

Sales Forecasting At “James & Sara Dairy Farm”

Posted on October 10, 2019

View this essay online:

<https://academic-master.com/sales-forecasting-at-james-sara-dairy-farm/>

Introduction

Sales forecasting helps James and Sara Dairy Farm convert three years of monthly records into decisions about production, feed, labor, packaging, refrigeration, transport, and cash flow. The thirty-six-month series described in the assignment contains a recurring seasonal pattern, with stronger sales generally occurring from October through March and weaker sales during late spring and summer. The statistical task is therefore to estimate seasonality, determine whether a time trend improves the model, and forecast months thirty-seven through forty-eight. Monthly dummy-variable regression is a reasonable approach because it makes seasonal differences easy to interpret, while adding a time variable allows the model to represent gradual growth or decline. Forecasting accuracy, however, should not be judged only by how closely a model fits historical observations. The farm must also verify data quality, compare models on unseen periods, inspect residuals, and communicate uncertainty so that a numerical forecast becomes a practical management tool rather than a false promise of precision.

Seasonality and the First Regression Model

The seasonal model uses eleven monthly indicator variables and treats December as the reference month. January equals one for January observations and zero otherwise, February is coded in the same way, and the remaining monthly indicators follow the same rule; December is represented when all eleven indicators equal zero. Omitting one month avoids perfect multicollinearity and gives the intercept a clear interpretation. The reported equation is $\text{Sales} = 279.666 + 3.333 \text{ Jan} - 15.667 \text{ Feb} - 40.667 \text{ Mar} - 65 \text{ Apr} - 97 \text{ May} - 131.333 \text{ Jun} - 136.667 \text{ Jul} - 109.333 \text{ Aug} - 77.667 \text{ Sep} - 43.667 \text{ Oct} - 22.333 \text{ Nov}$. Under this model, the intercept predicts about 279.7 units in December, January is about 3.3 units higher, and July is about 136.7 units lower. These coefficients describe recurring calendar differences, not causal effects of the months themselves, and they should be interpreted as averages across the three observed years. Interpreting dummy coefficients against an omitted reference category is standard regression practice and prevents redundant seasonal indicators (Newbold et al., 2013).

The seasonal-only model produces the same forecast for every January, every February, and so forth

because it contains no mechanism for gradual change across years. This can be appropriate when demand is stable, but it becomes restrictive if the series is trending. The residuals provide an important diagnostic: actual sales minus predicted sales should fluctuate around zero without a systematic pattern. If early observations are mostly below the fitted values while later observations are mostly above them, the model is probably missing an upward trend. A repeated wave in residuals suggests that the seasonal specification is incomplete, while unusually large isolated errors should be investigated for operational causes such as equipment failure, lost customers, weather, promotions, or supply constraints. Before modeling, the farm should also verify that all thirty-six observations use one consistent sales definition. Revenue, units, and liters are not interchangeable, and a rise in revenue may reflect price changes rather than stronger physical demand.

Adding Trend and Producing the Fourth-Year Forecast

The second specification adds a sequential time variable to the monthly indicators. Interpreting the reported equation as $\text{Sales} = 286.33 + 1.05444(\text{Time}) + 24.333 \text{ Jan} + 22.333 \text{ Feb} + 24 \text{ Mar} + 5 \text{ Apr} - 30.333 \text{ May} - 64.667 \text{ Jun} - 43.333 \text{ Jul} - 68.333 \text{ Aug} - 51.333 \text{ Sep} - 33.667 \text{ Oct} - 16.333 \text{ Nov}$, the trend coefficient represents an estimated increase of about 1.05 sales units per month while holding seasonal position constant. That is approximately 12.65 units per year. If observation one is January, period thirty-seven is January of year four. Its forecast is approximately $286.33 + 1.05444(37) + 24.333$, or 349.7 units. February at period thirty-eight is about 348.7 units, March at period thirty-nine is about 351.5 units, and December at period forty-eight is about 336.9 units because all monthly dummies equal zero. The calendar alignment must be confirmed before these values are used operationally. This is a deterministic trend-plus-seasonality model, so its usefulness depends on stable seasonal relationships across the forecast horizon (Montgomery et al., 2015).

Model Validation and Forecast Uncertainty

A lower in-sample mean squared error makes the trend-and-seasonal model fit the thirty-six observed months more closely, but adding a predictor often improves historical fit even when future forecasting does not improve. A stronger evaluation would estimate both models on the first twenty-four months and compare their predictions for months twenty-five through thirty-six using mean absolute error, root mean squared error, and, where appropriate, mean absolute percentage error (Hyndman & Athanasopoulos, 2021; Makridakis et al., 2022). Rolling-origin evaluation can create additional tests by repeatedly extending the training sample one month and forecasting the next. The farm should also compare the regression against simple benchmarks such as the seasonal naïve forecast, which sets each month equal to the same month in the preceding year, and Holt-Winters exponential smoothing, which estimates level, trend, and seasonality. A more complex model is worthwhile only when it produces consistently better predictions on observations that were not used to estimate its

parameters.

Point forecasts should be accompanied by forecast intervals because future demand is affected by random variation, model uncertainty, pricing, weather, customer changes, competitor behavior, seasonality shifts, and operational constraints. A forecast of 350 units is not a claim that exactly 350 units will be sold. Management should plan around a central estimate while recognizing a plausible range of outcomes. This matters especially for perishable inventory, where overproduction creates waste and underproduction creates lost sales. Forecast errors should also be separated from supply constraints. If customers wanted more product but the farm lacked milk, refrigeration, labor, or transport, recorded sales understate demand. Stockouts, unfilled orders, returns, and unusual discounts should therefore be tracked alongside completed sales. When sales are measured as revenue, the farm should model physical volume separately and apply expected prices afterward, because inflation or price changes can create an apparent trend that does not represent growing customer demand.

Using the Forecast in Farm Operations

The value of the forecast lies in the decisions it supports. High-season estimates can guide orders for packaging, feed, veterinary supplies, cold-storage capacity, transport, and temporary labor, while lower-demand months may provide better opportunities for maintenance or planned downtime. Biological production should not be adjusted mechanically to one forecast because herd size, lactation cycles, animal health, feed quality, and weather operate on longer time horizons than monthly sales. The farm should compare a demand forecast with a separate supply forecast and use scenarios when demand could exceed expected production. Cash-flow planning is equally important. Growing sales may require greater working capital because wages, feed, packaging, fuel, and utilities are often paid before customers settle invoices. A monthly forecast should therefore connect sales expectations with receipts, expenses, credit terms, inventory needs, and emergency reserves. The objective is coordinated planning across production and finance rather than producing one impressive spreadsheet number in isolation.

Data Governance and Monthly Updating

A forecasting system becomes useful when it is updated consistently. After each month, the farm should record actual sales, calculate forecast error, document exceptional events, and refresh the model. Persistent underforecasting may indicate that demand is growing faster than expected, while persistent overforecasting can cause waste and cash strain. The business should retain previous forecasts rather than overwriting them so that management can evaluate whether forecasting quality actually improves. The published assignment does not reproduce the complete thirty-six-month dataset, the regression output, coefficient standard errors, full MSE values, or the complete forecast table, so the reported coefficients cannot be independently recalculated from the article alone. A

complete working file should preserve all observations, formulas, definitions, and model settings. Transparent documentation is especially important in a family business because future decisions may be made by people who were not involved in building the first model and need to understand how each result was produced.

Conclusion

The James and Sara Dairy Farm series contains clear monthly seasonality and an apparent upward trend, making seasonal dummy regression with a time variable a reasonable forecasting framework. The monthly indicators show how each month differs from December, while the trend term allows the forecast to rise gradually across years. The reported trend-and-seasonal model produces a lower historical error and can generate forecasts for periods thirty-seven through forty-eight, but historical fit alone is not sufficient evidence of forecasting superiority. The farm should validate competing models on unseen months, compare regression with simple benchmarks, inspect residuals, and use forecast intervals rather than relying only on point estimates. Sales data should also be separated from price effects and capacity constraints so that the model represents demand as accurately as possible. When the forecasts are connected with herd planning, feed, labor, processing, cold storage, transport, product mix, and cash flow, the analysis becomes a repeatable management process rather than a one-time statistical exercise.

References

- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2022). The M5 accuracy competition: Results, findings, and conclusions. *International Journal of Forecasting*, 38(4), 1346–1364.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting* (2nd ed.). Wiley.
- Newbold, P., Carlson, W. L., & Thorne, B. M. (2013). *Statistics for business and economics* (8th ed.). Pearson.