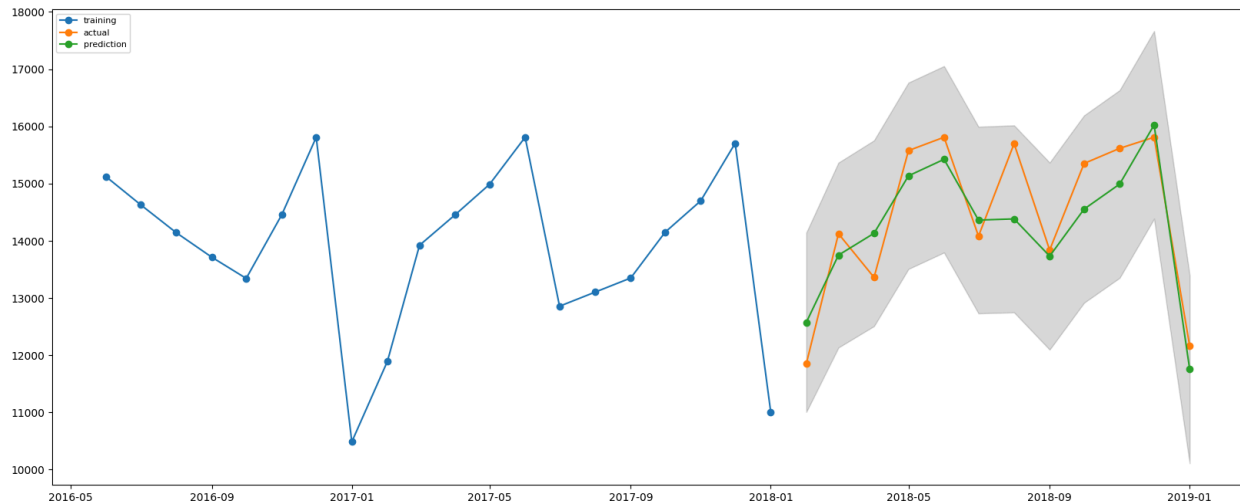# Time Series Analysis Lecture — 4

Confidence Intervals in Time Series Forecasting

- **Definition:** Range [A, B] indicating where the future value is likely to fall, considering forecast error and unpredictability.
- **Purpose:** Reflects uncertainty in predictions, providing a probable range rather than a precise value.
- **Importance:** Helps in assessing the reliability of forecasts, guiding decision-making with an expected range of outcomes.
- **Calculation:** Varies by model; in statsmodels, specified by 'alpha' parameter (e.g., 0.05 for 95% confidence).
- **Interpretation:** For a forecast ŷt with alpha = 0.05, the CI (m, n) implies a 95% probability the true value will fall within this range.
- **Visualization:** Plots can show actual values, forecasted values, and the CI, illustrating the forecast's accuracy and uncertainty.
- **Application Example:** Using statsmodels to get forecast CIs, illustrating how actual observations align with these intervals.
- **Adjusting Confidence Levels:** Varying alpha adjusts the CI's width, balancing between confidence and precision.
- **Instructor Note:**
  - Emphasize visual comparison between actual and forecasted values with CI ranges.
  - Highlight how adjusting alpha affects CI width, impacting decision confidence.

Exogenous Variables in Time Series Forecasting

- **Definition:** External factors not influenced by the system but affecting the output. Examples include holidays, product launches, weather conditions.
- **Impact on Predictions:**
  - Models like SARIMA, focusing on historical data, may miss out on patterns affected by these variables.
  - Incorporating exogenous variables can improve forecast accuracy, especially for events like holidays.
- **Example Analysis:**
  - Predictions vs. Actuals: Discrepancies during holidays indicate the model's inability to account for external impacts without exogenous inputs.
- **Performance Metrics:** MAE, RMSE, and MAPE highlight the overall prediction accuracy, with notable deviations during exogenous events.
- **Solution:** Integrate exogenous variables (like holiday flags) into models to better capture special event impacts.
- **Instructor Note:** Demonstrating the significance of exogenous variables through visual plots and performance evaluation enhances understanding of their role in accurate forecasting.

SARIMAX Model Overview

- **SARIMAX Model:** Extends SARIMA by incorporating exogenous variables (denoted by 'X'), allowing for a more detailed and accurate forecast by including external influences.
- **Key Features:**
  - **Incorporation of Exogenous Variables:** Adds external factors such as holidays into the forecast model, assigning them weights learned during model training.
  - **Hyperparameters:** Retains SARIMA's p, d, q, P, D, Q, s, with the addition of specifying exogenous variables through the exog parameter.
  - **Model Training and Forecasting:** Trains with both endogenous (time series) and exogenous (external) data to predict future values, adjusting for known external impacts.
- **Application:**
  - Demonstrated improvement in forecast accuracy when accounting for holidays as an exogenous variable, reflected in performance metrics (MAE, RMSE, MAPE).
  - The model better aligns predictions with actual peaks on holidays, showcasing the value of integrating relevant external factors.
- **Forecasting with SARIMAX:**
  - **Forecasting Process:** Involves fitting the model with historical data and exogenous variables, then predicting future values, potentially improving forecasts for periods with known external influences.

- **Visual Analysis:** Plots comparing actual vs. predicted values highlight the effectiveness of including exogenous factors. Red lines on plots indicate holidays, allowing visual assessment of forecast accuracy on these days.

- **Practical Considerations:**
  - **Performance Evaluation:** Enhanced model performance demonstrates the utility of exogenous variables in refining forecasts.
  - **Forecaster's Judgment:** Beyond metrics and models, the forecaster's expertise plays a crucial role in evaluating the plausibility and reliability of predictions.
- **Conclusion:**
  - SARIMAX offers a sophisticated approach to time series forecasting by integrating external factors, enabling more nuanced and potentially accurate predictions.
  - The model's effectiveness, especially in handling seasonality and external influences, underscores the importance of a comprehensive approach to forecasting.