Utilization of the American Truck Driver

by

Mei Qing Zhang

Bachelor of Science in Accounting, San Francisco State University, 2008

and

Adam Buttgenbach

Bachelor of Science in Ag Business, California Polytechnic State University, 2008

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

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

May 2020

Signature of Author: ____

Mei Qing Zhang
Department of Supply Chain Management
May 8, 2020

Signature of Author: ____

Adam Buttgenbach
Department of Supply Chain Management
May 8, 2020

Certified by: ____

Dr. David Correll
Research Scientist, Center for Transportation and Logistics
Capstone Advisor

Accepted by: ____

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

Utilization of the American Truck Driver

by

Mei Qing Zhang

and

Adam Buttgenbach

Submitted to the Program in Supply Chain Management
on May 8, 2020 in Partial Fulfillment of the
Requirements for the Degree of Master of Applied Science in Supply Chain Management

ABSTRACT

The electronic logging device mandate was implemented with the intention of keeping truck drivers in compliance with the hours of service regulations to reduce driver fatigue and trucking accidents. Two years after the electronic logging device mandate became law, there have not been many studies that use trucking operational data such as the newly available electronic logs to look for efficiency gain. Our team received six months newly available raw logging data. This paper aims to use different analysis techniques in machine learning on the raw electronic logging data to find areas of opportunity that can be used by management to control and improve driver utilization. The three significant factors that we investigated for on the amount of time a driver spends at each freight location are: the time of day the driver arrives at a shipper location, the impact from a specific location, and the frequency that the carrier visits a specific shipper. Each of the three factors were found to imply a statistically significant impact on the stop duration. This study shows the usefulness of using electronic logging data to identify the underlying factors on stop time so that managers can schedule truck drivers more efficiently. This will allow for higher driving hours during the day, which translates to higher income for the drivers. Since the raw electronic logging device data is readily available for all On The Road carriers, we hope to inspire further data analysis on electronic logging device data to help improve the lives of truck drivers.

Capstone Advisor: Dr. David Correll
Title: Research Scientist, Department of Supply Chain Management

ACKNOWLEDGMENTS

I would like to thank the Supply Chain Management team and my advisor, Dr. David Correll. I am forever gratefully for the experiences and friends I made during my time at MIT. I would also like to thank my family for their support and making this possible. In particular: Laura, Judy, Alex, Ken, Beth, and Beverly.

Adam

I would like to thank our capstone advisor, Dr. David Correll for all the guidance and support. I would also thank my family for their support.

Mei Qing Zhang

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ANOVA – Analysis of Variance

AOBRDs – Automatic On Board Recording Devices

ATA – American Trucking Association

ATRI – American Transportation Research Institute

BEG – Freight event where a loaded trailer is dropped at a transfer yard and picked up by a different driver to deliver (drop and hook)

ELD – Electronic Logging Devices

FMCSA – Federal Motor Carrier Safety Administration

Freight Point – Shipping and Receiving locations

HOS – Hours of Service

LLL – Freight event where the trailer is loaded while the driver waits

LPL – Freight event where trailer is already loaded (drop and hook)

OLS – Ordinary Least Squares

OTR – Over the Road

RODS – Record of Duty Status

US – United States

1. INTRODUCTION

1.1 *The American Trucking Industry*

The American trucking industry moves more than 70% of the freight in the US and employs approximately 3.5 million drivers (ATA, 2019), and out of the 3.5 million drivers there are 2 million heavy and tractor-trailer drivers (US Department of Labor 2019). According to a US Census Bureau 2019 article (Cheeseman Day, J and Hait, A 2019), this represents an all-time high in absolute terms; however, a study by the American Transportation Research Institute (ATRI) highlights an all-time high in the *shortage* of drivers in the industry as well. By one estimate the number was as large as 61,000 drivers in 2018 (ATA, 2019). The shortage was the most important concern among carriers for the past 3 years according to ATRI. The next most important concern for the industry was drivers' Hours of Service (HOS). Truck drivers are constrained by the 2015 Federal Motor Carrier Safety Administration (FMCSA) HOS rule which limits the number of hours a driver can work during the day and the total number of hours a driver can work during the week. The ATRI study shows that these issues have been growing annually as the largest concerns facing the industry, and with no apparent solution the shortage is expected to grow to over 105,000 drivers by 2023 (see *Figure 1)*.

For companies with multiple drivers, the driver shortage implies fewer available drivers than needed and fewer total hours of drive time available for their organization. With fewer total driving hours available, it can become difficult for carriers to meet the market capacity demand. Internal research by the MIT Center for Transportation and Logistics shows that at least some of the existing drivers are severely under-utilized when it comes to their available drive hours. Burks and Monaco (2019) concluded that drivers with higher earnings and hours driven are less likely to quit driving. The ATRI study also cites driver detention as a top concern for the industry and a further contributor to drivers not being able to fully take advantage of their available driving hours. This study will specifically look at truck driver HOS logs and ELD stop data from a large OTR carrier with approximately 1,500 tractor-trailers in order to identify where truck drivers' valuable time is lost and propose ideas for the first steps in regaining it.

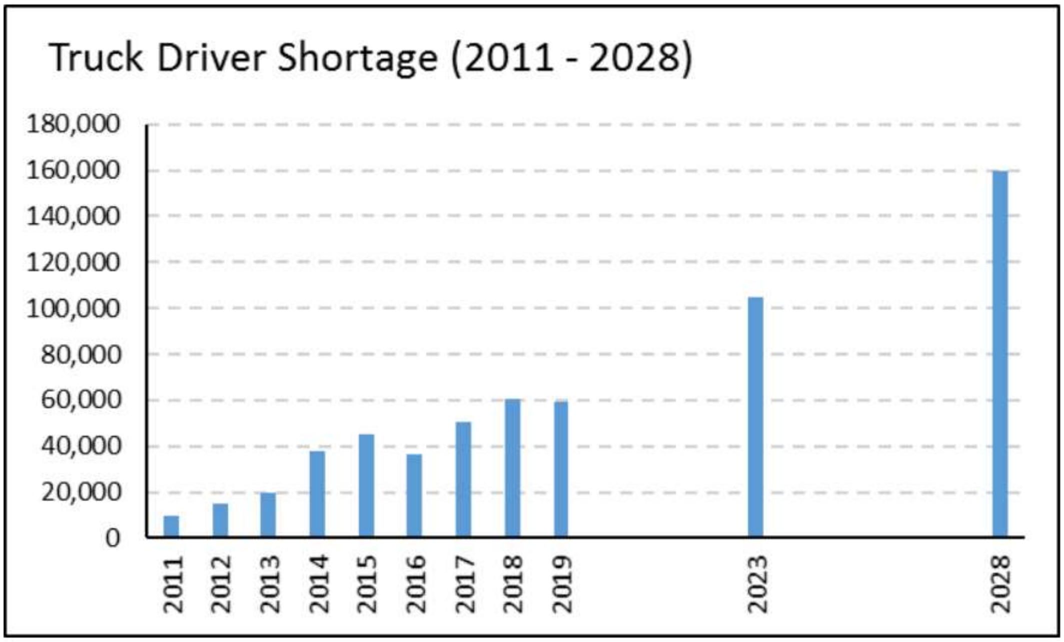

Figure 1: Truck Driver Shortage Analysis 2019, Adapted from ATA, July 2019.

1.2 *Driver Hours of Service (HOS)*

The FMCSA hours-of-service provide regulated guidelines on how long drivers can work and how long they need to be off before they can return to duty. OTR drivers can follow the 60/7 or 70/8 provision which allows them to work in an on-duty or driving status for up to 60 hours in 7 days or 70 hours in an 8-day period. Furthermore, drivers are restricted to working no more than 14 consecutive hours in an on-duty status with no more than 11 of those hours being driving (see *Figure 2)*. Drivers must also take a 30-minute rest break after 8 hours of being on duty. These provisions provide constraints on when drivers can work and for how long before needing 10 consecutive hours of off duty or sleeper berth status to reset their 14-hour clock. Drivers also need to be off duty for 34 consecutive hours to reset their total available hours for the week, i.e., 60/7 or 70/8. The Hours-of-Service rules are created to make sure drivers get the proper amount of rest and lower the chance of a crash due to driver fatigue. The FMCSA 78292 Final Rule (2015) estimates that the Hours-of-Service rules save 26 lives and $852 million annually due to crashes avoided.

In December 2017, the FMCSA implemented the phase-in of Electronic Logging Devices (ELDs) for commercial trucking hours-of-service record-of-duty status (RODS). The ELDs log and document date, time, vehicle miles, driver and vehicle identification, engine parameters, vehicle motion status, the motor carrier, and driver record of duty status. The final phase of the ELD mandate occurred at the end of 2019 with the conversion of the previously accepted Automatic On Board Recording Devices (AOBRDs) to the new ELD compliant devices. Using this newly available data, the team analyzes driver stops, shippers, and frequency of service to build a model, using machine learning techniques, to discover potential efficiency gains for drivers.

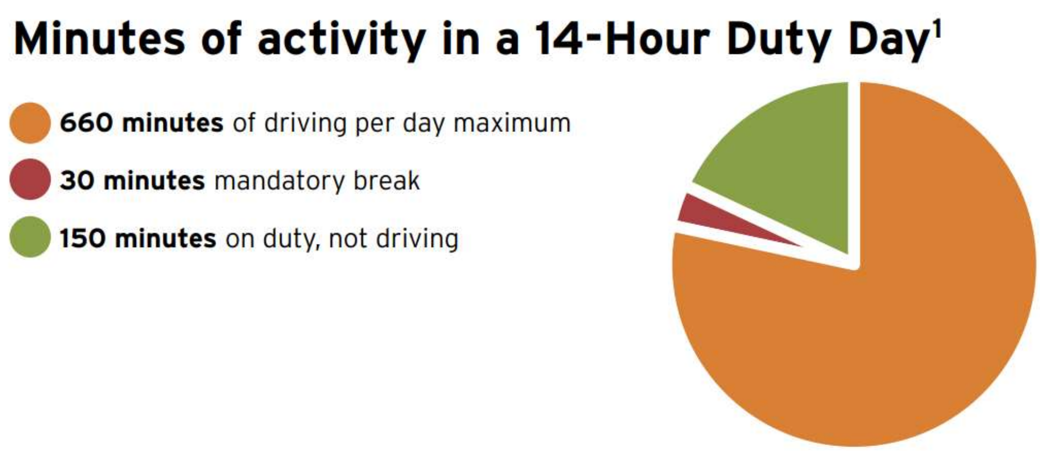

Figure 2: 660 Minutes: How Improving Driver Efficiency Increases Capacity (Adapted from JB Hunt, 2015).

1.3 *Operational Inefficiencies*

The growth in the industry has magnified the impact of the driver shortage and caused it to become the most important concern for carriers today. Most drivers are incentive-paid, based on miles driven, so they rely on available hours to drive those miles. Disruptions to those available driving hours limits driver pay and productivity for the industry. Dr. David Correll, a research scientist at MIT's Center for Transportation and Logistics, analyzed 2 years of company-released ELD data and found that some of the drivers are not driving the number of expected miles to the number of available hours (Correll 2019). This project aims to find ways to help increase the number of miles driven per trucker which could increase in the income of an average driver, and alleviate the driver shortage and reduce driver turnovers. While the driver shortage remains the primary concern for carriers, the driver turnover was approximately 94% in 2017 (Burks and Monaco 2019) and compounds the impact of the shortage. This means that the number of drivers that enter a company within one year is approximately equal to the number of drivers that leave the company every year. Driver demographics also factor into the concern for the industry with the average age of truckers being 46.4 (U.S. BLS, 2020). Low wages and the extended amount of time away from home are making young workers reluctant to enter the industry. A higher utilization of drivers and higher driver income could create a more attractive career for young workers, which would in turn reduce the driver shortage. With trucking continuing to dominate the movement of goods in the US and the American supply chain, the industry will need to rely on the increase in driver utilization to keep it moving forward. The present study will focus on the efficiency of drivers and look for ways to reduce driver downtime to increase the median number of hours driven per day. Our goal is to find opportunities to increase driver efficiency and reduce the driver shortage, by first determining where they are losing drive hours during their workday, and then suggesting what can be done to improve their utilization.

2. LITERATURE REVIEW

While driver shortage is often cited as the top industry concern, conflicting research suggests the shortage may be nothing more than "fake news". According to Burks and Monaco (2019), the overall market for truck drivers seems to work just as well as other blue-collar labor markets, and with enough time, driver supply should respond to price signals in the normal way. Some of the suggested paths to ease the driver shortage include using block chain technology to authenticate drivers (Dujak and Sajter 2019), lowering the minimum interstate driving age to increase the labor pool (Herrera 2019), and compensation increases to reduce driver turnover (Stinson 2019). This capstone looks at the issue of driver capacity as an under-utilization problem and uses truck driver log data in order to identify where truck drivers' valuable time is being lost and propose first-step ideas. We believe that improving utilization could also play a role in alleviating the driver shortage.

2.1 Trucking Industry Background

Turnover of the for-hire truckload carriers has long remained at high double digit and sometimes triple digit annual turnover rates (Min and Lambert 2002). Min and Lambert's 2002 study highlighted two significant variables that influence driver turnover: compensation and poor driver management. Other authors have noted that wages have failed to keep pace with inflation over the past 25 years (Belzer, 2000). In addition to lagging wages, poor management decisions that send drivers to shippers and receivers with irregular work and long detention times can be just as detrimental to driver retention (Min & Lambert, 2002). On the list of top industry concerns reported on ATRI's 2019 annual survey, the fourth overall (and apparently for the first time) was driver detention and delays at shipping and receiving facilities. The survey (see *Figure 3*) found that delays and detention at shipping and receiving locations caused a cascading effect on drivers' compensation, ability to get to a safe stopping place, and stay within their Hours-of-Service guidelines. Between 2014 and 2018, the industry saw a driver-reported increase of 27.4% in delays of 6 hours or more at shipper facilities (ATRI, 2019). According to a JB Hunt Paper (2015), on average a truck driver lost 2 hours per day on inefficient delivery, pickup and looking for parking.

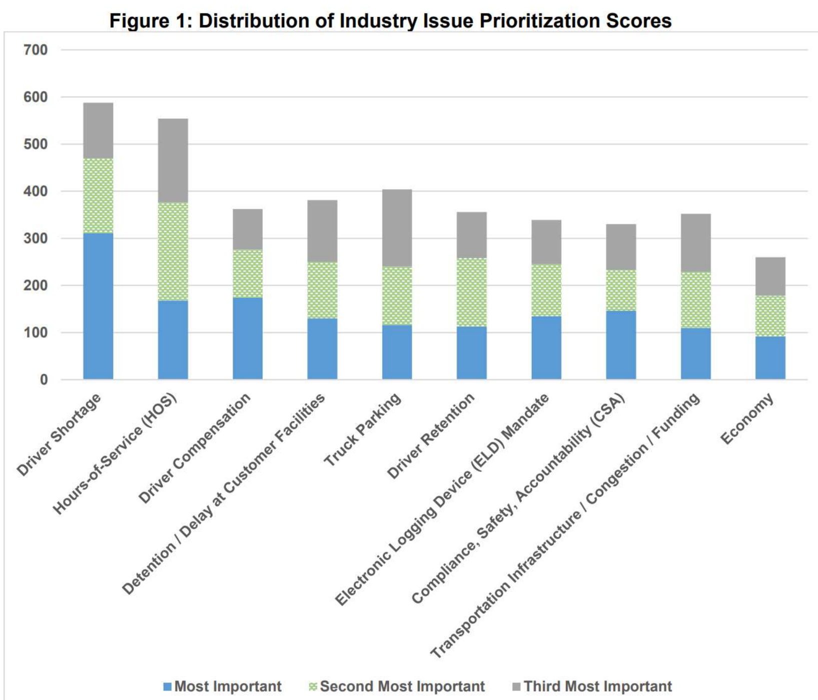

Figure 3: Distribution of Industry Issue Prioritization Scores. (Adapted from 'Critical Issues in the Trucking Industry -2019' issued by the American Transportation Research Institute.)

While the biggest concern for the trucking industry is the driver shortage (ATA 2019), the second largest concern is the drivers' Hours-of-Service inflexibility (see *Figure 3*). Burks and Monaco (2019) conclude that there is no evidence of a secular truck driver shortage, but drivers who earn more money in one period are less likely to quit in the next period. Rodriguez, Targa and Belzer (2006) explain that truck drivers are usually paid by the number of miles driven. Their study also found that as a driver's compensation increases, the chance of driver turnover decreases. Since an increase in driver wages would drive up supply chain costs, we will focus on opportunities for drivers to be more efficient as a way to increase driver total compensation and overall satisfaction.

2.2 Data Availability

In December 2017, the Federal Motor Carrier Safety Administration (FMCSA) mandate, requiring Electronic Logging Devices (ELDs) use in commercial motor vehicles, went into effect for the US trucking industry. Previously, the FMCSA allowed truck drivers to use paper logs to record their driving activities. In the event of a roadside inspection, a driver is now required to provide his or her record of duty status with an approved ELD. These new records are now in easy-to-report electronic format, making it easier to identify driver violations and eliminate falsification of duty status. Electronic logs can provide many benefits to carriers, including a reduction in clerical timekeeping errors, reduced HOS violations and fines. Cantor and Corsi (2015) found that using the electronic logging device data, firms can quickly find drivers who did not follow the Hours-of-Service rules issued by FMCSA, and potentially lower the risk of crashes due to driver fatigue. Suzuki, Crum and Pautsch (2009) used operational data from 2 truckload carriers to predict the likelihood of driver turnover. They tested the effects of operational work variables combined with operational work data and demographic data of drivers to find statistically significant variables that point to higher driver turnover. They also removed the variables that past studies showed were unlikely to affect driver turnover and other variables that were collinear. This study will follow Suzuki et al.'s approach to remove insignificant exogenous variables and focus on variables that have the greatest influence for truck driver efficiency. As a result of the ELD mandate, our study will use newly available data to identify where truck drivers' valuable time is being lost.

2.3 Literature Summary

This study focuses on the truck driver shortage in the United States as a utilization problem, not a pure labor shortage. Based on our review, previous studies have cited poor driver management and low wages as the primary factors of driver turnover (Min and Lambert 2002). Using carrier supplied operational ELD data, we will use machine learning techniques to identify where drivers are losing time. Based on the results, we will identify opportunities to improve driver working hours and miles, which in turn will lead to an increase in driver compensation and help reduce turnover. In a Medium article, Correll (2019) states that to eliminate the 3.4% shortage in the 1.8 million OTR truck driver population that drives an average of 6.5 hours per day, we would have to increase the current truck driving time per driver by 0.2 hours, or just 12 minutes. It is our goal to find ways to extract that additional 12 minutes of driving time.

3. DATA AND METHODOLOGY

3.1 Data Introduction

In this section we will evaluate the ELD HOS and truck driver stop data to better understand what is causing drivers to be in a non-driving status (on-duty, sleeper berth or off-duty) during their 14-hour window of on-duty time. We use data from a Midwest OTR national and regional trucking company with approximately 1500 trucks in the dry van trailer portion of the industry. This capstone will evaluate 6 months of ELD HOS data to review driver duty status throughout their workday. The 6 months of data will include operations from May through October of 2019. Six months is the amount of data we reviewed because it is the minimum required retention period for driver logs by the FMCSA. We will also review stop data over the same six-month period from the same company to gain additional information on the type of loads drivers are moving and the amount of time that is spent at a shipper. To validate some of the research and model assumptions, we interviewed the carrier providing the data to validate the initial data findings and assumptions, gain insights into possible hypotheses on lost time, and understand the most common issues drivers face during their day. The combination of this data and carrier insights will be used as the basis for identifying the current median driving hours per day, lost time during the day, and the first steps towards regaining that time.

3.1.1 ELD Data

The FMCSA ELD mandate has made traditional driver paper logs electronic and more detailed than before. The data for this study will include RODs and stop changes down to the minute and will be provided in multiple Excel files from the company's ELD provider, Omnitracs. The current electronic logging data include the driver code, tractor number, logging time, start time and end time, actual hours in each of the statuses, the type of activities, number of miles driven, plan and state of event (see *Figure 4 and Figure 5* for an example of the traditional paper logs and the visualization from the ELD data).

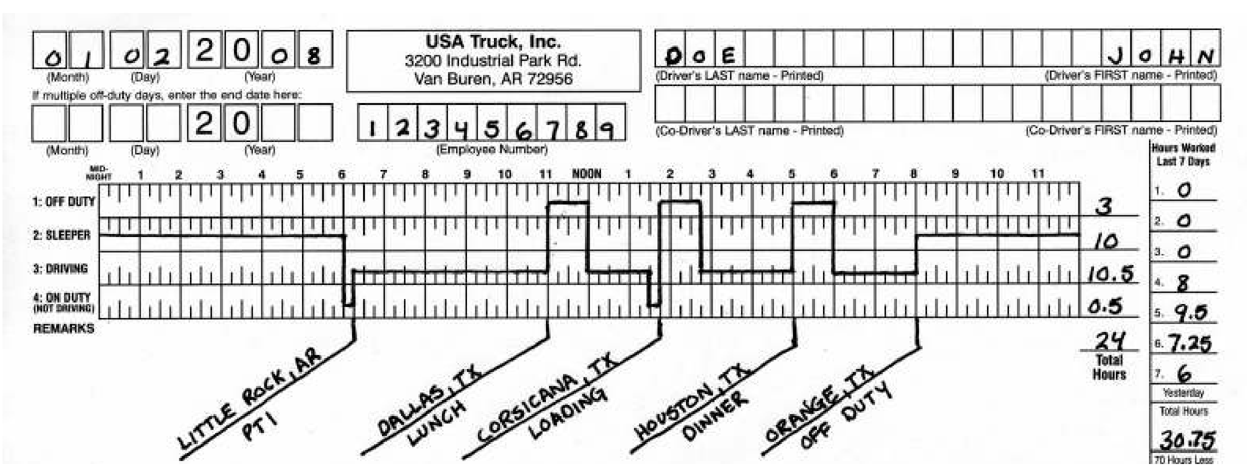

Figure 4: Traditional Paper Log for Driver's Hours of Service

	Driver ID	tractor_id	log_date	start_time	end_time	act_hours	name	miles	order_number	plan_number	plan_name	location	bdt_description	Driver ID.1
0		NaN	2019-05-01	2019-05-01 00:00:00	2019-05-01 15:18:00	15.300000	Off Duty	0.0	6734761.0	57717175.0	6734761	6 mi ENE of La Joya, BJ	OTR - Teams	NaN
1		104098	2019-05-01	2019-05-01 15:18:00	2019-05-02 00:00:00	8.700000	Off Duty	0.0	6734761.0	57717175.0	6734761	5 mi NNE of Forest Park, GA	OTR - Teams	NaN
2		104098	2019-05-02	2019-05-02 00:00:00	2019-05-03 00:00:00	24.000000	Off Duty	0.0	NaN	NaN	NaN	5 mi NNE of Forest Park, GA	OTR - Teams	NaN
3		104098	2019-05-03	2019-05-03 00:00:00	2019-05-03 06:32:00	6.533333	Off Duty	0.0	9425781.0	57715222.0	9425781-01	5 mi NNE of Forest Park, GA	OTR - Teams	NaN
4		104098	2019-05-03	2019-05-03 06:32:00	2019-05-03 09:45:00	3.216666	Off Duty	0.0	9425781.0	57715222.0	9425781-01	5 mi NNE of Forest Park, GA	OTR - Teams	NaN

Figure 5: ELD Data for Driver's Hours of Service

Figure 5 shows the data variables provided in the ELD Data and *Table 1* provides a description of those variables. The ELD HOS data frame includes a total of 14 variables and 1,614,399 entries.

Table 1: Driver ELD HOS variables

driver_code	A unique identifier for the driver.
tractor_no	A unique identifier for the tractor
log_date	The date of the entry in this row.
start_time	The time that that service status started
end_time	The time that that service status ended
act_hours	The difference between start time and end time
name	The HOS status logged by the driver
miles	Miles traveled under that status
Plan	A unique identifier for the load the driver is carrying
STATE	State level location data (actual data is a little more specific – but not lat long coordinates)

3.1.2 Stop Data

The truck driver stop data is data received from driver inputs based on their loads, arrival time, departure time, type of pickup, appointment windows, and *Freight Point* information. The study uses this data to evaluate detention time based on the time of day, the shipping location, and the frequency of service by location (see *Table 2* for Stop Data Variables and descriptions. *Table 3* lists the Freight Event codes with a brief description and company estimated stop time.)

Table 2: Driver Stop Data Variables

Plan	A unique identifier for the load the driver is carrying
Driver_	A unique identifer for the driver.
Trip_division	An internal classification of trip type
CONTROLLING_CUSTOMER	A unique identifer for the customer
FREIGHTPOINT	A unique identifer for the facility visiited at this appointment (i.e. warehouse, DC, or other facility)
TIMEZONE	Timezone in which entry was recorded
FREIGHT_EVENT	The type of freight appointment (see below)
APPT_WINDOW	The opening and closing date of the appointment. (Note, this could include re-schedules)
ARRIVAL	When the driver arrived at the facility (Gate In)
DEPARTURE	When the driver left the facility (Gate Out)
Division	Another internal classification of the trip.

Table 3: Driver Freight Event codes – Stop Data

Freight Event	Description	Estimated hours to complete
BEG	Pickup relay	0.5
CON	Continuous move, usually to pick up paperwork	0.5
END	Dropoff relay	0.5
LDA	Driver assists customer loading trailer	4.0
LLL	Customer loads trailer while driver is present	2.0
LLM	Third-party (lumper) loads trailer while driver is present	4.0
LPL	Trailer loaded when driver arrives	0.5
UDA	Driver assists in unloading trailer	4.0
UDT	Dropoff loaded trailer	0.5
UDU	Driver unloads trailer alone	4.0
ULL	Third-party (lumper) unloads trailer while driver waits	4.0
ULU	Customer unloads trailer while driver waits	2.0

3.1.3 Interview Data

The interview data comes from a discussion with the company's Continuous Improvement Manager that took place in February 2020. The contributions from this interview included hypothesis validation, soundness of initial results, expectations, data entry clarification, and a deeper understanding of the daily operations. The manager explained they currently use a Transportation Management System to track the total number of billable and empty miles. The company currently does not use the electronic logging device data outside of the regulatory compliance purpose. The company's current method of load planning relies on past delivery experiences passed down from dispatcher to dispatcher. One particular comment from the manager stuck out to our team was the company has been losing efficiency due to shorter distances between each delivery, so the drivers encountered the shipping location more frequently. With the rise of ecommerce sales, our team thinks trend of delivery distance getting shorter will only increase. This fact validated our paper's focus on stop time duration for each trucking stop. He also recommended using a minimum cutoff of 15 minutes for the length of a stop and a maximum cutoff of 8 hours because anything over 8 hours is most likely indicating that a driver is using the time as a DOT break. This cutoff information is important to our paper, because leaving incorrect data can skew our results, so we incorporation the cutoff in our data analysis. Lastly, the manager confirmed the three trucking stop hypotheses that our team decided to keep for our research are not actively tracked and managed with formal analytics, so our team felt our research goals has potential to provide new managerial recommendations to help reduce stop time duration.

3.1.4 Data Cleaning

The data for this study was received in 12 Excel data files for driver ELD HOS data and a single Excel data file for the driver Stop Data. Due to the size of the data, we use Python and Orange to concatenate the data into a single data frame so we can test our hypotheses and build our model (see *Figure 6*). Orange is a Python based program that uses widgets for data analysis and visualization. These programs will be used because of their capabilities and built in libraries that will assist with our descriptive statistics, ANOVA testing, and machine learning.

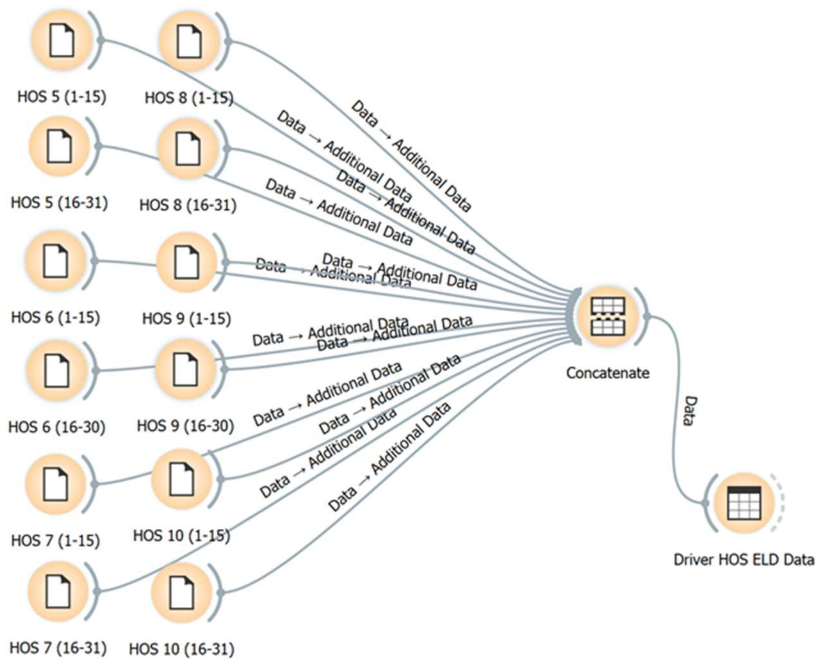

Figure 6: Visualization of concatenated HOS ELD Data process using Python and Orange

Both data sets fall into the *"Big Data"* category with over 1.6 million observations in our ELD Data set and approximately 60,000 observations in our stop data set. Due to the size of these data sets, we will be susceptible to outliers and data entry errors so we will need to clean the data to address these issues. To clean the data, we began by taking the following steps.

ELD HOS Data

- Remove missing entries and entries with null values
- Remove days with more than 14 hours of drive time because these would exceed the maximum allowable drive hours by law
- Remove days with no hours of drive time because we are looking for driver utilization opportunities on days where drivers are driving

Stop Data

- Remove missing entries and entries with null values
- Remove all Freight events other than (LLL, LPL, BEG), these events make up 99%+ of the data

- Remove stop times under 15 minutes – it is not feasible for drivers to be able to arrive at a shipper, unload, reload, and depart in less than 15 minutes. In addition, using traditional paper logs, 15 minutes was the minimum amount of time that could be used for a duty status
- Remove stop times over 8 hours – based on our interview with the company, they recommend removing stop times over 8 hours because they were most likely drivers taking a break

By removing outliers, missing data, and unrealistic values, we got a clearer picture of reality and were able to begin our analysis.

3.1.5 Data Distributions

To better understand the data we are dealing with we will begin with descriptive statistics, scatterplots, and data distributions for each of our key variables. Understanding the distributions of our data will allow us to make initial assumptions on the tests we will conduct and our final model. The initial data on the truck stop time shows aggregation with a very long tail of stop time. Based on the visual plots (*see Table 4, Figure 7, Figure 8, Figure 9, Figure 10)*, we will evaluate opportunities to categorize shippers based on their average stop time.

Table 4: ELD HOS Drive Hours per Day

	drivehr_perday
count	104380.000000
mean	7.049425
std	2.804361
min	0.016666
25%	5.250000
50%	7.583333
75%	9.249998
max	13.983331

Table 4 shows that truck drivers drove an average of 7.04 hours per day and a median of 7.58 hours per day over the 6 months of data we analyzed. The FMCSA rule states a driver can drive up to 11 hours in a 14 hour on duty period before needing a 10-hour restart. This indicates the drivers in the study are not fully utilizing all of their available hours. Our team decided further to analyze the detention time at delivery location as an opportunity to recover lost time.

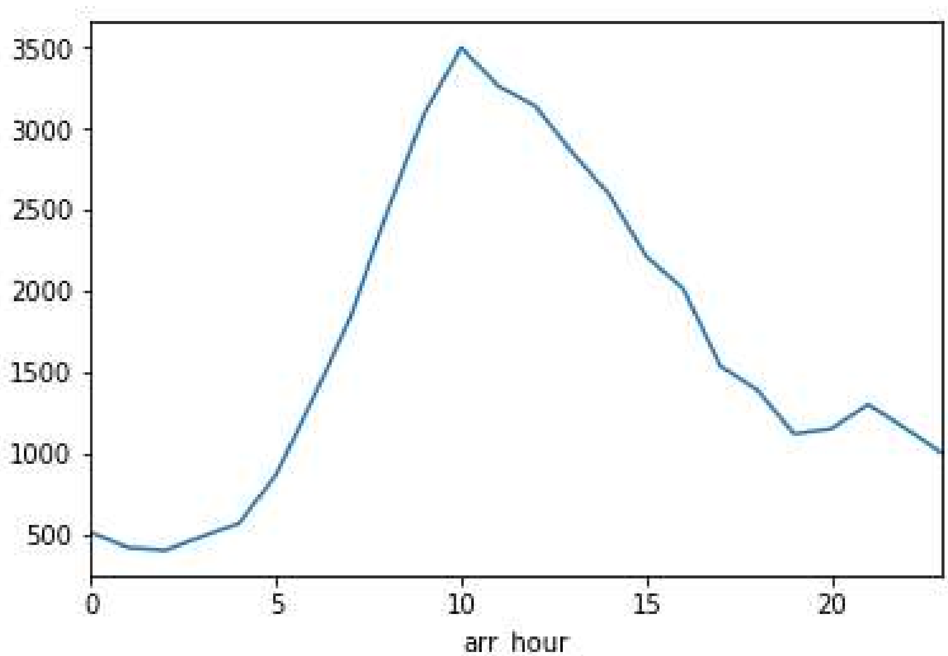

Figure 7: Plot of Truck Arrivals at Freight Point by Hour

Figure 7 is the count of the arrival hour of the date over the 6 months of arrival data. This figure shows the peak arrival count happens at 10 am. Our team was originally expecting the peak arrival time to be 8 am.

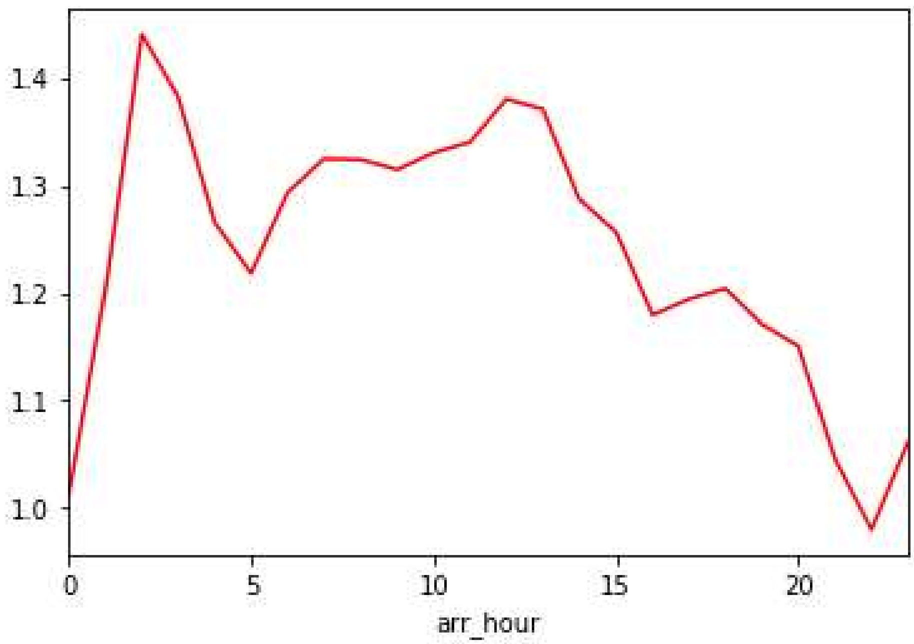

Figure 8: Plot of Average Stop Time by Arrival Hour

Figure 8 shows the peak stop time was the highest between the 1 am to 5 am and remains at lower level between 5 am to 10 am. This data indicates a potential strategy of changing the arrival time for the truck drivers.

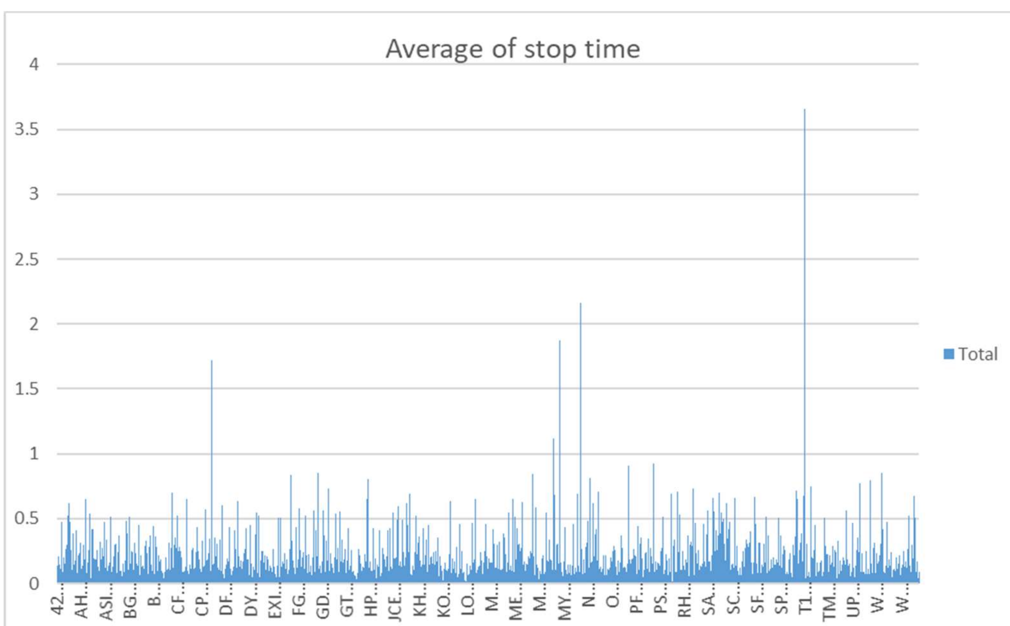

Figure 9: Plot of Average Driver Stop Time by Freight Point ID

Figure 9 shows there is a big discrepancy on the average detention time for drivers between each freight point. Our team took this data into account into developing our hypothesis of freight point detention time differences.

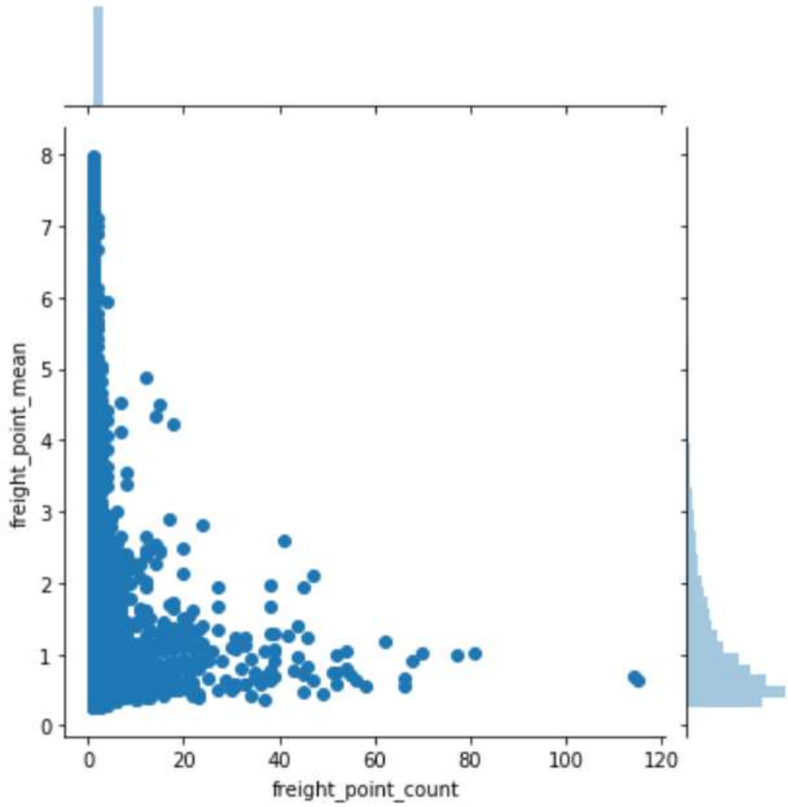

Figure 10: Plot of Average Driver Stop Time by Freight Point Frequency

Figure 10 is a joint plot of the freight point average time spent in comparison to freight point counts. This figure shows that the freight point mean detention time and freight point frequency are negatively correlated.

3.2 Hypothesis Introduction

We will be testing 3 hypotheses as we build our model for evaluating and solving driver inefficiencies. Our theories include testing for a difference in the detention time at a shipper based on the time of day, the difference in detention time between shipping locations, and the difference in detention time based on the frequency of service at shipping locations. Using an *Analysis of Variance* (ANOVA) test, we will test to see if there is a statistically significant difference in these variables. Based on the results, these variables will be used to model where time is being lost and initial changes to operations that could help reduce detention time and recapture some of the drivers' available drive time.

3.2.1 Shipping Pickup Times

This study will analyze driver arrival time at the shipper to determine if the time of day has a statistically significant impact on driver efficiency. The initial thought is that if a driver arrives during peak morning or afternoon times, they will be more likely to have to wait and incur more detention time. Before we add the time of day as a factor in our model, we will perform an ANOVA test to determine if there is a difference in the average detention time based on the driver time of arrival. For the ANOVA test, we will be looking for a high F-statistic and a p-value of <.05 to confirm that there is a statistically significant difference between the average wait time based on the time of day. Our hypothesis can be represented as,

$$H_o:\ \mu_1 = \mu_2 = \ \dots = \ \mu_n\ , \quad \mu_i = mean\ stop\ time\ at\ arrival\ hour\ i$$

$$H_a:\ not\ all\ means\ are\ equal$$

The null hypothesis for shipper pickup times is that average driver detention, depending on the hour that they arrive is the same for any hour throughout the day. The alternative hypothesis is that there is a difference in the average detention time depending on the driver's initial arrival time.

After testing for analysis of variance we performed a regression analysis using ordinary least squares (OLS) to test how much of the variability in our stop time data could be explained by the arrival time of a driver and the freight event type. To determine which variables to include in our model we used forward and backward stepwise regression to look for groups of time windows that showed to be statistically significant in explaining the variation of our stop time. To model the impact of the type of freight event on stop time, we grouped our events into two categories. The first freight events, coded (*LLL*), are loads where the customer loads the trailer once the driver is on site and present. The expected stop time for this type of load is 2 hours, based on information we received from the carrier. The second test will be for freight events, coded (*LPL* and *BEG*), both of these are drop and hook style loads and have an estimated stop time by the company of 30 minutes. These two groups make up more than 99% of the stop time data.

3.2.2 Freight Point Locations

The second hypothesis will test the difference in stop times between shipping and receiving locations. The assumption is that different shipping and receiving locations have different wait times based on how efficient they are, the amount of freight volume they process, etc. We hypothesize that not all shipping and receiving locations have the same amount of wait time. By confirming that there is a statistically significant difference in stop time between shipping and receiving locations, we will be able to add *Freight Point* as a variable in our model for finding opportunities to improve driver efficiency. We will be performing an ANOVA test on this variable to determine if there is a statistically significant difference in stop time between *Freight Points*. For the ANOVA test, we will be looking for a high F-statistic and a p-value of $<.05$ to confirm that there is a difference between shipping and receiving locations. Our Hypothesis can be represented as,

$$H_o\colon\ \mu_1 = \mu_2 = \ \dots = \ \mu_n\,, \quad \mu_i = mean\ stop\ time\ at\ Freight\ Point\ i$$

$$H_a\colon\ not\ all\ means\ are\ equal$$

The null hypothesis for *Freight Point* is that the average driver detention time is the same for all shipping and receiving locations. The alternative hypothesis is that there is a difference in the average detention time depending on the shipping and receiving location. We will perform this test for average stop times based on the *Freight Point* and evaluate the distribution of stop times by *Freight Points* to identify patterns and opportunities to categories the locations based on average wait times.

After testing for analysis of variance we will grade the *Freight Point* locations based on their average detention time. We will then perform a regression analysis using OLS to test how much of the variability in our stop time data could be explained by the different *Freight Points* and the freight event type. Our goal is to understand the amount of detention variation that can be explained solely by a *Freight Point's* average detention time. To determine which variables to include in our model we graded and grouped Freight Points into 4 groups based on their wait time (see *Table 5)*.

Table 5: Freight Point Grades on Average Stop Time

Stop Time (hrs)	Freight Point Grade	Count	Percentage
.25 - 1	A	692	29%
1 - 1.5	B	558	24%
1.5 - 2.5	C	703	30%
2.5 - 8	D	418	18%

3.2.3 Location Frequency

The third hypothesis looks at stop times based on driver frequency by *Freight Point*. The hypothesis is that drivers going back to the locations should have a good relationship with the shipper and receiver and be familiar with that location's processes to minimize delays. For example, a driver that goes back to a location multiple times a week will know where to check in, pick up his/her paperwork, which dock doors to back up to, and the procedures for picking up freight. In addition, we will be looking at the frequency for the company as whole and not individual drivers. The assumption is that companies will becoming familiar with frequently serviced locations, be more familiar with their processes, and share that information with their drivers. The current plan is to use an ANOVA test to get the F-statistic significance level on variables of count of frequency of *Freight Point* and the length of the average stop time per driver. For the ANOVA test, we will be looking for a high F-statistic and a p-value of <.05 to confirm that there is a difference in driver stop time based on how frequent the company services a location. Our Hypothesis can be represented as,

$$H_o:\ \mu_1 = \mu_2 = \ \dots = \ \mu_n\ , \quad \mu_i = mean\ stop\ time\ on\ frequency\ of\ Freight\ Point\ i$$

$$H_a:\ not\ all\ means\ are\ equal$$

The null hypothesis for *Location Frequency* is that the average driver detention time is the same for all shipping and receiving locations regardless of how often they are serviced. The alternative hypothesis is that there is a difference in the average detention time depending on the shipping and receiving location, and how often they are serviced by the carrier. We will perform this test for average stop times based on the *Freight Point* and the frequency that they are serviced by the carrier.

After testing for analysis of variance we performed a regression analysis using OLS to test how much of the variability in our stop time data could be explained by the different *Freight Points* and the frequency. To determine which variables to include in our model we graded and grouped Freight Points into 4 groups based on their frequency (see *Table 6)*.

Table 6: Freight Point Grades on Carrier Frequency

Frequency	Freight Point Grade	Count	Percentage
1-2	A	1374	55%
3-10	B	593	24%
11-100	C	406	16%
101-2359	D	136	5%

3.3 Modeling and Data Analysis

This study will model the median driver driving hours per day and determine how they are impacted by driver detention time at shipping and receiving locations based on the time of day, unique *Freight Points*, and *Freight Point* frequency. By confirming the difference in means among these variables, the next steps will include understanding their basic parameters, the ANOVA technique, and the first steps toward building a model to explains the variance in the median driver driving hours per day. To analyze driver utilization in the ELD HOS data, we will use supervised machine learning to determine whether a driver will be more efficient based on shipper arrival times, specific shippers, and shipper frequency. We will evaluate models for both live load stops and drop & hook type loads. The goal of the model will be to understand how the variables impact the median driver driving hours per day and how daily operational decisions can impact these times.

3.3.1 Descriptive Statistics

Once we have cleaned the data we began by looking at the distributions of each variable and their parameters. Some of the descriptive statistics we used are the mean, mode, median, standard deviation and percentiles to determine need to apply statistical methods. We decided to focus our statistical analysis on variables with high dispersion.

3.3.1.1 Analysis of Variance (ANOVA)

ANOVA is statistical test of whether two or more population mean are equal. The null hypothesis in ANOVA says the mean of all the population are the same. The ANOVA explains the variance between the groups and the variance within the groups. The formula for the F-statistic in ANOVA is:

$$F = \frac{Variance\ between\ the\ groups}{Variance\ within\ the\ groups}$$

The larger the F-statistic, the more likely the groups have different mean. High F-statistic value means the null hypothesis can be rejected.

3.3.2 Machine Learning Applications

Due to over 1 million data records received from the sponsor company, this project is considered to be using "big data". According to Grable and Lyons (2019), the term big data is used when the data set is so large and complex that older data processing methods cannot make sense of the data. Our data file crashed Microsoft Excel repeatedly. Grable et al. also said that the strength of the big data method lies in creating algorithms that find patterns that have predictive power. In Grable, et al.'s paper, International Finance Corporation used big data mapping to find the distributions of Digital Finance Services within each country and identify the largest potential users of services. Since our team has time-of-arrival and locations for trucks making each stop, we plan to use big data mapping techniques to find the drivers most likely to have to stop for a longer time than normal. With this information, dispatchers can devise a strategy to adjust the next delivery appointment ahead of time.

3.3.2.1 Multiple Linear Regression

Linear regression uses the independent data variables and the dependent variable of the data to fit a line into the data using the ordinary least squares method. This is a popular machine learning method. In our project, we have independent variables such as driver IDs, start time of the delivery and stop time of the delivery. We want to use the independent variables to predict the dependent variables. Below is the formula for multiple regressions:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon_i$$

Y_i is the estimated stop time per each stop
β_0 is the hypothesized constant
β_i is the weight of the variable
X_i is the independent variables
ε_i is the error term

3.4 Methods Summary

The availability of big data for this project will allow us to identify patterns and opportunities for managerial decision to change driver efficiency. Using the six months of ELD HOS and stop data to review the amount of time that is spent at a shipper, we can identify where drivers are losing their time. By categorizing variables and identifying relationships, the models used for this analysis can be used in the future to provide carriers insight to where time is being lost and what adjustments can be made in driver dispatching.

4. RESULTS AND ANALYSIS

4.1 Results of Hypothesis Testing

4.1.1 Hypothesis 1 Results

We tested our first hypothesis by looking at the impact of the arrival hours with the stop time. To visualize the impact of arrival hour and the average time spent, we charted the number of drivers arrived by hour and the average stop time (see *Data Distributions)*. Following the results of our visualization, we continued with an ANOVA analysis. Our initial results from the ANOVA test on driver arrival time of day show a statistically significant difference in mean stop time depending on the initial arrival hour.

ANOVA: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Hour 0	511	21.52983	1:00:40	0:03:57
Hour 1	423	21.10144	1:11:50	0:06:32
Hour 2	406	24.26859	1:26:05	0:06:48
Hour 3	492	28.19448	1:22:31	0:05:21
Hour 4	574	30.10502	1:15:31	0:04:11
Hour 5	872	44.18749	1:12:58	0:04:05
Hour 6	1342	72.28126	1:17:34	0:03:46
Hour 7	1880	101.8239	1:18:00	0:03:47
Hour 8	2514	137.6331	1:18:50	0:03:20
Hour 9	3103	169.5688	1:18:41	0:03:41
Hour 10	3547	194.5441	1:18:59	0:03:42
Hour 11	3286	182.5303	1:19:59	0:03:30
Hour 12	3148	180.7158	1:22:40	0:03:52
Hour 13	2886	163.3977	1:21:32	0:03:56
Hour 14	2615	139.3702	1:16:45	0:03:40
Hour 15	2219	115.9712	1:15:16	0:03:49
Hour 16	2040	99.24413	1:10:03	0:03:43
Hour 17	1546	76.73186	1:11:28	0:03:33
Hour 18	1393	69.76525	1:12:07	0:04:15
Hour 19	1142	54.92547	1:09:15	0:04:14
Hour 20	1155	55.31672	1:08:58	0:02:54
Hour 21	1306	56.74985	1:02:34	0:02:12
Hour 22	1160	47.24421	0:58:39	0:02:15
Hour 23	1003	44.58325	1:04:00	0:03:30

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.735097	23	0.031961	12.40213	6.63E-47	1.529505
Within Groups	104.4706	40539	0.002577			
Total	105.2057	40562				

Figure 11: ANOVA Test Results for Driver Arrival Hour on Stop Time

Based on the results in *Figure 11,* we can therefore reject the null hypothesis that the average stop time is the same regardless of driver arrival time. We will use this result as justification for using driver arrival time to model driver detention time.

We continued our analysis by using a multiple OLS regression model and used a stepwise forward and backward approach to find arrival time windows with a statistically significant impact on detention time. Our initial results produced a R-squared of .136, a statistically significant impact on detention by our *Freight Event* independent variable, and nine statistically significant time windows for driver arrival time. The nine statistically significant arrival times are 1am, 2am, 3am, 4am, 5am, 7am, 8am, 9am, and 10am. The base results show an intercept (β_0) of approximately 1 hour, or an estimated stop time of 1 hour for drop and hook type loads. For live loads (LLL), where the driver is present while the trailer is loaded, the stop time is expected to increase by approximately 1 hour which is given by our FREIGHT_EVENT_LLL (β_1) of .9928 hrs. For arrival hours of 1am, 2am, 3am, 4am, and 5am, our model found the coefficients to be positive, therefore indicating an increase in the estimated detention time. For arrival hours of 7am, 8am, 9am, and 10am, our model found the coefficients to be negative, therefore indicating a decrease in the estimated detention time. In addition, we can break these hours into two statistically significant time windows, one from 1am to 5am that showed an increase in wait time between .138 and .398 hours, and the second from 7am to 10am that showed a decrease in wait between .053 and .094 hours. The time windows information is an interesting result and counter intuitive as the lowest freight arrival volume occurs between 1am and 5am, and freight arrival volume increases and peaks during the 7am to 10am window. With a high F-value of 632.4, we can conclude overall significance of the model with approximately 13.6% of the variation in detention time being explain by the load type and arrival window (*see Figure 12*).

OLS Regression Results

Dep. Variable:	stop.spend_time_hours	**R-squared:**	0.136
Model:	OLS	**Adj. R-squared:**	0.136
Method:	Least Squares	**F-statistic:**	632.4
Date:	Sun, 29 Mar 2020	**Prob (F-statistic):**	0.00
Time:	00:17:50	**Log-Likelihood:**	-62315.
No. Observations:	40269	**AIC:**	1.247e+05
Df Residuals:	40258	**BIC:**	1.247e+05
Df Model:	10		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
Intercept	0.9685	0.008	124.284	0.000	0.953	0.984
FREIGHT_EVENT_LLL	0.9928	0.013	79.171	0.000	0.968	1.017
arr_hour_1	0.1648	0.056	2.947	0.003	0.055	0.274
arr_hour_2	0.3982	0.057	6.978	0.000	0.286	0.510
arr_hour_3	0.3508	0.052	6.750	0.000	0.249	0.453
arr_hour_4	0.2085	0.048	4.325	0.000	0.114	0.303
arr_hour_5	0.1380	0.039	3.518	0.000	0.061	0.215
arr_hour_7	-0.0561	0.027	-2.039	0.041	-0.110	-0.002
arr_hour_8	-0.0921	0.024	-3.849	0.000	-0.139	-0.045
arr_hour_9	-0.0943	0.022	-4.344	0.000	-0.137	-0.052
arr_hour_10	-0.0531	0.021	-2.589	0.010	-0.093	-0.013

Omnibus:	21662.843	**Durbin-Watson:**	1.901
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	158080.331
Skew:	2.547	**Prob(JB):**	0.00
Kurtosis:	11.263	**Cond. No.**	10.8

Figure 12: Hypothesis 1 Regression Model for Freight Event and Arrival Window

4.1.2 Hypothesis 2 Results

We tested our second hypothesis that *Freight Point* location will have impact on stop time required by driver. The stop data contains over 2500 unique shipping and receiving locations, so we grouped locations by performance and gave them a graded ranking based on the average detention time. Our goal is to first confirm these groups are statistically significant, which we expect to see as these groups are ranked on their average detention time. After confirming there is a statistically significant difference based on the detention groups we predetermined, our goal is to understand how much of the detention variation can be explained by grouping *Freight Points* based on their average detention time. We continued with an ANOVA analysis and the results showed a statistically significant difference in mean stop time between each group (*see Figure 13*).

```
In [108]:  stats.f_oneway(stop['spend_time_hours'][stop['FRT_PT_RANK'] == 'A'],
                          stop['spend_time_hours'][stop['FRT_PT_RANK'] == 'B'],
                          stop['spend_time_hours'][stop['FRT_PT_RANK'] == 'C'],
                          stop['spend_time_hours'][stop['FRT_PT_RANK'] == 'D'])

Out[108]:  F_onewayResult(statistic=4459.282686052509, pvalue=0.0)
```

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
A	B	0.3568	0.001	0.3242	0.3894	True
A	C	1.0427	0.001	1.007	1.0784	True
A	D	2.3779	0.001	2.3166	2.4391	True
B	C	0.6859	0.001	0.647	0.7249	True
B	D	2.0211	0.001	1.9578	2.0843	True
C	D	1.3352	0.001	1.2703	1.4	True

Figure 13: ANOVA Test Results for Freight Point by Group Ranking

Based on the results in *Figure 13,* we can therefore reject the null hypothesis that the average stop time is the same for each of our *Freight Point* groups. We will use this result as justification for using *Freight Point* groups to evaluate estimated stop time based on a given location. We continued our analysis by using a multiple OLS regression model to determine the significance of our groups on stop time. Our initial results produced a R-squared of .249 and statistically significant group rankings. With a high F-value of 4459, we can conclude overall significance of the model with approximately 24.9% of the variation in detention time being explain by the *Freight Point* group (*see Figure 14*). The results show an intercept (β_0) of approximately .814 hours for group 'A', an additional detention time of .356 hours for group 'B', an additional 1.042 hours longer for group 'C' than group 'A', and an additional 2.377 hours longer for group 'D' than group 'A'. By ranking *Freight Point* locations, carrier dispatchers can use a similar technique to forecast estimate stop time and better load plan for drivers.

OLS Regression Results

Dep. Variable:	stop.spend_time_hours	**R-squared:**	0.249
Model:	OLS	**Adj. R-squared:**	0.249
Method:	Least Squares	**F-statistic:**	4459.
Date:	Sat, 28 Mar 2020	**Prob (F-statistic):**	0.00
Time:	23:28:10	**Log-Likelihood:**	-59476.
No. Observations:	40269	**AIC:**	1.190e+05
Df Residuals:	40265	**BIC:**	1.190e+05
Df Model:	3		
Covariance Type:	nonrobust		

	coef	**std err**	**t**	**P>|t|**	**[0.025**	**0.975]**
Intercept	0.8147	0.008	103.519	0.000	0.799	0.830
FRT_PT_RANK_B	0.3568	0.013	28.106	0.000	0.332	0.382
FRT_PT_RANK_C	1.0427	0.014	75.135	0.000	1.016	1.070
FRT_PT_RANK_D	2.3779	0.024	99.703	0.000	2.331	2.425

Omnibus:	21327.749	**Durbin-Watson:**	1.938
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	171422.932
Skew:	2.451	**Prob(JB):**	0.00
Kurtosis:	11.839	**Cond. No.**	4.97

Figure 14: Hypothesis 2 Regression Model for Freight Point Group

4.1.3 Hypothesis 3 Results

We tested our third hypothesis for *Freight Point* location frequency by the carrier and how it would have impact on stop time. Due to the number of unique locations we categorized them into four groups based on frequency (*see Table 6*). We continued with an ANOVA analysis and the results showed a statistically significant difference in mean stop time between each group (*see Figure 15*).

```
In [110]:  stats.f_oneway(stop['spend_time_hours'][stop['FRT_PT_FRQ'] == 'A'],
                          stop['spend_time_hours'][stop['FRT_PT_FRQ'] == 'B'],
                          stop['spend_time_hours'][stop['FRT_PT_FRQ'] == 'C'],
                          stop['spend_time_hours'][stop['FRT_PT_FRQ'] == 'D'])

Out[110]: F_onewayResult(statistic=629.2400627198929, pvalue=0.0)
```

```
Multiple Comparison of Means - Tukey HSD, FWER=0.05
====================================================
group1 group2 meandiff p-adj   lower   upper  reject
----------------------------------------------------
     A      B   0.0323 0.8381 -0.0719  0.1365  False
     A      C  -0.2206  0.001  -0.305 -0.1361   True
     A      D  -0.6834  0.001 -0.7651 -0.6017   True
     B      C  -0.2529  0.001 -0.3263 -0.1795   True
     B      D  -0.7157  0.001 -0.7859 -0.6455   True
     C      D  -0.4628  0.001 -0.4975 -0.4282   True
----------------------------------------------------
```

Figure 15: ANOVA Test Results for Freight Point by Carrier Frequency

Based on the results in *Figure 15,* we can therefore reject the null hypothesis that the average stop time is the same for each of our *Freight Point* frequency groups. While at least one group is statistically significantly different from the others, group 'A' and 'B' are not statistically significantly different from each other.

We will use this result as justification for using *Freight Point* frequency groups to evaluate estimated stop time based on how often the carrier has been to that given location. We continued our analysis by using a multiple OLS regression model to determine the significance of our groups on stop time. Our initial results produced a R-squared of .045 and statistically significant group rankings. With a high F-value of 943.6, we can conclude overall significance of the model with approximately 4.5% of the variation in detention time being explain by the *Freight Point* frequency (*see Figure 12*). The results show an intercept (β_0) of approximately 1.777 hours for group 'A' and 'B', for group 'C' we see an expected decrease in stop time of .239 hours, and a decrease in .702 hours from our intercept for group 'D'. This confirms our initial hypothesis that locations that are visited more frequently have shorter stop times.

OLS Regression Results

Dep. Variable:	stop.spend_time_hours	R-squared:	0.045
Model:	OLS	Adj. R-squared:	0.045
Method:	Least Squares	F-statistic:	943.6
Date:	Sun, 29 Mar 2020	Prob (F-statistic):	0.00
Time:	11:12:40	Log-Likelihood:	-64330.
No. Observations:	40269	AIC:	1.287e+05
Df Residuals:	40266	BIC:	1.287e+05
Df Model:	2		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	1.7774	0.020	88.790	0.000	1.738	1.817
FRT_PT_FRQ_C	-0.2393	0.023	-10.436	0.000	-0.284	-0.194
FRT_PT_FRQ_D	-0.7022	0.021	-32.838	0.000	-0.744	-0.660

Omnibus:	20462.718	Durbin-Watson:	1.916
Prob(Omnibus):	0.000	Jarque-Bera (JB):	125832.367
Skew:	2.443	Prob(JB):	0.00
Kurtosis:	10.150	Cond. No.	7.34

Figure 16: Hypothesis 3 Regression Model for Freight Point Frequency

5. DISCUSSION

5.1 Conclusion

Even though the adoption of the ELD mandate was intended for reducing trucking accidents, the ancillary benefits of the mandate with the newly available data have proven invaluable in deriving insights into the daily operations of drivers. With the use of data analytics, it is our goal to help management find strategies that help drivers spend less time waiting and more time driving. Our project achieved the goal of finding factors that negatively impact the number of hours driven per day. Based on our analysis, we have found opportunities for management to dispatch drivers that would help limit their stop time at shipping and receiving locations. We found statistically significant variables in arrival time, *Freight Point* locations, and *Freight Point* frequency. Developing dispatch strategies based on driver arrival time, load *Freight Points*, and the frequency of service to those locations can help make improvements on the amount of time drivers spend at shipper and receiver locations and in turn give them more available time to drive. Our goal was to help alleviate the industry driver shortage by finding ways to improve driver utilization, specifically by 3.4% or 12 minutes of drive time per day. While our results did not identify a full 12 minutes for every driver, we were able to identify time saving activities that can work toward finding the full 12 minutes.

5.2 Limitations of the Model

We used one carrier data, and the specific data may not representative of trucking industry as a whole. We did not exam broad factors such as congestion and variation of equipment performance due not having data on those factors. A further follow up analysis of using finer-grained measures (e.g. change in stop time over course of a week) are needed to measure the available stop time factors in greater detail.

Our analysis was based on the assumption that trucking firms can change their shipping and receiving arrival windows and influence their shipping and receiving locations. In reality, the benefits of this analysis are limited to the ability to change parts of the daily operation of the trucking firm. Without the ability to change operational strategy in the short term, the proposed benefits will not materialize.

5.3 Managerial Implications

Our team thought by giving the insights on detention variability, it will allow the company to predict hours of driving time availability. Having the ability to better predict the hours of available drive time will allow the company to reduce the time gap between deliveries which can result in better service and utilization of available driving hours.

5.3.1 Time of Arrival Impact on Detention Time

The results from our first hypothesis would suggest that the time of arrival has a statistically significant impact on detention time. Having drivers avoid delivery times between 1 am and 5 am could save an average of 15 minutes per stop. According to our model, deliveries between 7am and 10 am could save an additional 5 minutes of detention time per stop. Helping drivers plan their delivery times can help them save time at their stops which would give them more time to drive during the day.

5.3.2 Detention Time Variability at Different Locations

Due to past experiences, the company is aware there are variation of amount of wait time at customer locations. However, the company currently does not have insight on the amount of time its drivers spent on each shipping and receiving location. The results from our second hypothesis show that there are statistically significantly different wait times between shipping and receiving locations. By grading carriers based on the average detention time dispatchers can use this as a guide to estimate the amount of time a driver will spend at a shipper and be able to use this information for future load planning. We found that creating grades for shippers solely based on their average detention time, can account for approx. 25% of the variation in detention time at that stop.

5.3.3 Location Frequency Impact On Detention Time

When we first interviewed the sponsor company, they thought the frequency of visits would not have a significant impact on detention time. Our third hypothesis found that carrier frequency to a shipping location does have a statistically significant impact on stop time. We believe that keeping drivers more regionally based and servicing freight points they are familiar with can help reduce detention time at shipper locations.

We hope these simple load planning changes can help our company and others adjust the strategies in their daily operations. By managing these variables, we hope to be able to reduce the time drivers spend at shippers and regain their valuable drive time.

REFERENCES

American Transportation Research Institute (2019), “Critical Issues in the Trucking Industry”, 4-12.

American Trucking Association (2019), “Reports, Trends & Statistics”, Trucking.org https://www.trucking.org/News_and_Information_Reports_Industry_Data.aspx

American Trucking Association (2019), “Truck Driver Shortage Analysis 2019”, 2-15.

Belzer, M. (2000). “Sweatshops On Wheels: Winners and Losers in Trucking Deregulation”. Oxford University Press: 891 – 893

Burks, Stephen V., Monaco, Kristen (2019). “Is the U.S. labor market for truck drivers broken?”. Bureau of Labor Statistics Monthly Labor Review March 2019

Cantor, David E., Thomas M. Corsi, and Curtis M. Grimm. "Do electronic logbooks contribute to motor carrier safety performance?" *Journal of Business Logistics* 30.1 (2009): 203-222

Correll, David (2019) “Is There a Truck Driver Shortage or a Utilization Crisis?” Medium.com https://medium.com/mitsupplychain/is-there-a-truck-driver-shortage-or-a-capacity-crisis-73a7d7a5bcb

Cheeseman Day, Jennifer C. and Hait, “Number of Truckers at All-Time High”, US Census Bureau, https://www.census.gov/library/stories/2019/06/america-keeps-on-trucking.html

Federal Motor Carrier Safety Administration (2015), “Rule 78292”, 49 CFR Parts 385, 386, 390, and 39

Min, Hokey and Lambert, Thomas (2002) “Truck Driver Shortage Revisited”, *Transportation Journal;* Winter 2002; 42, 2; ABI/INFORM Collection

Overdrive (2019), “A Goldrush for ELD Data”, 26 -30

Rodriguez, Daniel A., Targa, Felipe and Belzer, Michael H. (2006), “Pay Incentive and Truck Driver Safety: A Case Study”, ILR Review

Suzuki, Yoshinori, Crum, Michael, and Pautsch, Michael (2009), “Predicting truck driver turnover”, Transportation Research Part E, 538 – 550

U.S. Bureau of Labor Statistics (2019), U.S. Department of Labor, Occupational Outlook Handbook, Heavy and Tractor-trailer Truck Drivers, https://www.bls.gov/ooh/transportation-and-material-moving/heavy-and-tractor-trailer-truck-drivers.htm

U.S. Bureau of Labor Statistic (2020), Labor Force Statistics from the Current Population Survey. https://www.bls.gov/cps/cpsaat11b.htm

U.S. Census Bureau (2019), America Counts: Stories Behind the Numbers, America Keeps on Truckin’, https://www.census.gov/library/stories/2019/06/america-keeps-on-trucking.html