

Data Analysis Report

Main Insights for the Container Service Price Series The `CONTAINER_SERVICE_PRICE` dataset contains daily service prices associated with container handling, with `APPOINTMENT_CREATION_DATE` serving as the timestamp for each recorded price. The series comprises 17040 non-null entries after data cleaning, covering the time period from January 2, 2024, to July 31, 2024. The values range from 410 to 1169, with various distinct prices and some outliers that were initially present, such as `boorra`, being filtered out.

Seasonal decomposition of the time series reveals the following insights (see Figure 1 for the decomposition plot):

- 1. **Observed Values:** The average service prices vary daily, showing fluctuations with no consistent increasing or decreasing trend over the observed period.
- 2. **Trend:** The trend analysis shows no strong upward or downward movement, indicating a relatively stable average price for container services, with minor short-term variations.
- 3. **Seasonality:** The seasonality component exhibits small cyclic patterns without a dominant seasonal effect.

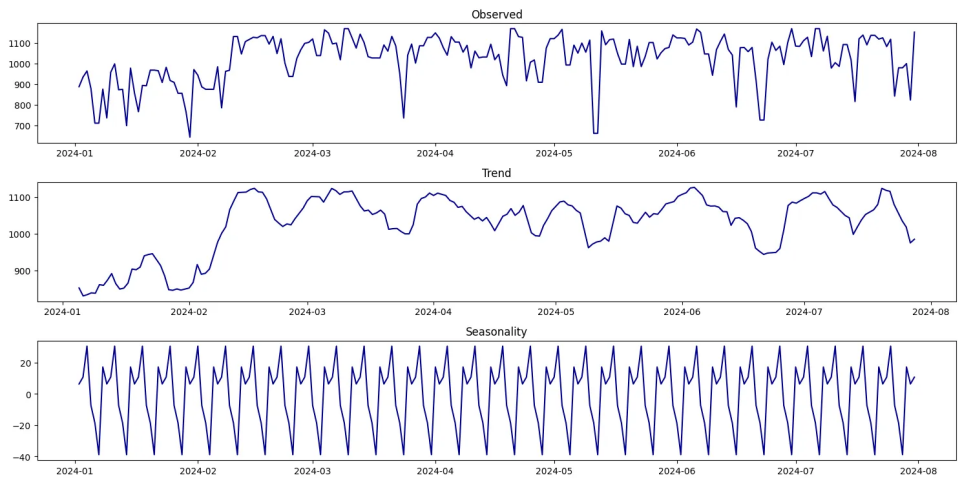


Figure 1. Container service price series decomposition plot.

Time series visualization reveals fluctuations throughout the observed period, with several peaks and valleys, but no distinct seasonal pattern is evident. The `APPOINTMENT_CREATION_DATE` factor was analyzed on a monthly and yearly basis, indicating more consistent activity in certain months without extreme deviations in service price occurrences.

The dataset also underwent time series augmentation, where the missing dates were filled using forward-filling, ensuring consistency and avoiding data gaps.

Experimentation Tracking Report

The experimentation process involved using Bayesian inference and a Neural Network approach to forecast `CONTAINER_SERVICE_PRICE` values. Here are the key metrics and results for the Bayesian inference model used during backtesting:

- **Refit Value:** The Bayesian model was refit iteratively, increasing the training window by 15 days in each iteration. The average refit value was 184 days, allowing for a larger training sample as the experiment progressed.

- **Execution Time:** The average execution time per refit was approximately 0.00061 seconds, indicating a fast computation rate for each refitting iteration.
- **Score Metric (MAPE):** The mean absolute percentage error (MAPE) averaged at 8.77%. This score suggests that the model captures the overall trends in the price well, but there remains some variance from the true values, possibly due to fluctuations without apparent seasonality.
- **Degradation Metric:** The degradation metric, which indicates the stability of the forecast accuracy, averaged at 6.67. This indicates that the number of days where the forecast remained within an acceptable deviation threshold was relatively consistent across iterations.

Figure 2 shows the results of the Bayesian forecasting, comparing the predicted values with the actual service prices.

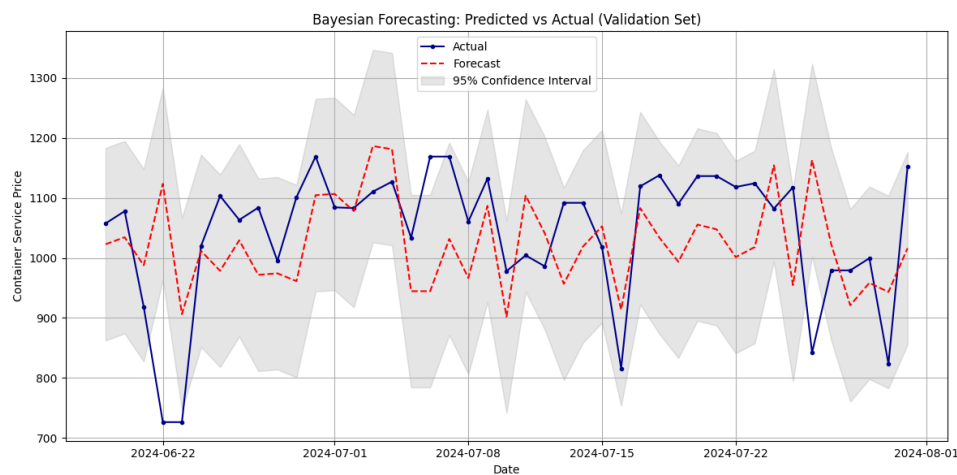


Figure 2. Predicted series using Bayesian Forecasting.

For the neural network model, MAPE after training was observed to be 8.20%, which suggests improved accuracy compared to the Bayesian model, with the degradation metric showing a change of 3.32%, indicating slight performance decay over time. The results of the Convolutional Neural Network forecasting are depicted in Figure 3, showing the predicted versus actual values.

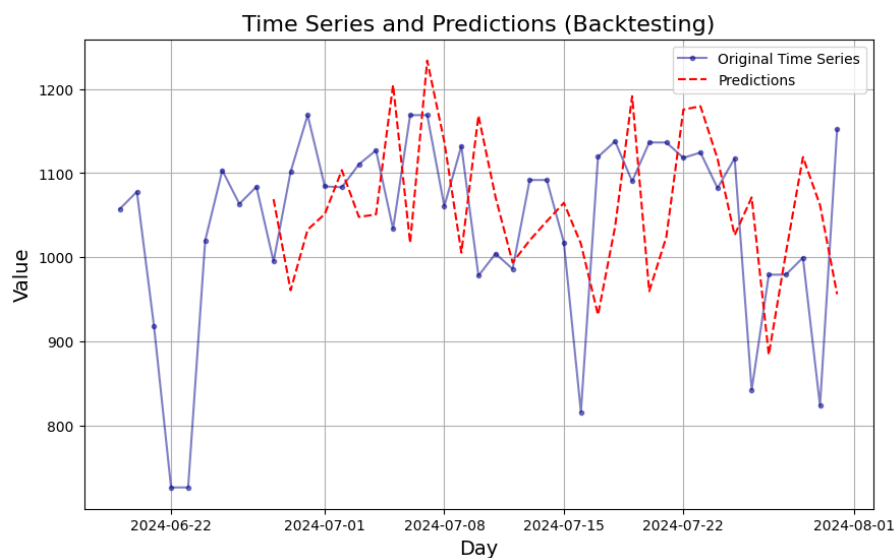


Figure 3. Predicted series using Neural Network Forecasting.