<h1 style="text-align:center">Trinity Rail Data Science Problem</h1>

Problem:

Shopping cars (maintenance events) is a large logistical problem for rail car owners and managers. One pain point is not knowing how long a car will be out of service while in a repair facility (shop). Attached is a list of shopping events from 2018 to 2019 for various shops. These cars are in shop for the same reason but some cars may require additional attention.

Goal:

Create a regressor to predict how long a car will be in shop (Cycle Time). Use cars completed 1/1/2018 to 5/31/2019 as your training set and provide performance metrics (Mean Absolute Error & Root Mean Square Error) for cars completed 6/1/2019 to 12/31/2019. Compare your performance metrics with a baseline (Avg Shop Cycle Time).

* Hint: $CycleTime_{days} = DateCompleted - ArrivalDate$

Deliverable:

Jupyter notebook containing the following sections:

- Exploratory Data Analysis (EDA)
- Model Training
- Model Performance: MAE and RMSE on test set

Dataset: problem_dataset.csv -> '|' delimited file

- ArrivalDate, DateCompleted – Datetime ☐   Age – Years since car was built

- ShopRecordID – Shopping event Identifier
- ShopID – Shop location identifier
- CarModelID – Car model identifier
- CommodityID – Commodity Type Identifier
- IsHeavy – Heavy work for repair event identifier