Problem Statement

Torqata, a wholly owned subsidiary of American Tire Distributors (ATD), is in prime position to help coordinate a new sustainable future for the automotive aftermarket. As a connector with data visibility and analytics capability across the supply chain, Torqata can help manufacturers, distributors, and retailers accomplish sustainability goals through a coordinated effort to recycle scrap tires.

However, there are obstacle in the way – even for the largest tire distributor in North America. Despite ATD being a juggernaut in the industry, delivering over 44MM tires annually across 80K retail locations through a network of 120 Distribution Centers (DC), ATD's foray into scrap tire recycling logistics faces two challenges:

- **Scrap Tire Forecast:** How many tires of a given size_code will be collected by a given Distribution Center every day?
- **Logistics Optimization:** How to efficiently move the scrap tires downstream to the recycler?

More importantly, ATD needs to accomplish this *while minimizing* those Green House Gas (GHG) emissions *per tire*! Are you up to the task?

Sustainability is the future of the automotive aftermarket. And to help tackle these challenges, we have invited some of the best data science teams from across the US and the world to compete in this year's ***Reinvent the Wheel 2.0*** Hackathon.

You will have 24 hours to find a solution to both challenges. May the best team win!

1. Scrap Tire Forecast

1.1 OVERVIEW

When a customer visits a tire shop for tire replacement, the tires removed (known as scrap tires) must be disposed of. ATD is exploring the logistics of scrap tire recycling – acquiring and delivering them to downstream recyclers. To effectively do this, ATD needs to know the future demand for scrap tires at their warehouses on any given day. Predicting this volume is the goal of the first challenge. We are challenging you to build a model that will forecast the number of scrap tires in each size that a DC can expect to receive from its customers daily, one to seven days out.

1.2 THE QUESTION

Given a distribution center and size_code combination, can you tell us the number of tires that will be received at the distribution center for any given day, seven days out?

1.3 DATA

1.3.1 Historic Tires Received (Challenge1_train_data.csv)

A CSV containing details about all the tires received at the distribution centers. You are given their historical data from September 20, 2020, to September 19, 2022.

date	dc_name	zip_code	size_code	retail_price	total_tires
2020-12-05	Sacramento	81235	2253521	143.56	5
2020-12-05	Sacramento	81235	2653519	281.59	14
2020-12-05	Bakersfield	31256	2253521	145.62	2
……	……	……	……	……	……
2020-12-07	San Jose	56734	2253521	151.07	12

Feel free to use any other external data sources (for example, weather data), on top of the data sources provided.

1.4 RESULTS FORMAT

Your task is to forecast the number of tires received at the distribution centers for each size_code for each of the days from September 20, 2022, to September 26, 2022. As an example, your predictions should look like:

date	dc_name	size_code	total_tires
2020-09-20	Sacramento	2253521	1
2020-09-20	Sacramento	2653519	5
2020-09-20	Bakersfield	2253521	12
……	……	……	……
2020-09-22	San Jose	2253521	2

Make sure your column names are the same as shown above when submitting, including case. Also, don't include any additional columns in your submission. Note that there is no time component for the date (e.g., the date is "2022-09-20", not "2022-09-20-2022").

1.5 EVALUATION

We will calculate the WAPE (Weighted Absolute Percentage Error) as our metric to calculate the final score. WAPE is calculated across DC-Size Code for each horizon, i.e., a single WAPE number per day. The final score will be an average WAPE across the horizons; lower is better. Here's the equation to calculate the score.

$$\text{Score} = \frac{1}{7} \sum_{September\ 20}^{September\ 22} \left(\frac{\sum(abs(actual - predicted))}{\sum(abs(actual))} \right)$$

Here is an example of how WAPE is calculated. We will also provide you with a python function to calculate WAPE. This will be available along with the dataset you will receive.

A WAPE score of 10 implies an error of 1000%. If your score is more than 10, we will cap it to 10 to normalize the scores for the leaderboard.