[Heartbeat Forecast Overview](https://confluence.kroger.com/confluence/display/8FCEF/Heartbeat+Forecast+Overview)

* Created by [DANIEL ROGERS](https://confluence.kroger.com/confluence/display/~DR21060), last modified on [May 09, 2023](https://confluence.kroger.com/confluence/pages/diffpagesbyversion.action?pageId=316739922&selectedPageVersions=2&selectedPageVersions=3)

Overview

The "Heartbeat" forecast is a store/item/day forecast used in the Kroger supply chain.  It's primary purpose is to drive store orders.  It is also sent to the Demand Planning team at Kroger for inclusion in their Oracle RDF system where it is "blended" with other forecasts to create a forecast that is used in DC replenishment.  This page provides a high-level overview of the forecast.  More details can be found at the links below:

* [Internal FCOE Documentation](https://confluence.kroger.com/confluence/display/8FCEF/Heartbeat) - more detailed, technical documentation
* [Heartbeat Forecast Files](https://confluence.kroger.com/confluence/display/8FCEF/Heartbeat+Forecast+Files) - how to access Heartbeat files in Azure

History

The Heartbeat forecast was developed by Paul Helman, Hank Vaccaro and team at KSS Retail.  The original system was designed for real-time out-of-stock detection and alerting.  It was even deployed in other companies (outside Kroger) for that purpose.  It would monitor sales as they occurred and look for significant deviations in observed sales from expected sales patterns.  In 2014 it was modified to also produce forward-looking forecasts.

KSS Retail was acquired by dunnhumby prior to the split of dunnhumbyUSA.  When dunnhumby split, both of the new companies maintained ownership of the Heartbeat forecast, although some of the SMEs came to dunnhumbyUSA and some (specifically Mike Dolan) stayed with dunnhumby proper.

Heartbeat was evaluated in 2014 as a potential replacement for the CAO forecast used in store ordering.  Multiple forecasts were measured (including CAO, a forecast from Kroger's Operations Research team, and an Oracle RDF forecast).  Heartbeat was found to have the best performance and rollout to the company began in 2015.

The forecast was maintained by the Science Team at 84.51 and a team at Kroger Technology that supported the technical system, servers, and file integrations.  In 2022, the Heartbeat science team at 84.51 joined FCOE (Forecast Center of Excellence).

Algorithm

The Heartbeat forecast consists of two parts: a "traffic" forecast that predicts total store traffic (essentially total store sales) by day, and the core Heartbeat model that estimates what percentage of store sales should be allocated to each item.

  This design is an artifact of Heartbeat's original purpose of detecting out-of-stocks in real time.  In this original setup, actual sales data would be fed to Heartbeat throughout the day.  It did not need to predict total store traffic (sales).  This was instead available.  It could then focus on understanding how each individual item in the store should be represented as a function of total store sales.  In order to use Heartbeat as a true, forward-looking forecast we need to first forecast total store traffic and provide this to Heartbeat as an input.  The core Heartbeat model will then be able to examine the mix of items that are sold.

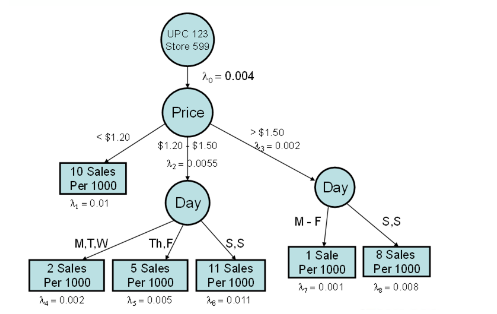
**Traffic Model**

As of 2023, Heartbeat's traffic model is a combination of a Linear Model (LM) and RNN (Recurrent Neutral Network) model.  The LM model was deployed with Heartbeat originally.  It is a fairly simple model of store traffic as a function of historical traffic with factors to adjust for holidays and events.  It also has a clustering component to identify similar stores for newer stores that lack sufficient history to proper model holidays and to use these as proxies for the newer stores in the model.  The RNN component was added later as a Science Research project.  It is more of a "black box" that models historical traffic patterns in some way and produces forecasts.  This was found to be superior to the linear model in the original assessment, especially in terms of holiday forecasting.  The RNN was used to produce traffic forecasts for the first 14 days of the forecast.  Beyond those 14 days the linear forecast will still be used.

In January 2023, the RNN model had trouble due to snowstorms in the prior year.  The model focused on these snowstorms, expecting them to repeat, and produced forecasts 50-100% higher than they should be.  The FCOE team turned off the RNN in response to this and used the Linear Model for everything.  In May, the RNN was turned back on just for Senior Day events in Fry's and Fred Meyers, since these were significant misses in the linear model.  In 2023, the FCOE team prioritized research into the Heartbeat traffic model, looking for ways to improve the forecast while maintaining its stability.

Tree Model

Heartbeat models the mix of products sold within the store using a tree-based model.  Technically, this is a Bayesian "Option Tree", built using a custom model developed by Paul Helman and Hank Vaccaro.  The graphic below provides some idea of how this works:



A separate tree is built for each item in every store.  This will model the rate of sale for that particular item with respect to overall store traffic.  The simplest model might just say that the item sells X units for every 1,000 units of store traffic.  The tree will then explore how this sales rate changes in response to prices, promotions, and temporal features (such as holidays and events).  In the sample above, the tree is split first based on item price and then by day-of-week.  It is important to note that the total store traffic pattern captures overall changes in traffic, especially as it varies by day-of-week and for holidays.  This allows the tree to focus only on cases where the overall mix of items changes.

The details of the tree model are pretty complicated.  It uses an advanced search and pruning algorithm to build the trees.  The final trees are then used in a multi-level ensemble mode, making predictions at each level of the tree and combining those forecasts in a statistical way to create its final prediction.