

Executive Summary

A Pricing Analytics Tool for the Planning & Optimization Team

Market Survey

There are many periods during which a subset of competitors omit a particular product or rack type from their offering. [Client] stands to profit by pushing unbranded premium product during periods of higher index cost, and branded regular/diesel otherwise.

Historical analysis has revealed that [...] can serve as a benchmark competitor for regular and premium branded product and [...] for branded diesel. [Client] can reference these when competitors when developing strategies for pricing unbranded product.

Modeling Capabilities

Time Series Analysis

Decompositional models have been built to quantify the predictability of historical prices by breaking into a seasonal index (low in Winter, high in Summer for these data) and overall trend component (developed over long periods). These measures can be used to highlight anomalous periods for further investigation.

Forecasting models have been used to project margin profiles and index prices into the future. In particular, a Holt-Winters forecast with additive effect applied to month-level data has yielded the best results. These predicted values are used later in the modeling pipeline but can also be analyzed directly for any desired time series.

Bid Outcomes

Models developed to predict the probability of winning a given bid for commercial contract. Best accuracy achieved is 85.7% by Random Forest Classifier which most heavily weights values for expected margin, merged adder, and existing gallons per month. A logistic regression is also presented for simplicity/transparency; coefficient values can be used to interpret model behavior, but more observations are needed in order to make definitive conclusions.

Palantir Workflow

Data, visualizations, transforms, and models have all been built in Palantir cloud environment. Structure has been built out for building and evaluating commercial contract bids, tying in the above features. This work also provides a foundation for further analysis/implementation in PySpark within Palantir. Detailed documentation has been prepared on how to navigate this environment.

Next Steps

The most important component to this project's long-term success is the collection of more bid outcome data. Evaluation of the notebook interface will need to be completed and alternative options for input/output researched if notebook proves insufficient.