



# Atrium - Crossroads Classic Analytics Datathon

This document breaks down the details of the Atrium truck data set and problem. This problem is open-ended and can be solved with a variety of techniques, including data visualization, summarization, modeling, or other potential techniques.

## The Problem: Truck Sales Predictions

Truck sales can be a difficult thing to predict as sale prices are often determined based on a large number of factors, including market conditions, truck characteristics, and economic factors like competition, supply, and demand. Much of the time, the process of selling a truck can involve subjective decision making.

Atrium is interested in being able to automate truck pricing based on a variety of factors, and provide their salespeople with a better starting point in sales negotiations.

Output from this analysis could be used in several ways:

- 1) Inform **Top Factors** that determine the price of a truck
- 2) Forecast

## The Data

To build this model, you will have access to a single data file containing information about a sample of sale truck prices from recent transactions.

Field Name	Type	Description
Truck ID	Categorical	Identification number assigned to each truck sale
Qtr	Categorical	Quarter and Year of truck sale
Quarter	Categorical	Quarter of truck sale
Year	Categorical	Year of truck sale

Sale_Price	Numerical	Final sale price of each truck
List_Price	Numerical	Original listed price of each truck
Pct_of_List	Numerical	List Price percentage of Sale Price
Book_Value	Numerical	Booked value of each truck
Type	Categorical	Indicates a Truck vs. a Tractor
Truck_Manufacturer	Categorical	Manufacturer of the truck
Region_Nm	Categorical	Region truck is sold from
Condition	Categorical	Condition the truck is in at time of sale
Model_Year	Categorical	Model year of truck for sale
Odometer	Numerical	Miles driven at time of sale
Model	Categorical	Model name of truck for sale
Transmission	Categorical	Transmission type
Engine_Manufacturer	Categorical	Manufcaturer of truck's engine
Fuel_Type	Categorical	Type of fuel the truck uses
Wheelbase_Size	Numerical	Distance between the cneters of the front and rear wheels of a vehicle
Gear_Ratio	Numerical	Ratio of the number of rotations of a driver gear to the number of rotations of a driven gear
Num_Axles	Numerical	Number of axles on the truck
Num_Wheels	Numerical	Number of wheels on the truck
Front_Tire_Size	Categorical	Size category of front tire
Rear_Tire_Size	Categorical	Size category of rear tire
Color	Categorical	Paint color of truck for sale
Delivery_Region	Categorical	Region the sold truck was delivered to
Channel	Categorical	Sales path that was taken to get truck to end

		customer
Delivery_Location	Categorical	Location the sold truck was delivered to
Age_Yrs	Numerical	Truck age
Rear_Axle_Weight	Numerical	Weight limit for the truck's rear axle
front_axle_wt	Numerical	Weight limit for the truck's front axle
Horsepower	Numerical	Unit of measurement for power of the truck
Engine_RPM	Numerical	Unit of measurement for speed of the truck's engine
eng_est_life	Numerical	Estimated miles left of the road for the truck
Engine_Type	Categorical	Type of engine in the truck

## The Details - Modeling

Please write your code in R or Python.

Outputs could consider any or all of the following:

1. Exploratory data analysis
2. Model training
3. Model selection and evaluation
4. Model explanations (e.g., feature importance) and interpretations

If you have any questions, please feel free to email Paul Harmon at [paul@atrium.ai](mailto:paul@atrium.ai).

## The Details - Dashboarding

Please build your dashboard in either Tableau or Power BI.

Outputs could consider any or all of the following:

5. Data Cleanliness Analysis
6. Sales Analysis Dashboard



If you have any questions, please feel free to email Tyler Pollard at [tyler@atrium.ai](mailto:tyler@atrium.ai).