Transportation Late Load Analysis

Michael A. Paris

DCS 530

Bellevue University

Abstract

This paper will detail the analysis of over-the-road transportation loads of a major U.S. retailer. The analysis will include descriptions of all variables including definitions, descriptive statistics including mean, mode, spread, and tails where applicable. Analysis tools will include cumulative distribution functions, analytical distributions, and scatter plots of the variables. Regression analysis will be performed and the hypothesis will be tested.

Keywords: Supply Chain Analysis, Transportation Loads, Late Delivery Analysis, Regression Analysis

Transportation Late Load Data Analysis

Walmart’s Transportation Division is responsible for moving freight from vendors to distribution centers (DCs) and from DCs to stores and clubs. Delivering freight on time and in full as efficiently as possible is one of the most important goals for the Walmart Transportation division. On-time delivery reduces out of stocks, helps reduce lost sales, and reduces wages in the form of idled unloading teams. This project will analyze transportation load data to determine what significant correlations exist in the data to help formulate actionable process changes that will reduce the late delivery of loads.

# The Analysis Outcome

The outcome of my analysis was inconclusive. While I was able to gain some insight into the variables, I was ultimately unable to correctly perform the logistic regression test against the group of variables that I chose in my analysis. The two regression tests that I was able to get partially through showed current function values of .618 and .608 which I interpret to be an approximate match of 61% which is not very good.

## Analysis Misses

I spent more time during my analysis in cleaning and understanding the data. I spent more time on this portion of the project than I had anticipated and that additional time spent here took time away from the rest of the analysis. This was a big lesson learned for me and I look forward to the next class DCS540 for increasing my skillset in this topic.

### **Additional helpful variables**

One aspect of the analysis was visualizing where the transportation loads were coming from and where they were going. The dataset I have has both location numbers and address information. I chose to use the location numbers as I expected individual locations to be relevant to determining late loads. This decision made visualizing the data very difficult as there are over five thousand origin locations. In retrospect, I could have broken the dataset into smaller locations related to geographic areas. I will continue working on this project over the summer to s

### **Assumptions**

As I mentioned above, I assumed that late loads will be related to specific origin locations or groups of origin locations. I don’t think this assumption is incorrect, I have just not been able to verify the results.

### **Challenges**

One challenge I had was having a mix of continuous and categorical data to analyze. I honestly forgot about logistic regression for the first part of the project and had to study that section again. I would like to read additional work on analyzing sets of data that are made up of multiple forms of variables.