Using time series forecasting to predict grocery store sales

Food loss and food waste is a major global concern. Food is wasted at three levels:

- At the production level, where food is damaged or spoiled
- At the retail level, where food is thrown out due to overbuying or deviations from what is considered optimal (bruising, discoloration, abnormal shape, etc.)
- At the consumer level, when customers buy more than they need and throw out unused or spoiled food.

According to the World Food Program USA, nearly one third of all food produced each year is squandered or spoiled before it can be consumed and approximately $1 trillion dollars' worth of food is lost or wasted yearly. According to this article published in 2013 by the Central European Journal of Engineering, more than 95% of food waste end up at landfills where they rot, emitting methane, carbon dioxide, and other greenhouse gases, thus contributing to global warming and climate change.

A possible solution to decreasing food waste at the retail level is to help grocery stores find the delicate balance between providing just enough inventory for customers and restocking at the appropriate time to avoid shortages. Accurate forecasting can help ensure retailers please customers by having just enough of the right products. In this project, I will predict store sales on data from Corporación Favorita, a large Ecuadorian-based grocery retailer. The data is readily accessible through Kaggle and includes dates, store and product information, whether the item was being promoted, as well as the sales numbers. I will use time-series forecasting to build a model that more accurately predicts unit sales for thousands of items sold at different Favorita stores.

Food waste amounts to a major squandering of water, land, energy, labor, and capital resources, producing greenhouse gas emissions, and contributing to catastrophic climate change. To combat this problem, a model can be built to predict the unit sales for items sold at groceries stores so that consumers can be provided with just the right amount of inventory without over or understocking.

Extra note:

Going forward, to reduce food waste at the consumer level, we can try to implement this to every household where they input what they want to cook and how many servings of it. We use this information to calculate the precise amount of groceries needed to make the dishes to prevent overbuying groceries which can lead to food spoilage and food waste.
Current subjective forecasting methods for retail have little data to back them up and are unlikely to be automated. The problem becomes even more complex as retailers add new locations with unique needs, new products, ever-transitioning seasonal tastes, and unpredictable product marketing.