



V 2.0.4

Jedox Alssisted™ Planning

CHEAT SHEET ON AISSISTED™ PLANNING SERVICES

Time Series Prediction

STATISTICAL



Linear Model

Data Needs: Trend & Seasonality

Speed & Quality: Fast speed; High quality (Highly data dependent)

Uses **trend and seasonality** to create a linear model for prediction



Holt Winters

Data Needs: Periodicity & Seasonality required, at least 2 periods of data

Speed & Quality: Fast speed; High quality (Highly data dependent)

Uses **trend, season** and level to predict Time Series Forecast; recommended for larger prediction ranges (> 1 year)



Seasonal Naive

Data Needs: Seasonality required

Speed & Quality: Fastest speed; Intermediate quality (Heavily data dependent)

Basic ARIMA modelling assuming **seasonality**



Exponential Smoothing

Data Needs: Periodicity, at least 2 periods of data

Speed & Quality: Intermediate speed; High quality (Highly data dependency)

Uses **trend** and **season** to predict Time Series Forecast; recommended for larger prediction ranges (> 1 year)



ARIMA

Data Needs: No specific data characteristics needed

Speed & Quality: Intermediate speed; Highest quality (Lowest data dependency)

Uses **trend, periodicity, seasonality**, cyclic nature of time to create optimized ARIMA model series; recommended for larger prediction ranges (> 1 year)



Random Walk with Drift

Data Needs: No specific data characteristics needed

Speed & Result Quality: Fastest speed; Intermediate quality (Heavily data dependent)

Uses **automated drift (error)** calculation

Time Series Prediction

INNOVATIVE



STL

Data Needs: Seasonality required

Speed & Quality: Intermediate speed;
High quality (Highly data dependent)

Decomposes **non-linear relationships** in season and trend, bias adjustment (Box Cox transformation); recommended for larger prediction ranges (> 1 year)



STLF

Data Needs: Seasonality required

Speed & Quality: Intermediate speed;
High quality (Highly data dependent)

Decompose **non-linear relationships** in season (**automatically optimized**) and trend, bias adjustment (Box Cox transformation); recommended for larger prediction ranges (> 1 year)



TBATS

Data Needs: Seasonality required

Speed & Quality: Slow speed;
High quality (Highly data dependent)

Uses **multiple, shifted seasonalities**, bias adjustment (Box Cox transformation); recommended for larger prediction ranges (> 1 year)



Neural Network

Data Needs: No specific data characteristics needed

Speed & Quality: Slow speed;
Highest quality (Lowest data dependency)

Univariate **Neural Network model** (layers, weights), bias adjustment



Croston's Method

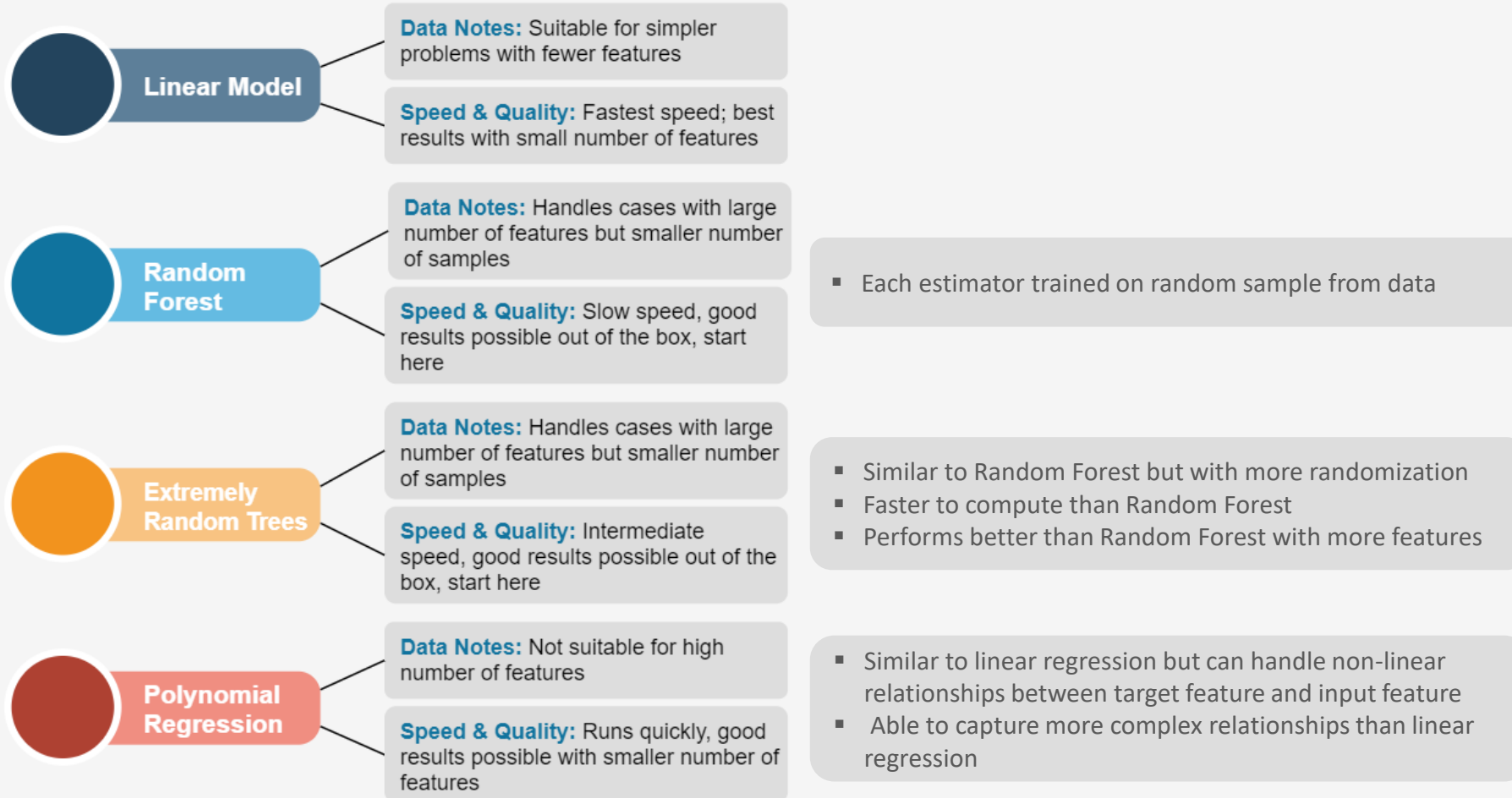
Data Needs: Periodicity, at least 2 time series, for discrete time series (e.g. counts), no negative values

Speed & Quality: Fast speed;
Intermediate quality (Highly data dependent)

Combines **exponential smoothing** estimates with **average demand intervals** to create a model for prediction

Driver-Based Forecasting

MULTIVARIATE ALGORITHMS



Multivariate Forecasting

DRIVER BASED

Gradient Boost

Data Notes: Sensitive to data with outliers

Speed & Quality: Fast speed; Intermediate quality (overfitting can be a problem)

- Continuously improve estimator by learning from errors of previous one
- Predict based on weighted average, more likely to overfit

SVM

Data Notes: Works well with large number of features, requires extensive hyper-parameter search for best results

Speed & Quality: Slower to get highest quality results with hyper-parameter search

Support Vector Machines

- Can model more complex non-linear relationships
- Works well with a large number of features
- Training time can increase heavily with bigger database

Nearest Neighbors

Data Notes: Handles more complex non-linear relationships and large amount of features

Speed & Result Quality: Speed increases heavily with larger datasets, good results possible with little tuning

Jedox Alssisted™ Planning – THE AI

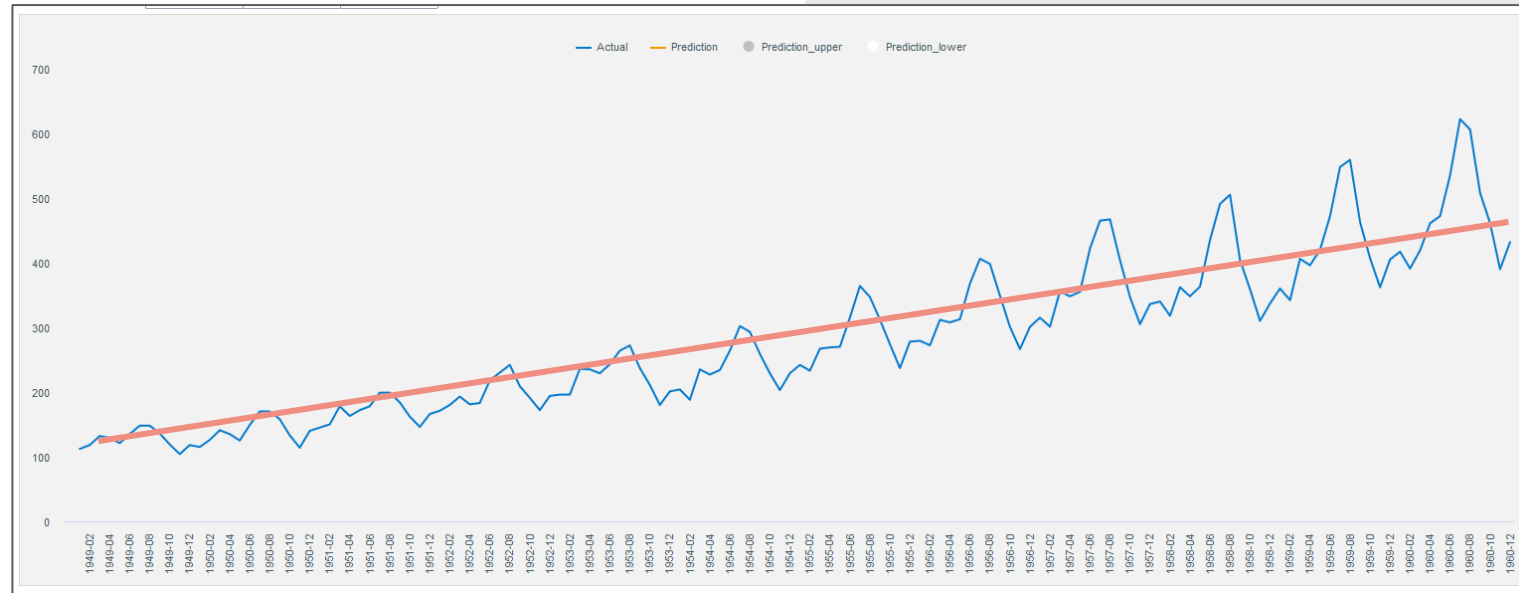


What the Algorithms Do



Trend

Decrease, Increase, Static...



Always Look at Data Before the Project

Saves time, gives you better idea of what level to predict



Simplify Planning, Reporting & Analytics

Jedox Alssisted™ Planning – THE AI



What the Algorithms Do



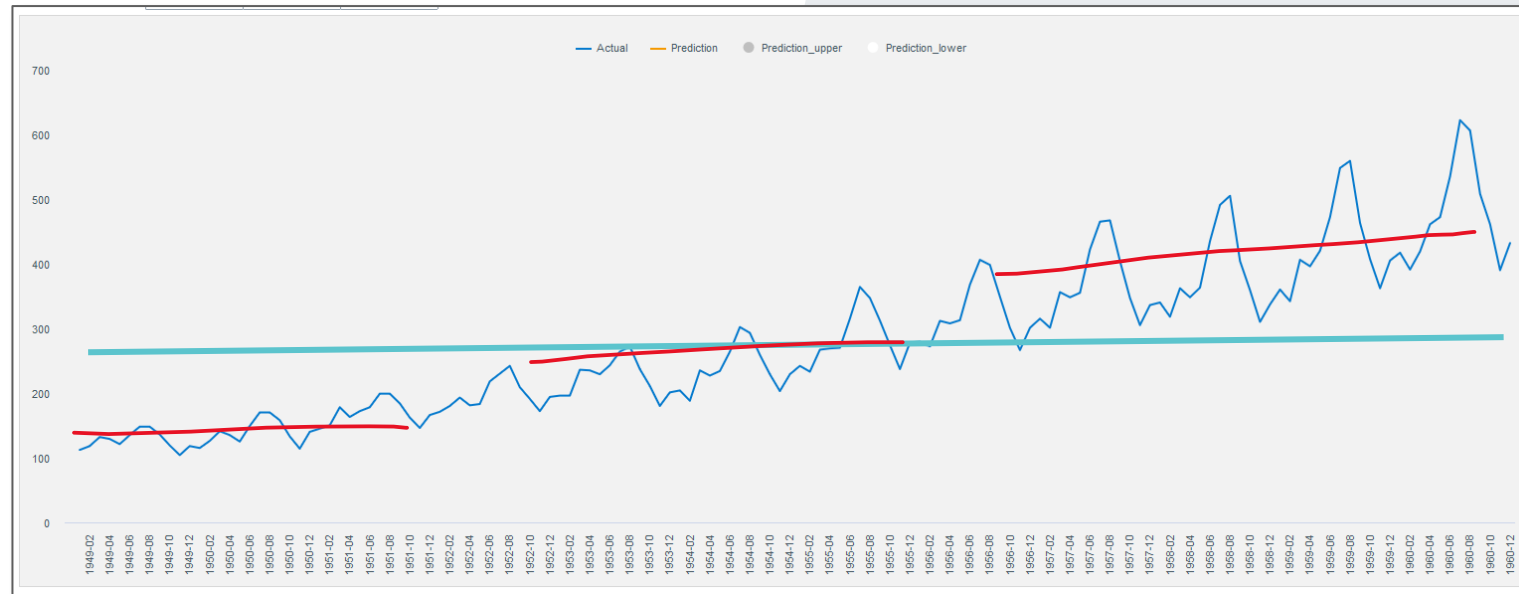
Trend

Decrease, Increase, Static...



Level

Average of Value



Always Look at Data Before the Project

Saves time, gives you better idea of what level to predict



Simplify Planning, Reporting & Analytics

Jedox Alssisted™ Planning – THE AI



What the Algorithms Do



Trend

Decrease, Increase, Static...



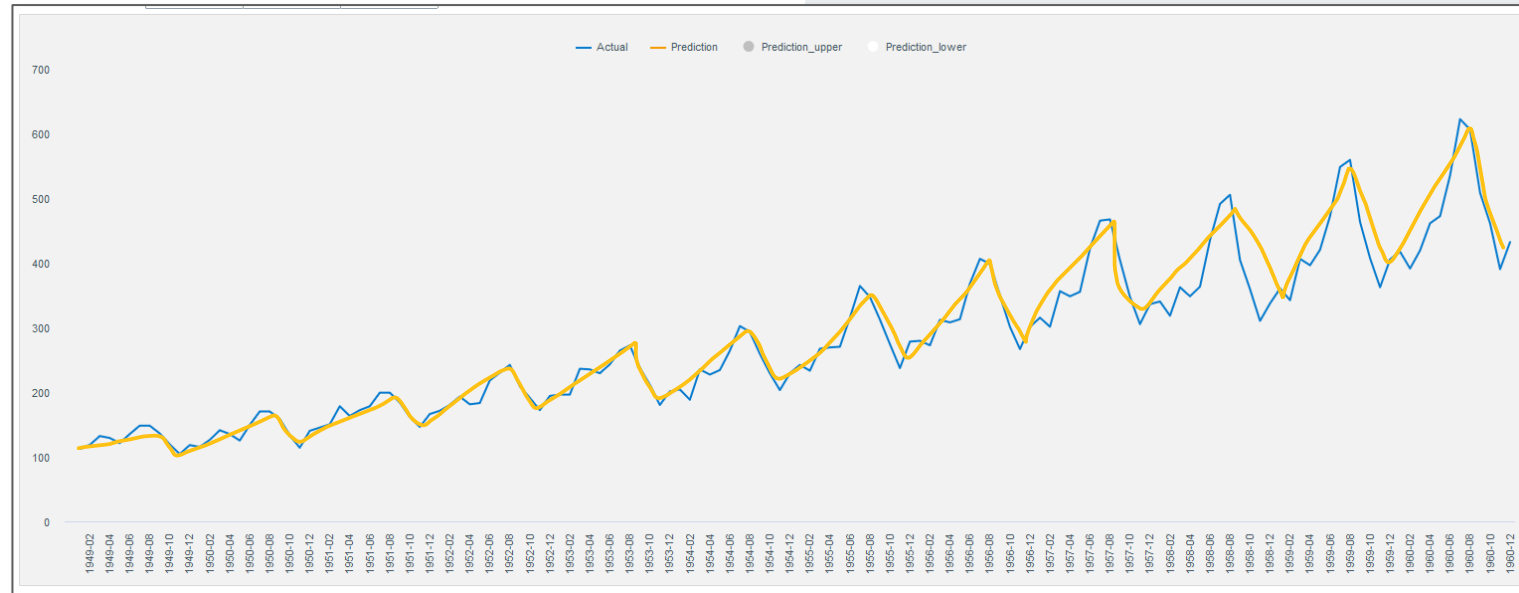
Level

Average of Value



Seasonality

Peaks, valley, etc.



Always Look at Data Before the Project

Saves time, gives you better idea of what level to predict



Simplify Planning, Reporting & Analytics

Jedox Alssisted™ Planning – THE AI



What the Algorithms Do



Trend

Decrease, Increase, Static...



Level

Average of Value



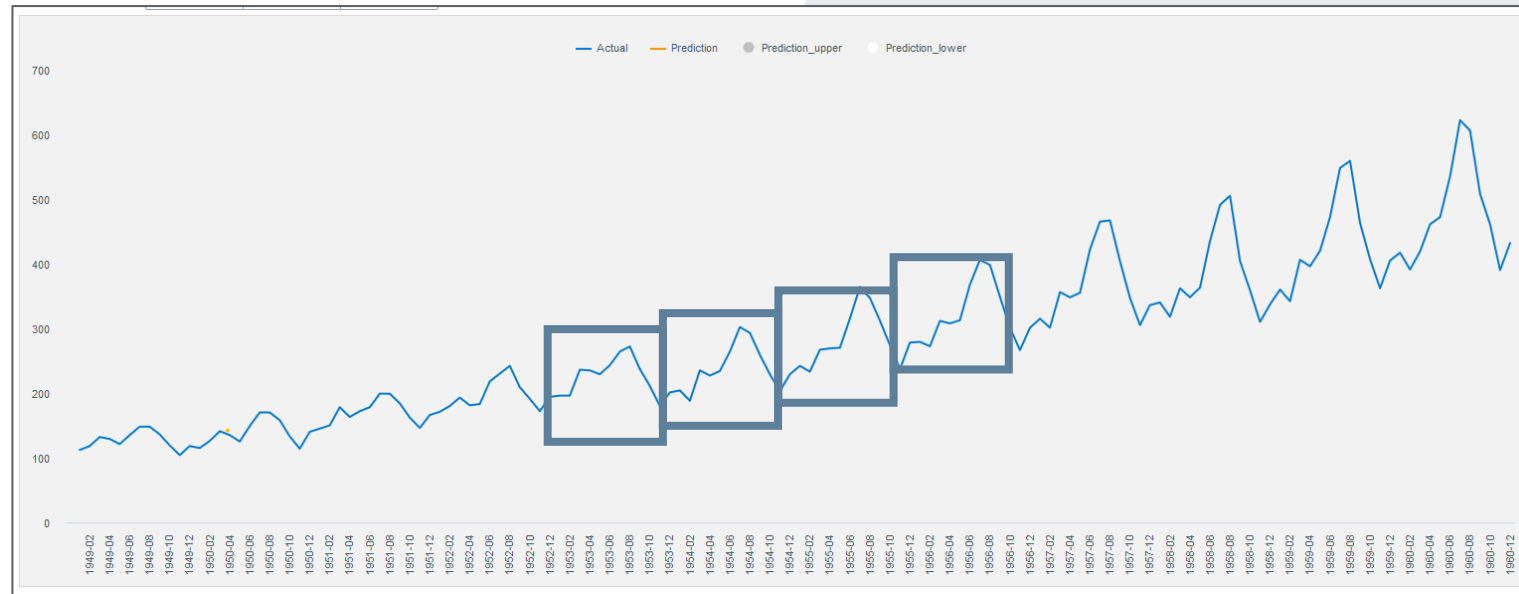
Seasonality

Peaks, valley, etc.



Period

e.g. 1 year



Always Look at Data Before the Project

Saves time, gives you better idea of what level to predict

Jedox Alssisted™ Planning – THE AI



Algorithms take all of this into account to calculate Prediction



Trend

Decrease, Increase, Static...



Level

Average of Value



Seasonality

Peaks, valley, etc.



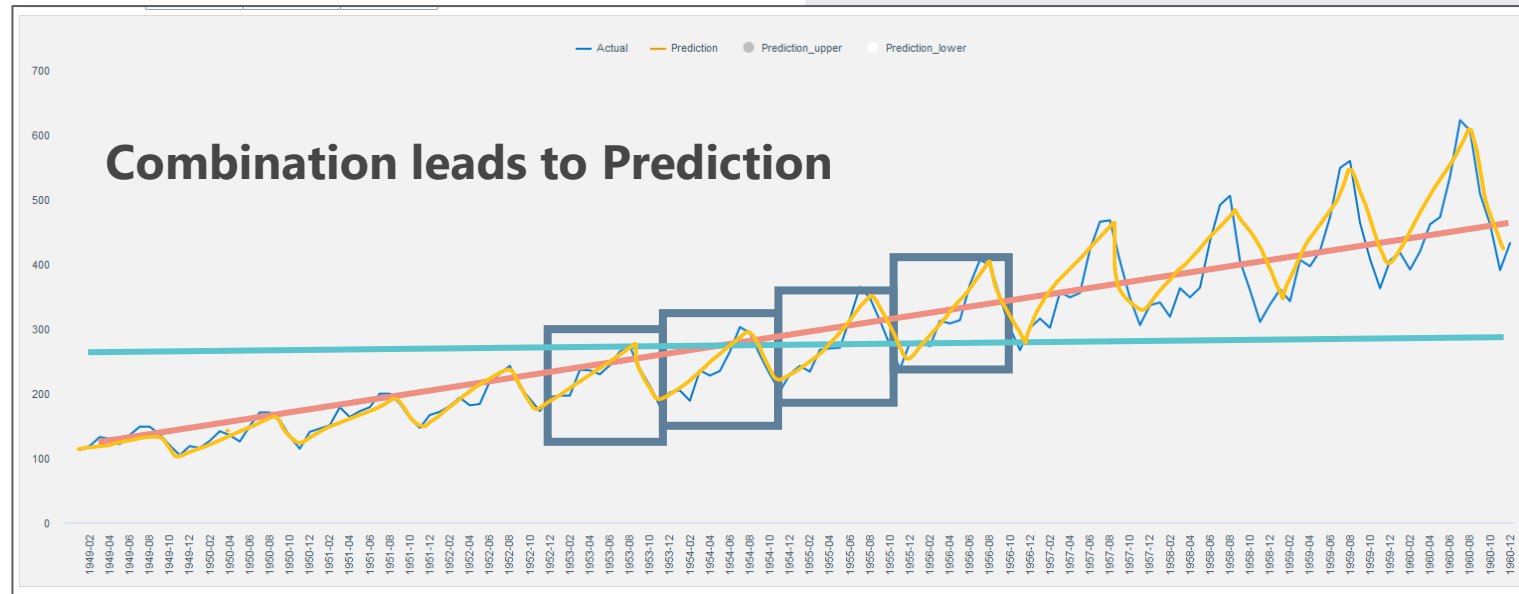
Period

e.g. 1 year



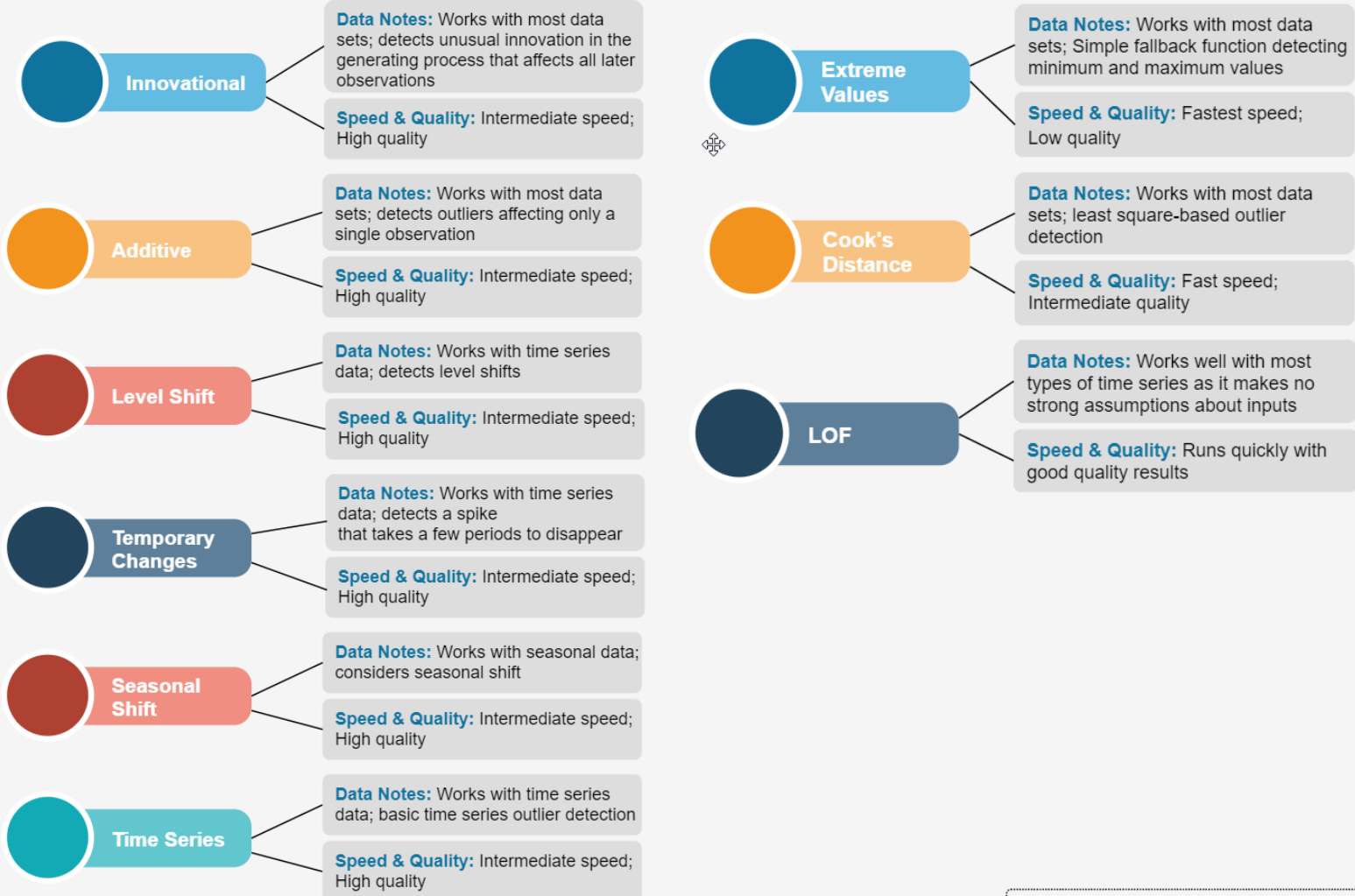
Always Look at Data Before the Project

Saves time, gives you better idea of what level to predict



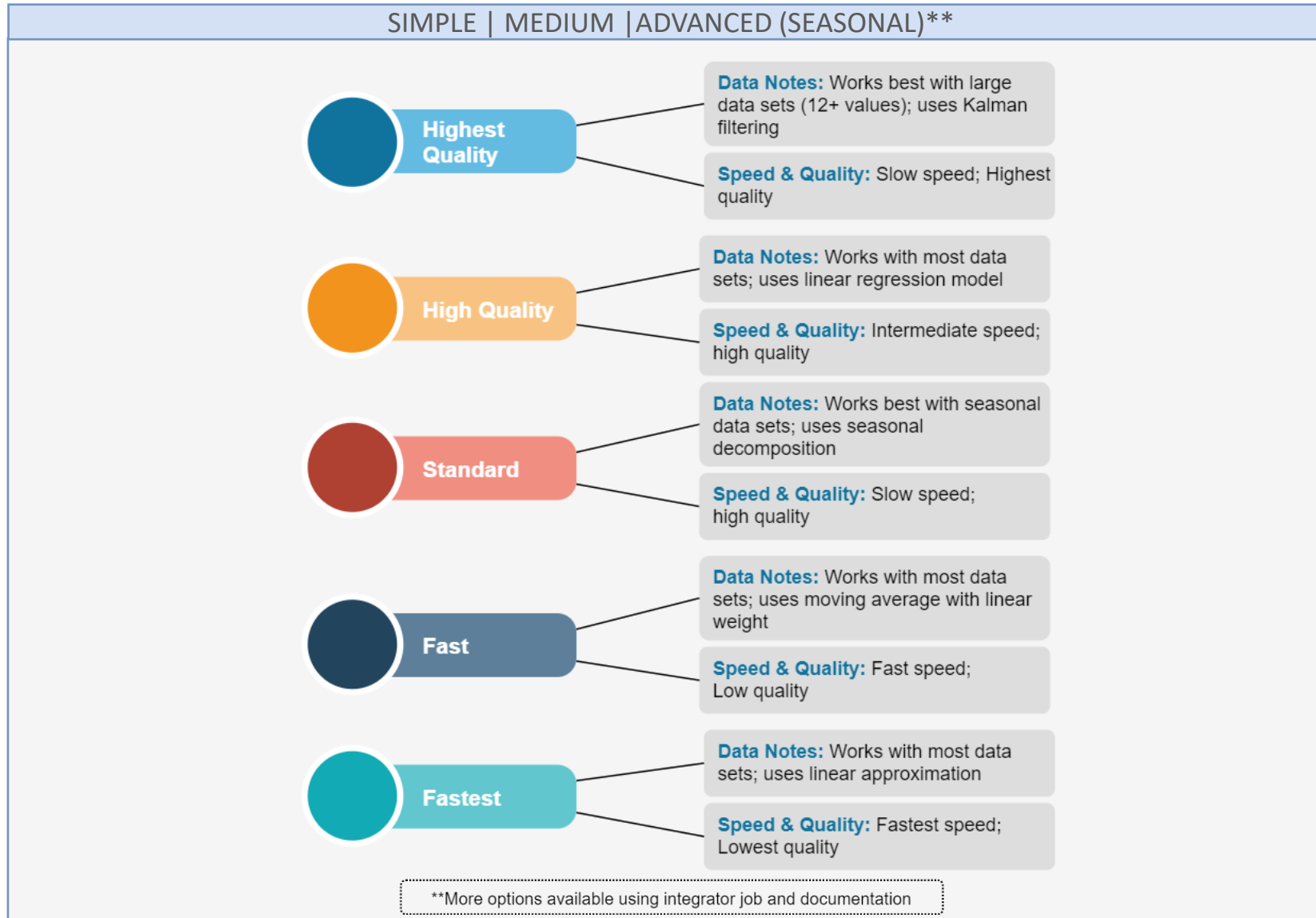
Data Preparation: Outlier Detection

SIMPLE | MEDIUM | ADVANCED (SEASONAL)*



*Fail safe mechanism can affect speed and result quality

Data Preparation: Interpolation



Classification

ALGORITHMS

Logistic Regression

Data Notes: Suitable for simpler problems with fewer features

Speed & Quality: Fastest speed; best results with small number of features

- Good to start here if a linear relationship between the target and input features is expected and if number of features is small relative to number of samples

Random Forest

Data Notes: Handles cases with large number of features but smaller number of samples

Speed & Quality: Slow speed, good results possible out of the box, start here

- Each estimator trained on random sample from data

Extremely Random Trees

Data Notes: Handles cases with large number of features but smaller number of samples

Speed & Quality: Intermediate speed, good results possible out of the box, start here

- Similar to Random Forest but with more randomization
- Faster to compute than Random Forest
- Performs better than Random Forest with more features

Support Vector Machines

Data Notes: Works well with large number of features, requires extensive hyper-parameter search for best results

Speed & Quality: Slower to get highest quality results with hyper-parameter search

- Can model more complex non-linear relationships
- Works well with a large number of features
- Training time can increase heavily with bigger database

Classification

ALGORITHMS

Gradient Boost

Data Notes: Sensitive to data with outliers

Speed & Quality: Fast speed; Intermediate quality (overfitting can be a problem)

- Continuously improve estimator by learning from errors of the previous one
- Predict based on weighted average, more likely to overfit

Naive Bayes

Data Notes: Suitable for all data sets, including smaller sample sizes, usually used for text classification

Speed & Quality: Fast speed; also works well with small number of features

K Nearest Neighbors

Data Notes: Handles more complex non-linear relationships and large amount of features

Speed & Result Quality: Speed increases heavily with larger datasets, good results possible with little tuning