## 1. Problem statement:

At Rumford Station railyard, shipments destined for this station or ones nearby are loaded, unloaded or stored according to an non-automated process that considers a railcar’s ownership, contents and destinations. Anticipating the duration of railcar dwell times at a railyard can help both the railroad company and their customers more accurately estimate shipment delivery times and car hiring costs. Often, dwell time targets of 48 h are missed, delaying shipments and costing railroads additional car usage fees. Using historical data, this project aims to create a model that can to predict how long an incoming train will sit idle at Rumford station in Rumford, Maine. In the creation of this model, key features that most influence train car idle times can be identified that may be applicable for the development of similar models for other train stations.

## 2. Hypothesis:

Car attributes such as car type features, hours at the station and final destination most strongly influence the likelihood and frequency of car turnover. Car turnover can be optimized from its current process to accelerate delivery, provide more accurate delivery estimates and reduce car hire costs.

# 3. Limitations:

* + 1. The dataset used is from a fairly small dead end station for simplicity.
    2. The span of time covered in the data set is limited to one year

# 4. Access to Data:

Data used for this model is accessible from RailConnect by Railrcar Management Inc. The Rumford Station dataset consists of approximately 80,000 lines of train car inventory over the past year, each with information regarding its destination, type, loading, accessibility and more.