Table 4.2.

Figure 4.2

Elbow Method of Model 1

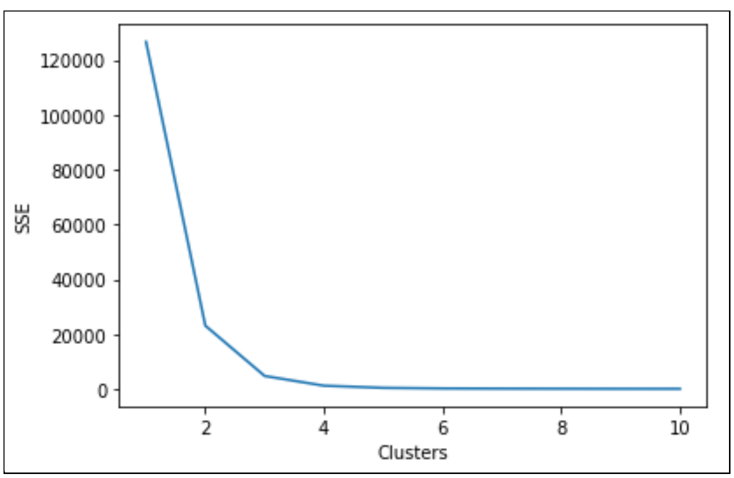


Figure 4.3

Data Distribution in Model 1

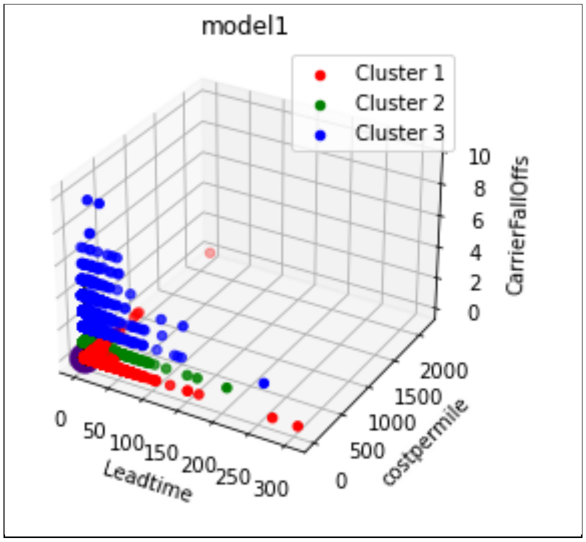


Table 4.2

Data Size in Model 1

Cluster	Data size
Cluster 1	387272
Cluster 2	73444
Cluster 3	14402

4.2.2 Model 2

In Model 2, we considered CarrierFallOffs, costpermile, PrebookingTime, BookingTime, and LeadTime and gave them the same weight. Based on the results of the elbow method, as shown in Figure 4.4, we chose to use 3 clusters. The three data clusters' spot plot and size after doing K-Means are shown in Figure 4.5 and Table 4.3.

Figure 4.4

Elbow Method of Model 2

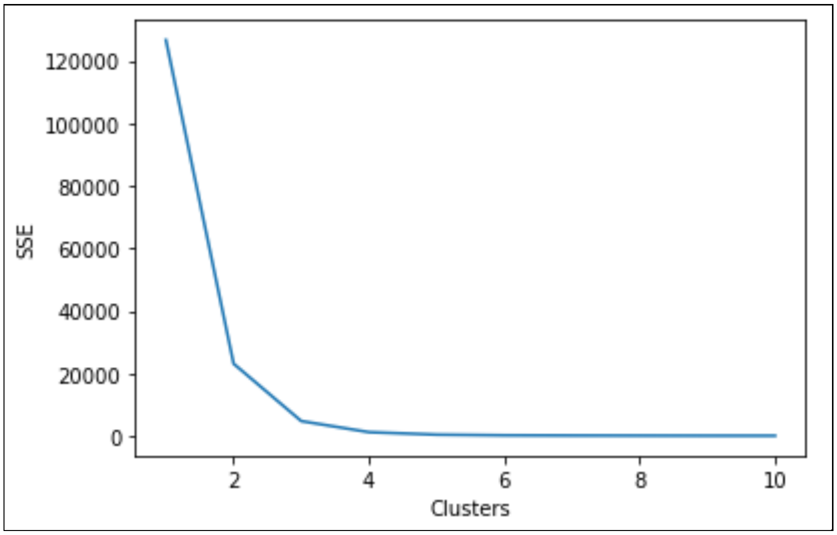


Figure 4.5

Data Distribution in Model 2

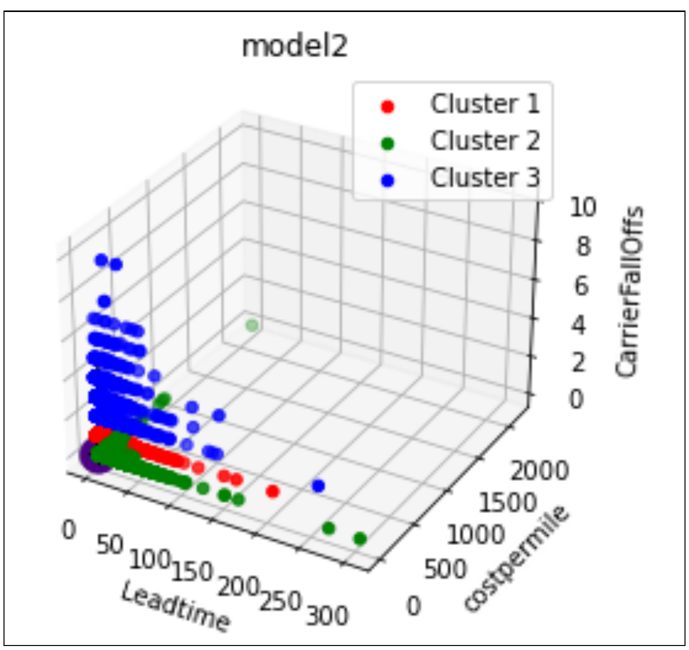


Table 4.3

Data Size in Model 2

Cluster	Data size
Cluster 1	73444
Cluster 2	387271
Cluster 3	14402

4.2.3 Model 3

In Model 3 we considered the same variables as in model two and weighted all variables the same. This time, we chose to use 4 clusters instead of 3. The four data clusters' spot plot and shape after doing K-Means are shown in Figure 4.6 and Table 4.4.

Figure 4.6

Data Distribution in Model 3

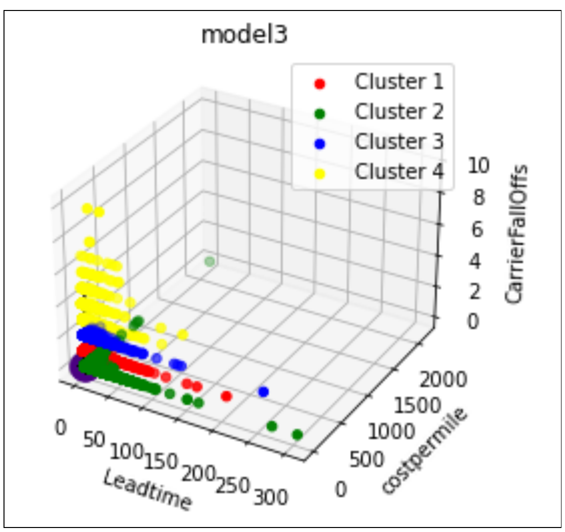


Table 4.4

Data Size in Model 3

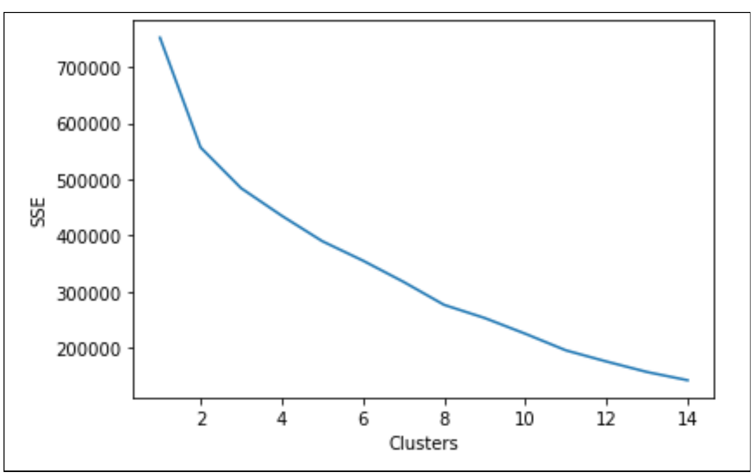
Cluster	Data size
Cluster 1	387271
Cluster 2	73444
Cluster 3	11768
Cluster 4	2634

4.2.4 Model 4

In Model 4, we considered CarrierFallOffs, costpermile, PrebookingTime, BookingTime, LeadTime, Pickuptime, and ProductCategory, giving all features the same weight. But as shown in Figure 4.7, the elbow method indicates that the data only needs to be divided into two clusters, which is not consistent with the sponsoring entity's real business needs. For this reason, we won't consider using Model 4.

Figure 4.7

Elbow Method of Model 4



4.2.5 Model 5

In Model 5, we considered CarrierFallOffs, costpermile, PrebookingTime, BookingTime, and LeadTime, weighting each variable differently. After experimenting with several different combinations of weights for each variable, we landed on a model that multiplied the scaled data for PrebookingTime, BookingTime and LeadTime by 10, multiplied the costpermile by 100, and left the value for CarrierFallOffs as-is. The elbow method result for these features is K=3, as shown in Figure 4.8. The three data clusters' spot plot and size after doing K-Means are shown in Figure 4.9 and Table 4.5.

Figure 4.8

Elbow Method of Model 5

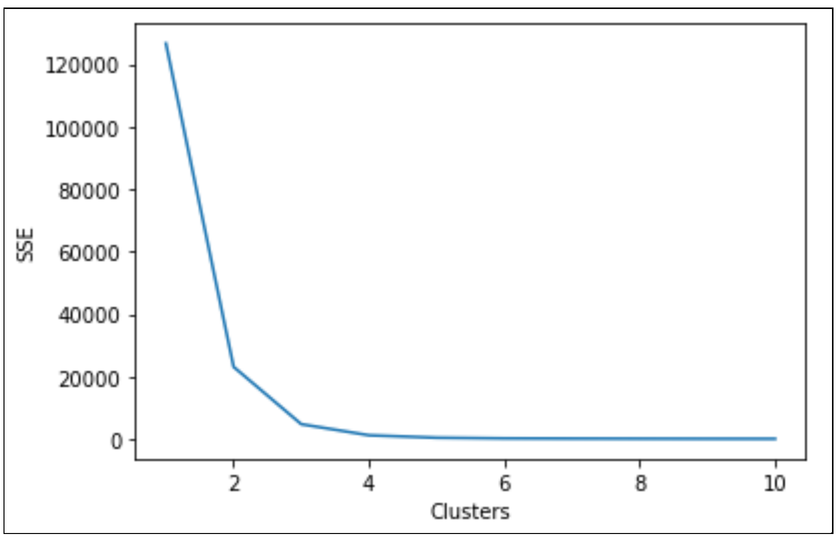


Figure 4.9

Data Distribution in Model 5

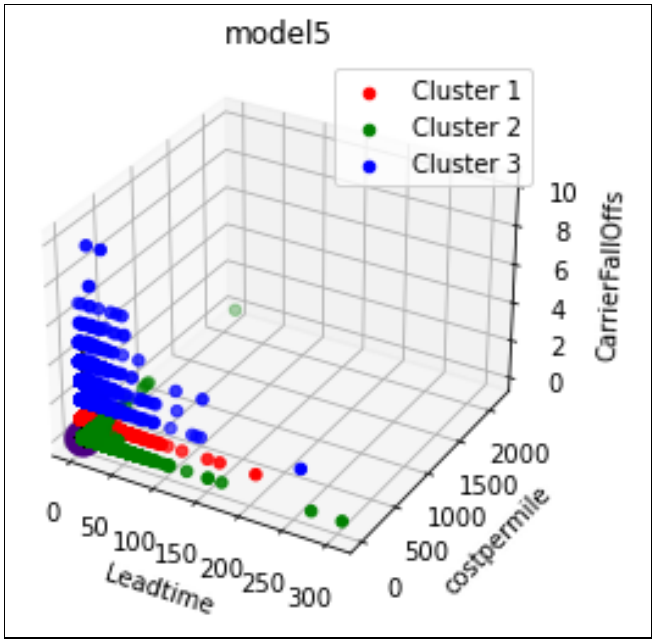


Table 4.5

Data Size in Model 5

Cluster	Data size
Cluster 1	73444
Cluster 2	387271
Cluster 3	14402

When comparing the data distribution plot and data size of the different models, we found Model 1, Model 2, and Model 5 to have similar clusters. That means that when there are three clusters, Pickuptime and ProductCategory do not have a significant influence on the clusters. After comparing the results from Model 2 and Model 3 with the needs of the business, we identified Model 3 as the best option for further analysis.

4.3 Difficulty Score Model

Using the results of clustering in Model 3, we analyzed the main features' characteristics and tried to find the salient features of each group. This allowed us to find the most suitable parameters for our difficulty score model.

First, we collected the mean value of each of the main features in each cluster. The results are provided in Table 4.6. Because these eight features have different scales, we transferred the data into change ratios. The results are shown in Table 4.7.

Table 4.6

Eight Features' Mean Value in Different Group

	Model 3			
	Group1	Group2	Group3	Group4
CarrierRate	1630.38	1679.48	1626.43	1642.00
CargoWeight	32784.99	32370.00	32819.08	32821.19
Miles	656.46	675.47	654.17	656.58
CarrierFallOffs	1.00	0.00	2.00	3.28
Costpermile	3.76	3.75	3.93	4.12
LeadTime	4.44	3.31	5.95	8.85
Bookingtime	0.68	1.36	0.49	0.45
Prebookingtime	5.02	2.73	6.86	9.81

Table 4.7

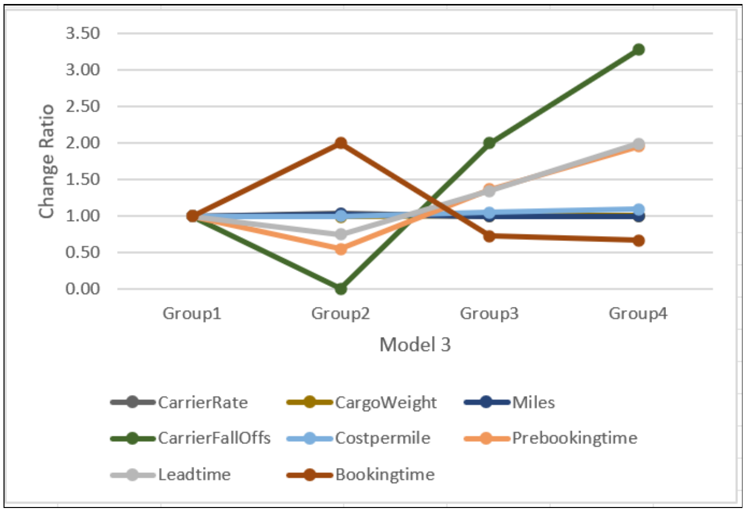
Eight Features' Change Ratio in Different Groups

	Model 3			
	Group1	Group2	Group3	Group4
CarrierRate	1.00	1.03	1.00	1.01
CargoWeight	1.00	0.99	1.00	1.00
Miles	1.00	1.03	1.00	1.00
CarrierFallOffs	1.00	0.00	2.00	3.28
Costpermile	1.00	1.00	1.05	1.10
LeadTime	1.00	0.75	1.34	1.99
Bookingtime	1.00	2.00	0.72	0.66
Prebookingtime	1.00	0.54	1.37	1.95

To make it easier to analyze, we graphed the trend lines for each variable by cluster. Figure 4.10 shows the features' trend in different groups in Model 3. CarrierFallOffs, costpermile, PrebookingTime, and Leadtime vary significantly across groups, and the change trends in the same direction. Bookingtime also varies significantly across groups, but its change trend is the opposite of that of the other four features. The remaining three features, CarrierRate, CargoWeight, and Miles seem not to change too much across groups. For this reason, we chose to use only CarrierFallOffs, costpermile, Leadtime, Bookingtime and Prebookingtime in the final difficulty score model.

Figure 4.10

Feature Change Ratio Line Chart



For the basis of the model, we created four categories of difficulty: Easy, Normal, More Attention, and Hard by considering the relative values of CarrierFallOffs, costpermile, Leadtime, Bookingtime and Prebookingtime. The actual value of each feature corresponds to a difficulty score. The sum of difficulty scores of each of the 5 features equals the total difficulty score of the shipment. The starting difficulty score is 0 points, which would sort the shipment into the "Easy" category; every 25 additional points will move the shipment to the next level of difficulty, with a maximum difficulty value of 100 points. Detailed information regarding the point values for each range of features is shown in Table 4.8.

Table 4.8

Difficulty Score Calculation Table

	Total Score Range	Feature	Feature Range	Feature Score
Easy	**0-25**	Leadtime	<=3.3	0
		Bookingtime	>=1.4	0
		Prebookingtime	<=2.7	0
		Costpermile	<=3.7	0
		CarrierFallOffs	0	0
Normal	**26-50**	Leadtime	3.3-4.5	5
		Bookingtime	0.8-1.4	5
		Prebookingtime	2.7-5.0	5
		Costpermile	<=3.7	0
		CarrierFallOffs	1	5
More Attention	**51-75**	Leadtime	4.5-5.9	10
		Bookingtime	0.5-0.7	10
		Prebookingtime	5.0-6.9	10
		Costpermile	3.8-3.9	10
		CarrierFallOffs	2	10
Hard	**76-100**	Leadtime	>=6	20
		Bookingtime	<=0.4	20
		Prebookingtime	>7.0	20
		Costpermile	>=4	20
		CarrierFallOffs	>=3	20

5 Discussion

The results revealed that there are some shipment characteristics that can provide important insights into the amount of internal resources that a 3PL or freight brokerage will need to manage a particular shipment. The Leadtime, Bookingtime, Prebookingtime, Costpermile, and CarrierFallOffs all have the ability to affect the difficulty of a shipment, depending on the magnitude of each of those values. We found it difficult to create a model that could anticipate a difficulty score based solely on features that would be identifiable at the exact time of shipment request; however, there is some element of prediction that can be found by examining the correlation coefficients between certain time-related features, and by comparing features of similar levels of difficulty, according to our matrix.

5.1 Managerial Insights

When deciding where to allocate limited internal resources and deciding which shipments to focus more effort on, 3PLs should continuously assess for features indicative of difficulty and re-allocate resources in a way that maximizes their service levels. Allocation of staffing has the ability to impact 3PL performance, which ultimately affects firm profitability. A 3PL would benefit from paying attention to any marked increases in difficulty for certain shipments and re-allocating their internal resources dedicated to shipment management accordingly.

The most important managerial takeaway from this research is the combined impact that having several contributors of “difficulty” can have on the number of resources required for a 3PL to manage a shipment. Individually, most features defined as “difficult” might not have a significant effect on a firm; however, when multiple features that contribute a high degree of difficulty are combined within one shipment, the overall difficulty of that shipment increases significantly. Our research also identified what types of shipment characteristics do not contribute to difficulty. For example, late night pickup requests or commodities requiring refrigeration are intuitively regarded as being more difficult, and might drive a 3PL to allocate more resources to those types of shipments; however, our analysis concludes that these features do not correlate with quantitative measures of shipment “difficulty,” implying that additional resources would be better allocated elsewhere. Similarly, a longer lead time between the time of shipment request and the requested pickup date is actually indicative of a more difficult shipment. This was a surprising finding, and underlined the importance the use of quantitative measures more frequently when making strategic decisions, such as with personnel and resource allocation.

5.2 Importance of each Indicator of Difficulty

The results of clustering indicated that having a long lead time between the time of shipment request and the time of pickup can make a shipment more difficult to manage. This was a surprising finding, as the intuitive assumption would be that a customer providing more days of notice prior to a pickup would make the shipment easier to manage because of higher carrier availability for dates further into the future. Similarly, a long lead time between the booking of a carrier to the time of pickup is also indicative of a shipment that was more difficult to manage. This also contradicts the intuitive assumption that having a carrier already booked far in advance is indicative of a shipment that was easy to locate a carrier for and to manage; the data presents otherwise, possible due to a higher likelihood of carrier cancellation over longer periods of time.

Some of our results did align with more traditional, intuitive assumptions. A long lead time between the time of shipment request to the time a carrier was booked means that a shipment is more difficult to manage. This is consistent with the intuitive assumption that a 3PL requiring a long time to locate a carrier also required more resources to do so. Other non-time related indicators of difficulty also follow traditionally accepted assumptions. A low costpermile understandably indicates a shipment of low difficulty, which aligns with traditional assumptions that a low carrier charge relative to the distance traveled means that the carrier also views the shipment as one of relatively low difficulty. This metric correlates with a low number of carrier cancellations, another understandable indication that the shipment has not been a difficult one to manage.

5.3 Predicting Difficulty Based on Early Identifiable Features

Although our model is not predictive in the sense that the final difficulty score can be anticipated solely from features identified from the start of the shipment request, there is evidence that some level of difficulty can be anticipated early on by looking at features that correlate with one another. There is a moderate, positive correlation between Leadtime and Prebookingtime, meaning that a customer providing notice of a shipment further out from the desired delivery date correlated with the 3PL requiring longer to locate and book a carrier. We also found some correlation between Leadtime and Bookingtime, meaning that providing notice of a shipment very far in advance of the desired pickup date is somewhat related to having the carrier booked further in advance of the pickup date. This correlation was also confirmed when we clustered the features through unsupervised machine learning, and indicates that our difficulty matrix can be used not just to compile a total difficulty score, but also to make general predictions about the magnitude of certain features based on features that fall within the same category of difficulty.

Similarly, the CarrierFallOffs metric can provide some insights for firms that do not track this metric through their TMS. By finding the category of difficulty that a shipment falls under based on the approximate magnitude of other features, a 3PL can estimate the number of carrier cancellations for a shipment based on the CarrierFallOffs value for the category of difficulty of the other features. In this way, our difficulty score model can serve as a basis to predict how difficult certain shipments will be to manage, helping to determine how internal resources should be allocated.

6 Conclusion

While nearly all of the shipment characteristics included in our analysis are outside of the control of most 3PLs, understanding the anticipated workload based on quantifiable shipment characteristics provides important insights into where 3PLs should allocate more internal resources to help maintain and improve service levels. Because service is often a differentiator for 3PLs, knowledge of what makes certain shipments difficult to manage and consumes proportionately higher resources can help firms to allocate resources in a way that optimizes their service level. Our model serves as a guideline by providing a frame of reference for the amount of effort a certain shipment might require to manage against other types of shipments, so that firms can make informed decisions when staffing and choosing which shipments to take on.

Most importantly, this model provides a quantitative measure by which to assess the level of difficulty that different types of shipments present. Although the results occasionally contradict intuitively accepted ideas of what makes a shipment "difficult", they prove that quantitative analysis can be an effective tool for challenging traditional operational processes and driving organizational change. Effective organizational structure will be an important driver of differentiation and service level for 3PLs as goods movement continues to grow and carrier capacity becomes more and more constrained. While our difficulty score model certainly isn't the first in the transportation industry to quantify traditionally qualitative ideas, it represents part of a growing movement to make transportation more efficient and competitive by examining and putting numbers to areas of transportation management that had previously not been explored. Applying quantitative modeling to drive operating efficiencies has the power to not only improve operations within transportation firms themselves, but to also improve the overall experience for all of those who participate in goods movement.

7 References

2021 Top Freight Brokerage Firms. (n.d.). Transport Topics. Retrieved October 15, 2021, from https://www.ttnews.com/top50/freightbrokerage/2021

Aoaeh, B. L. A. (2016, December 9). Industry Study: Freight Trucking (#Startups). *Medium*. https://brianlaungaoaeh.medium.com/industry-study-freight-trucking-startups-191c8acbbcda

Apostolides, A. D. (n.d.). *Estimation and Analysis of Multifactor Productivity in Truck Transportation: 1987-2003*. 48.

Ban, D., Tanabe, M., Ito, H., Otsuka, Y., Nitta, H., Abe, Y., Hasegawa, Y., Katagiri, T., Takagi, C., Itano, O., Kaneko, H., & Wakabayashi, G. (2014). A novel difficulty scoring system for laparoscopic liver resection. *Journal of Hepato-Biliary-Pancreatic Sciences*, *21*(10), 745–753. https://doi.org/10.1002/jhbp.166

Carroll, B. (n.d.). *The Economics of Sales Commissions for Freight Brokers: Part 1 – 3PL Perspectives*. Retrieved November 4, 2021, from https://3plmagazine.tianet.org/2020/03/02/the-economics-of-sales-commissions-for-freight-brokers-part-1/

Compiled SCMr Capstone Scope Documents .pdf: SCM.800 Capstone Project in SCM. (n.d.). Retrieved October 15, 2021, from https://canvas.mit.edu/courses/11578/files/1618467?module_item_id=416925

Correll, D. (2019, November 9). Is There a Truck Driver Shortage or a Utilization Crisis? *MITSupplyChain*. https://medium.com/mitsupplychain/is-there-a-truck-driver-shortage-or-a-capacity-crisis-73a7d7a5bcb

Council of Supply Chain Management Professionals. (2013). *Supply Chain Management Terms and Glossary*. https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms.aspx

CTL, M. (2019, November 9). Is There a Truck Driver Shortage or a Utilization Crisis? *MITSupplyChain*. https://medium.com/mitsupplychain/is-there-a-truck-driver-shortage-or-a-capacity-crisis-73a7d7a5bcb

Davis, B. A., & Figliozzi, M. A. (2013). A methodology to evaluate the competitiveness of electric delivery trucks. *Transportation Research Part E: Logistics and Transportation Review*, *49*(1), 8–23. https://doi.org/10.1016/j.tre.2012.07.003

Economics and Industry Data. (n.d.). American Trucking Associations. Retrieved October 15, 2021, from https://www.trucking.org/economics-and-industry-data

Federal Motor Carrier Safety Administration. (2020, September 28). *Summary of Hours of Service Regulations*. https://www.fmcsa.dot.gov/regulations/hours-service/summary-hours-service-regulations

Freight Shipments by Mode | Bureau of Transportation Statistics. (n.d.). Retrieved October 15, 2021, from https://www.bts.gov/topics/freight-transportation/freight-shipments-mode

Harrison, D. (2021, October 15). Retail Sales Rise, Showing Strong Consumer Demand, Higher Inflation. *Wall Street Journal*. https://www.wsj.com/articles/us-economy-september-2021-retail-sales-11634246791

Lee, S. Y., Goh, B. K. P., Sepideh, G., Allen, J. C., Merkow, R. P., Teo, J. Y., Chandra, D., Koh, Y. X., Tan, E. K., Kam, J. H., Cheow, P. C., Chow, P. K. H., Ooi, L. L. P. J., Chung, A. Y. F., D'Angelica, M. I., Jarnagin, W. R., Peter Kingham, T., & Chan, C. Y. (2019). Laparoscopic Liver Resection Difficulty

Score—A Validation Study. *Journal of Gastrointestinal Surgery*, *23*(3), 545–555. https://doi.org/10.1007/s11605-018-4036-y

McReynolds, J., Boyle, A., Ekberg, H., Hawkins, D. D., & Leathers, D. (2021). *Critical Issues in the Trucking Industry—2021*. 29.

Procter, D., & Sousa, P. J. (2021). *Goldilocks and the Three Dispatchers: Quantifying the Impact of Dispatcher Management on Truck Driver Performance*. https://dspace.mit.edu/handle/1721.1/130951

Shen, S., Fowkes, T., Whiteing, T., & Johnson, D. (2009). *ECONOMETRIC MODELLING AND FORECASTING OF FREIGHT TRANSPORT DEMAND IN GREAT BRITAIN*. 21.

Smith, J. (2021, September 3). Chip Shortage Curtails Heavy-Duty Truck Production. *Wall Street Journal*. https://www.wsj.com/articles/chip-shortage-curtails-heavy-duty-truck-production-11630661401

US Department of Transportation. (2020). *Transportation Statistics Annual Report 2020*. US Department of Transportation, Bureau of Transportation Statistics. https://doi.org/10.21949/1520449

What is a TMS? A Beginner's Guide to Transportation Management Systems. (n.d.). Retrieved November 4, 2021, from https://resources.coyote.com/source/what-is-a-tms