Kinaxis Case Study: Forecasting the Unforeseeable

Rohan Chintalapati, Sindhoora Prakash

Mechanical & Industrial Engineering, University of Massachusetts Amherst

MIE 697SC: Supply Chain Management

Dr. Michael Prokle

December 13, 2022

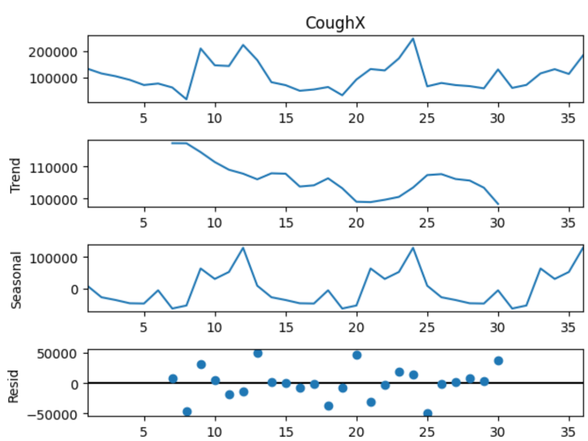
Executive Summary

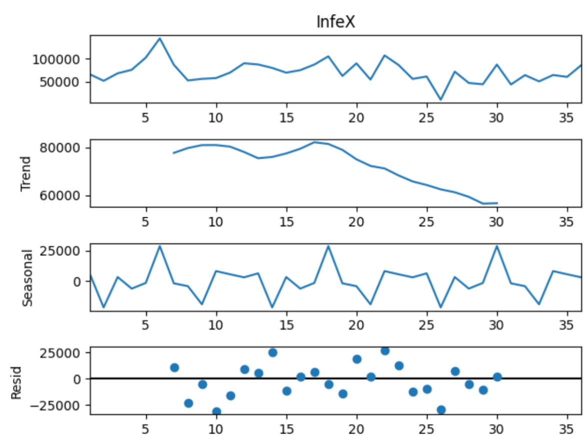
L'Apothicaire produces the products CoughX and InfeX, which are composed of ingredients that need to be carefully stored, and their suppliers require advance notice of demand to prepare for production. Due to these constraints, L'Apothicaire needs to be able to precisely determine their demand (and eventually, their production plan) to ensure that they are able to meet all stakeholder requirements.

The short-term goal for this project is to be able to predict the demand for 2019 and subsequently, 2020 for both products. L'Apothicaire would also like to understand how to make predictions for 2021 and beyond, to ensure that unexpected changes in demand can be anticipated and hence accommodated.

The results from the forecasts for 2019 and 2020 suggest that there could be improvements in the modeling approach as well as some of the operational decisions in L'Apothicaire w.r.t. supplier lead times and production planning time periods.

Demand Pattern





The demand for CoughX has a prominent pattern with a dip in the end of Summer, followed by two peaks in Fall and Winter, and no visible trend at first glance. However, plotting the seasonal decomposition of the 2017-2018 demand shows that there is a slight decreasing trend along with strong seasonality.

On the other hand, the demand for InfeX does not show any noticeable seasonal patterns nor does it have visible trend. There is a spike in mid-2017 with extremely high demand when compared to the rest of the 2017-2018 data. The seasonal decomposition plot of the demand shows that there is weak seasonality, with a dip at the end of Summer and peaks roughly every six months.

Methodology

Part 1

The dataset given has historical demand for 2017 and 2018 (24 data points in total). Since the dataset is very small, it is not possible to use sophisticated statistical models such as ARIMA or SARIMA for either product. So, simpler techniques such as moving averages, weighted moving averages, or regression models should be used to predict the demand for the next 12 months.

Since exponential smoothing can be utilized for both small datasets as well as datasets that do not show significant seasonality, it has been chosen for the first part of the project. No specific parameters were chosen for the ExponentialSmoothing function in the statsmodels Python package apart from multiplicative decomposition (as the initial 2-year dataset had lower residuals from multiplicative decomposition), and the default values were applied to all other parameters.

As a control, a regression model has also been trained using lag values ranging from 1 to 12 as the predictors (features filtered based on correlation >= 0.5 with the current month's demand).

For each product, the data has been split for training and testing in an 80:20 ratio, and both models were evaluated through the mean absolute percentage error. MAPE has been chosen as the metric because the scale of the demand is very high and RMSE may not be the easiest to interpret.

The regression models for CoughX and InfeX had MAPE of 13.31% and 26.51% respectively, and the exponential smoothing models for the same had 11.95% and 26.09%. Since the exponential smoothing forecasts in for the test sets had slightly better MAPE values, the same model was fitted with the complete dataset to predict 2019 demand for both products.

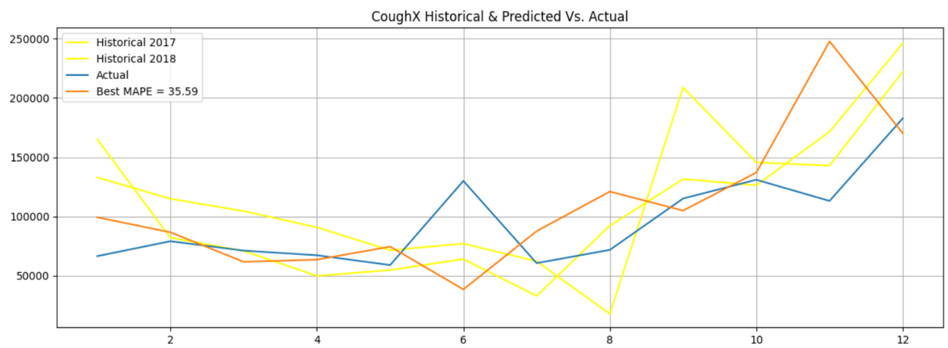
Part 2

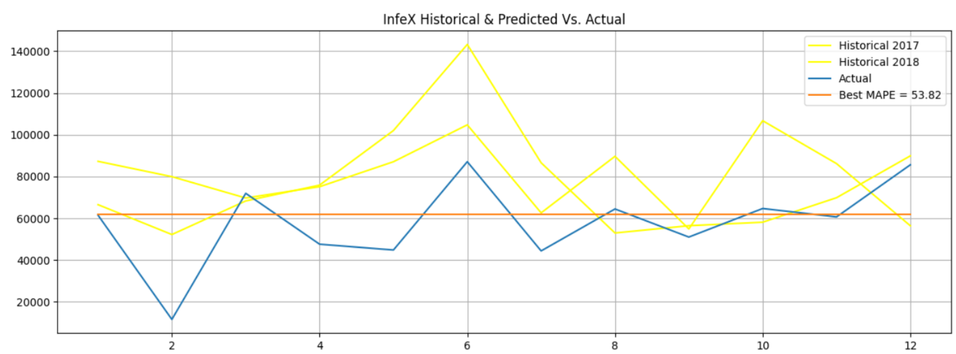
Comparing the forecasts made in step 1 using exponential smoothing with the actual 2019 data, the MAPEs were 37.06% and 71.23% for CoughX and InfeX respectively. This shows that the model (or the default parameters used) did not explain the true patterns in demand and was overfitting the small training data that was available.

In the second step, the methodology was improved by implementing the following:

- Removed (and replaced) outliers in InfeX before fitting the models.
- Performed grid search through a manually chosen subset of the target parameter space.
- Switched from 80:20 train:test ratio to 66:33, so that 2 full cycles of the data can be used for training and 1 full cycle for testing.

The grid search combinations of trend, damped, seasonal, use_boxcox, remove_bias, smoothing_level for the model for CoughX were strictly assuming that the seasonal and trend parameters could not be ignored. The same for InfeX were created assuming that there is no seasonality in the data, so the combinations were varying use_boxcox, remove_bias, smoothing_level. The best models after performing grid search achieved MAPE of 35.59% and 53.82% for CoughX and InfeX respectively.





Best Parameters:

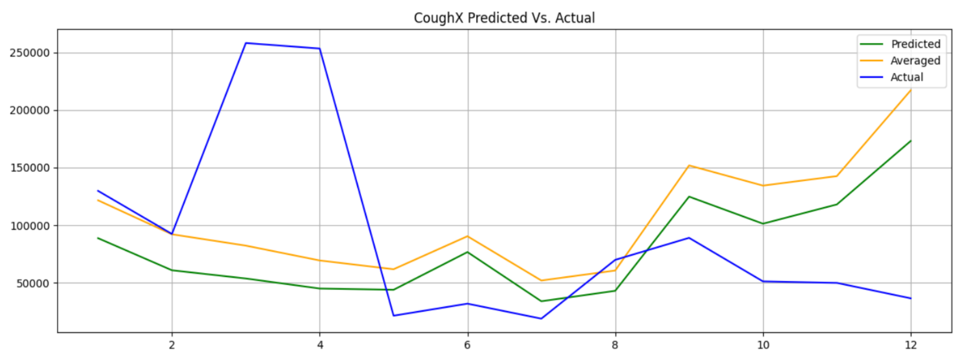
Product	trend	damped	seasonal	use_boxcox	remove_bias	smoothing_level
CoughX	mul	False	mul	False	False	0.8
InfeX	None	-	None	False	False	0.8

In Step 1, an assumption was made that complex models such as ARIMA would not perform well when the dataset provided is small. But since the MAPE values were high with the first model, the ARIMA model was also explored for both products. The MAPEs achieved were much higher than the exponential smoothing models (CoughX: 40.87%, InfeX: 94.93%), so this model was not used for demand predictions for 2020.

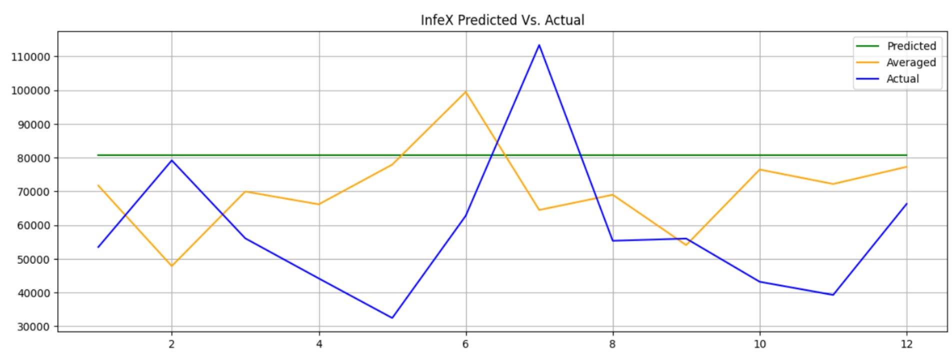
A simple average of the demand in the same period in the previous years was also calculated as a control for the new models explored in part 2. While this performed better for 2019 demand than the exponential smoothing for CoughX (MAPE 32.75%), it was not used for 2020 because there is a downward trend in the data in the 3 years, which a simple average can not sensibly account for.

Precision of Forecasts

Both the exponential smoothing and the averaged values over the same period each year were not able to forecast the CoughX demand for 2020 accurately. The MAPE values (135% for averaged values, 103% for exponential smoothing) were exorbitantly high due to the spike in demand in early 2020, which can be attributed to the panic buying associated with the announcement of COVID-19 lockdowns.



On the other hand, the MAPE values for InfeX demand forecasts (49.59% for averaged values, 57.53% for exponential smoothing) were better than the 2019 forecasts. It must be noted that InfeX is a medicine used for bacterial infections, and hence there are no COVID-19 related spikes in demand. Since there was more representative data (2 vs. 3 full seasonal cycles) to fit the models, the results in part 2 were better than that in part 1 for InfeX.



Effect on Supply Chain

The demand for CoughX in the first half of 2020 was more than twice the value that was forecasted. Since the orders to suppliers had to be made in advance, there was no way to quickly modify the production plan to match demand and backorders in that period. The latter half of 2020 saw the opposite, where due to COVID-19 lockdowns and increased use of masks in public areas, the demand for general cough medicine itself had come down. This means that in case L'Apothicaire followed their initial production plan, they would be stuck with an excess in inventory, which is very expensive.

Hence, it is clear that they should try to choose suppliers that have shorter lead times so that the forecasts need not be made at the beginning of each year. They should give importance to seasonality but at the same time, be vigilant of any oncoming global health crises such as the COVID-19 pandemic, as pharmaceuticals will be one of the first industries to be impacted.

Future Work

An easy way to improve forecast accuracy is by obtaining more observations for the models to learn from. L'Apothicaire could provide data at a more granular level and also plan its production in smaller time periods so that the model can pick up changes in demand patterns quickly and the production plans can be adjusted accordingly.

The methodology followed in this project focused purely on the univariate time series (i.e past demand), however in reality there could be various other predictors that may be able to explain the demand better. A deep dive into the different actors in the pharmaceutical industry and the demographics of L'Apothicaire's customers may help in identifying these potential predictors.

In the future, the forecasting team shouldn't blindly rely on the values generated by models but also review their methodology and forecasts with subject matter experts so that any foreseeable issues can be caught beforehand.