FORECASTING · TIME SERIES · DEMAND PLANNING

Demand Forecasting: *the seasonality, the accuracy, and the variance you monitor*.

A monthly demand series forecast with a seasonal exponential-smoothing model, backtested on a rolling origin against a naive baseline, then wrapped in a forecast-versus-actual monitor that flags when the number needs a corrective override.

Author S. Ize-Iyamu **Audience** Demand planning & merchandising **Length** 2 pages
Method Holt-Winters · rolling-origin backtest **Stack** Python · pandas · statsmodels

The Problem

A forecast is decision infrastructure, not a number on a slide. Replenishment, assortment, and financial targets all read off the same demand number, so the job is three connected things: capture the seasonal shape, get the level right enough to trust, and know the moment the forecast drifts from what actually sells.

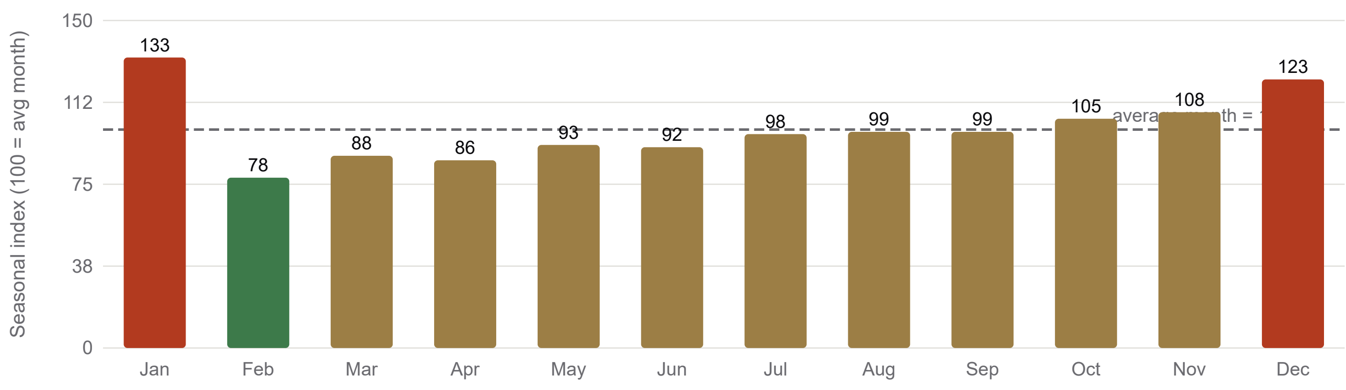
This brief works a real monthly demand series[1] (204 months, Jul 1991 to Jun 2008) end to end, from the seasonal profile through a backtested model to the variance monitor that sits on top.

BACKTEST WMAPE	NAIVE BASELINE	ACCURACY GAIN	FORECAST BIAS	HORIZON
6.5%	10.9%	40%	-2.2%	12 mo

The Seasonal Shape

The first thing a demand forecast has to respect is the calendar. A ratio-to-moving-average decomposition[2] pulls out a repeating seasonal index: this series peaks hard in January (index 133) with a second lift into December (123) and troughs in February (78), a 55-point swing from top to bottom. Underneath the season the level grows about 12% a year, so the model has to carry a trend and a multiplicative season at the same time, which is exactly what Holt-Winters is built for.[3]

FIGURE 1 · SEASONAL PROFILE



Average seasonal index by month (100 = a typical month). January and December run well above trend; February is the trough. A forecast that ignores this shape is wrong by more than half a typical month at the peak.

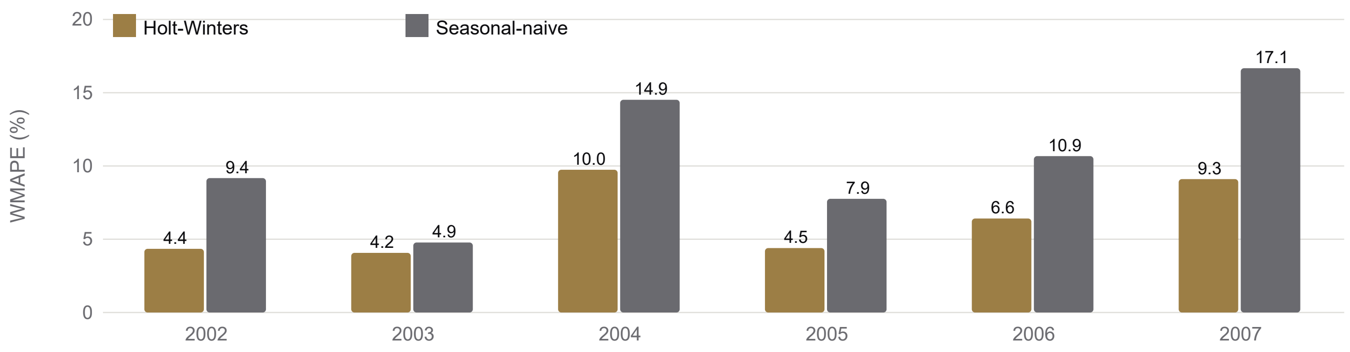
> **WHY IT MATTERS FOR MERCHANDISING**
> The same swing shows up in any retail category that sells on a season or a promotion. Get the shape right and stock is in place for the peak and off the shelf for the trough, so the seasonal index is the part of the forecast that touches inventory first.

Sources. 1. Monthly demand series: total monthly scripts of a subsidised pharmaceutical (a10), Australia, Jul 1991 to Jun 2008, in AUD millions, from R. J. Hyndman & G. Athanasopoulos, **Forecasting: Principles and Practice** (3rd ed.). https://otexts.com/fpp3/ 2. Classical seasonal decomposition by ratio-to-moving-average: Hyndman & Athanasopoulos, FPP3, ch. 3 (Time-series decomposition). https://otexts.com/fpp3/decomposition.html 3. Holt-Winters (trend + multiplicative seasonal) exponential smoothing: FPP3, ch. 8; implemented with statsmodels `ExponentialSmoothing`. https://otexts.com/fpp3/holt-winters.html

Does the Model Actually Forecast?

Accuracy is measured the honest way, out of sample. A rolling origin[4] refits the model at each January and forecasts the next twelve months, so every number is a real forecast the model had not seen. Averaged over six annual origins the Holt-Winters forecast lands at **6.5% WMAPE** (6.6% MAPE), against **10.9%** for a seasonal-naive baseline that repeats last year. That is **40% less error**, the comparison that says the model adds signal rather than dressing up last year.[5]

FIGURE 2 · ROLLING-ORIGIN BACKTEST, ERROR BY FORECAST YEAR

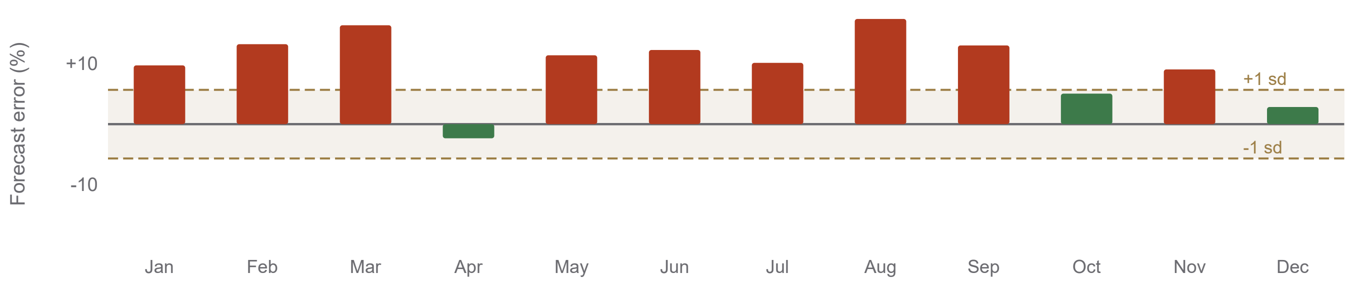


Weighted MAPE for each 12-month forecast, Holt-Winters against the seasonal-naive baseline. The model beats the baseline in every origin year.

Monitoring the Variance

A forecast is not done when it ships. The monitor reads each month's actual against forecast as a percentage error, with control limits at plus or minus one standard deviation of the backtest residual.[6] In the 2007 holdout almost every month prints above the forecast: demand ran ahead of the model all year, a persistent one-sided miss rather than noise. That signature, several months breaching the same limit in the same direction, is the trigger to override the forecast up rather than wait for the annual refit.

FIGURE 3 · FORECAST-VS-ACTUAL VARIANCE MONITOR (2007 HOLDOUT)



Monthly forecast error, actual minus forecast as a percent of forecast. Bars outside the shaded plus/minus one-standard-deviation band (red) are exceptions worth a planner's attention; the run of positive misses signals the forecast is biased low and needs a corrective lift.

Limitations & Method

Honest scope. One series demonstrates the workflow; a retail book runs thousands of item-location forecasts and needs promotion and price effects and reconciliation up the category hierarchy, which a single univariate model does not cover. The discipline is what carries into that setting: a seasonal model, an out-of-sample backtest against a baseline, and a monitor that reads bias. **Method:** ratio-to-moving-average decomposition; statsmodels Holt-Winters (additive trend, multiplicative season); six-origin rolling-origin backtest at a 12-month horizon; WMAPE, MAPE, signed bias; residual control limits at plus or minus one standard deviation.

Sources. 4. Rolling-origin (time-series cross-validation) evaluation: Hyndman & Athanasopoulos, FPP3, sec. 5.10. https://otexts.com/fpp3/tscv.html 5. Accuracy metrics (MAPE, scaled error / MASE) vs a seasonal-naive benchmark: Hyndman & Koehler, "Another look at measures of forecast accuracy," **Int. J. Forecasting** 22(4), 2006. https://doi.org/10.1016/j.ijforecast.2006.03.001 6. Control limits on forecast residuals (statistical process control): D. C. Montgomery, **Introduction to Statistical Quality Control**. https://otexts.com/fpp3/accuracy.html