

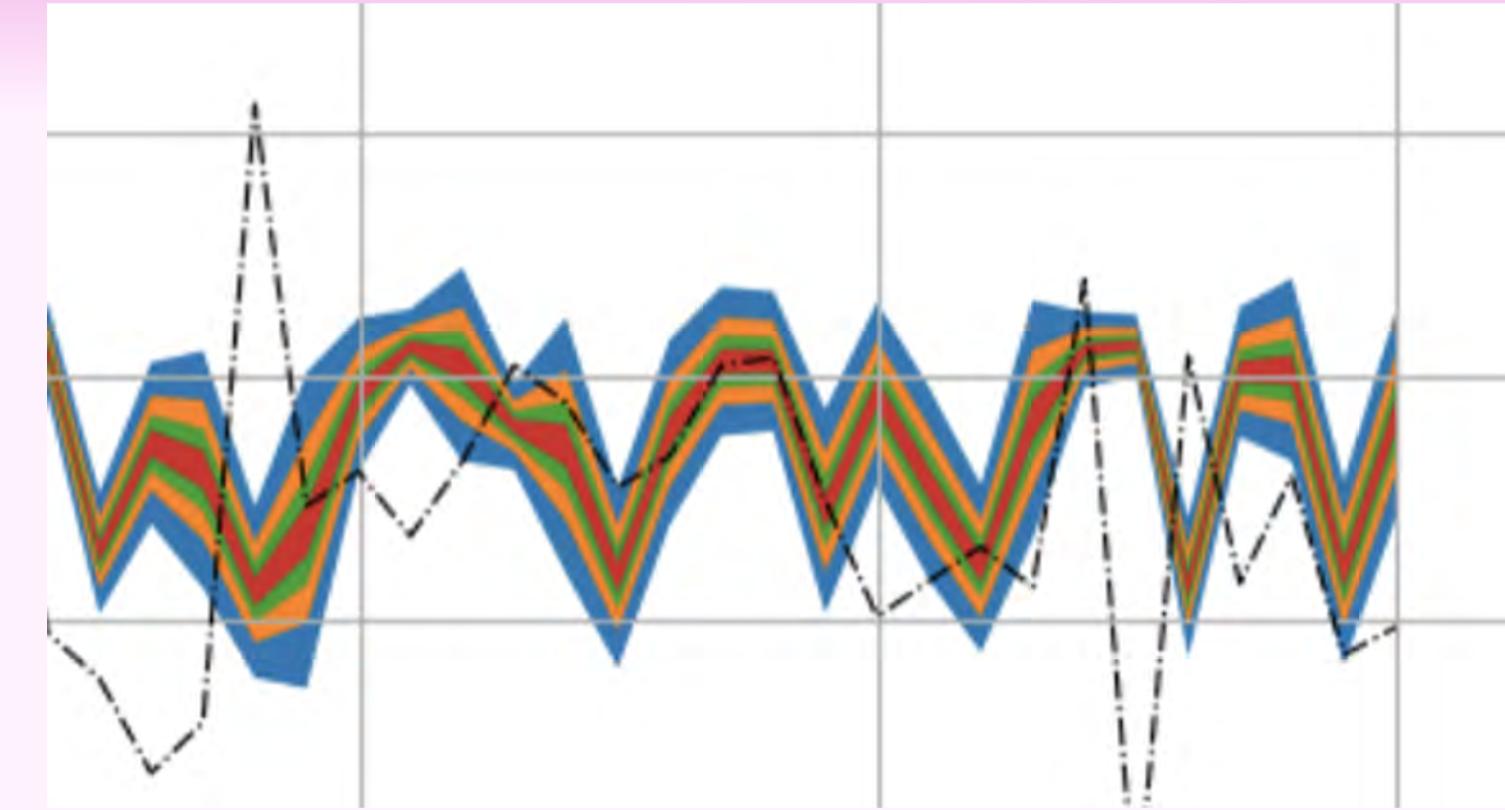
Advanced topics

INTRO

- Confidence intervals
 - confidence intervals from machine learning models
 - use cases
- Other techniques for sequence data
 - Prophet, deep learning, DTW,



Confidence intervals

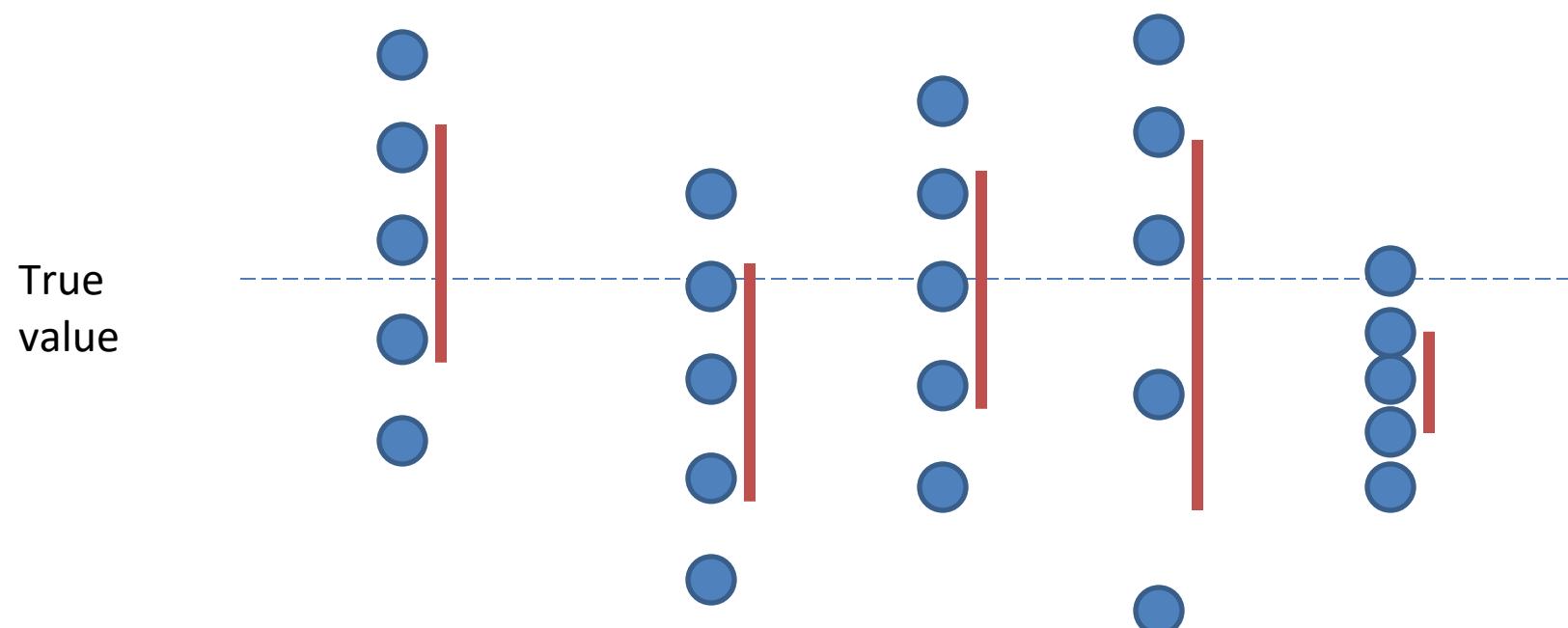


Confidence interval depicts the confidence of our prediction.

Confidence interval

“80% confidence interval”

If similar experiments are conducted 100 times,
in 80 cases the true value will lie within each
experiment's calculated interval



Confidence interval

“80% confidence interval”

If similar experiments are conducted 100 times,
in 80 cases the true value will lie within each
experiment's calculated interval

80% chance that the true value will lie within
~~this range~~.

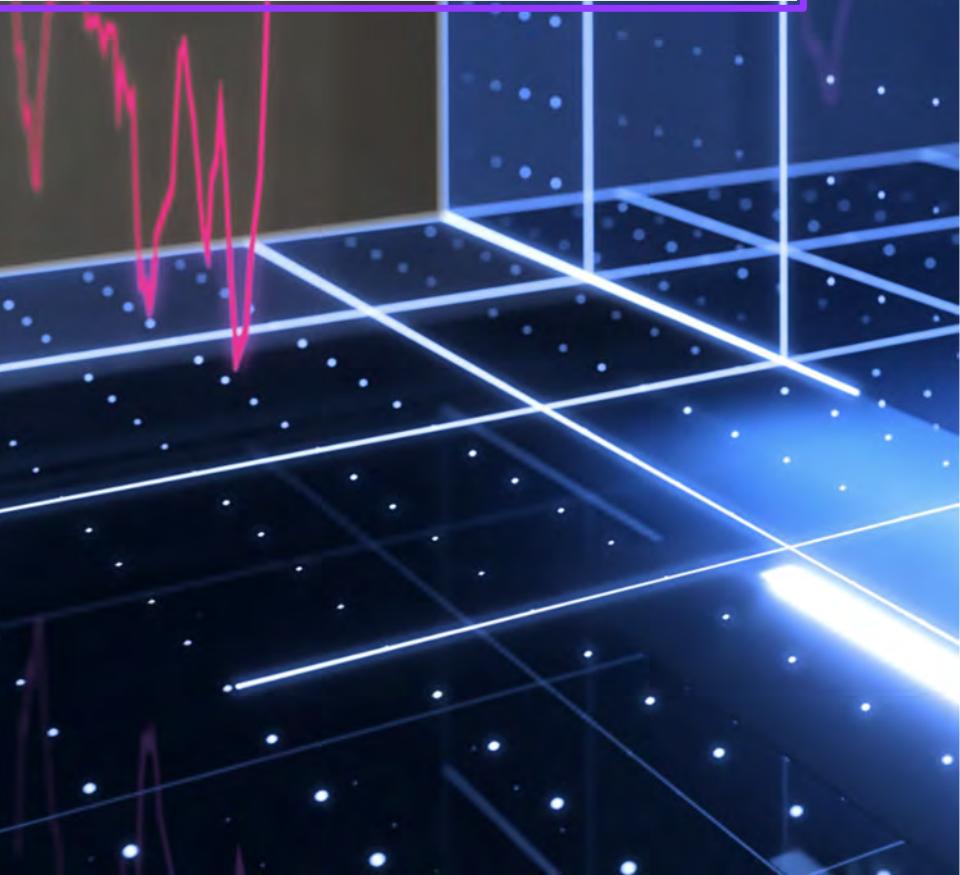


Confidence interval

“80% confidence interval”

If similar experiments are conducted 100 times,
in 80 cases the true value will lie within each
experiment's calculated interval

~~80% chance for another experiment's
prediction to be in this interval~~



Confidence interval calculation

Confidence interval is often of the form

$$(p - c \sigma, p + c \sigma)$$

where p is the predicted value, σ is the standard deviation, c is some constant

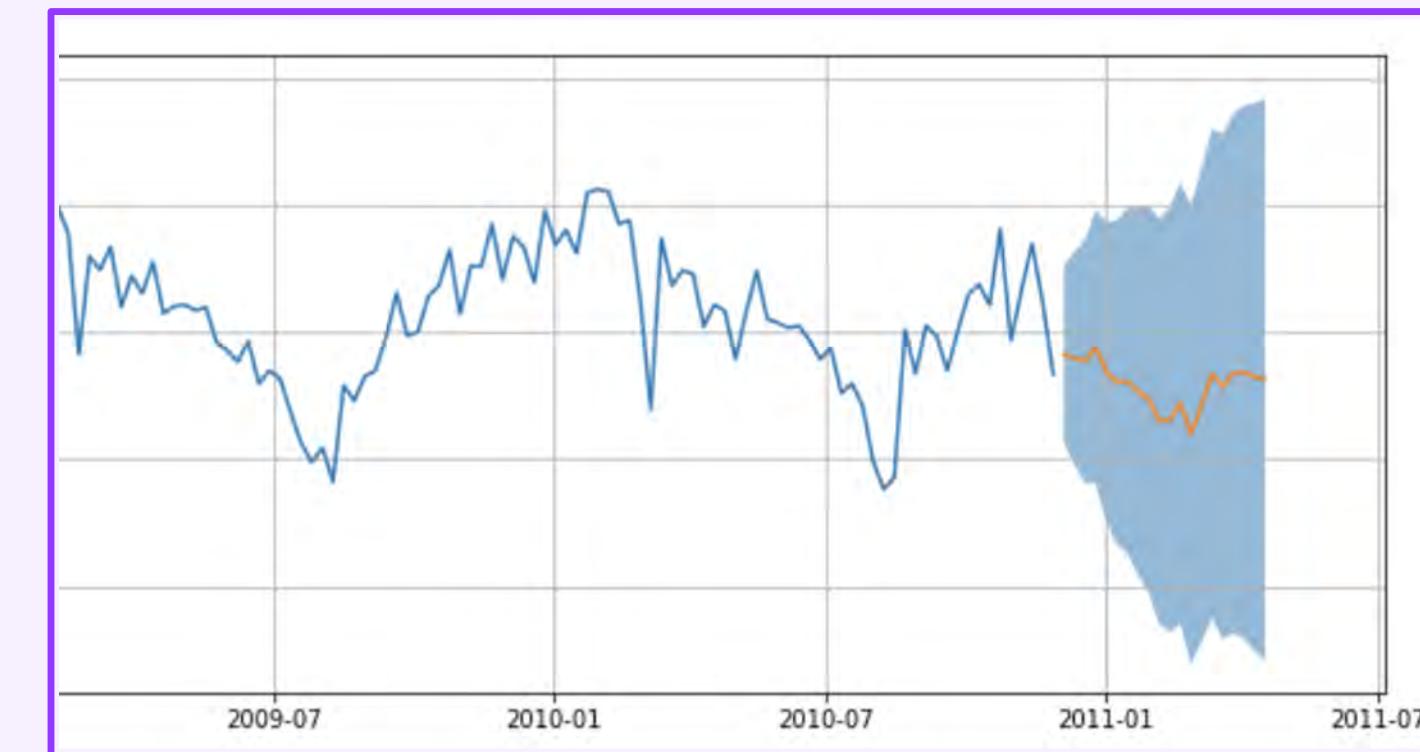
The exact form depends on the assumption about the underlying data or modeling used.

Implications: for real world use cases, confidence interval should be interpreted with care.



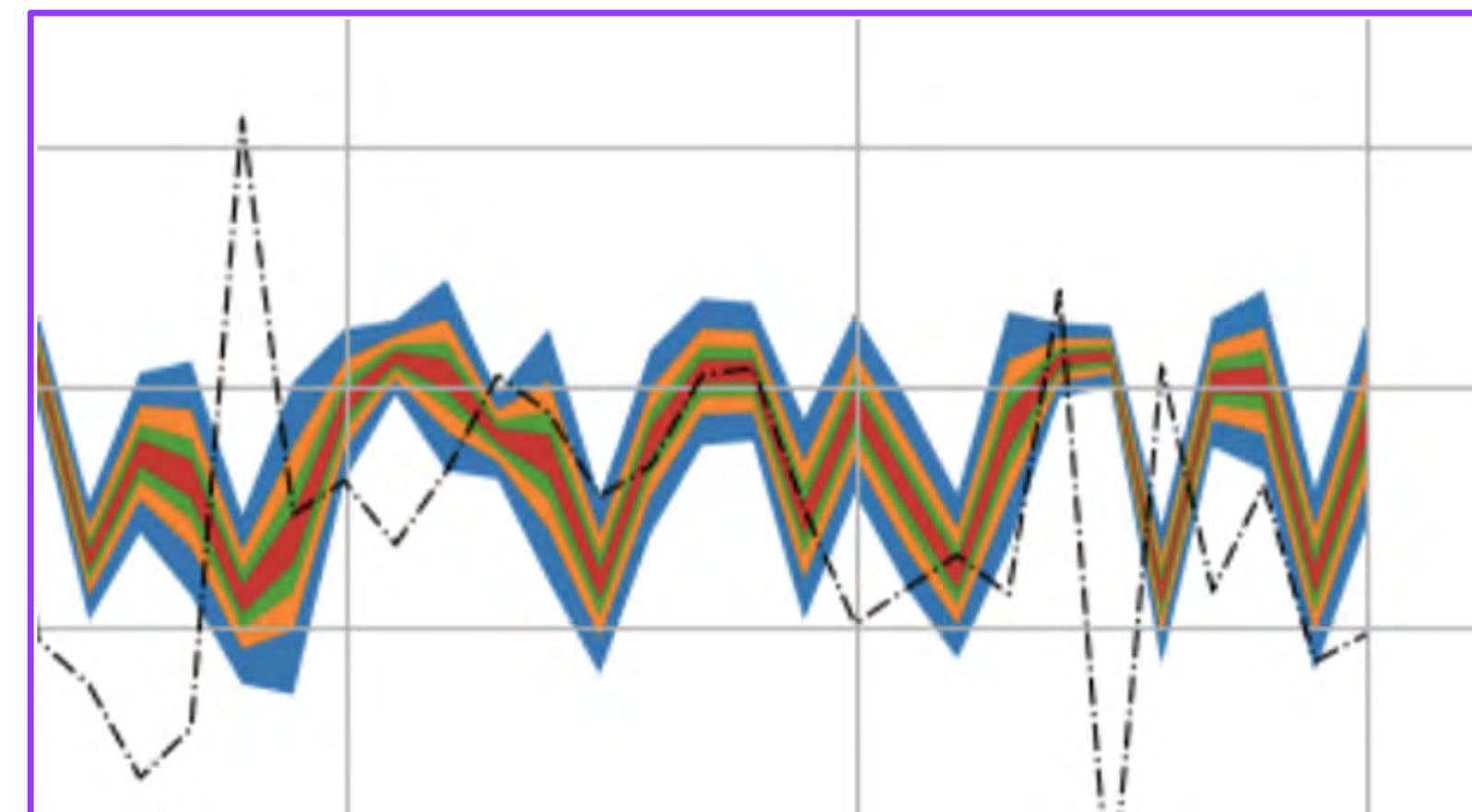
Confidence intervals in ARIMA models

- Assumes the error is a Gaussian distribution with 0 mean
- The standard deviation of the error widens as the forecast goes further in time
 - Errors accumulate

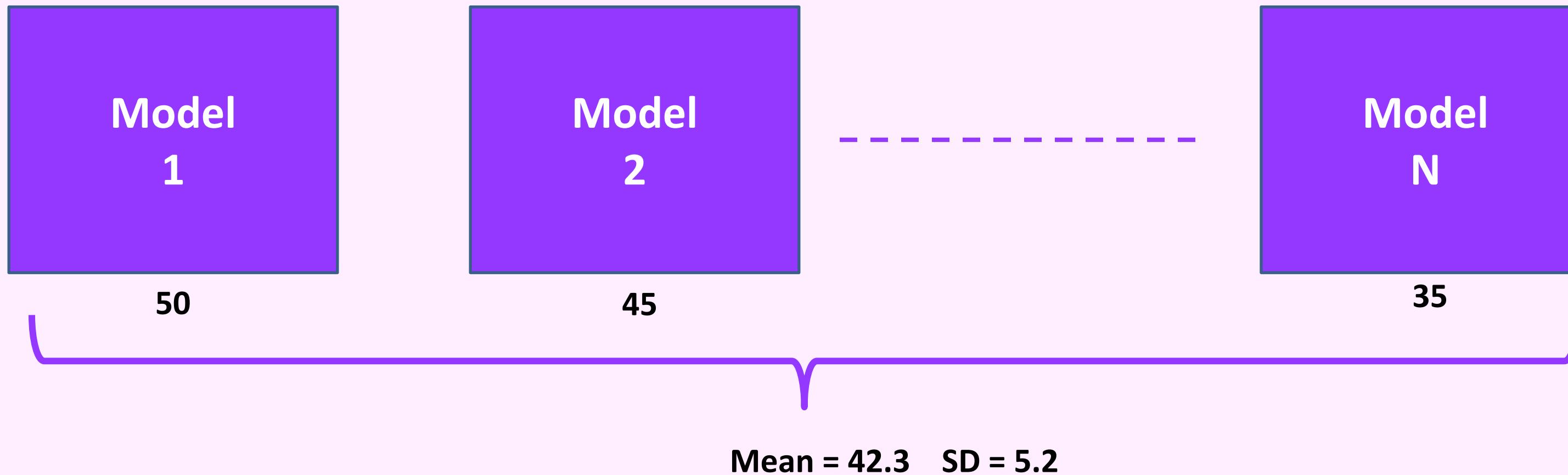


Confidence intervals for machine learning models

- Simple confidence intervals can also be estimated for machine learning models by finding the variance of the prediction.



Confidence from ensembles



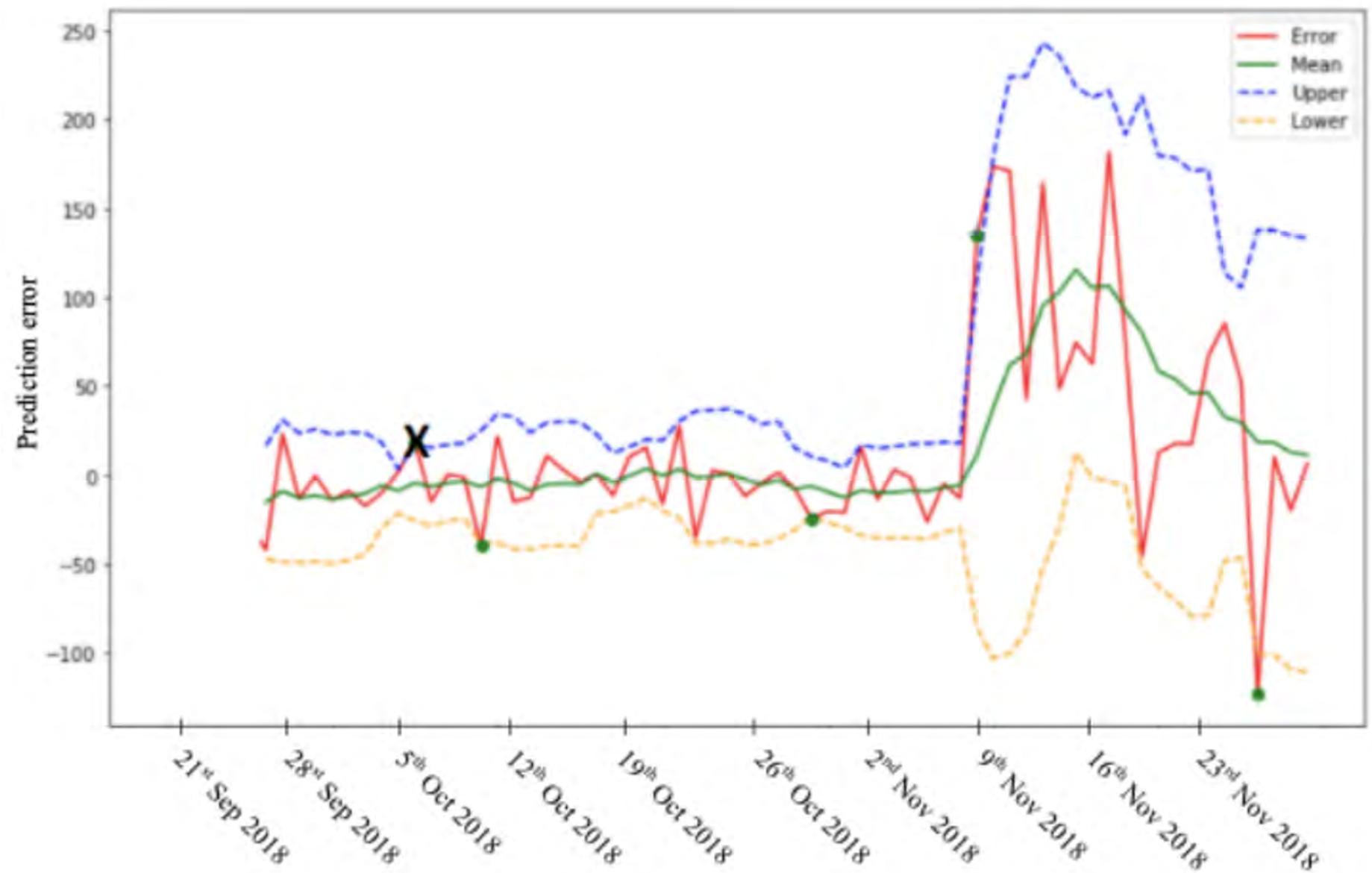
Then, assuming Gaussian we can find the range that captures 80% of the probability to get the 80% interval.

Anomaly detection with confidence interval

To monitor for unusual behavior, one can have a system that alerts when the actual value deviates from the confidence interval.



Case study from real estate website

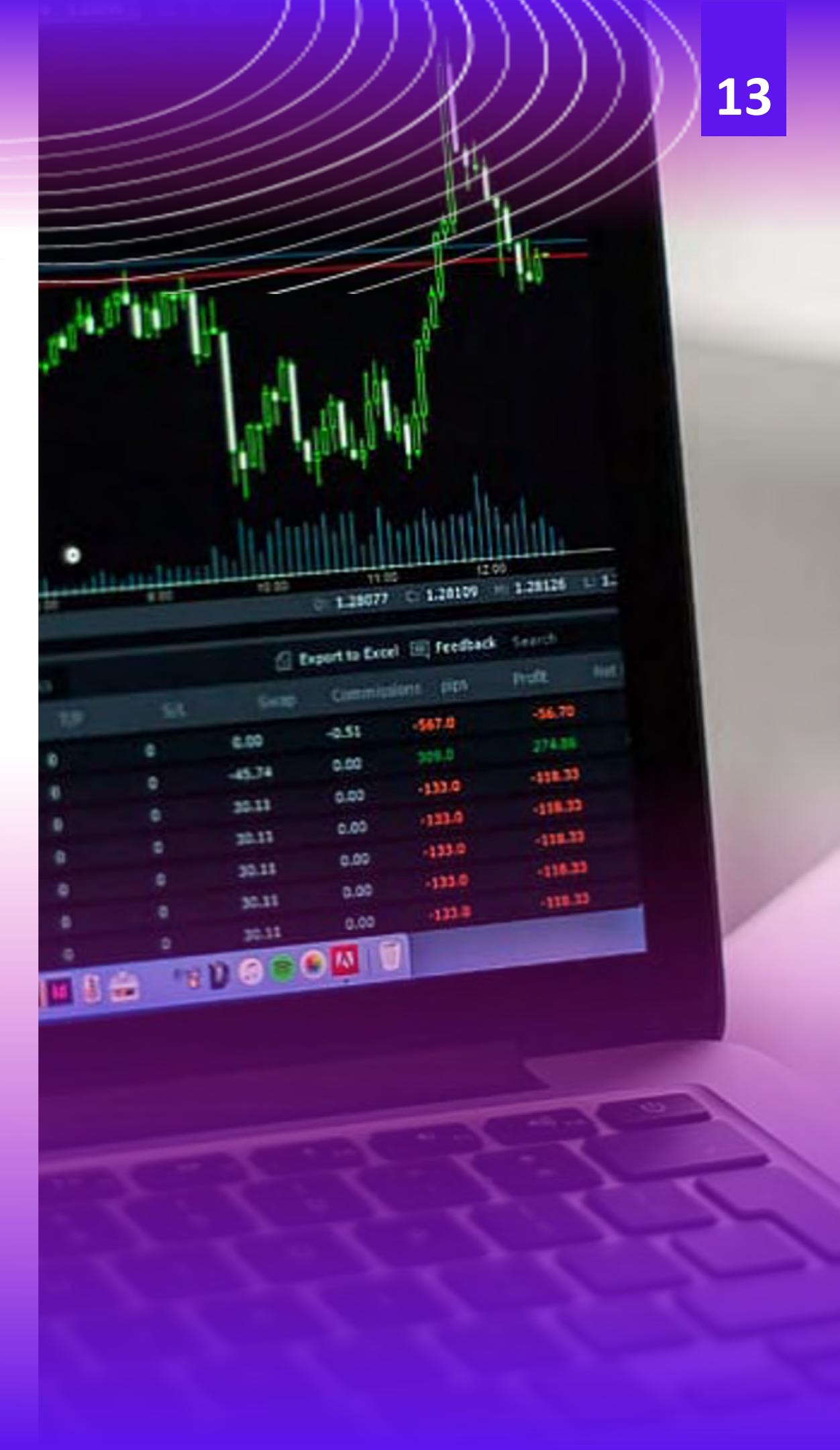


Choukuljaratsiri, Arcchaporn, et al. "Anomaly Detection for Online Visiting Traffic as a Real-Estate Indicator: The Case of HomeBuyer." *Behavioral Predictive Modeling in Economics*. Springer, Cham, 2021. 291-301.



Other methods for time series analysis

- Machine learning models
 - ❑ Support Vector regressor
 - ❑ Deep learning (LSTM, GRU)
- Pattern matching
- Facebook Prophet



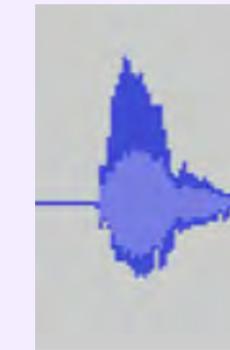
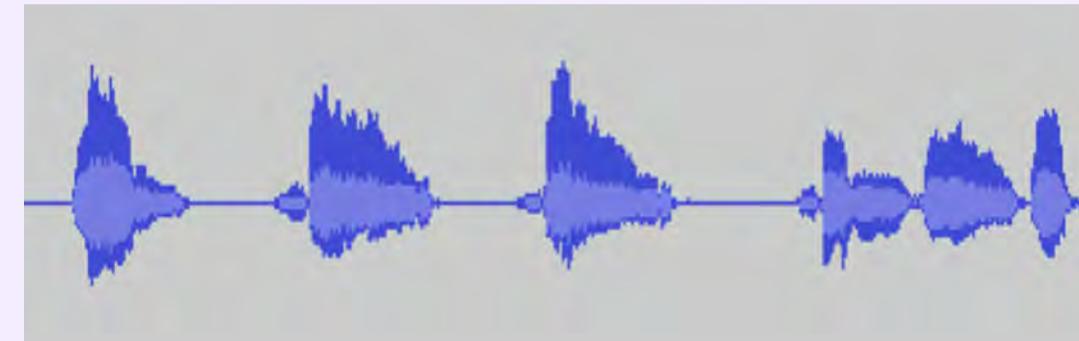
Pattern matching

Search for a similar pattern in the database

Crosscorrelation or Dynamic Time Warping (DTW) can help find similar regions

Can model with one example

Pencil and pens are on the table



pen

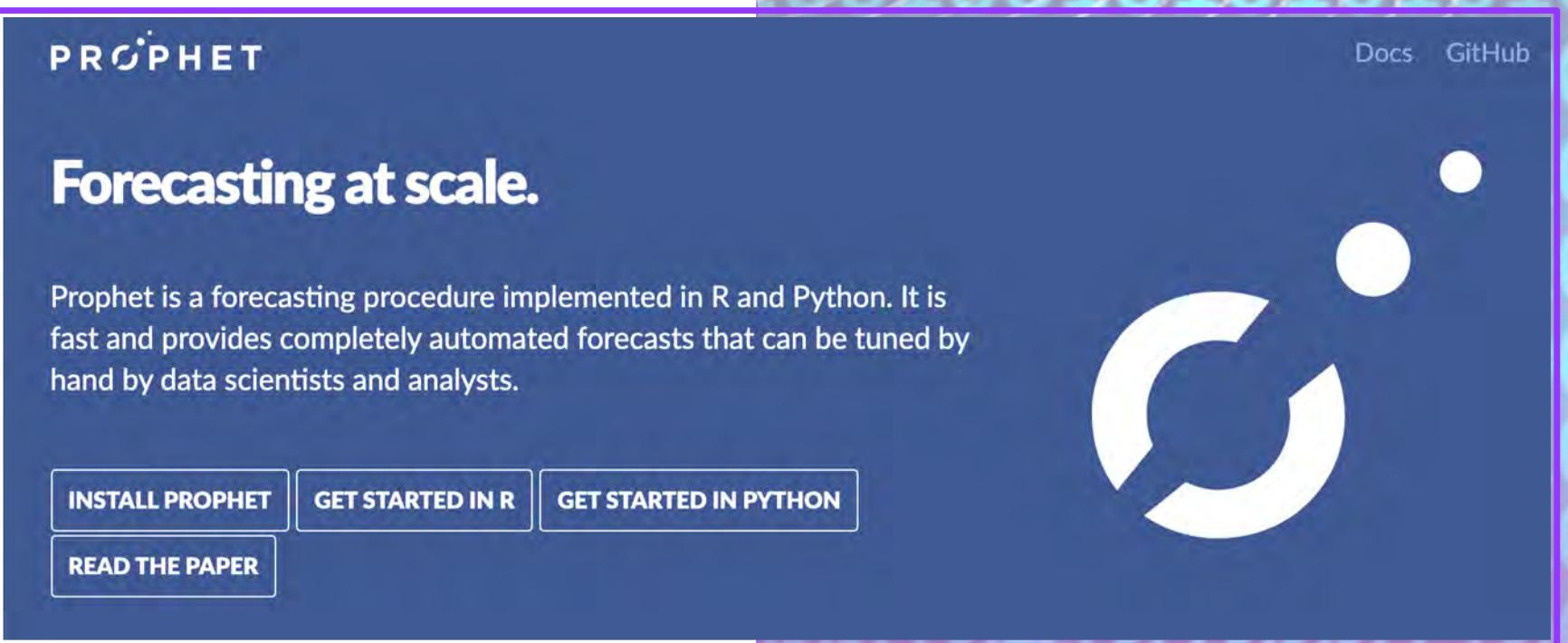
Facebook Prophet

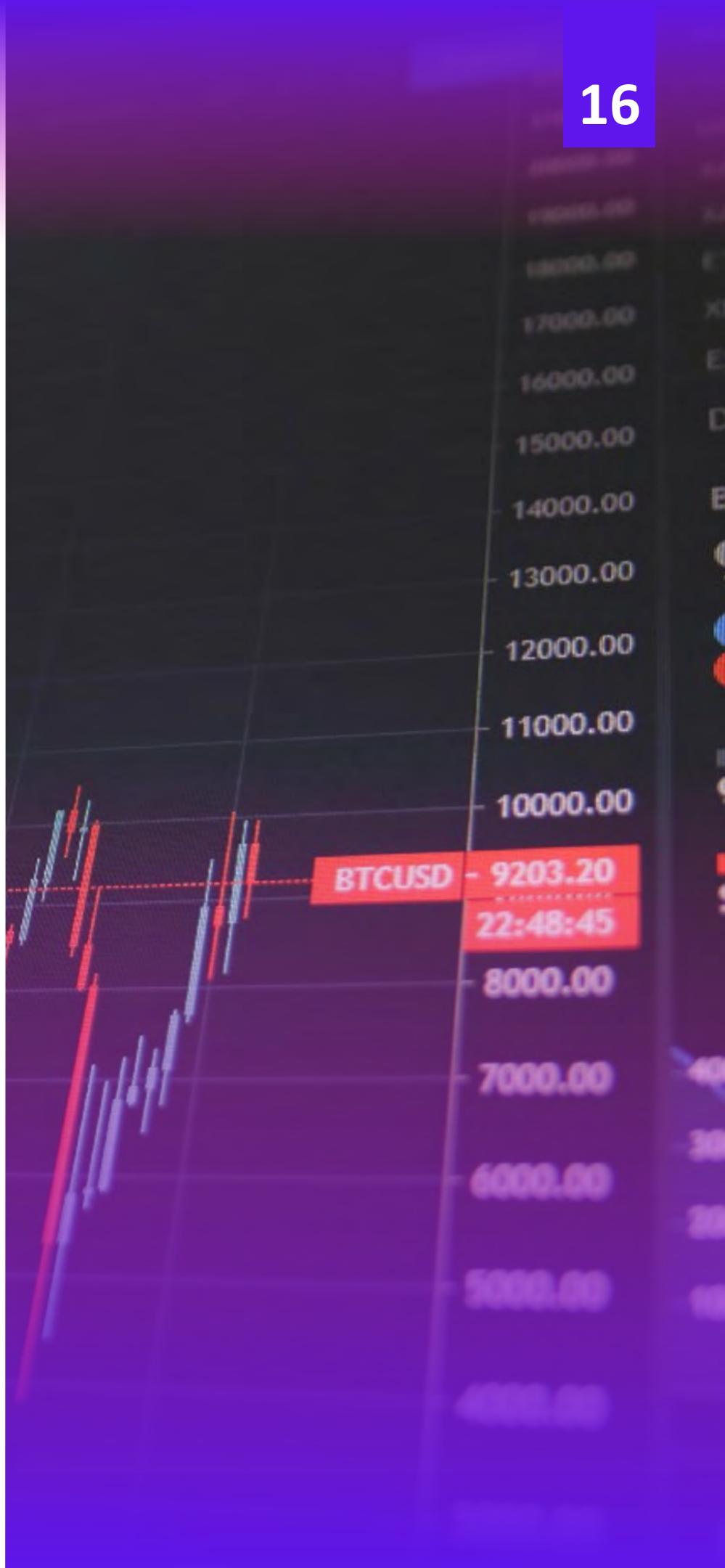
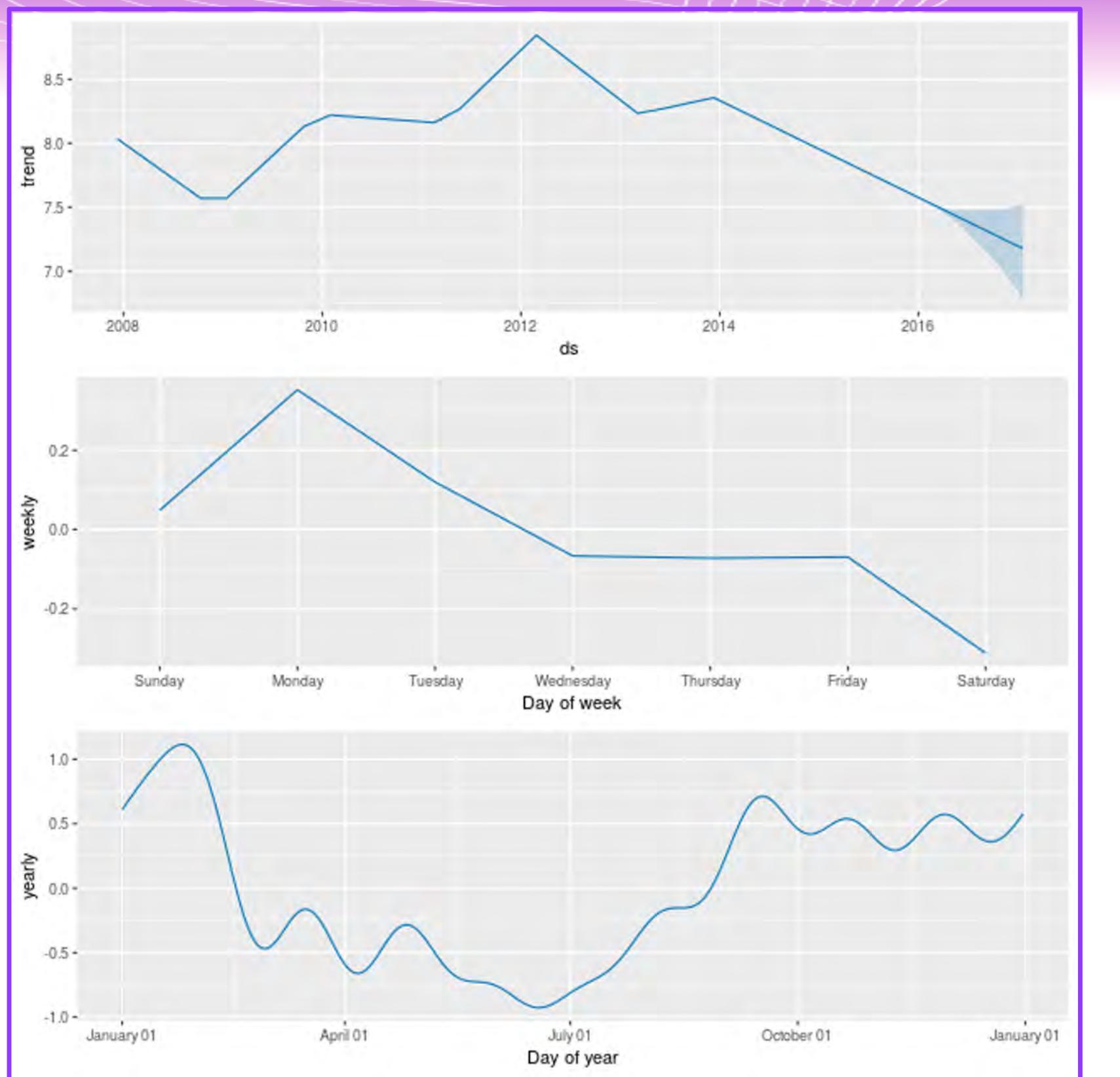
An extension of SARIMA models with additional features. Put humans in the loop to guide the model with external info

- Special events (H) can be modeled. Good for holidays, sales events, etc.
 - T + S + H + error
 - can cap the trend
 - maximal potential market share reached
- Easy to use, can decompose different seasonality effects

Tutorial

https://facebook.github.io/prophet/docs/quick_start.html#python-api





Conclusion

