A Study of Freight Performance and Carrier Strategy

Caroline C. Bleggi
Bachelor of Arts, International Studies, The Johns Hopkins University, 2011
Master of Business Administration, Enterprise Risk Management, The Johns Hopkins University, 2016

And

Frederick (Qian) Zhou
Bachelor of Mechanical Engineering, Tongji University, 2004

SUBMITTED TO THE PROGRAM IN SUPPLY CHAIN MANAGEMENT
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING IN SUPPLY CHAIN MANAGEMENT
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUNE 2017

Signature redacted

Signature of Author............ ....
Master of Engineering in Supply Chain Management Program
May 12, 2017

Signature redacted

Signature of Author...
Master of Engineering in Supply Chain Management Program
May 12, 2017

Signature redacted

Certified by.......................... ....
Dr. Chris Caplice
Executive Director, Center for Transportation and Logistics
Thesis Supervisor

Signature redacted

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

A Study of Freight Performance and Carrier Strategy

By
Caroline Bleggi

And

Frederick Zhou

Submitted to the Program in Supply Chain Management
on May 12, 2017 in Partial Fulfillment of the
Requirements for the Degree of
Master of Engineering in Supply Chain Management

Abstract

This research analyzed freight performance to determine the groupings of attributes that influence carrier performance. Binary logistic regression and hierarchical clustering were used to identify individual and groupings of freight attributes that impacted performance success in terms of on time delivery, on time pick up, and first tender acceptance rate. From the analysis, three main performance groups of carriers were identified and their subsequent underlying attributes and strategies were analyzed. This research confirmed industry belief that differing strategies and freight profile roles result in different performance, specifically that more focused carriers tend to provide better service than unfocused carriers. Insights for shippers were gleaned from the analysis and comparison of a different shippers' carrier portfolios. From this, diversified portfolios with a higher proportion of more focused carriers were shown to have stronger performance. The significance of this research is that it offers a strategic review of groups of freight attributes that contribute to performance outcome. Within this strategic review, carriers were shown to have different underlying roles within shippers' portfolios which may suggest the need of different ways of measuring their performance.

Thesis Supervisor: Dr. Chris Caplice
Title: Executive Director, Center for Transportation and Logistics

Acknowledgements

We would like to sincerely thank Chris Caplice, our advisor, for his valuable guidance and time dedicated to navigating the thesis process. Also, this research would not have been possible without the insights and resources from our sponsor, C.H. Robinson and TMC. We are grateful to Pamela Siska and Tammy Lai for the generous dedication of their time to editing and helping to structure this thesis. Finally, thank you to our families, Yike Jin, Xiaofei Lu, and especially our parents, for their support.

TABLE OF CONTENTS

LIST OF FIGURES

1 INTRODUCTION

As inventory costs rise and consumer service level expectations grow, transportation efficiency is increasingly becoming a critical component of business strategy for shippers. Shippers and carriers within the freight industry are increasingly seeking to improve their efficiency and profitability in this competitive market. With the trend of digitization of the industry, key industry players can gain new insights to improve business results through data analytics. However, these insights tend to be focused on single attributes of freight rather than considering how groups of attributes might impact freight performance.

The goal of this thesis was to determine what attributes and groups of attributes have the most significant impact on the performance success or failure of a truckload shipment. For this purpose, the data was provided by TMC. TMC is a division of C.H. Robinson, a third-party logistics and multimodal supply chain management provider that offers transportation management systems (TMS) software to customers. Attributes of freight performance were chosen through analysis of a TMC dataset. The three-year dataset consisted of information from truckload shipments within the contiguous United States from January 2014 to December 2016.

This research sought to develop a statistically significant model, in the form of linear regression algorithms and clustering analysis, to quantitatively describe how different combinations of attributes impact final performance results. Thus, the challenge was to determine what attributes in the dataset influenced whether a carrier would be successful. For this thesis, success was defined as a set of attributes indicative of high performance as an industry standard that was determined through a literature review. In contrast to previous research, the research in this thesis investigated beyond the impact of individual attribute relationships, to interactive attribute impacts, and did not include cost in the metrics to define success. This is a study of what types of carrier portfolios lead to better transportation level of service. In turn, this research profiled shippers by their portfolio of clustered carriers to gain insights into what types of carrier portfolios lead to stronger shipper performance. This thesis pushes research to further consider that multiple attributes combination or more holistic attributes strategy may provide a superior indicator of "success" as defined by two factors: reliable on-time performance and high first tender acceptance ratio. On time performance is composed of on time delivery (OTD) and on time pick up (OTP). These metrics measure whether a shipment was delivered (OTD) or picked up (OTP) before on at the time that is was scheduled to be delivered. The first tender acceptance rate represents a carrier accepting a shipment that has been offered by a shipper when it is first tendered meaning that the shipper does not need to move down through their routing guide to offer the shipment to other carriers in order to have the shipment accepted.

2 LITERATURE REVIEW

This review is divided into three distinct research areas. First, freight industry dynamics, general insights, and performance metrics are discussed. Second, the relevant data analytical methodologies and how these methodologies can be applied to a specific industry problem rather than the general theory of a specific methodology are reviewed. Lastly, an overview of the analysis previously done on the TMC dataset including relevant previous research on specific carrier and tender related attributes and their impacts on carrier performance is given.

2.1 PERFORMANCE METRICS IN THE FREIGHT INDUSTRY

This area of research explains why certain attributes are commonly used as performance metrics. The common service performance metrics studied were on time delivery, on time pick up, and first tender acceptance rate.

In terms of defining performance metrics, Zsidisin, Voss, and Schlosser (2007) focused on the impact of carrier reduction and relationship building on success. One attribute used in this research was on-time performance, which includes on-time delivery and on-time pick up. This was the main overlapping attribute identified between the research of Zsidisin et al. (2007) and the data available in the TMC dataset. The attribute of on-time performance was also common throughout much of the other research, particularly for studies done on freight inside the contiguous United States. To illustrate, in addition to on-time performance reliability, Whyte (1992) emphasized the importance of an electronic link between the shipper and the carrier, as well as a more personalized and tailored needs-based carrier relationship with the shipper. Yet again, McGinnis (1990) cited the importance of on-time performance reliability, but added other auxiliary attributes like transit time and tracking to develop his performance measures.

In contrast, Cheng (2003) did not find on-time performance to be one of the most indicative factors of success. Cheng's (2003) research, which used an analytical hierarchical process, found that cost and tracking were the most desirable measures in choosing a carrier. However, Cheng (2007) acknowledged that geographical segmentation might be needed; in the United States, in contrast to China, service was likely more highly valued than direct cost.

Varma (2008) examined freight performance in Minnesota through the lens of government research. This research developed a group of over 40 attributes to measure performance with the focus on cost, safety, and access. The use of so many different metrics in this research highlights one of the issues prevalent in the body of research of freight performance: there is no generally accepted set of performance metrics for success or failure of a carrier. Instead, many attributes contribute to performance outcome and can be grouped to target a specific goal.

Another attribute frequently analyzed was carrier awarded volume from a shipper. This was relevant as it was known to be related to tender commitment, which is measured as acceptance rate. Research by Armstrong and Associates (2009) demonstrated that more volume does not necessarily result in a high acceptance rate due to the cost structure of the carrier combined with its network balance. Shippers generally assume that a lower rate should be given with the assignation of higher volume. However, if the load does not improve the carrier's load balance in a lane, the rate tends to be higher (Armstrong & Associates, Inc., 2009).

From the carrier's perspective, lane balance is important to performance: the actual rate shippers will pay to move the freight fully depends on it. Instead of loading all the volume to one carrier and trying to drive for "economies of scale," the right shipper procurement strategy should be to appropriate the right mix of carriers and distribute volume in the way that the carriers can have balanced load within their own trucking network. If a carrier's load can be balanced, it will be able to bid a sustainable rate for both itself and the shipper. A shipper that distributes its loads correctly will have a bell curve shaped load distribution by the rate per mile. This distribution means the shipper will receive stable rates per mile and its aggregated rate for all loads will also tend to be lower than a more dispersed and less normal distribution. Factoring into model development, this implied that a linear relationship between rate and performance is not likely to be present (C.H. Robinson 2013).

2.2 Analytical Methodologies

This thesis employed regression and machine learning techniques. First, regression is a well-established technique and a powerful tool for representing the functional relationship between variables. Dagiasis (2012) stated that regression was a strong tool for logistics companies to use due to its simplicity. To illustrate, a dataset very similar in content and structure to the TMC dataset was used by Dagiasis to show the strength of least-squares regression analysis. This test case, in conjunction with the analysis discussed in the following section, supported the use of regression analysis on the TMC freight dataset. One drawback of regression was that it showed correlations that many interpret as causation. Regression has predictive uses but cannot be used as a definitive prediction itself without external validation. Regression cannot prove one variable causes another to change; rather, it describes the relationship from an existing dataset.

The second methodology employed was clustering analysis, which falls under the umbrella of machine learning. Within machine learning, there are supervised and unsupervised methodologies. Supervised machine learning algorithms use labeled training data while unsupervised algorithms use unlabeled training data to create their models. Shmueli, Bruce, Stephens, and Patel (2017) showed multiple examples of classification application through machine learning algorithms during which a key success was defined as binary result. This success treatment was very similar to the performance outcome variables used within this research on the TMC dataset. (Shmueli, Bruce, Stephens, and Patel, 2017).

In terms of machine learning tools, the unsupervised algorithms of k-means and hierarchical clustering are very useful for mining unknown patterns in the industry data for insights. These methods were particularly applicable to this research as common patterns in critical variables can be identified for the industry leaders who have better service performance compared against the rest. With inspiration from Shmueli et al.'s (2017) application of k-means, finding previous unknown patterns to give success estimations, clustering was chosen as a method in variable selection and in later stages of model building.

2.3 Prior Analysis on TMC Dataset

This section explores industry performance indicators, mainly for carriers, and important industry variables for carrier and shipper, and the relationships among those variables. Significant research has been completed on freight data provided by TMC. This research was completed on load data 1 year prior to the set used within this thesis, but the variables and the data source remain the same. Much of this prior research included regression as one of the primary methodologies for analysis. On a TMC dataset with data up to 2015, Chen and Tsai (2016), studied multi-stop carrier behavior to decrease costs and increase acceptance ratios. They used acceptance ratio and cost as their performance metrics and employed logistic regression as their methodology. The complications they found was clustering of performance results by region. Similarly, Caldwell and Fisher's (2008) research on lead time's impact on costs to weigh the tradeoffs of lead-time lengthening included acceptance rate as a performance metric and used the same regression methodology on a TMC dataset.

Another research project using regression on TMC data by Amiryan and Bhattacharjee (2015) focused on the price to performance relationship to determine whether paying higher prices results in better performance. Metrics for performance were on-time pickup, on-time delivery, and acceptance ratio; these three metrics were also used in this thesis. They also used least squares regression as their main methodology. Defining performance as a set of attributes by grouping them into a single target variable is common throughout the research on TMC datasets.

As an overview of the set of specific attributes overlapping through the majority of the prior research done with TMC datasets, on time delivery (OTD) and on time pick-up (OTP) are the most straightforward measurements of carrier performance. However, it is important to understand that some waivers exist in certain conditions to exclude delays from the OTD measurement; typically those delays are not the fault of the carrier. Acceptance Rate (AR), measures the percentage of total load accepted out of the total load awarded to the carrier by one shipper. Normally, shippers have used aggregated loads over a set period, typically one month, to measure its carriers.

A common industry assumption is that high price (carrier rate) results in better performance. However, a quantitative correlation study by Amiryan and Bhattacharjee (2015) demonstrated that no correlation can be found between carrier rate and OTP and AR. As discovered in TMC data, there was correlation between rate and OTD, but the relationship showed a stepwise change at the point of market rate, which means rate was a more significant threshold measure rather than a numerical variable for performance prediction (Amiryan and Bhattacharjee, 2015).

Another important attribute of the freight industry is tender lead-time, measured as the time tender offered to the market till the required time load needs to be lifted. Caldwell and Fisher (2008) used statistical significance analysis on a TMC dataset to validate the industry belief that longer tender lead-time results in an increased probability of acceptance rate of loads, measured by the depth in the routing guide at which a load will be accepted. Their statistical methods used to provide a quantitative model predicted how much more a shipper must paid (depth in the routing guide) by the length of lead time (reverse correlation). Their process of model building and prediction variables selection provide practical insight for the model building process in this thesis. (Caldwell and Fisher, 2008).

A commonality across these research projects discussed and this thesis is that there is a generalization of results over the entirety of the United States. All have acknowledged that there is significant regional sensitivity within the dataset, but this is consistent with the industry overall and is to be expected. To account for geographic sensitivity, performance was investigated on a regional level.

2.4 Summary

Multitudes of metrics are available to quantify the performance of carriers. Without including cost, two of the most widely accepted and accessible metrics to assess quality within the United States are on-time performance and acceptance rate. The TMC dataset lent itself to regression analysis and history of its data being meaningfully used. Although much of the prior research continued a pattern of analysis of the cost aspect of freight, this thesis does not.

However, regression analysis to determine the relationship variables have with quality of performance remains a viable methodology to use.

Ample research has identified many performance metrics for the freight industry. In theory, these metrics can be combined to provide a basis for benchmarking carriers against each other. However, comparisons can become more complicated than useful because many are metrics available, and many are tied to differing definitions of success. This tailoring means that comparison is rarely apples to apples as good performance is composed of different variables for different research or business driven purposes. This literature review provided the basis for the selection of the three target performance variables, on time delivery and pickup and acceptance rate, to be targeted in Chapter 3.

3 METHODOLOGY

This chapter is divided into four sections: Data Preparation, Data Mapping, Data Analysis, and Modeling. Data preparation included the cleaning and manipulation of the dataset, while data mapping and analysis gave an overview of the variables used and their relationships with each other on a one-to-one basis. Data modeling explained the regression techniques used as well as the machine learning tools that were applied.

3.1 DATA PREPARATION

The first step with the TMC dataset was to pull only the records needed to address the research question. Specifically, this meant retaining for analysis linehauls that were greater than 250 miles in length and dry van trips, as opposed to refrigerated trips. Trips greater than 250 miles constitute long-haul trips, as they require greater than a single truck driver's shift to complete the trip. Rates for hauls less than 250 miles became significantly more volatile and thus were excluded. Long haul and dry van data was retained as the clear majority of the dataset was dry van. The choice to restrict the data by these two constraints was reasonable and valid because short haul and refrigerated freight tend to behave very differently than long-haul dry van freight and could skew the results of the models. Another constraint placed on the data was that it should not include dedicated runs, meaning that it should only have route guide tenders rather than fixed rate runs. Additionally, due to the sheer size of the initial data set and the differing treatment needs of the data, the data was split into single stop and multi stop loads. The analysis listed within this section refers to the single stop data, as the multi stop data only represented ~6% of the initial data and required different treatments than the single stop data.

Prior to analysis, the dataset was cleansed to remove invalid values. Extreme outliers were eliminated as they were assumed to be input error and would skew the results. Outliers included but not limited to were trips outside of the contiguous 48 states, outside the reach and framework of this thesis, or could be per mile rates far above or far below the averages seen and expected within the dataset. A listing of the records that were cleaned from the dataset prior to analysis is given in Appendix A. Missing values were treated as null values to maintain as much of the original dataset as possible to reduce bias towards complete records. There was a significant percentage of incomplete scattered fields throughout the dataset records that did hold valuable data.

Following the cleaning of the dataset, many of the fields required manipulation to be ready for analysis, e.g. the dataset had a field for rate and a field for miles but for analysis the calculation of rate per mile field was needed. Similarly, the tender lead time and rate age had to be calculated as the difference between ship date and tender date and the difference between expire date and effective date, respectively. The full data overview including data availability, definitions, and manipulations can be found in Appendix B.

3.2 Data Mapping

Post cleaning and initial manipulation, the dataset was profiled through visual and statistical methods. The initial analysis explored relationships between the independent variables with first tender acceptance rate, on time delivery, and on time pick up as the dependent variables. Based on the patterns identified and the subject knowledge from industry practitioners from C.H. Robinson and TMC, correlations between variables within the dataset and variables that would be valuable in the final modeling stages of the research were identified. This step also helped to identify potentially correlated variables that could skew model results due to multicollinearity. From these variables, several noteworthy overview statistics were found within the data. Nearly 80% of the loads were on 10% of the lanes used, and the loads were scattered across the country with concentrations on the East Coast, West Coast and Midwest (by descending order of lane volume). The statistical relationships found between the variables are discussed in section 3.3, Data Analysis. The distribution of loads to lanes has a high concentration of loads on few lanes with a long tail on the distribution.

The data kept for analysis was that of the 48 contiguous states. Figure 3.3 shows the distribution of loads by origin zip code as compared with the average population in those areas. There was a higher concentration of origins and destinations within the dataset on the East Coast of the United States; this was consistent with the expectation that more densely populated areas harbor more origins and destinations.

Figure 3.3: Origin and Destination Zip Code Mapping. Origin and Destination Concentrations by Zip Code

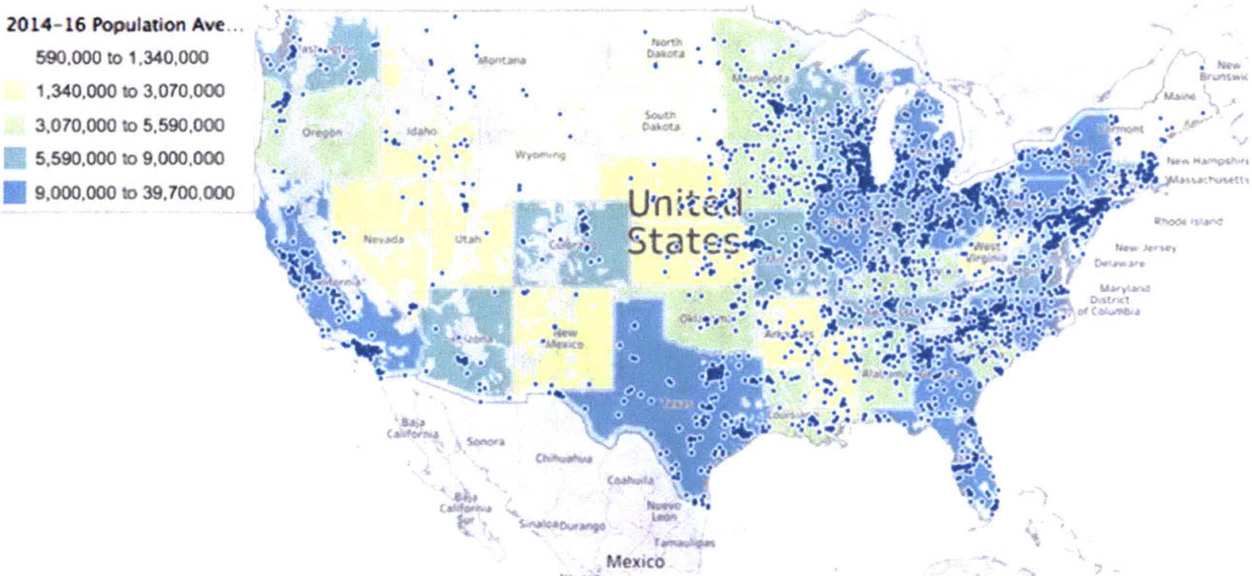


Figure 3.3.1. Origins displayed by geographical region.

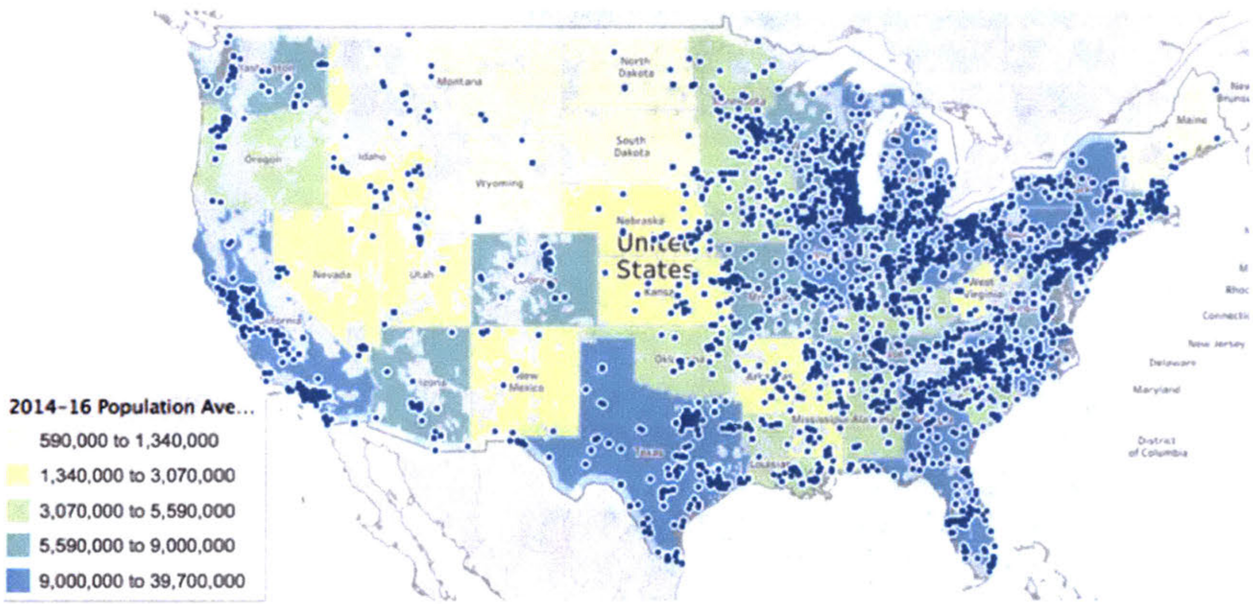


Figure 3.3.2. Destinations represented by zip code and geographical region.

The final dataset overview statistics relevant to this analysis is the number of loads by fleet size available in the dataset as shown in Figure 3.4 and the number of carriers by fleet size. As can be seen, large and small fleets, reflecting the number of trucks, as well as non-asset carriers are represented in the data; this allowed the model to offer a holistic view of the industry. Additionally, the most common fleet size was large, at 1000+ trucks. Within the data, there are 226 non-assets based Carriers which take 17% of total loads (173K) and 607 asset based carriers which represent the remaining 83% of loads (901K). Non-asset carriers are heavily represented within the data as shown in Figure 3.5.

Figure 3.4. Number of Loads per category of Fleet Size (Number of Trucks)

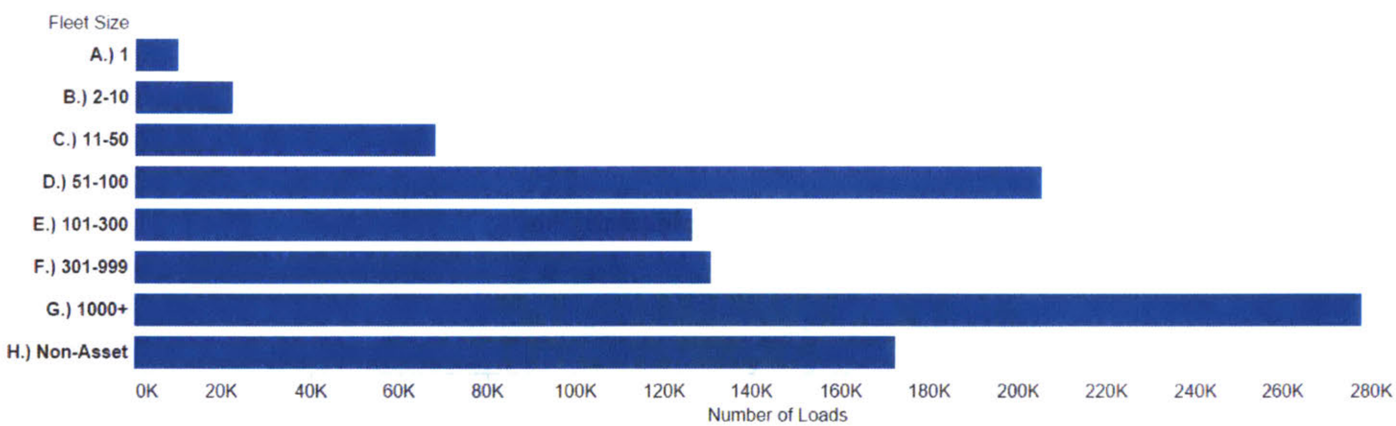
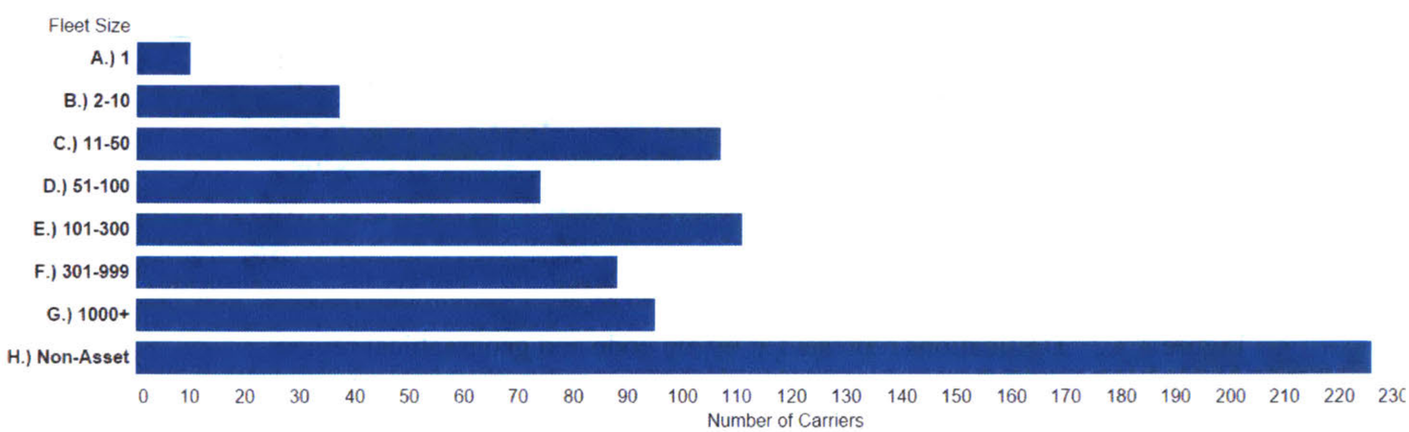
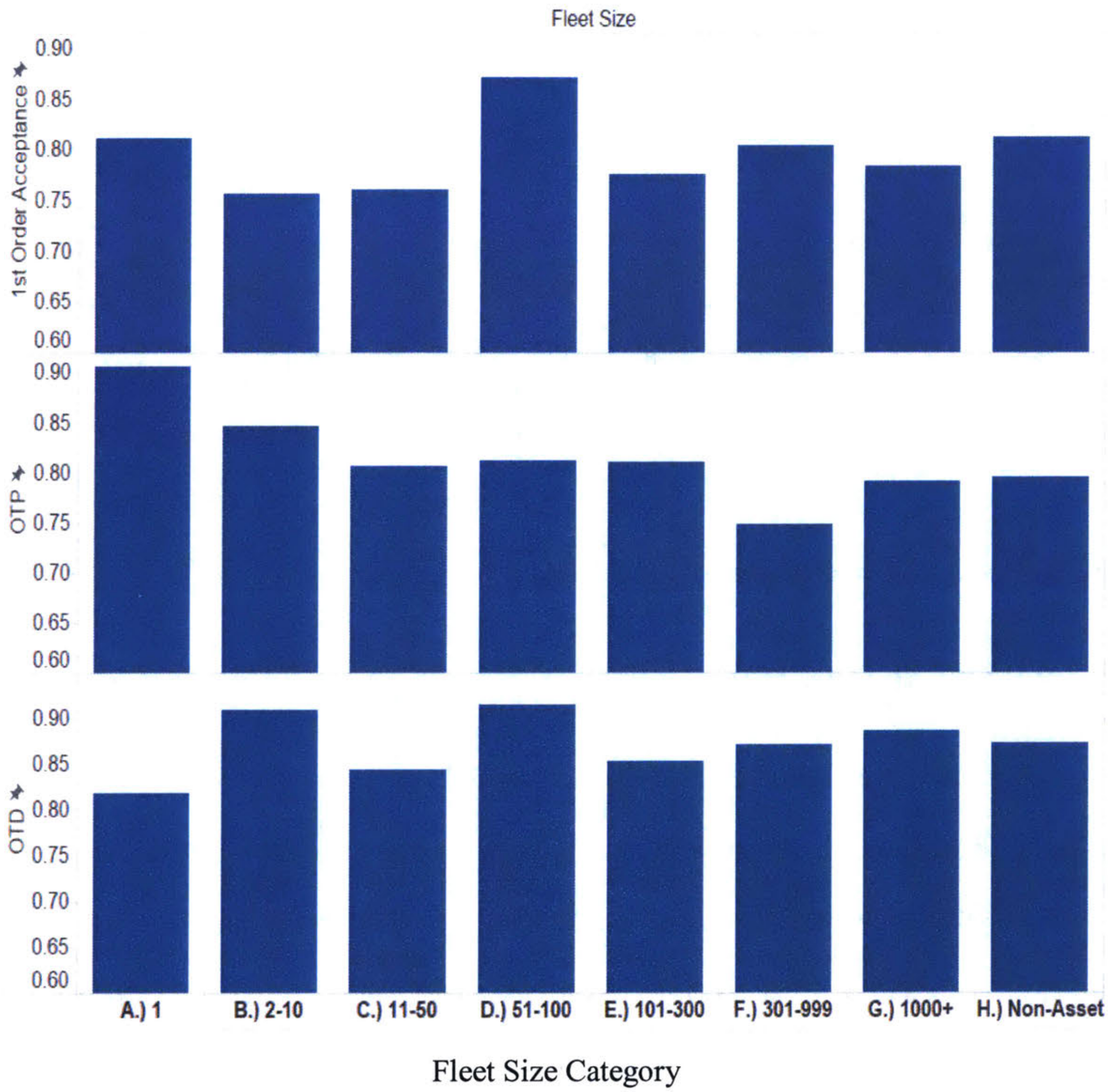


Figure 3.5 Number of Carriers per category of Fleet Size



To further delve into fleet size, Figure 3.6 shows the performance of each category of Fleet Size against the three performance metrics: OTD, OTP, and 1st Tender Acceptance. From this figure, category D with a fleet size of 51-100 trucks has the highest average performance for OTD and 1st Tender Acceptance while category A with a fleet size of 1 has the highest average OTP.

Figure 3.6 Performance by Fleet Size



An overview of the content of the dataset showed that on time delivery outperformed on time pickup every year, and that the price per mile was in decline as shown in Figure 3.7. The measures of OTD, OTP, and Price per Mile illustrated that there were differences over time but the performance percentages for OTP remained relative constant despite significant swings in first tender acceptance, price per mile, and OTD.

Figure 3.7. Calendar Year-by-Year General Overview

Measure	2014	2015	2016	**Overall**
OTD	84%	88%	87%	**87%**
Price/Mile	$2.47	$2.28	$2.10	**$2.19**
OTP	78%	80%	79%	**80%**
1st Tender AR	71%	76%	85%	**80%**

As Figure 3.7 indicated, even on a calendar year-by-year broad analysis freight was not a consistent performance industry and was susceptible to significant performance swings. This could have been the result of anything from weather to economic market fluctuations. This finding implied that in modeling it was essential to consider change over time.

3.3 Variable Relationships

This section gives an overview of the individual relationships discovered between the variables that influenced the variables selected in the modeling stages of this research.

The first variable investigated was carrier type, meaning asset or non-asset based carriers. The first finding of note was that asset-based carriers, carriers that have more than 1,000 vehicles, receive more loads than small asset-based carriers in terms of loads/carriers; the load distribution reflects carrier capacity. This was common sense but provided a logic check on the data set. Many non-asset carriers, 33% of the dataset, account for a relatively small number of loads, roughly 15% of the dataset. Discussion with TMC and C.H. Robinson revealed that some of this distribution explanation could be due to small carriers appearing and disappearing over time. On average, asset-based carriers have better acceptance performance than non-asset based carriers. Also, within the asset-based carrier group, larger carriers (>300) have better average acceptance rate performance than smaller ones (<300). Interestingly, some small carriers' (<50) acceptance rates were even lower than non-asset carriers. Additionally, the position in the route guide, which was an ordering of the carriers that a load was offered to, had a significant impact on acceptance rate for the top 3 carriers. However, it makes no significant impact on the acceptance rate for the carriers further down on the tender list.

To delve a little deeper into Acceptance Rate based findings, Extra Long Haul (>800 miles) loads have better than average first tender acceptance rates. Speculatively, the improved acceptance rate could reflect the preference of carriers towards long haul to avoid frequent repositioning of their assets. No significant difference was identified between asset and non-asset carriers in terms of load tender lead-times.

Shipper industry was found to be a potential impacting variable as different shipper industries showed different average tender lead-times with different shipper industries demonstrating significantly different average acceptance rate performances. For example, the automotive industry had the best acceptance rate average, while the paper and packing industry had lowest acceptance rate average as measured by first tender acceptance. Another variable that showed significant one-to-one results was tender lead-time's impact on first tender acceptance; it was positive and almost linear with acceptance rate when the lead-time was under three days. The price age had the opposite relationship with first tender acceptance; the younger the price age, the higher that the acceptance rate generally was. These variable insights served as the basis for model building.

3.4 Modeling

As discussed in Chapter 2, the primary methods used for analysis were binary logistic regression (binary logit) and clustering analysis.

3.4.1 Regression

Logistic regression allows for an estimation of the probability that a characteristic is present or not. It is a predictive model that explains the relationships independent variables have to a single discrete dependent variable. This method estimates the log odds of an event; it provides an estimation of a multiple linear regression function. The output is in the form of coefficients for each independent variable, which is particularly applicable in this research case, where it is searching for independent predictors of freight service success or failure. The binary portion of the regression refers to the dependent variable being dichotomous, meaning having only 2 possible types. One of the differences between regular regression and logit regression is that for regular regression, parameters are selected based on the goal of reducing the sum of the squared errors. For logit regression, the parameters are selected with the goal of maximizing the probability of observing the dependent variable as present.

For this thesis, the binary variables were on time (1) or not on time (0) for OTD and OTP and accepted first tender (1) or rejected first tender (0) for acceptance rate. All the regressions were run on a by-load basis. The models were created using a stepwise approach; variables were added dependent on significance. Then their impact was checked sequentially; independent variables were not included in the model if they were not significant enough. Much of the initial variable significance testing stemmed from the insights gained from the variable research process explained in Section 3.3. Significance was judged on the standard of a p-value of $<.05$.

For this research, service success was defined as (a) on-time delivery and pickup and (b) acceptance rate. The initial models identified the predictors of each of these three characteristics individually. The insights gained from the individual variable relationships discussed in Section 3.3 guided the selection of variables to be included in the regression analysis. Four logit regression models were developed. The first, second, and third models served to investigate the relationship between the selected independent variables and on-time delivery, on-time pick up, and acceptance rate, respectively. The fourth model combining all chosen performance metrics was a perfect shipment model, meaning that the dependent variable is perfect performance on all three metrics. Perfect performance indicated that a load was picked up and delivered on time and was accepted on the first tender.

The results of each of the regressions showed the degree to which the independent variables increase or decrease the likelihood of the presence of the dependent variable as measured against a base case. P-value, or significance, test performance percent, and a comparable measure to residual sum of squares (R^2) were used as values to compare the models. The software used to run the regression models was JMP, which calculates R^2 as the following for a logistic regression model:

$$R^2 = 1 - \ln(L_M) / \ln(L_0)$$

In this function, L_0 represented the likelihood function for a model without predictors, and L_M represents the likelihood for this research model. This R^2 is not the same as that found in linear regression; the natural logarithm of L_0 is an estimation that is akin to the linear regression calculation of R^2 and, for this research's purposes, will serve as a proxy R^2 value. As it is only a proxy, the R^2 value provided by the JMP software was not the most important measure of fit for the models. Rather, it was used as one of several measures for comparison from model to model; in terms of fit, the focus was on ensuring that the variables included in the model were significant and that the model had strong predictive potential in testing. Significance was measured as the p-value and the significance threshold was at the norm value of 5%.

Log odds is another way to express likelihood for these models. It is the log of the odds ratio which is the likelihood that the outcome will be present given the presence of a certain variable compared with the likelihood that the outcome will not be present in the absence of a certain variable. The outcome for the binary logit regressions was given as OTD, OTP, 1st Tender Acceptance Rate, and Perfect Shipment.

In building each model, a random subset of the data set was used henceforth referred to as the training set, which represented 80% of the date. The data was split into these subsets to have a means to measure and compare the performance of the models. Once the models were established and run, their performance was tested against the remaining data henceforth referred to as the test set, which represented 20% of the data. The models' prediction effectiveness was compared on several measures, including the sum of squared residuals and their error rates resulting from the test set. The results and interpretations of each of the regression models are provided in Chapter 4.

3.4.2 Carrier Clustering

Clustering analysis was employed to better grasp common characteristics among carrier groups with distinct performance levels. It was a useful tool identifying groupings of industry leaders and laggards in terms of the strategies that influence carriers that were doing well on the service performance metrics vs those who were falling short. The tool chosen for this analysis was hierarchical clustering. Unlabeled profiling was used to create groups of carriers with distinctive behaviors in terms of selected explanatory variables without take information of their performance metrics. Once grouped, the dataset was labeled with each grouping's service performance result to see if their performances are significantly different.

For unlabeled profiling, the existing unsupervised machine-learning algorithm fit well. The cluster selection was less biased and the difference of explanatory variables behaviors were statistically distinctive due to the nature of the underlying algorithm. It was an iterative process to verify whether the grouping result corresponded with performance differences as there was no prior knowledge of how many distinct groups should be created. Also, the hierarchical clustering algorithm required greater computing power and longer run times. This approach was used to achieve more objective and accurate results by minimizing selection bias.

The hierarchical clustering employed was based on a Euclidean distance calculation. It clusters, step by step, the nearest two data points into a single cluster based on Euclidean distance. By selecting the number of clusters to analyze through the dendrogram, visualization balanced the significance in the differences among different clusters. The algorithm was set to recognize patterns within the dataset and was useful for regression and classifications. The clusters were grouped by performance in terms of the perfect shipment into three main groups: Leaders, Laggards, and Major Players. The output for hierarchical clustering was in the form of dendrogram and a visualization of the clusters graphed against the first two principle components, as found in Chapter 4.2, Clustering Results. A dendrogram hierarchically partitioned the dataset. In partitioning the dendrogram two conditions were considered:

1. Number of Clusters. There could not be too many as over fit would occur.

2. Maximize the difference. Ensure that each cluster was distinctive by maximizing the difference of explanatory valuables' behavior so that a clear pattern could be identified

Two sets of clustering analysis were done by carrier: (a) on asset based carriers and (b) on non-asset based carriers. Asset and non-asset carriers were split before clustering as they cater to different needs across a shipper's freight portfolio. It was important to highlight the different strategies that led to their respective performance as this difference could account for some variance in first tender acceptance and on time performance. The clustering analysis over the entirety of the dataset was also done only on carriers that had greater than 200 loads as this would ensure that smaller carriers that might not have a representative strategy sample were excluded. The analysis was by carrier and was clustered by the following set of variables:

1. Fleet Size (number of trucks)
2. Geographical Coverage (number of states covered)
3. Number of Lanes Served
4. Number of Customer Served
5. Industry Coverage
6. Load Density per Lane
7. Load Density per Customer
8. Total Number of Loads

Although some of these variables could be considered shipper attributes rather than carrier attributes, they were chosen as a depiction of strategy, i.e. If a carrier worked exclusively with manufacturing shippers, this was part of their strategy and should be considered in the clustering process. Each cluster was measured for service performance in terms of its likelihood to have perfect shipments. The results of the clustering model and the regression models can be found in Chapter 4, Results.

3.4.3 Shipper Profiling

The results from the regression analyses and the carrier clustering analysis were combined and compared against the performance in terms of perfect shipment of shippers. The performance of shippers was also measured against load consistency, portfolios of clustered carriers, lane density, and the individual performance metrics of OTD, OTP, and first tender acceptance rate. This comparison allowed for the significant underlying performance differences to be identified and stratified by shipper performance. These results are given in Chapter 4.4.

4 RESULTS

This chapter provides the results of the regression and machine learning models used to address the question of what characteristics influence service performance success for freight. These models highlight structural and systemic trends that lead to strong performance. Each regression model was evaluated based on the significance of the variables included in the model. Comparisons were made between the models to fully grasp which aspects of the models are useful for industry guidance. Industry guidance was discussed in the chapter 5, Discussion. Finally, shipper profiling is discussed in relation to shipper's portfolios of carriers.

4.1 Regression Results

As described in the Methodology chapter, three individual performance target variable logit regression models were completed in addition to a regression targeting the combination of all three performance metrics.

4.1.1 Single Target Performance Variable Regressions

Three binary logit regressions with a single target variable were completed on the full dataset spanning the January 2014 to December 2016 timeframe. The dependent target variables were first tender acceptance, on time delivery, and on time pickup. The base case, the case against which the results of the binary logit regression are compared, was as follows:

1. Fleet Size: Non-Asset
2. Industry Type: Paper & Packaging
3. Tender Date: Monday – Thursday
4. Ship Date: Monday – Thursday

First Tender Acceptance

For first tender acceptance, the variables found to be significant, according to their p-values, were Carrier Size, Shipper Industry, Day of the Week, and Tender Lead Time; Tender Lead Time was the strongest predictor of the chosen variables. Asset based carriers were significantly more likely to have first tender acceptance than non-asset carriers. Within shipper industry, manufacturing had significantly better acceptance likelihood than the remainder of the industries. For day of the week, loads tendered on a Friday had better acceptance performance likelihood while loads tendered on a Saturday or Sunday had lower first tender acceptance likelihood. In terms of tender lead time, with a lead-time of less than three days first tender acceptance was significantly less likely. The potential underlying cause behind this result could be that as a carrier received requests they utilize their fleet; with shorter notice the carrier's fleet has higher inefficiencies as they may not have additional capacity or could not plan their routes effectively and in time to accept the load quickly. The results of the regression including coefficients are available in Appendix C.

When this model developed on a training dataset, was compared against the test data set, it had a prediction performance of 88.49% but an r-square value of .41. As discussed in Chapter 3, this R^2 value should not be taken at face value, as it serves as a proxy to iteratively measure the model's predictive improvement in conjunction with the individual variable p values in the model building process. The predictive performance of this model was relatively strong but did not explain all the variation within the dataset.

On time Delivery

In terms of significant variables, destination dwell, shipper industry, carrier size, and spot bids all showed significance in the probability of a load having an on-time delivery. For destination dwell, the longer the load dwelled at the destination, the higher probability of late delivery. For the shipper industry, Paper & Packaging, the base case, loads were associated with better OTD. The carrier size of 51-100 was associated with lower OTD performance and Spot Bid load delivery tended to be more likely to be on time. Upon testing, this model had a relatively strong prediction accuracy performance of 88.36%. The coefficients and full results of this regression can be found in Appendix D.

On Time Pick-Up

From the OTP regression results, CPG and manufacturing industries had a stronger likelihood of having on time performance. For carrier size, smaller carriers increased the likelihood of a late pick-up, while carriers with fleets of 300-999 trucks had the strongest likelihood of an on time pick up. Spot Bid loads and loads shipped on a weekend had better OTP likelihoods. Additionally, like the other regressions, shorter price ages improved the likelihood of better performance on this metric. The full results of the on time pick up regression are available in Appendix E. Comparatively, the OTP model was the least predictive of the four regression models run.

4.1.2 Perfect Shipment Regression

The fourth logit regression model was coined the "Perfect Shipment" model. Rather than having a single characteristic as the dependent variable, OTD, OTP, and first tender acceptance were combined to be the single dependent variable. The results of this model showed the characteristics that have the strongest influence on creating an order that was picked up and delivered on time and accepted in the first tender.

The significant variables for this regression were Carrier Size, Shipper Industry, Day of the Week Tendered, Spot Bid, and Dwell Time. In general, the asset based carriers were more likely to perform better on this perfect order shipment than non-asset carriers. Also, manufacturing had a higher likelihood of perfect shipments than the rest of the industries. Loads tendered on a Friday had better acceptance performance likelihood while the opposite was try for loads tendered on a weekend. Consistent with the other regressions, longer tender lead-time increased better performance likelihood. Finally, Spot Bid loads and loads with origins with long dwell times lower perfect shipment likelihood. The full results with coefficients for this model can be found in Appendix F.

4.1.3 Regression Profile Comparisons

The magnitude of the impact of each of the variables within the regressions is given by their coefficients. The coefficients for each of the regressions are listed in Appendix C, D, E, and F. The p values and coefficients from each of the individual regression results helped to create a profile of a load that would be more likely to have better First Tender Acceptance, OTD, OTP, and Perfect Shipment. The successful profiles include the most significant variables from each of the regressions and allow for comparisons of the models. The success profiles for each of the models are shown in Figure 4.1. This gives an overview of the significant variables that increase the likelihood of success for each of the models i.e. for First Tender Acceptance, OTP, and Perfect Shipment, asset based carriers were more likely to have first tender acceptance than non-asset carriers. The implication of these findings is that generally longer hauls (>700 miles) with younger price ages and longer lead times in with shippers from the manufacturing industry generally performed better.

Figure 4.1. High Performing Profile Regression Comparison

	First Tender Acceptance	OTD	OTP	Perfect Shipment
Carrier Type	Asset Carrier	Not Significant	Asset Carrier	Asset Carrier
Tendered On	Weekday	Not Significant	Weekday	Weekday
Shipper Industry	Manufacturing	Paper & Packaging	Manufacturing	Manufacturing
Bid Type	Non-Spot	Spot	Spot	Non-Spot
Length of Haul	>706 miles	>723 miles	Not Significant	>716 miles
Tender Lead Time	>1.3 days	Not Significant	Not Significant	>2.4 days
Price Age	<152 days	<151 days	<152 days	<148 days

4.2 Clustering Results

Hierarchical clustering was used to analyze carrier performance and identify groups, or clusters, of homogenous carrier strategies. Carrier strategy is defined by to how a carrier selectively provides its tucking service across different shippers, different shipping lanes, or chooses to accept or reject load in a certain manner. The clustering analysis offers insight into influence of these strategies on a carriers' OTD and OTP performance for the loads taken. Identifying the carrier's strategy or strategies allows for characterization of how high performing carriers behave. Further, it is a step towards carrier understanding from the shippers' perspective with the end goal of improving overall service level received. Specific clusters were composed of carriers who shared similar strategies in terms of the characteristics. Clustering was completed for asset based and non-asset based carriers separately as each holds different roles and in turn followed different strategies within the freight industry. It was important to consider the potential impact of their separate strategy performances to gather more valuable and applicable insights.

4.2.1 Asset Based Carrier Clustering

The asset based carrier clustering revealed different strategies employed by carriers. The initial clustered strategies focused on the following:

1. Geographical focus – measured by number of loads per lane and number of loads per state. This represents whether a carrier provides service in a limited number of lanes and geographical areas or if the service it provides has a wide coverage of both.
2. Industry focus – measured by number of shipper industries that the carrier served. This represents whether a carrier focuses its service in a specific industry or it carries loads from any industry.

3. Customer focus – measured by number of customers the carrier served. This represents whether a carrier builds relationships and familiarizing itself with a limited number of shippers or vice versa.
4. Fleet Size of the carriers- measured by the number of trucks the carrier has. For the asset-based carrier community, fleet Size is the most straightforward representation of its capacity.

Figure 4.2 gives a visualization of the three major clusters identified within the asset based carrier group. This figure represents all three years of data. Each of the clusters was measured by the strategies listed above and against the metric of perfect shipment rate which led to their given titles: Leader, Laggard, and Major Players.

Figure 4.2. Constellation Plot showing Asset Based Carrier Performance Clusters

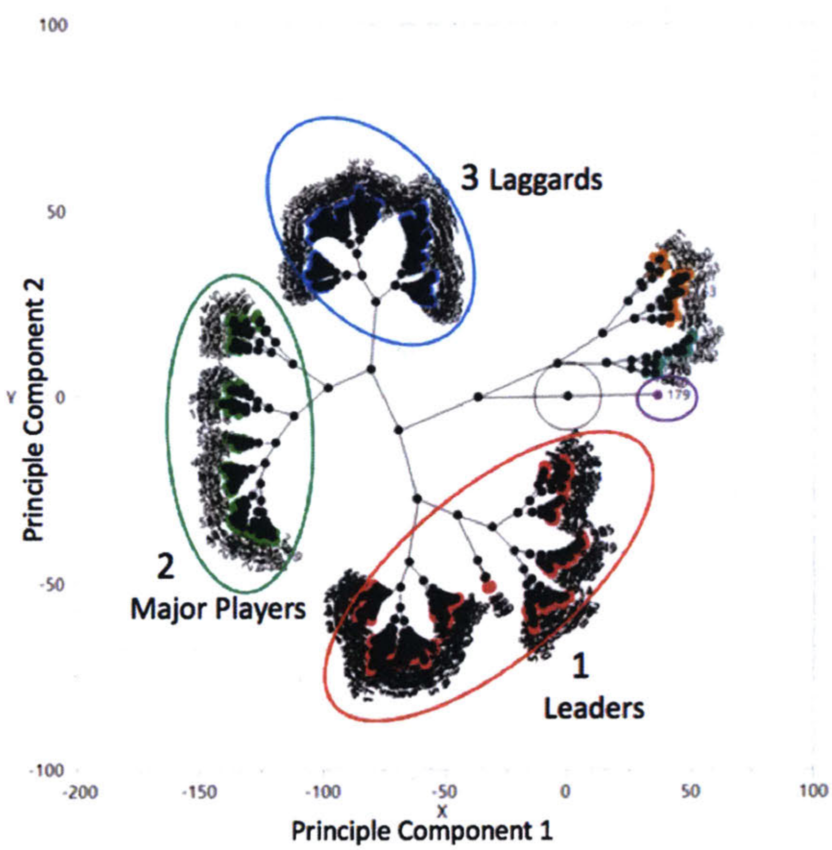
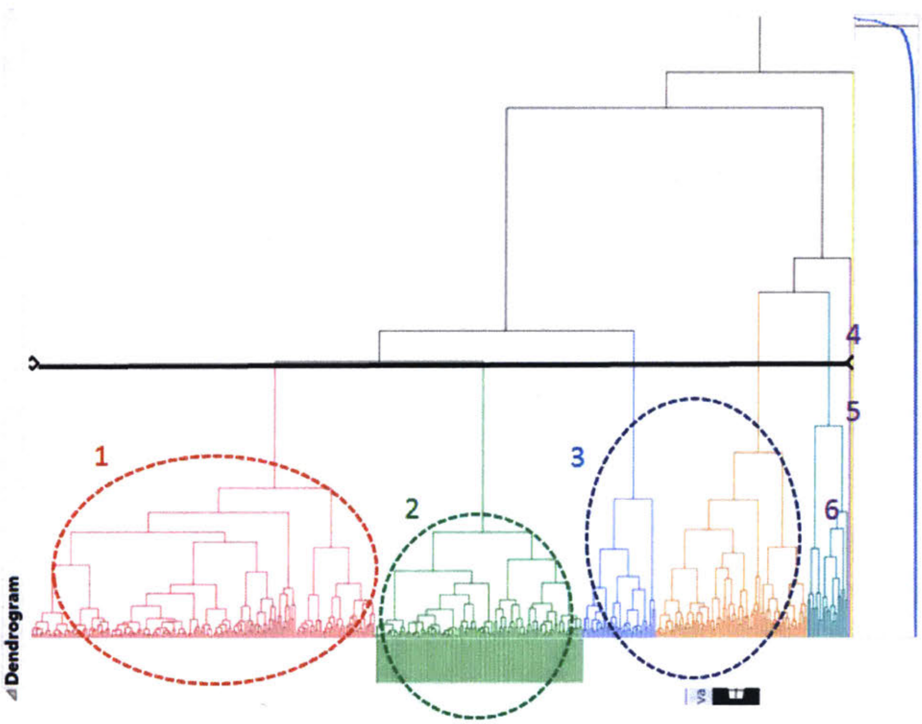


Figure 4.3 provides a dendrogram showing where each of the clusters was cut. It also shows the relative size of each cluster. Cluster 1 had 209 carriers, Cluster 2 had 45 carriers, and Cluster 3 had 184 carriers. In this dendrogram, the selection line was chosen to maximize the clusters' distinctiveness. Then the clustering result was labeled back to the original dataset of loads records to compute the service performance difference of each cluster. The selected cluster number was six, and Clusters 1, 2, and 3 accounted for 89% of the total records. Clusters 4 and 5 each only had one record, thus considered outliers of different kinds. Cluster 6 contained 9% of data records. In evaluation, the focus was on cluster 1, 2, and 3 as they contain the majority of the data and maximized the distinctiveness of the profiling result.

Figure 4.3. Dendrogram showing the hierarchical splitting leading to the 3 main clusters



The clustering result revealed 3 main clusters of carriers with distinct service strategies, the difference in these strategies is shown in Figure 4.4. Each of the different service strategies corresponded to different levels of service in terms of perfect shipment rate.

Leader Cluster:

This cluster is composed of medium sized carriers who only carry loads from a relative few shipper in a single industry and tend to have higher focus on specific lanes for those respective shippers. This focus gives this group its leading performance of a 76% perfect shipment rate.

Major Player Cluster:

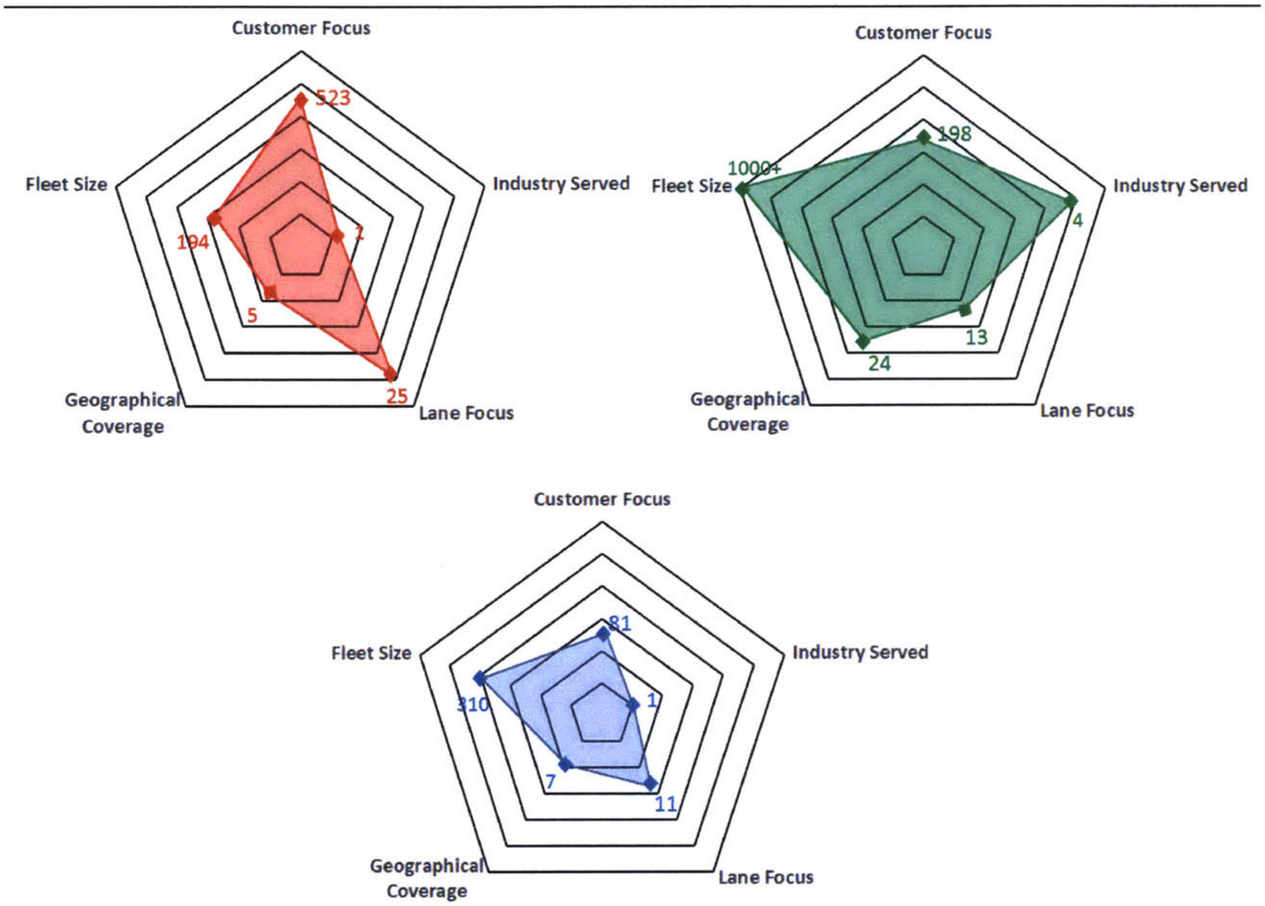
This cluster is composed of the larger fleet size carriers who have 1000+ trucks on average. They serve a much wider geographical area and carry loads from many shippers regardless of their industry. This group has a mediocre performance of 56.3% perfect shipment and a much wider strategic spread than the Leader cluster carriers.

Laggard Cluster:

Like the Leader cluster, the carriers within the Laggard cluster are medium fleet sized carriers. The difference between these clusters is that this Laggard cluster seems use their limited capacity opportunistically. They tend to serve many different lanes across a relatively large number of different shippers. Their perfect shipment performance is the lowest at 31%.

Figure 4.4. Asset Based Carrier Strategy Profiles by Cluster

1. Leaders | 2. Major Players



3. Laggards

To further show clear delineations between the clusters, the tables in Figure 4.5 give the average values for Customer Focus, Industry Focus, Lane Focus, Fleet Size, and Geographical Coverage for each of the 3 identified major clusters. This measure is helpful in recognizing the spread of strategy used within each cluster. The leader and Laggard clusters have a relatively small standard deviation across all variables, especially in lane focus. This shows the robustness of the clustering process for distinctively characterizing this two groups. While the Major Players cluster has a higher standard deviation almost across all the variables. This might demonstrate that more variety of service strategies exist within this cluster.

Figure 4.5: Cluster Summary Statistics

Cluster 1: Leaders

Dimensions	Mean	Std. Dev.	Min.	Max.
Customer Focus	523	210	281	1604
Industry Served	1.1	0.1	1	1.4
Lane Focus	25.4	3.2	10	51
Geography	5	3	3	9
Fleet Size	194	221	110	899

Cluster 2: Major Players

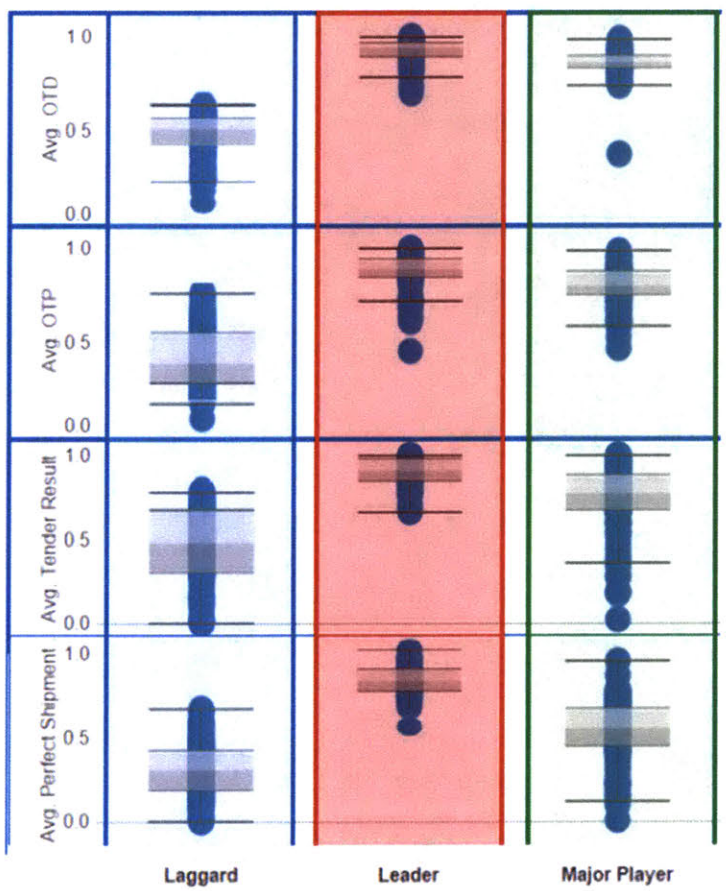
Dimensions	Mean	Std. Dev.	Min.	Max.
Customer Focus	198	382	11.3	554
Industry Served	4	1.28	1	5
Lane Focus	13	14.7	1.5	30
Geography	24	7.3	12	46
Fleet Size	>1000	1473	110	1521

Cluster 3: Laggards

Dimensions	Mean	Std. Dev.	Min.	Max.
Customer Focus	81	210	23	300
Industry Served	1.4	0.5	1	3
Lane Focus	11	4.3	2.3	25
Geography	7.2	3	3	9
Fleet Size	310	221	98	1113

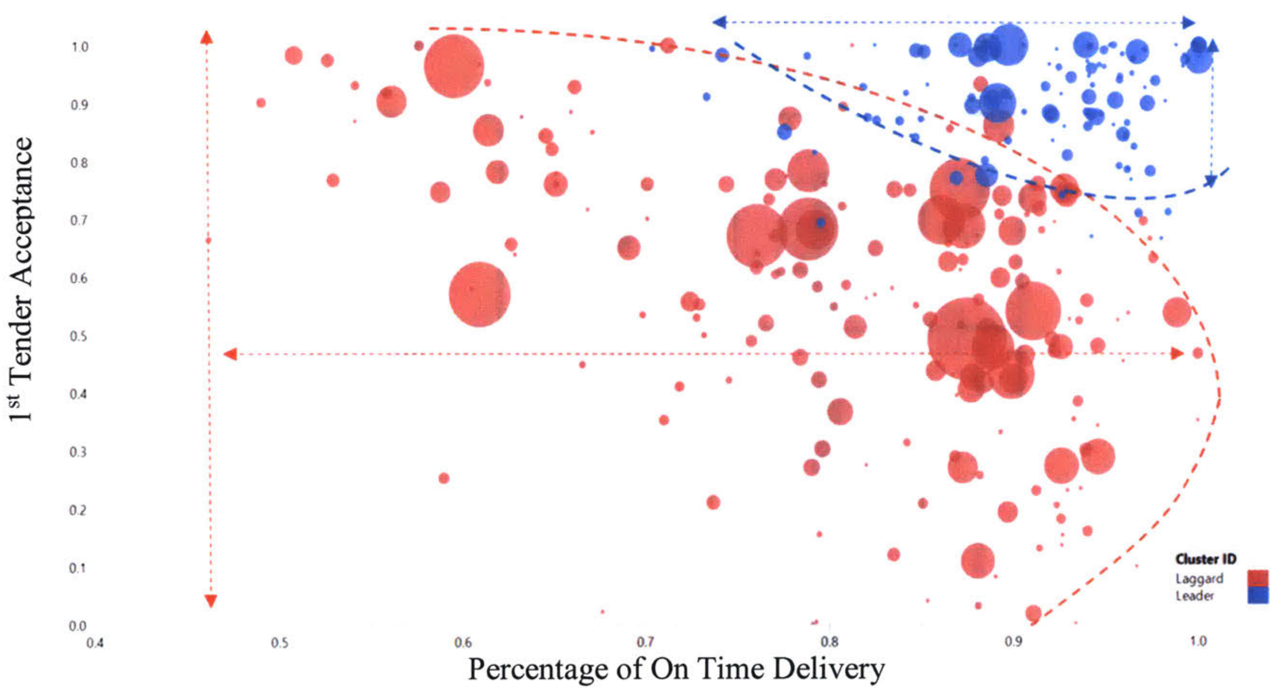
The clustering is based on perfect shipment result to reflect comprehensive carrier performance. However, after identifying clusters with performance differences, it is also important to see where carriers in the low performing laggard cluster lag. This question is if laggards have low performance in all metrics or low in a particular area? In profiling the larger identified groups of Leaders, Laggards, and Major Players, performance was consistent within the performance measurements for Leaders and Laggards. This consistency implied that Leaders generally performed well on OTD, OTP, First Tender Acceptance, and Perfect Shipment (the combination of the three). Alternatively, Laggards performed consistently poorly across these same metrics. Figure 4.6 below shows this consistency of performance across OTD, OTP, First Tender Acceptance, and Perfect Shipment within the major carrier performance groups. This figure also shows that the spread of OTP for Leaders is slightly wider compared with the other measurements. To interpret the spread shown in Figure 4.6, the closer the horizontal lines are to each other, the smaller the spread or standard deviation is within each metric. Compared with the relatively consistent performance in Leaders and Laggards as shown through the narrow spread in Figure 4.6, Major Players have relatively wider spread of performance. This spread could indicate inconsistent carrier strategies within the Major Player performance cluster. Additionally, within the Major Players group, First Tender Acceptance was the major driver of perfect performance.

Figure 4.6. Performance Metric Variation within Cluster Profiles



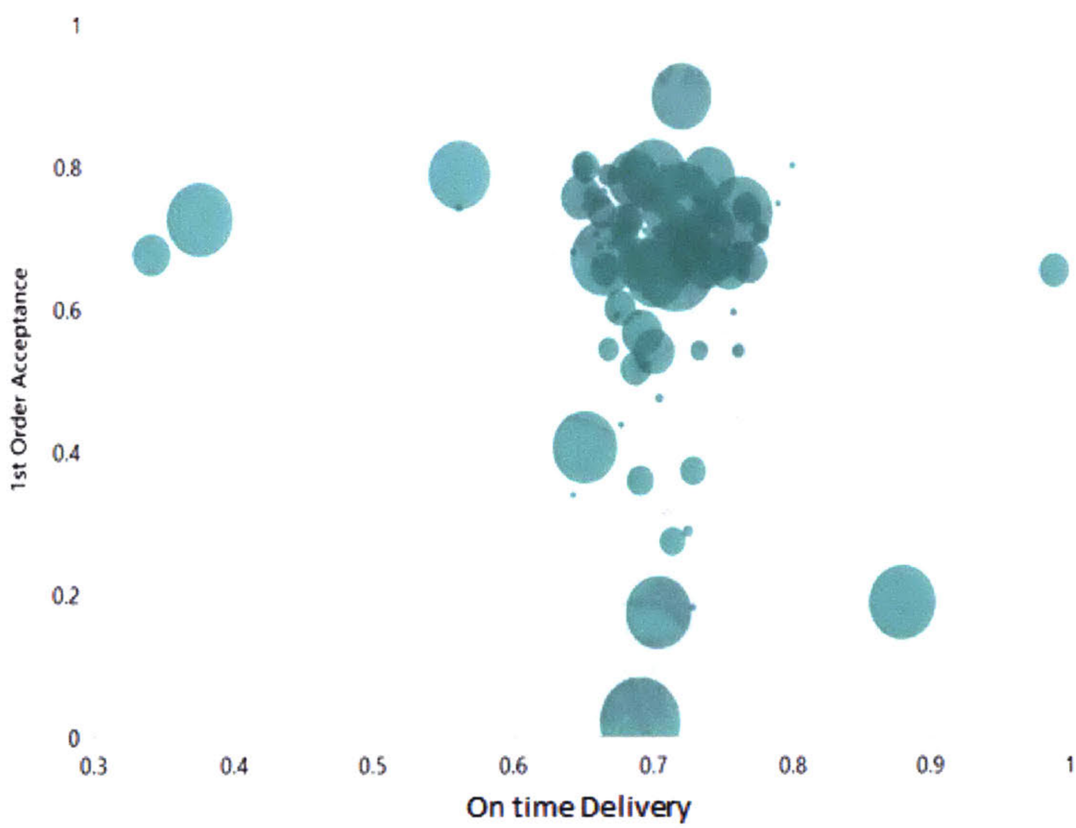
Considering the internal performance variation of Leaders and Laggards gave some additional strategy insights. Leaders are more consistent than Laggards regarding performance in both OTD and First Tender Acceptance. Figure 4.7 provides a visualization of this variation. Each bubble corresponds to a carrier's OTD and 1st tender acceptance rate. The carriers in the Laggard cluster have a range of on-time-delivery from 100% to 46% which is a spread of 54%, and 1st order acceptance spread range of almost 100%. The spread and range of both metrics are much smaller for the Leader cluster showing higher focus. This consistency could be a result of a more consistent and homogenous strategy in terms of the characteristic grouping within the Leader group. The spread of Laggard group in terms of OTD and AR is wide and polarized, showing that while there was higher performance in OTD in this group, there was poorer performance on First Order Acceptance and vice versa.

Figure 4.7. OTD and First Tender Acceptance Variation within the Leader and Laggard Groupings



Comparatively, within the Major Player group, performance ranged from 80% to 60% for the carriers, with several outliers displayed as dispersed bubbles in Figure 4.8. This means the major players are less polarized in general than Laggards, but there are some outliers in the groups who have high performance in one metric combined with low performance in another contributing to their categorization within this cluster. The size of the bubbles for this figure are representative of the truck count or fleet size of the relative carriers within this cluster. The Major player spread is shown separately in Figure 4.8 to avoid overlap of bubbles with the other clusters shown in Figure 4.7.

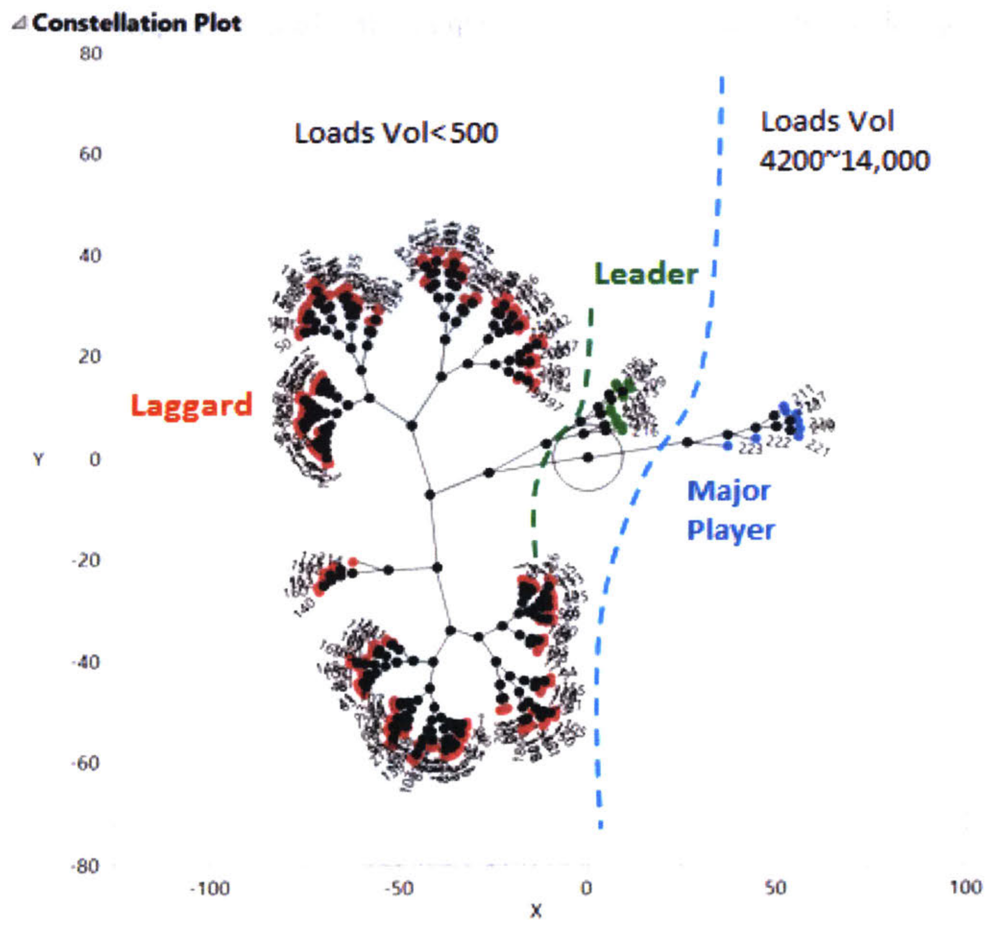
Figure 4.8: Dispersion within the Major Player Group for First Tender Acceptance and OTD



4.2.2 NON-ASSET BASED CARRIER CLUSTERING

Three performance clusters were also identified for the non-asset based carriers. The top performers in this group, non-asset Leaders, included 135 carriers with a perfect shipment rate of 81% while the Lagging group had 140 carriers and a perfect shipment rate of 23%. The group in the middle, non-asset Major Players had a perfect shipment rate of 51% and a carrier count of 43 carriers. Figure 4.9 is a visualization of the labeled clusters plotted against the first two principle components.

Figure 4.9. Constellation Plot showing Non-Asset Based Carrier Performance Clusters



To compare this non-asset clustering with the asset based carrier clustering, three variables were changed. First, fleet size was replaced by total number of loads as non-asset based carriers do not own their own fleets. This replacement was done to represent the non-asset based carrier's capacity as overall number of loads which is a proxy value. Second, lane focus was replaced by the number of lanes the carrier serves. Third, customer focus was replaced by the number of customers the carrier provides trucking service to. The reason for this replacement is to gain insight from the relatively small number of loads and higher lane spread seen in the non-asset based carrier community as compared with the asset based carrier community. The value of these variables is in creating more substantive variables to identify differentiating strategies for the non-asset carrier community which is important for effective clustering. The clustering result revealed 3 clusters of non-asset based carriers with distinct service strategies. Of the three non-asset based carrier clusters, the top carrier performance group, Leaders, took less loads overall which is also shown in Figure 4.10 through the load volume delineator depicted.

Leader Cluster:

This cluster is composed of relatively small capacity carriers who takes only lower number of loads from single shipper's lanes. This hyper lane and shipper focus gives this group its leading performance of an 86% perfect shipment rate.

Major Players Cluster:

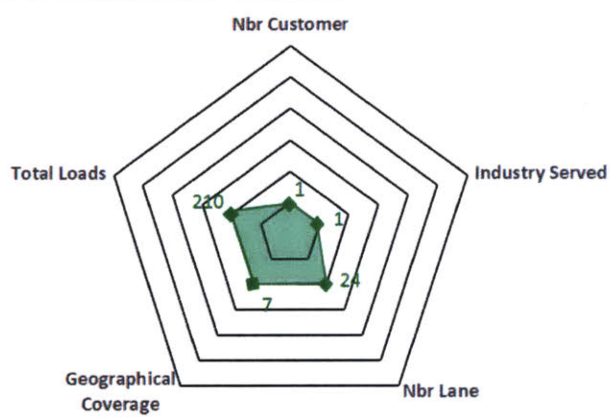
Carriers in this clustered group tend to hold relatively larger capacity with much higher volume of loads accepted. They also serve more shippers and have relatively more lanes to cover. They have a mediocre perfect shipment performance of 66%

Laggard Cluster:

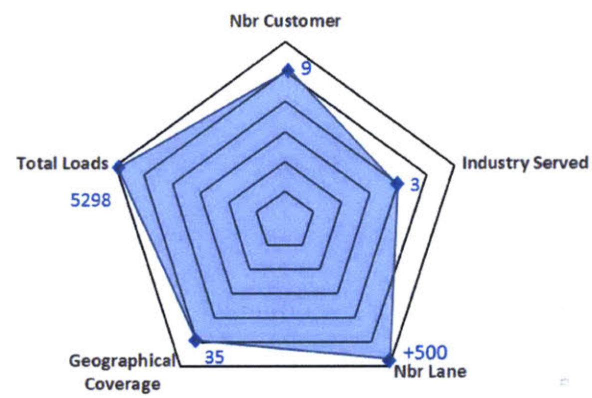
The laggard cluster is composed of small capacity carriers demonstrating less focus than the carriers in the leader cluster. Their lack of focus is apparent from the wider range of shippers served and significantly wider lane coverage. This might reflect a more opportunistic way of taking load tender instead of focusing on one shipper and focused lanes and cultivating capability in those limited lanes. This cluster has the lowest performance within non-asset based group of carriers with a 43% perfect shipment rate.

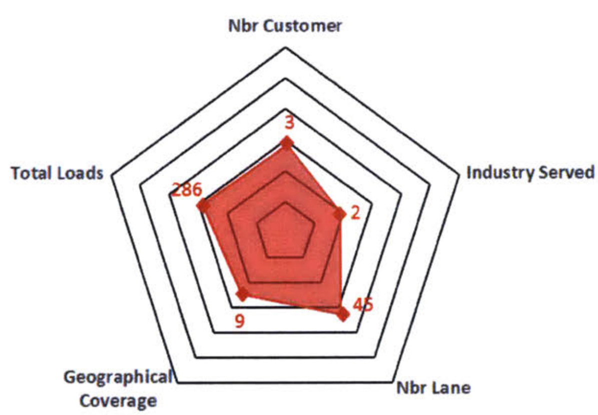
Figure 4.10: Non-Asset Based Carrier Strategy Profiles by Cluster

Cluster 1: Leaders



Cluster 2: Major Players





Cluster 3: Laggards

To enumerate the differences between the clusters, the following tables in Figure 4.11 give the average values for Customer Focus, Industry Focus, Geographical Coverage, Load

Volume, Customer Volume, and Lane Focus for each of the three identified major clusters. The differences in the strategies are highlighted by the significant variation between the clusters in Customer Focus, Lane and Load Volume, and Geographical Coverage. In comparing each of the tables, more focused carriers are more successful in their performance result.

Figure 4.11 Non-Asset Carrier Cluster Statistics

Non-Asset Leaders (44 carriers)

Dimensions	Mean	Std. Dev.	Min.	Max.
Customer Focus	269	201	45	949
Industries Served	1	0	1	1
Lane Focus	9	10	1	45
Geography	7	5	1	16
Loads	210	289	45	1,027
Customers	1	0	1	2
Lanes	24	24	1	78

Non-Asset Doers (19 carriers)

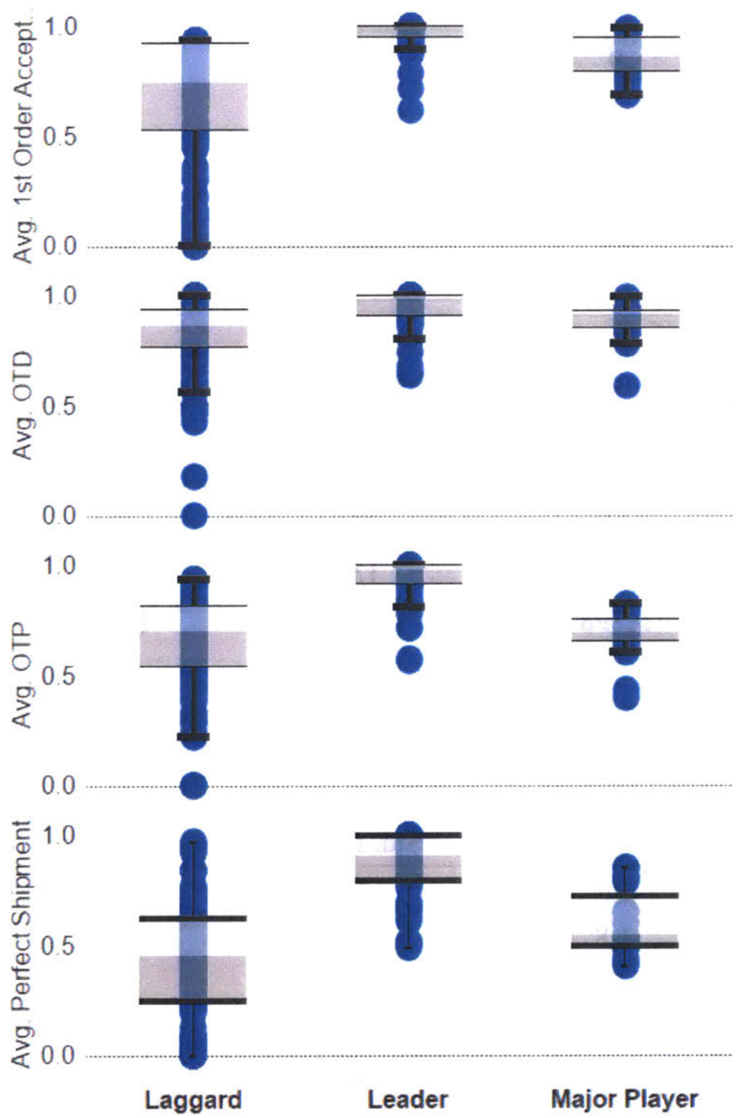
Dimensions	Mean	Std. Dev.	Min.	Max.
Customer Focus	1,134	965	260	3,961
Industries Served	2.6	1.6	1	5
Lane Focus	15	24	2	103
Geography	35	13	6	48
Loads	5,298	4,185	1,040	16,882
Customers	9	11	3	46
Lanes	799	580	23	2,113

Non-Asset Laggards (146 carriers)

Dimensions	Mean	Std. Dev.	Min.	Max.
Customer Focus	98	148	1	357
Industries Served	1.8	0.7	1	4
Lane Focus	13	19	1	102
Geography	9	9	1	38
Loads	286	345	1	1,920
Customers	3	1.6	2	11
Lanes	45	61	1	326

To expand on the argument for correlating focus with improved performance, the performance variation between the clusters in each of the performance metrics is given in Figure 4.12. As can be seen from the horizontal lines with a smaller spread between them, the Leaders' performance is shown to be more consistent, while Laggards' performance has a wider spread. From Figure 4.12, Major Players are very similar in performance to Leader in terms of OTD and acceptance rate but lose their performance strength in the perfect shipment in OTP.

Figure 4.12. Non-Asset Carrier Performance Variation



To further discuss the spread, Figure 4.13 shows that the Leaders are more consistent than Laggard group in terms of performance on both OTD and 1st Order Acceptance. The spread of Laggards in OTD and AR is wide and polarized. Of note, there is range of overlap between Leaders and Laggards that is less cleanly separated in comparison to the asset based carrier clusters.

Figure 4.13 Non-Asset Leader (Blue) and Laggard (Red) Bubble Plot showing OTD and 1st Order Acceptance Spread

Within the Major Players group, performance was more concentrated and less polarized. This could be, in part, due to the smaller size of this cluster rather than the spread correlating with performance. This concentration is shown in Figure 4.14 separately to avoid overlap of bubbles with the other groups.

Figure 4.14. Non-Asset Carrier Bubble Plot with Concentrated OTD and 1st Order Acceptance

4.2.3 Clustering and Regression Result Comparisons

Clustering and Regression

The regression analysis provided insights into the attributes of freight correlated with better performance on each of the three performance metrics and on the combination of all three. The clustering analysis grouped carriers and measured the performance of their strategies and roles within the freight industry against the same combination of three performance metrics. Combining the results from both analyses offered insights on both strategy and individual attributes. Further, these combined results allow for shipper profiling and analysis as discussed in the following Section 4.4, Shipper Profile.

While the regressions offered insights of the specific variables that increase the likelihood of better performance, specifically longer hauls (>700 miles) with younger price ages and longer lead times with asset based carriers performed better; the clustering analysis offered a different perspective specifically per the asset based carriers. Although the regressions pointed to asset based carriers improving the likelihood of better performance, the clustering analysis showed some non-asset carriers with different strategies outperforming asset based carriers and vice versa. This suggests that the underlying performance driver is not asset or non-asset, rather it is a set of strategies that serve to bolster performance.

Asset vs Non-Asset

Further delving into the asset vs non-asset carrier based, overall, non-asset carriers on average serve less customers and lanes than asset carriers. However due to their lower overall volume loads, their focus is not necessarily higher than asset carriers. Meanwhile, spread of the non-asset carriers was wider than that of asset based carriers likely due to the polarization of carriers by more differentiated geographical coverage and role filled within the industry. Common for both the asset and non-asset based carriers, focus improved the performance. This is an important insight as it implies that it is not what assets a carriers has that influence their performance, rather it is how their assets and or strategy is employed. This is shown in both the asset and non-asset based clustering results that imply that a large spread in terms of lane, shipper, and geographical strategy result in poor performance regardless of asset base. This is shown in both the asset and non-asset based clustering results in that in both categories of carriers, a leader cluster can be identified whose performance is consistently higher than the rest.

4.3 LOAD DENSITY ANALYSIS

Industry experience suggests that Load Density Consistency over time impacts carrier performance. This can be measured as how the variation in demand impacts performance. Two calendar years, 2015-2016, were used on a quarterly basis to view the volume of loads offered to a carrier over time. To measure variation in demand over time, two variables were constructed:

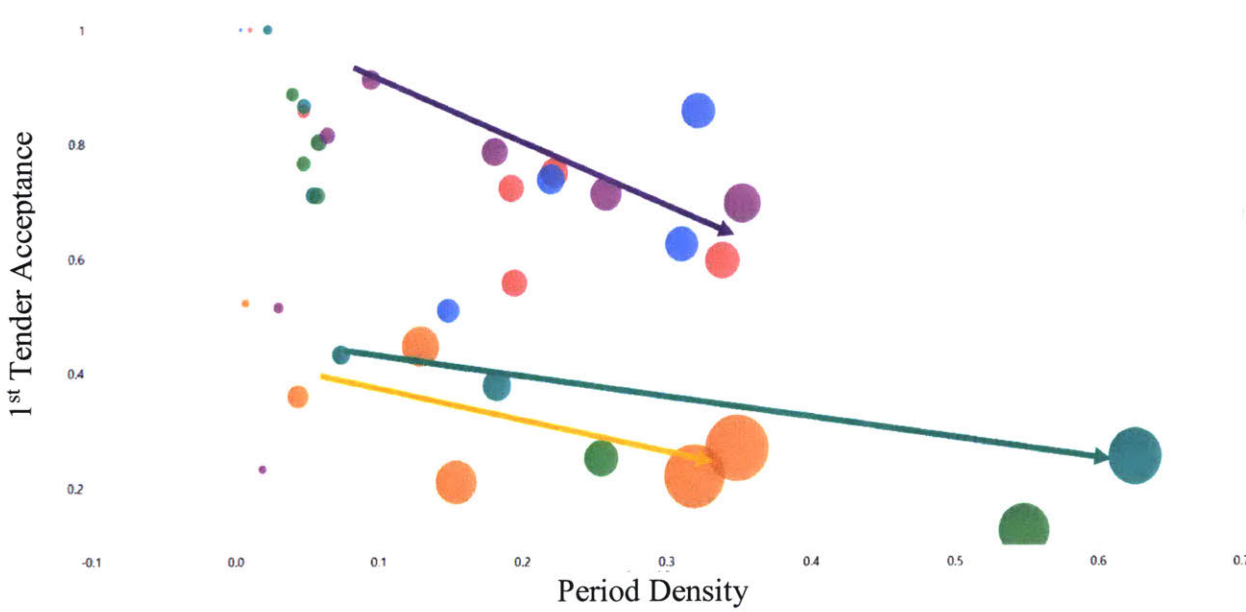
1. Period density – measured as the number of loads in each period as a % of total loads accepted by the carrier. The period was defined as a quarter of a year.
2. Standard deviation of Period Density- measured as the standard deviation of period density across two years to measure the demand volatility

Industry believe suggests that due to higher load volume in a period with high period density, acceptance will be lower in those periods than in other periods. A large standard deviation of period density means that the volume of loads taken by a carrier are not consistent across different periods, hence predictability and the overall performance should be lower. Two hypotheses were tested:

1. Higher standard deviation of period density across periods result in lower overall acceptance performance.
2. Higher period densities for a particular period result in lower acceptance performance in that period.

Each hypothesis was tested on each of the three main asset based clusters discussed in 4.1.1. For the first hypothesis, significant results of negative correlation of standard deviation of period density with average 1st tender acceptance were identified with P-value < 0.05 in cluster 2, Major Players. However, the results for Leaders and Laggards were insignificant due to very high p-values. The second hypothesis was also vindicated for the Major Players cluster by a negative correlation with a p-value < 0.05 between period density and 1st order acceptance for each period. Figure 4.15 shows an example of 5 representative carriers, shown in different colors for each carrier, from within the Major Player cluster to show the relationship witnessed between period density standard deviation and 1st Tender Acceptance. As shown, the standard deviation of period density increases for a carrier, the average first tender acceptance decreases. Detailed correlation results for period density and each of the performance metrics are provided in Appendix G.

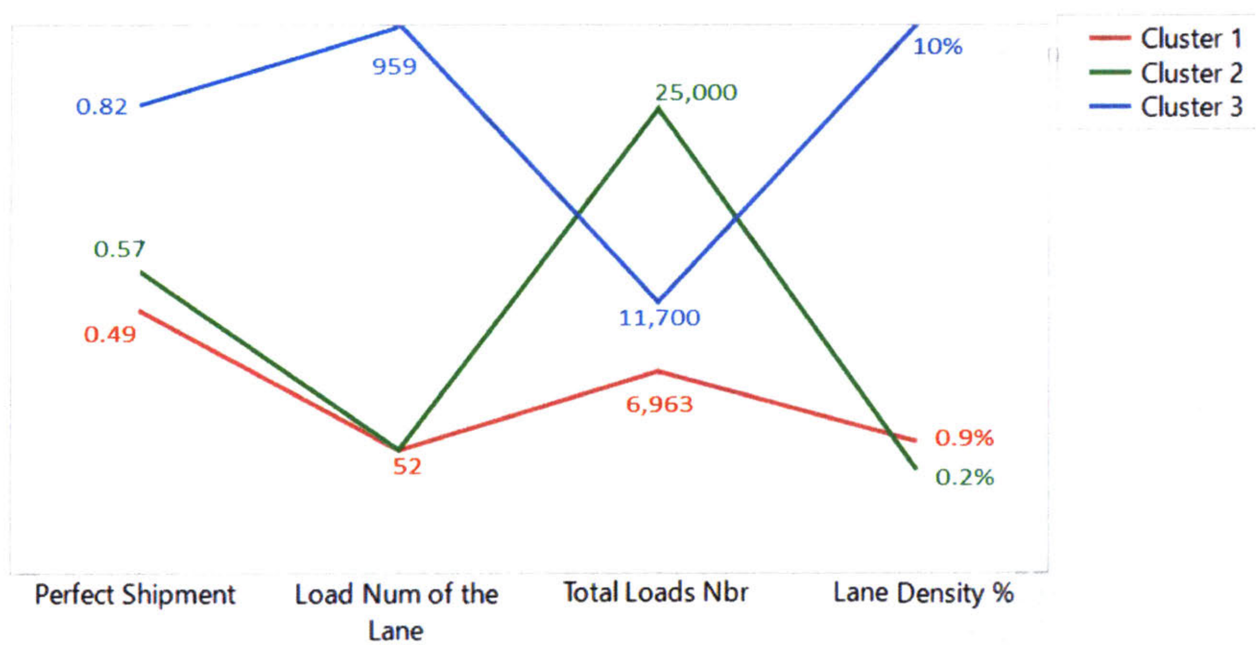
Figure 4.15 Decreasing First Tender Acceptance with Increasing Period Density Standard Deviation



Another industry belief is the existence of the "regional-champion phenomenon", that a carrier performs much better than average in certain lanes where it concentrates its resources. To investigate this, a lane density variable was established to organize a dataset by the load density level in each carriers' lane. Lane Density was measured as the number of loads in each lane for a carrier as percentage of the total loads accepted by the carrier. For example, a 10% lane density means one lane has 10% of total loads a carrier accepted out of the hundred lanes a carrier provides its service on.

For the lanes with high lane density in the major player cluster, Figure 4.16 shows three sub-clusters of carrier-lane combinations measured by perfect shipment rates. In this figure, cluster 3 show a sub-group of carrier's lanes with higher lane density and higher perfect shipment performance. This demonstrates that same carrier performs much better in those lanes with relatively more loads than the lanes with less loads. The more loads in those lanes might reflect that a carrier is more familiar with certain lanes or deploys more assets in certain lanes than the others, hence the performance is also higher. This is another proxy for the argument that higher focus results in better performance. Notably, cluster 3 with carrier-lanes with high density were not common within the dataset, representing only 2% of the data. Also, most of the major carriers have their loads evenly distributed across lanes they service.

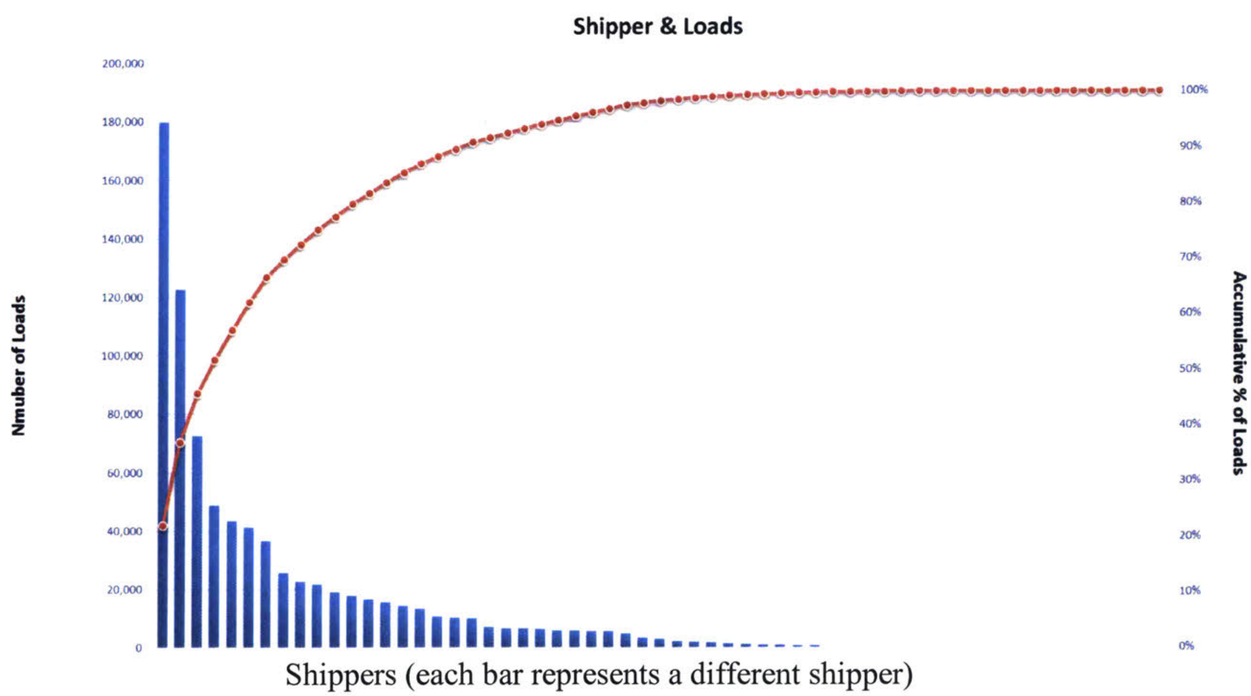
Figure 4.16 Clustered Lane-Carrier Combinations against Performance



4.4 SHIPPER PROFILE

To extend the regression analysis and the carrier clustering analysis to shippers, shippers were profiled by (a) performance and (b) their respective portfolios of carriers. The entire dataset was used for this analysis. The proportions of clustered carriers that each shipper used compared with their own measured performance elicited valuable insight. Before profiling the shippers, it was necessary to acknowledge skewed load distribution; strategies were not evaluated by load, they were measured by individual shipper performance. Not weighting by load volume meant that a smaller shipper's portfolio of carriers was weighted the same as a larger one to measure the impact of portfolio performance unbiased by one or two large shippers carrying significant load volume. To illustrate, Figure 4.17 shows that roughly 80% of the asset based loads represented in the dataset were offered by 12 of the 61 shippers analyzed. Non-asset based carrier loads showed the same distribution pattern, with a small subset of the brokers taking the majority of the loads.

Figure 4.17 Load Distribution by Shipper

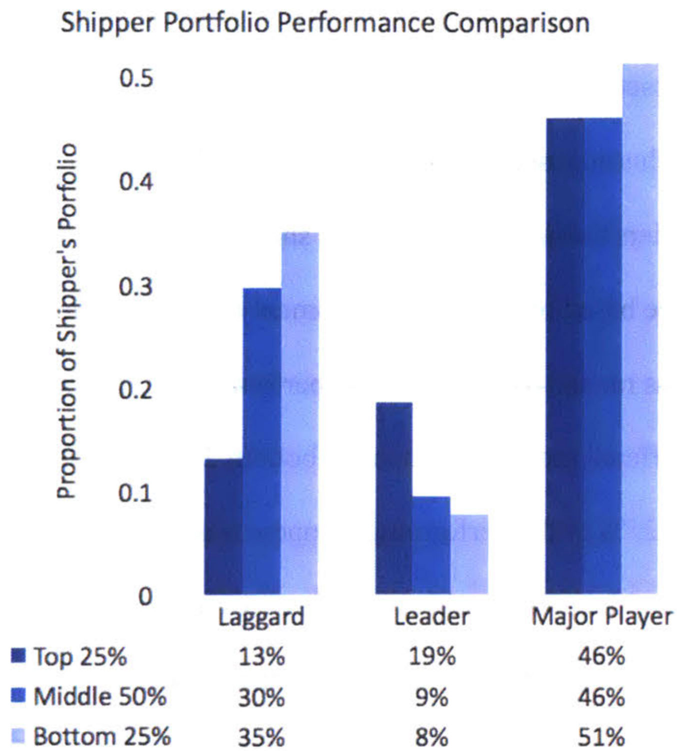


4.4.1 SHIPPER'S PORTFOLIO OF CARRIERS

Shippers who performed better than average in terms of perfect shipment rate had a more balanced portfolio of asset-based Leaders, Major Players, and well performing non-asset players in terms of the carrier clusters used for their loads. They also kept a narrower group of carriers, using fewer than 4 carriers to take their loads. As shown in Figure 4.18, shippers were split into quarters by performance based on the perfect shipment metric. In terms of perfect shipment rate, each of the shippers was ranked by their average performance. The top performing 25 percent of shippers, on average, offered more loads than the bottom 25 percent of shippers. The average load volume of the top 25% of the performing shippers was nearly double that of the lowest 25% of shippers. Also, one of the most significant differences in the composition of the shipper's portfolio of carriers was in the proportion of Leader and Laggard carriers used. As shown, the proportion of Leader and Laggard carriers was very different in the leading and lagging shipper groups.

Shippers who performed in the middle 50% had a high percentage of Major Players in their portfolios and 63 carriers on average. However, they did not offer high average load volume to each of these carriers, making shipper-carrier relationships and higher lane density more difficult to establish. High performing shippers used a higher proportion of leader carriers, while low performing shippers used a higher proportion of laggard carriers. Figure 4.18 represents asset-based carriers,

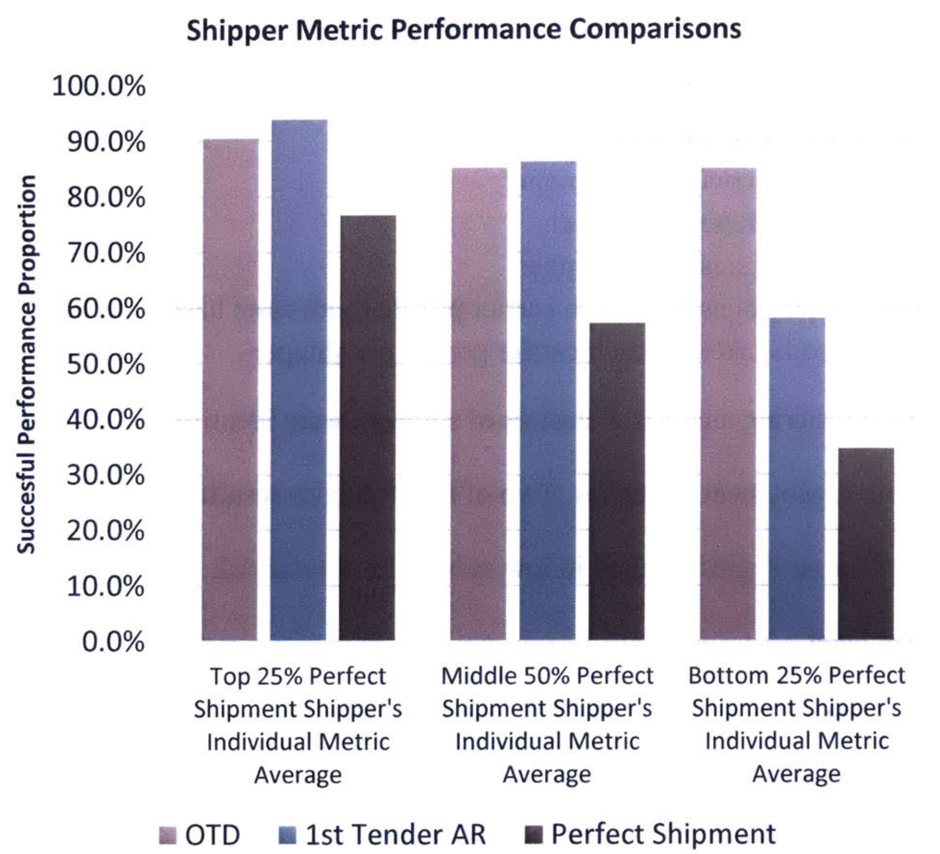
Figure 4.18 Shipper's Asset Based Carrier Portfolios compared against Shipper Performance



Lane focus was implied as one of the underlying factors in the different performance shown. Since lane focus was one of the more significant differentiators of high performing carriers, it followed that shippers using carriers that have better lane focus are likely to see better performance. For example, one of the shippers with the highest performance used SLDC, a carrier with a very high 80% perfect shipment rate, for 20% of its loads. SLDC only worked with the shipper in question, which means that there was a stronger relationship, higher lane density, more carrier focus, and subsequently better shipper performance.

In first tender acceptance rate, the top shippers are significantly outperforming the bottom 25 percent for the shipper community represented in the dataset as shown in Figure 4.19. The insights taken from the regression analysis in 4.1.1 shows that the highest correlated attribute with first tender acceptance is lead time. This correlation implied that the top shippers excel in giving ample lead time in tenders to their carriers. When the average lead times were calculated for the top and bottom 25% of perfect shipment performance for shippers this differential was confirmed.

Figure 4.19 Perfect Shipment Shipper Comparison to Individual Performance Metrics



4.4.2 Clustered Shipper Profiles

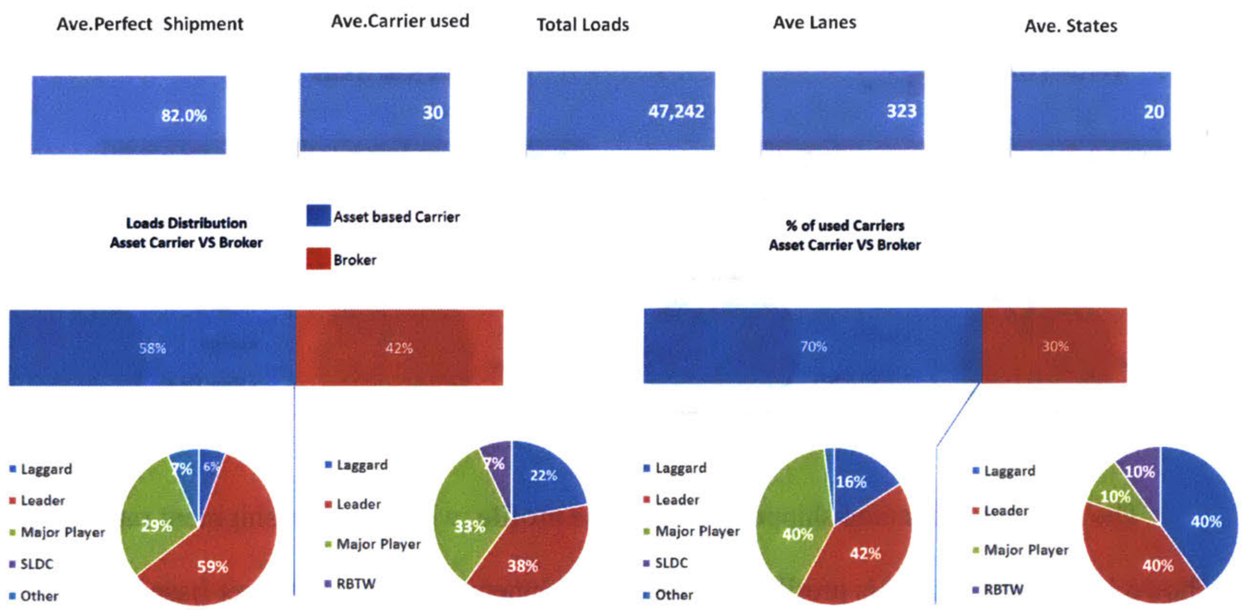
To further explore if a certain carrier deployment strategy can be identified for shippers who tend to receive higher service levels, another clustering was completed for shipper profiles. Shippers were clustered based on the following characteristics:

1. Number of Loads per shipper
2. Number of Lanes per shipper
3. Number of States covered per shipper
4. Perfect Shipment Rate per shipper
5. Number of Carriers used per shipper
6. Number of Carriers used by each carrier profile (both asset based and broker) per shipper.
7. Number of loads taken by each carrier profile per shipper

Through the clustering analysis 4 clusters of shippers were identified with different respective carrier deployment strategies. Two of the strategies resulted in high service level received, one of the strategies resulted in low service level received, and the last resulted in medium service level.

When clustered instead of ranked, four performance groups were identified with differing underlying group profiles. Two shipper types had high performance levels, high performing shipper Group 1 and high performing shipper Group 2. The last two clusters were composed of shippers with profiles measured with medium performance and low performance in terms of perfect shipments. Figure 4.20 shows the profile of the first high performing shipper group representing six shippers and includes the breakdown between asset and non-asset carriers within their profile. Of note, there is a near even load distribution between asset and non-asset carriers. As mentioned before in the regression analysis, asset carriers had a higher likelihood of having a perfect shipment. What this profile implies is this group of shippers use both broker and asset base carriers non-discriminately, but they extensively use the leader carriers to move the majority of their loads. They tend to supplement their portfolio with major carriers in those lanes not covered by leader carriers. There are 6 shippers in this cluster who received an 82% perfect shipment rate on average. This group of shippers used an average of 30 carriers in total and use both asset-based carriers and non-asset carriers. In both groups, they used primarily carriers within the Leaders cluster & Major Players cluster to carry most of their loads. They also used some carriers from the laggard group. However, carriers in the laggard group did not take a significant portion of loads from the shippers, as shown in Figure 4.20. The interpretation if this is that this group of shippers enlists certain laggard carriers to ensure certain lane coverage or as capacity back-up, but does not put them in the top positions in their route guide.

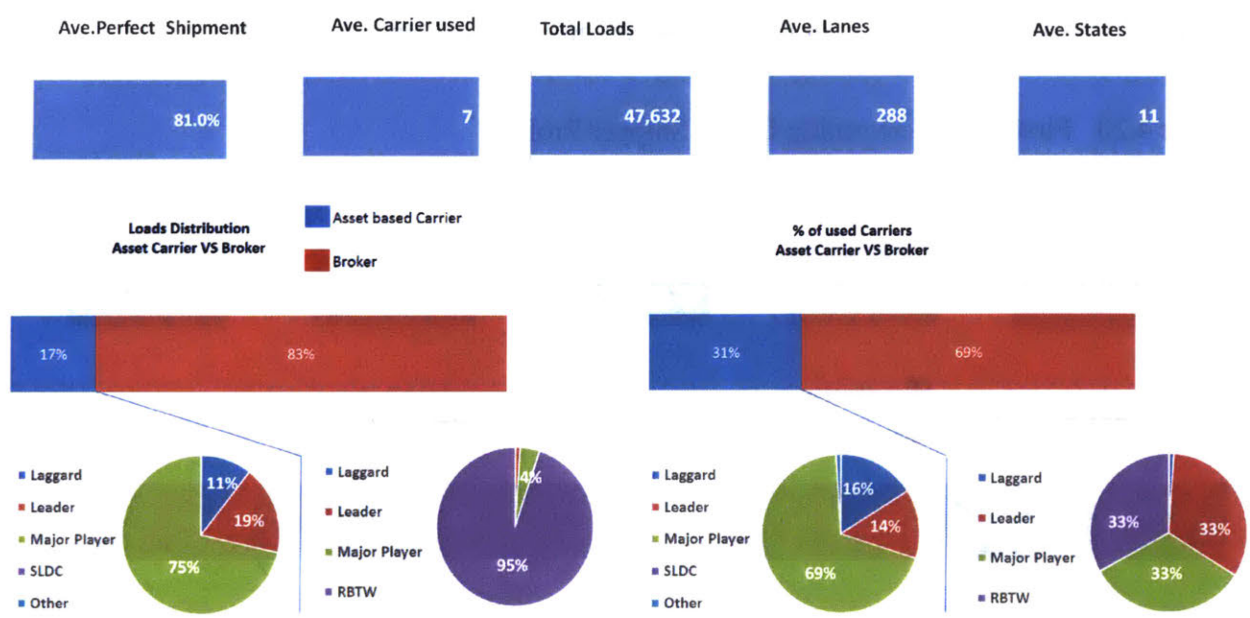
Figure 4.20. First High Performing Group Shipper Profiles



The second highest performing group of clustered shippers showed a significantly different allocation of loads to asset and non-asset carriers. The profile representing 13 shippers, as shown in figure 4.21, also performed very well with a perfect shipment rating of 81% but strongly favored brokered business. When the underlying carriers used in the brokered business allocation of this cluster were assessed they were the better performing non-asset carriers.

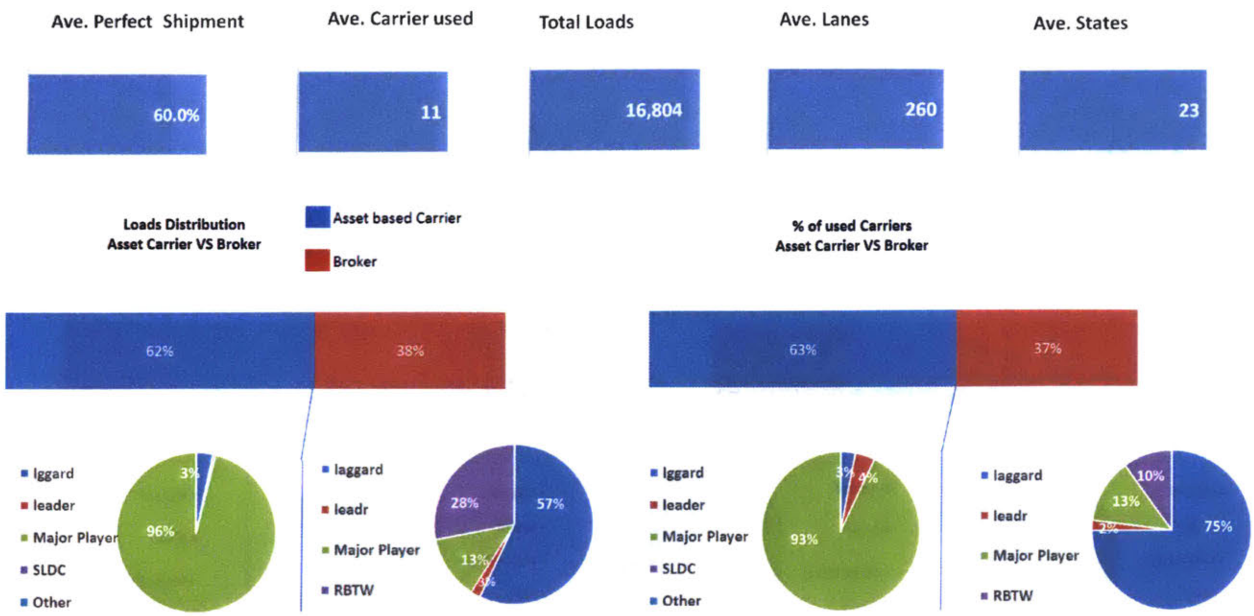
Although this group used brokered business significantly, they only used the higher performing non-asset carriers in their brokered business making them much more selective in their choice of brokers (Such as RBTW, which is a special outlier carrier with high performance found in the carrier clustering exercise) than the subsequent clusters of shippers. Also, this group used fewer carriers, carriers that maintained higher lane focus over a narrower geographic region. This potentially may have influenced the more successful performance results seen.

Figure 4.21. Second High Performing Group Shipper Profiles



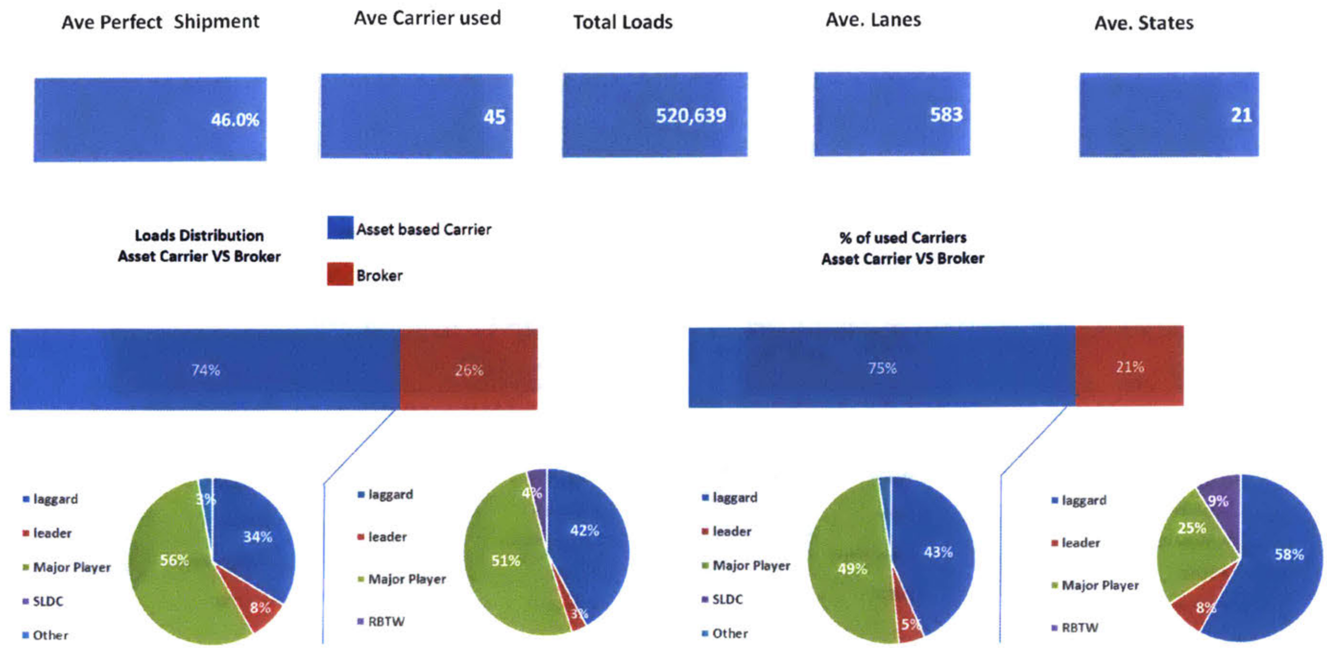
The mediocre clustered shipper performance profile, with a perfect shipment rate of 60%, is shown in Figure 4.22. This profile represents 14 shippers. The usage of asset-based carrier and non-asset carriers are quite balanced. But their carrier base used is not composed of many leader carriers in either category. Within the non-asset group of carriers, a large volume of the loads are shipped by laggard non-asset carries. In the asset-based carrier group, a few major player carriers covered majority of their loads.

Figure 4.22. Medium Performing Shipper Profiles



The final clustered group of shippers represented 23 shippers and showed even heavier use of the asset-based community in comparison to brokers (Figure 4.23). This suggests that although the asset-based community saw better performance than non-asset on average, there was a better way to leverage their resources. Using an asset based carrier blindly, meaning on a lane they do not typically run, in a geographic area they are not familiar with, or without appropriate lead time for their own planning purposes, did not improve a shipper's service performance average. The shippers in this cluster on average received a 46% perfect shipment rate. Two main differentiators are present in this cluster. The first is that this cluster of shippers use more carriers overall, 45 carriers per shipper on average. The second is that this cluster of shippers used significantly less leader carriers in both their asset-based carrier group and non-asset based carrier group, and they extensively employed major player carriers and laggard carriers to cover their loads.

Figure 4.23. Low Performing Shipper Profiles



These clustered shipper profiles suggest that higher performing shippers always have a combination of asset carriers and brokers in their profile. Although overall asset carriers perform better than brokers, there are high performance carriers in broker group and vice versa for asset groups. Using one or the other did not guarantee strong or poor performance, but no shipper at the top of the performance spectrum exclusively used one or the other type of carrier. Additionally, shippers with high service extensively use niche carriers in both the asset carriers group and broker group suggesting carrier focus on lanes, area, and industry is very important. Two examples of high performing non-asset carriers came to light; RBTW and SLDC were two non-asset carriers found to be as high-performance carriers that drove high service for the that shippers used them. From this research, the strategy to receive good service would be to use asset and non-asset niche players to cover a many lanes as possible. It is of less importance for a carrier to be asset based than it is to have familiarity with the geography and lanes and to have a focused strategy. Thus, it can be concluded that consistent use of a carrier and relationship development between a shipper and a carrier on a set of lanes is important to performance success.

As a means of comparison between the shipper performance profiles, figure 4.24 shows that both the highest and lowest performing shipper groups heavily relied on asset based carriers while the portfolios of the second highest performing shipper group and medium performing shipper groups are near inverses of each other. This implies that choosing a carrier simply based on their asset base will not result in stronger performance, rather, there are more underlying performance measures that should be considered in the measurement and choice of the tender list of carriers on the part of the shipper community. This figure also shows that shippers receiving high performance tend to have a combination of asset-based carriers and brokers. Although there is a view in the industry that asset-based carrier performs better than non-asset, the results show that no significant difference in usage of this two categories of carriers to support this opinion. These findings support the view that there are high performing carriers in both groups, and shippers who can leverage both asset and non-asset capacity strategically receive good service.

Figure 4.24 Shipper Profile Comparisons: Asset vs Non-Asset Carriers

Performance	Perfect Shipment Rate	Asset Based Proportion of Carriers	Non-Asset Based Proportion of Carriers
High Performance Profile 1	82%	70%	30%
High Performance Profile 2	81%	33%	67%
Medium Performance Profile	60%	63%	37%
Low Performance Profile	46%	79%	21%

The Service level shippers receive is positively correlated with the proportion of leader carriers they can utilize to fulfil their demand. This relationship is shown in figure 4.25.

Following in this view, a shipper should identify leader carriers in its shipping lanes, and put those leader carriers on top of their route guide for those lanes.

Figure 4.25 Shipper Profile Comparisons: Leader Carriers

Performance	Perfect Shipment Rate	% of Leader Carriers (both asset carrier & broker)	% of Loads by Leader Carrier (both asset carrier & broker)
High Performance Profile 1	82%	42%	51%
High Performance Profile 2	81%	51%	82%
Medium Performance Profile	60%	25%	30%
Low Performance Profile	46%	5.4%	6.7%

Major player carriers are the backbone the shipping capacity offered in the market as they are the largest carriers serving the broadest market, all major shippers use Major Players. This usage is shown in figure 4.26. Although they do not guarantee top performance for all their loads accepted, they provide wide capacity coverage.

Figure 4.26 Shipper Profile Comparisons: Major Player Carriers

Performance	Perfect Shipment Rate	% of Major Player Carriers (both asset carrier & broker)	% of Loads by Major Player Carriers (both asset carrier & broker)
High Performance Profile 1	82%	31.0%	30.6%
High Performance Profile 2	81%	44.2%	16.0%
Medium Performance Profile	60%	63.4%	64.5%
Low Performance Profile	46%	42.0%	5.7%

4.4.3 Shipper Clustering with Lane and Carrier Load Density

In addition to altering a carrier deployment strategy to optimize a shippers' received service level, industry experience suggests a shippers' behavior can influence its carrier's performance and subsequently its service level from the respective carrier. For example, a common belief is that if a shipper offers enough loads to a carrier in a lane, the carrier's performance in that lane will improve. One of the explanations for this is that the carrier might adjust its asset deployment to capture more predictable demand. To verify this belief, the data was prepared into each shipper-lane combination and the load density per carrier in each shipper-lane was calculated. Then the same clustering method was used to classify shipper-lanes into clusters with different service levels. In this case, for performance the focus was on on-time-delivery and 1st tender acceptance, as the clustering process revealed a significant difference between these two indicators.

Load density per carrier-lane is correlated with shipper performance. Low performance for shippers was strongly correlated with low density carrier-lanes. Figure 4.27 shows the results of clusters of shippers (clustered by lane volume) with density per lane and density per carrier-lane. This highlights that high performance was not as strongly dependent on density per carrier-lane as low performance. However, high carrier-lane density consistently presented with higher performance. This further supports the argument for higher focus on lanes tied to better performance. The full clustering results including the corresponding dendrogram and sensitivity values tied to figure 4.27 can be found in Appendix H.

The figure reveals three distinct behavior/service level combinations:

1. High Service level with high load density, shown in Cluster 2 and 5
2. High Service level with low load density, shown in Cluster 1
3. Low Service level with low load density, shown in Cluster 3 and 4

The results show that higher load density per carrier is associated with higher service levels, but low density shipper-lanes have both high and low service levels. One of the explanations could be that this data does not take into consideration loads from different shippers offered to the same carrier in a geographical lane. This means that the load density from this shipper is low, but not low to the carrier after all the shippers' loads are combined. Additionally, when the load density per carrier is low, the load density per shipper-lane is also low. As the load density per lane reflect the shipper's business reality, this shows there might be limited opportunity to improve a carrier's service to those shippers by allocating more loads to one carrier. Further, low service level occurred in few shipper-lanes. For example, there is very low OTD in cluster 3, and very low 1st tender acceptance in cluster 4. This might give direction for a focused effort to identify low performance root causes and improve the overall service level for this shipper community.

Figure 4.27 Clustering of Shipper Lanes by Density

Cluster	Volume of Shipper Lanes	OTD	1st Tender Acceptance Rate	Load Density Per Lane	Load Density by Carrier - Lane
1	31,999	92%	93%	8.6	4.1
2	2,281	88%	90%	174	73
3	7,551	10%	91%	10.5	4.5
4	2,937	75%	4.2%	2.9	2.2
5	189	92%	88%	975	386

Further, when using the percentage of loads in different carrier-lane groups to categorize shipper clusters, different groups of shippers have different combinations of Carrier-Lane Density. These differences manifested in some of the shippers with low carrier-lane density having lower perfect shipment performance. This performance differential and the full details of the clustering results are available in Appendix I. Additionally, it should be noted that shipper lane density reflects the business reality in which a shipper cannot change and the shipper attributes used in this analysis cannot cleanly predict whether a shipper will receive a high or low service level. This is because high performance exists in both high density and low density lanes. There are likely other factors driving the service level received by shippers beyond those used in this shipper clustering analysis. These other attributes are explored in the carrier clustering and subsequent shipper profiles of carriers developed earlier in this analysis and discussed within the prior section, 4.3. These attributes, in conjunction with the way in which a shipper uses different carriers, link to the suggested carrier deployment strategy discussed in the following chapter 5.

5 DISCUSSION

This chapter provides insights and industry guidance drawn from the successful performance profiles of the results of the regression and clustering analysis. It also responds to the question of what makes a leader or a laggard in the freight industry in terms of characteristics gleaned from the collective models used in this thesis. The limitations and sensitivities of these results, models, and dataset are evaluated.

5.1 INDUSTRY GUIDANCE

As discussed in the regression results, there are clear relationships between individual attributes and each of the three-success metrics used. This is not new knowledge but it does provide some guidance as to what variables have the most significant impact on a specific performance metric. More applicably in terms of industry guidance, the insights from the clustered attributes are more predictive and arguably actionable than those of regression in terms of maintaining lane balance in a carrier's portfolio to become more successful as they consider grouped attributes rather than individual attribute impact. From the clustering analysis, the focus of a carrier surfaced as one of the stronger indicators of performance. However, more focused carriers have a smaller geographical coverage offering. Meanwhile, the carriers in the major players group, who have a much wider geographical coverage and mediocre performance overall, have certain lanes with more focus to make them a "regional leader" and an appropriate strategic choice for those lanes. This Implies a carrier deployment strategy for a shipper to take advantage of different clusters of carriers to optimize the service level while fulfilling all the truckload demand from its different geographical areas. The actual carrier portfolios employed by different shippers, as revealed in Section 4.4.2, also provide detailed empirical evidence for this strategy. The strategy drawn from this would suggest the pecking order of carrier selection follow the three guidelines below:

1. Identify leader carriers in the lanes a shipper needs truckload service on and maximize leader carrier's available capacity.
2. Identify Carriers in Major players group, maximize their capacity in the lanes they are "regional leaders" on.
3. Complement the remaining loads using carriers in the major players group.

The first of the guidelines, identifying more focused carriers on the lanes a shipper needs stems from the focus based performance findings from the asset and non-asset clustering analyses completed. Leader carriers implies that the carriers selected for the routing guide would likely have an established relationship with the shipper, familiarity with the lanes used, and have free capacity assets or access to capacity on a relatively consistent basis within the specified geographic region. Along with this recommendation, the shipper should follow the guidance gleaned from the regression analysis; they should offer ample lead time to their carriers, have younger price ages, and attempt to control their shipment volatility.

The second guideline is to use regional leaders when a shipper needs a broader coverage. These leaders come from the clustered Major player group but play a strong role in the higher performing shipper portfolios and fill a significant market need. Regional leaders also stem from a specific strategy of focus, a shipper should look to find a carrier with familiarity on their respective lanes and region.

The final guiding principle for shippers is to complement the remaining loads they are offering to carriers following a strategy from the Major Player cluster as opposed to the Laggard Cluster of shippers. Not every lane will have constant volume and carriers with both capacity and familiarity with that lane especially in an industry highly susceptible to seasonal and meteorological fluctuations. However, in selecting an unknown carrier this research suggests finding a larger carrier with a clear strategy beyond universal load acceptance.

This research also supports the concept of relationships inherit in the lane focus measurement. Higher lane focus correlated with stronger performance giving strategic value to the niche carriers with lane and shipper familiarity. This idea of focus builds upon the research of Zsidisin, Voss, and Schlosser (2007), who studied the positive impacts of relationship building on success as discussed in the literature review. These close relationships are a proxy for focus. Additionally, this research also supported and built upon the research of Armstrong and Associates (2009), who showed that more volume is not correlated with a high acceptance rate. This research further strengthens the idea of consistency, in volume and lanes, as well as focus as the keys to success in the freight industry.

In terms of consistency, one approach that proved useful in interpreting some of the underlying potential causes of the regression results was comparing the results against the volatility of demand. For example, in the first order acceptance model, asset based carriers were found to have significantly better first tender acceptance than non-asset carriers. However, this difference may not be due to non-asset carriers being less reliable. Rather, when the underlying demand trends for those carriers were considered, the lanes and corresponding volumes were not as consistent over time as the lanes and volumes given on the asset based first tender lanes. The difference could be accounted for by the trend of shipments which were offered first to non-asset carriers; these tended to be more volatile and less predictable, making it more difficult to guarantee first tender acceptance. This difference indicated that having pre-planned shipments would serve to stabilize demand and could in turn increase first tender acceptance.

5.2 Limitations and Sensitivities

Although clear implications and industry guidance are offered in this thesis, there are limitations and sensitivities inherent in the research. The first is a question of data representativeness. The data was pulled from TMC, a single source; in that regard, the data is biased towards only the carriers and shipper with whom TMC interacted on a by-load basis. While this thesis highlighted potential strategies derived from the available dataset, these strategies may not be broad enough to extrapolate and apply the recommendations to the entire freight industry network.

The second sensitivity was the model limitations that could have impacted the results. Clustering was completed on two years of data and regressions were completed on three years of data. The practice of using a larger time span of data usually implies less biased data, as there is a larger sample, but comes at the price of smoothing the results as performance is averaged over that amount of time. This results in analysis that represents an average and is less volatilely impacted by factors like seasonality. The final disclaimer to this research is the likelihood of external factors impacting the results. Although the data had 23 variables this was not a robust variable group. There are likely other factors that could have an impact on the results seen which are not represented in the data and therefore did not surface as significant variables within either the clustering or regression analysis.

6 CONCLUSION

6.1 SUMMARY

This body of research confirms industry belief that choosing the right carriers that match focus with the shipper will improve shipper performance. Choosing the right strategy in using focused and planned freight will positively impact service performance. This research showed that there are groups of attributes that work together to improve freight performance. Some of which were found to be longer lead times, consistency of load volume, geographic and lane focus, younger price ages, and certain mixes of types of both asset and non-asset carriers within a shipper's portfolio.

6.2 FUTURE RESEARCH

This thesis provides insights into the freight industry through carrier strategies and profiles but by no means should be considered a comprehensive guide for carriers or shippers in the industry. From the insights gained from the research completed, some opportunities were found for future research. Specifically, there is ample opportunity to delve deeper into geographical performance clustering and to study market and meteorological impacts on clustered strategy performances. Comparing location, economic seasonality, and weather driven industry shifts to clustered carrier and shipper performance over time could lead to improved forecasting of freight demand, behavioral insights, and eventual performance.

Additionally, the research offered within this thesis suggests that carriers cater to different market needs and the differences in strategy discovered within the clustering analyses for both asset and non-asset carriers suggests that the uniform scorecards used to evaluate shipper performance may not be the most appropriate way to rate carriers. There is a research opportunity in developing more strategy specific key performance metrics (KPIs) and corresponding scorecards in order to give shippers a better understanding of the performance of the carriers relative to their specific market needs.

6.3 BROADER SIGNIFICANCE

The broader significant of this research is in the suggestion of building relationships with carriers where possible. Rather than focusing on universal blanketed metrics and changing carriers based on their short term performance, to improve overall service it is important to holistically consider a full profile and ensure that it is simultaneously balanced and composed of focused carriers.

REFERENCES

Abramson, M., Sawant, A., Massachusetts Institute of Technology. (2012). Impact of risk sharing on competitive bidding in truckload transportation. (Unpublished master's thesis). Massachusetts Institute of Technology, Massachusetts

American Trucking Association. (2016). American Trucking Trends 2016. Washington DC: American Trucking Association.

Amiryan, Bhattacherjee, & Massachusetts Institute of Technology. Engineering Systems Division. (2015). Relationship between Price and Performance: An Analysis of the US Trucking Market. (Unpublished master's thesis). Massachusetts Institute of Technology, Massachusetts

Armstrong & Associates, Inc. (2009). Carrier Procurement Insights: Trucking Company Volume, Cost, and Pricing Tradeoffs – 2009 (n.p.).

Bardi, E., (1973). Carrier Selection from One Mode. Transport Journal. 12(4), 23-29.

Caldwell, E. R., Fisher, B. C., & Massachusetts Institute of Technology. Engineering Systems Division. (2008). The impact of lead time on truckload transportation rates. (Unpublished master's thesis). Massachusetts Institute of Technology, Massachusetts.

Chen, Tsai, & Massachusetts Institute of Technology. Engineering Systems Division. (2016). Multi-Stop Trucking: A Study on Cost and Carrier Acceptance. (Unpublished master's thesis). Massachusetts Institute of Technology, Massachusetts

Cheng, W., & Massachusetts Institute of Technology. Engineering Systems Division. (2009). Measuring Performance of Transportation Carriers. (Unpublished master's thesis). Massachusetts Institute of Technology, Massachusetts

Martens,B. & Suzuki, Yoshinori. Iowa University (2011). Stale Rates Research: Benefits of Frequent Transportation Bids.

McGinnis, M. A., (1989), A Comparative Evaluation of Freight Transportation Choice Models, Transportation Journal, 29 (2).

Shmueli, G., Bruce, P. C., Stephens, M. L., & Patel, N. R. (2017). Data mining for business analytics: concepts, techniques, and applications in JMP Pro. Hoboken, New Jersey: John Wiley & Sons.

Varma, A. (2008). Measurement Sources for Freight Performance Measures and Indicators. Retrieved from http://www.dot.state.mn.us/ofrw/PDF/2008PerfIndic.pdf

Whyte, J. L., (1992), Characteristics of Transport Managers which Influence Hauler Selection. International Journal of Physical Distribution and Logistics Management, 22(5).

Winter. (2011). Deriving Strategic Advantage from Truckload procurement. CSCMP.

Zsidisin, G. A., Voss, M. D., and Schlosser, M. (2007). Shipper-Carrier Relationships and Their Effect on Carrier Performance. Transportation Journal, (2), 5.

APPENDICES

Appendix A. Values cleaned from the Dataset

- Rate per mile <$0.50
- Origin or Destination outside of the 48 contiguous states
- Special Character input errors
- Missing tender or pick up date fields
- Missing rate or mileage fields
- Customers with <52 loads per year
- Less than Truckload (LTL) trips
- Mileage <250 miles
- Multi Stop trips
- Dedicated contractual runs
- Anything other than Dry Load
- Year outside of 2014 to 2016

Appendix B. Variable Definition Overview

No.	Variable	Description
1	Tender Lead-Time	Time between requested delivery time and tendered time
2	Load Volume	Volume of load by weight (kg)
3	Linehaul Length	Distance of linehaul measured in miles
4	Rate	Price of the load measured in dollar value
5	Rate Age	Age of the active price measured by the time differential between tendered time and price effective time
6	Fuel Surcharge	A component of total rate measured by dollar value
7	Spot Quote	Shipper's industry for the concerned load
8	Origin Dwell Time	If the load is tendered for spot market or not
9	Destination Dwell Time	Carrier's time of stay in pick-up location measured by hour
10	Tender Date	Carrier's time of stay in destination location measured by hour
11	Ship Date	Actual day of the week the load is tendered to the carrier
12	Shipper Industry	Actual day of the week the load is shipped by the carrier
13	Asset or Non-Asset Carrier	If the carrier is an asset-owning carrier or a broker
14	Carrier Fleet Size	Size of the carrier measured by the number of trucks the carrier owns
15	Customer Focus	How many loads per customer for the concerning carrier
16	Lane Focus	How many loads per shipping lane for the concerning carrier
17	Geographical Coverage	How many states the concerning carrier carries loads from and to
18	Period Density	Load in the concerning period (quarter) as % of total loads in the time frame for a carrier
19	Period Consistency	Variation of number of loads over periods measured by standard deviation of Period Density
20	Lane Density	Load in the concerning Lane as % of total loads the carrier Taken
21	Lane Consistency	Variation of number of loads over lanes measured by standard deviation of Lane Density

Appendix C. First Tender Acceptance Regression Results

In the results shown below Day of Ship [1] is Friday; Day of Ship[2] is Saturday and Sunday; Day of ship [3] is weekday is the baseline This is the same for the Day of Tender. Tender lead time is shown in days.

Source	LogWorth	PValue
Industry Type	1502.784	0.00000
Fleet Size	1127.862	0.00000
Total Rate (Actual)	341.371	0.00000
Miles	259.867	0.00000
Price Age	65.675	0.00000
Fuel	47.679	0.00000
Date of Ship	43.574	0.00000
Tender Leadtime	35.380	0.00000
Date of tender	32.531	0.00000
ActualWeight	3.157	0.00070

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.8560	0.0609	928.9	<.0001*
Total Rate (Actual)	0.0006	0.0000	1564.3	<.0001*
Fuel	-0.0013	0.0001	213.7	<.0001*
Miles	-0.0010	0.0000	1189.2	<.0001*
Tender Leadtime	-0.0213	0.0017	157.4	<.0001*
Price Age	-0.0007	0.0000	296.3	<.0001*
Fleet Size[A.) 1]	0.2671	0.0413	41.7	<.0001*
Fleet Size[B.) 2-10]	0.3819	0.0247	239.0	<.0001*
Fleet Size[C.) 11-50]	0.1441	0.0178	65.2	<.0001*
Fleet Size[D.) 51-100]	-1.0474	0.0162	4163.8	<.0001*
Fleet Size[E.) 101-300]	0.3155	0.0146	467.9	<.0001*
Fleet Size[F.) 301-999]	-0.0637	0.0155	17.0	<.0001*
Fleet Size[G.) 1000+]	0.1259	0.0115	119.6	<.0001*
Industry Type[Automotive]	-0.4333	0.0140	961.1	<.0001*
Industry Type[F&B / CPG]	-0.1902	0.0125	232.8	<.0001*
Industry Type[Manufacturing]	0.9724	0.0149	4266.1	<.0001*
Industry Type[Other]	-0.5824	0.0366	253.8	<.0001*
Date of Ship[1]	0.1437	0.0104	192.7	<.0001*
Date of Ship[2]	-0.1432	0.0149	92.1	<.0001*
Date of tender[1]	0.6541	0.0572	130.7	<.0001*
Date of tender[2]	-1.3445	0.1133	140.8	<.0001*
ActualWeight	0.0000	0.0000	11.5	0.0007*

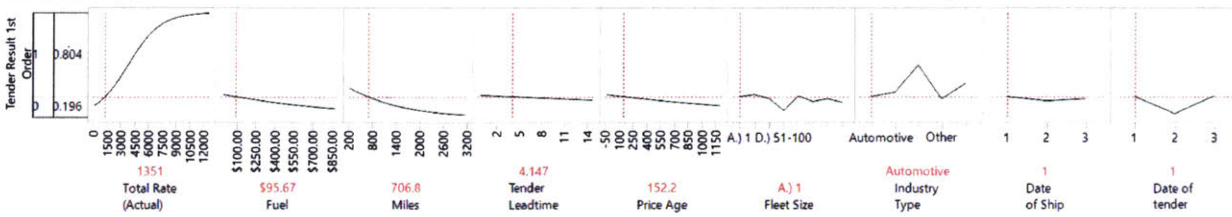
<3 Days

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Tender Leadtime	0.0919	0.01072	73.41	<.0001*

>3 Days

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Tender Leadtime	-0.0178	0.0028803	38.20	<.0001*

▾ **Prediction Profiler**



Appendix D. On Time Delivery Regression Results.

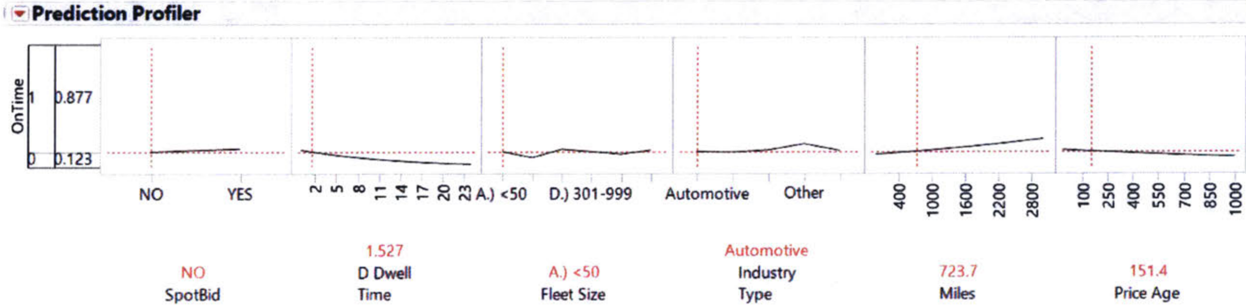
In the results shown below tender lead time is shown in days.

Source	LogWorth	PValue
Fleet Size	887.295	0.00000
D Dwell Time	317.500	0.00000
Miles	209.021	0.00000
Industry Type	123.693	0.00000
ActualWeight	120.995	0.00000
Fuel (Load)	57.072	0.00000
SpotBid	43.353	0.00000
Price Age	36.702	0.00000
Rate	25.534	0.00000
O Dwell Time 2	11.923	0.00000
Tender Leadtime	7.617	0.00000
Day of Ship	2.191	0.00644

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-1.7041173	0.0177151	9253.6	<.0001*
SpotBid[NO]	-0.099272	0.0071288	193.92	<.0001*
Fuel (Load)	-0.0007487	4.6719e-5	256.82	<.0001*
ActualWeight	5.51134e-6	2.3491e-7	550.44	<.0001*
D Dwell Time	-0.0826236	0.0021665	1454.4	<.0001*
Fleet Size[A.) <50]	0.08000226	0.0097902	66.78	<.0001*
Fleet Size[B.) 51-100]	-0.4678616	0.0084914	3035.8	<.0001*
Fleet Size[C.) 101-300]	0.24216334	0.0093308	673.57	<.0001*
Fleet Size[D.) 301-999]	0.08505448	0.0093871	82.10	<.0001*
Fleet Size[E.) 1000+]	-0.1162227	0.0067791	293.92	<.0001*
Industry Type[Automotive]	-0.1233525	0.0071942	293.99	<.0001*
Industry Type[F&B / CPG]	-0.1686718	0.0081923	423.91	<.0001*
Industry Type[Manufacturing]	-0.0165409	0.0111349	2.21	0.1374
Industry Type[Other]	0.36157718	0.0181931	394.99	<.0001*
Miles	0.0003188	1.0315e-5	955.26	<.0001*
Rate	-7.725e-5	7.2865e-6	112.40	<.0001*
Tender Leadtime	-0.0107866	0.0019333	31.13	<.0001*
Price Age	-0.0006008	0.000047	163.46	<.0001*
Day of Ship[Fri]	-0.0239658	0.0077303	9.61	0.0019*
Day of Ship[Mon to Thu]	-0.0096478	0.0066734	2.09	0.1483
O Dwell Time 2	-0.0055345	0.0007788	50.50	<.0001*

Prediction Profiler



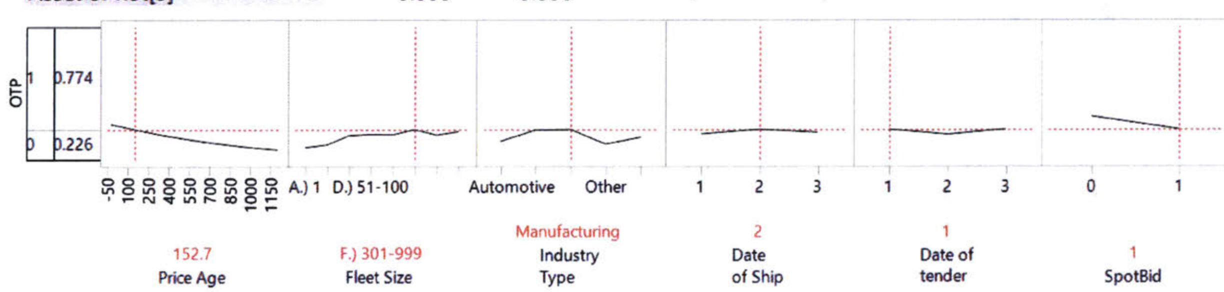
Appendix E. On Time Pick Up Regression Results.

In the results shown below Day of Ship [1] is Friday; Day of Ship[2] is Saturday and Sunday; Day of ship [3] is weekday is the baseline This is the same for the Day of Tender. Tender lead time is shown in days.

Source	LogWorth	PValue
Industry Type	555.020	0.00000
Price Age	218.953	0.00000
Fleet Size	168.013	0.00000
Fuel	156.930	0.00000
Miles	67.323	0.00000
ActualWeight	44.873	0.00000
Total Rate (Actual)	37.133	0.00000
Date of Ship	21.899	0.00000
Date of tender	2.773	0.00169
SpotBid	2.312	0.00487
Tender Leadtime	2.122	0.00755
Asset or not	.	.

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-2.046	0.092	492.36	<.0001*
Total Rate (Actual)	0.000	0.000	165.43	<.0001*
Fuel	0.002	0.000	715.66	<.0001*
Miles	-0.000	0.000	303.86	<.0001*
Tender Leadtime	0.004	0.002	7.14	0.0075*
Price Age	-0.001	0.000	1001.0	<.0001*
Fleet Size[A.) 1]	-0.697	0.066	110.74	<.0001*
Fleet Size[B.) 2-10]	-0.467	0.036	172.67	<.0001*
Fleet Size[C.) 11-50]	0.093	0.021	19.40	<.0001*
Fleet Size[D.) 51-100]	0.156	0.015	101.68	<.0001*
Fleet Size[E.) 101-300]	0.134	0.017	59.31	<.0001*
Fleet Size[F.) 301-999]	0.380	0.016	550.58	<.0001*
Fleet Size[G.) 1000+]	0.123	0.014	75.00	<.0001*
Industry Type[Automotive]	-0.247	0.012	402.93	<.0001*
Industry Type[F&B / CPG]	0.336	0.011	950.46	<.0001*
Industry Type[Manufacturing]	0.365	0.016	507.50	<.0001*
Industry Type[Other]	-0.442	0.030	217.97	<.0001*
Date of Ship[1]	-0.095	0.010	95.94	<.0001*
Date of Ship[2]	0.116	0.013	81.89	<.0001*
Date of tender[1]	0.075	0.024	9.95	0.0016*
Date of tender[2]	-0.157	0.045	12.16	0.0005*
ActualWeight	0.000	0.000	200.88	<.0001*
SpotBid[0]	0.244	0.087	7.93	0.0049*
Asset or not[0]	0.000	0.000	.	.



Appendix F. Perfect Shipment Regression Results.

In the results shown below Day of Ship [1] is Friday; Day of Ship[2] is Saturday and Sunday; Day of ship [3] is weekday is the baseline This is the same for the Day of Tender. Tender lead time is shown in days.

Effect Summary

Source	LogWorth	PValue
Fleet Size	731.275	0.00000
Industry Type	599.059	0.00000
Price Age	453.955	0.00000
SpotBid	217.538	0.00000
Miles	182.990	0.00000
Total Rate (Actual)	143.709	0.00000
O Dwell Time	35.540	0.00000
Date of tender	32.346	0.00000
Fuel	24.556	0.00000
Tender Leadtime	3.451	0.00035
Date of Ship	3.240	0.00058
ActualWeight	1.839	0.01450
D Dwell Time	0.061	0.86915

Parameter Estimates

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Intercept	-0.2131	0.0267	63.77	<.0001*
Total Rate (Actual)	0.0003	0.0000	654.9	<.0001*
Fuel	0.0006	0.0001	107.9	<.0001*
Miles	-0.0006	0.0000	835.5	<.0001*
Tender Leadtime	0.0045	0.0013	12.76	0.0004*
Price Age	-0.0015	0.0000	2082	<.0001*
Fleet Size[A.) 1]	-0.1202	0.0323	13.87	0.0002*
Fleet Size[B.) 2-10]	-0.1413	0.0208	46.31	<.0001*
Fleet Size[C.) 11-50]	0.0720	0.0140	26.61	<.0001*
Fleet Size[D.) 51-100]	-0.4493	0.0108	1727	<.0001*
Fleet Size[E.) 101-300]	0.3917	0.0121	1042	<.0001*
Fleet Size[F.) 301-999]	0.1795	0.0117	234.6	<.0001*
Fleet Size[G.) 1000+]	0.0677	0.0089	57.56	<.0001*
Industry Type[Automotive]	-0.2068	0.0090	526.9	<.0001*
Industry Type[F&B / CPG]	-0.0206	0.0086	5.76	0.0164*
Industry Type[Manufacturing]	0.4808	0.0109	1929	<.0001*
Industry Type[Other]	-0.3698	0.0216	292.2	<.0001*
Date of Ship[1]	-0.0003	0.0079	0.00	0.9675
Date of Ship[2]	0.0260	0.0114	5.20	0.0226*
Date of tender[1]	0.2259	0.0209	117.2	<.0001*
Date of tender[2]	-0.4749	0.0401	140.3	<.0001*
SpotBid[0]	-0.2517	0.0080	994.4	<.0001*
ActualWeight	0.0000	0.0000	5.98	0.0145*
D Dwell Time	-0.0002	0.0015	0.03	0.8692
O Dwell Time	-0.0126	0.0010	158.1	<.0001*

Term	Estimate	Std Error	ChiSquare	Prob>ChiSq
Tender Leadtime	0.07664278	0.0081465	88.51	<.0001*

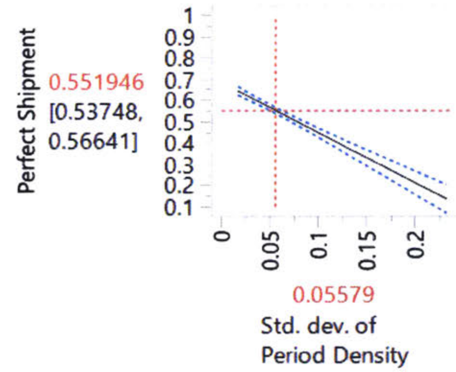
Appendix G. Period Density Correlations for Major Players

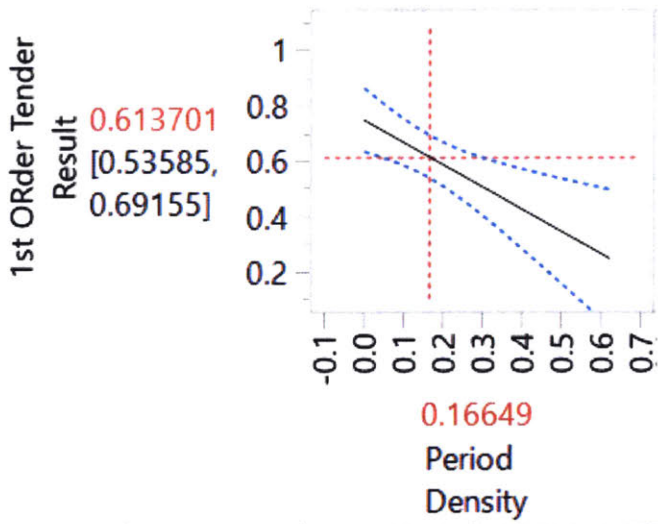
Correlations

	Std. dev. of Period Density	Period Density
OTD	-0.1245	0.0218
OTP	-0.2671	-0.0760
Perfect Shipment	-0.6739	-0.0868
1st ORder Tender Result	-0.2946	-0.1645

Correlation Probability

	Std. dev. of Period Density	Period Density
OTD	0.0166	0.6762
OTP	<.0001	0.1446
Perfect Shipment	<.0001	0.0956
1st ORder Tender Result	<.0001	0.0015

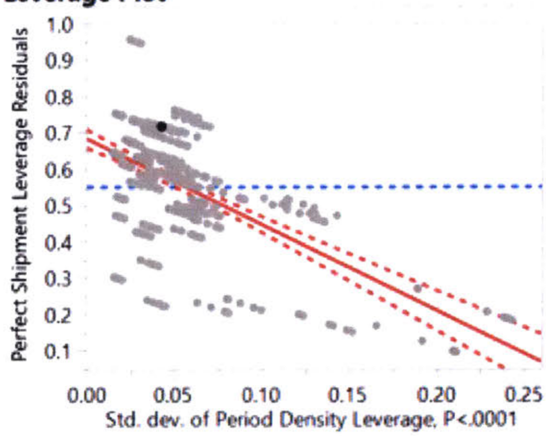




Effect Summary

Source	LogWorth	PValue
Std. dev. of Period Density	30.151	0.00000
Period Density	0.266	0.54220

Leverage Plot



Appendix H. Shipper Clustering by Lane and Carrier Density

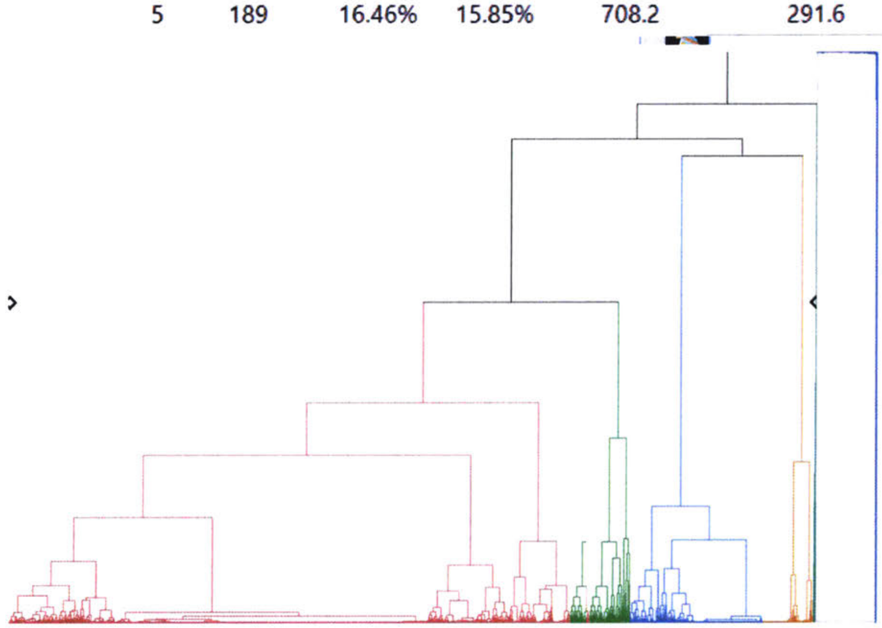
Cluster Summary

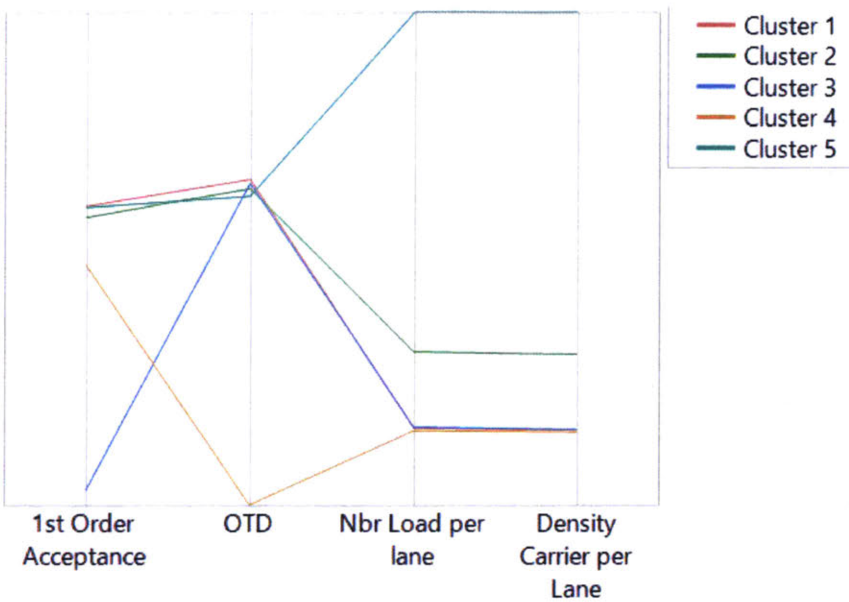
Cluster Means

Cluster	Count	1st Order Acceptance	OTD	Nbr Load per lane	Density Carrier per Lane
1	31999	92.02%	92.54%	8.6	4.1
2	2281	88.76%	90.06%	174.5	73.1
3	7551	10.82%	91.49%	10.5	4.5
4	2937	75.23%	4.19%	3.0	2.2
5	189	91.64%	87.98%	915.7	386.0

Cluster Standard Deviations

Cluster	Count	1st Order Acceptance	OTD	Nbr Load per lane	Density Carrier per Lane
1	31999	15.63%	13.98%	14.0	6.1
2	2281	14.41%	11.10%	142.6	50.1
3	7551	15.07%	15.00%	23.7	8.2
4	2937	40.91%	10.06%	6.1	4.1
5	189	16.46%	15.85%	708.2	291.6





Cluster Means

Column Summary

Column	RSquare
1st Order Acceptance	0.7309
OTD	0.7128
Nbr Load per lane	0.5833
Density Carrier per Lane	0.6138

Portion of total variation in each column absorbed by clustering

Appendix I. Shipper Clustering by Percentage of Loads on Different Lanes

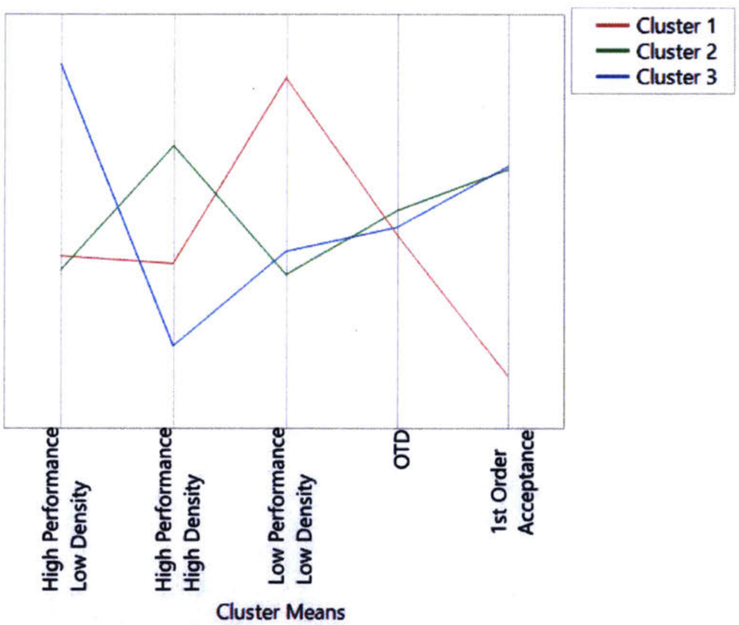
Cluster Summary

Cluster Means

Cluster	Count	High Performance Low Density	High Performance High Density	Low Performance Low Density	OTD	1st Order Acceptance
1	10	30.8%	38.9%	30.0%	85.2%	57.1%
2	21	27.5%	67.8%	3.9%	87.1%	88.9%
3	9	73.8%	18.6%	7.0%	85.8%	89.3%

Cluster Standard Deviations

Cluster	Count	High Performance Low Density	High Performance High Density	Low Performance Low Density	OTD	1st Order Acceptance
1	10	15.7%	17.9%	15.2%	5.6%	8.9%
2	21	13.0%	14.8%	3.3%	8.3%	7.2%
3	9	12.3%	12.7%	5.3%	10.3%	8.0%



Column Summary

Column	RSquare
High Performance Low Density	0.6785
High Performance High Density	0.6626
Low Performance Low Density	0.6553
OTD	0.0109
1st Order Acceptance	0.7717

Portion of total variation in each column absorbed by clustering