RapidMiner

Séries Temporais - Tendência

*Rui Humberto Pereira*

*rui.humberto.pereira@gmail.com*

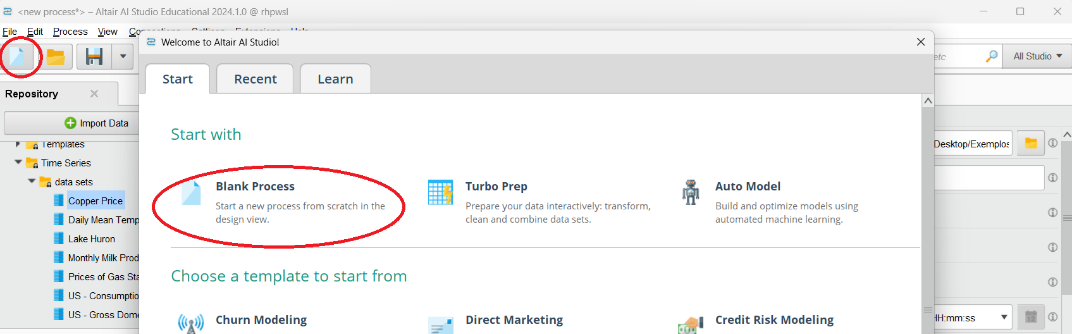
*Fevereiro de 2025*

**Tendência**: é a direção geral que os dados seguem ao longo do tempo, ou seja, o seu comportamento de longo prazo, mostrando se há crescimento, declínio ou estabilidade nos mesmos. Podemos identificar a tendência visualmente ou através de métodos estatísticos como médias móveis, regressão linear ou decomposição de séries temporais.

Neste exercício iremos utilizar o RapidMiner para observar a tendência de uma série temporal.

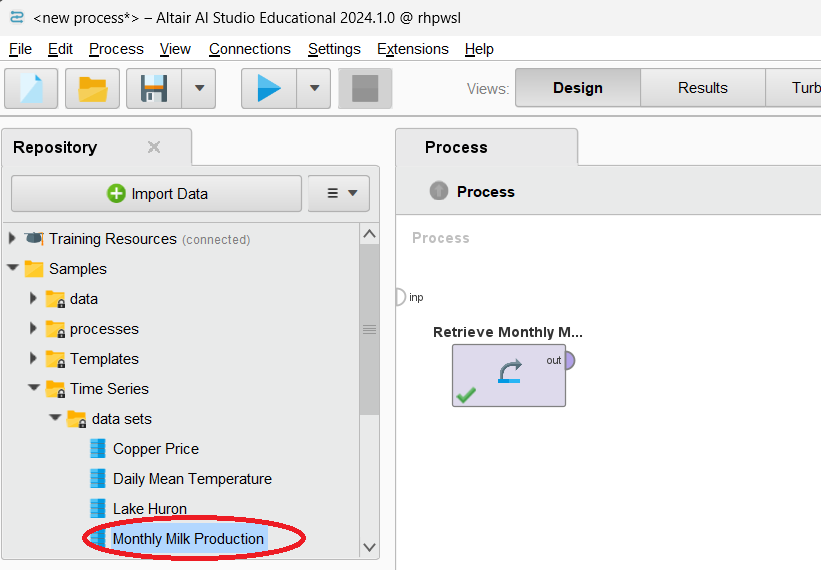
## Inicie um novo projeto no AI Studio

* Menu File -> New Process



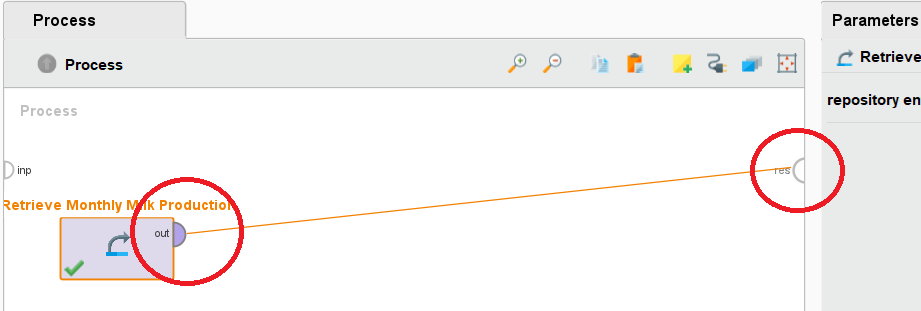
## Obtenção de dados

* O RapidMiner disponibiliza um conjunto de dados de exemplo que vamos utilizar neste exercício. No explorador do Repositório, selecione o dataset de demonstração “Monthly Milk Production” e arraste-o para a área de design do processo.

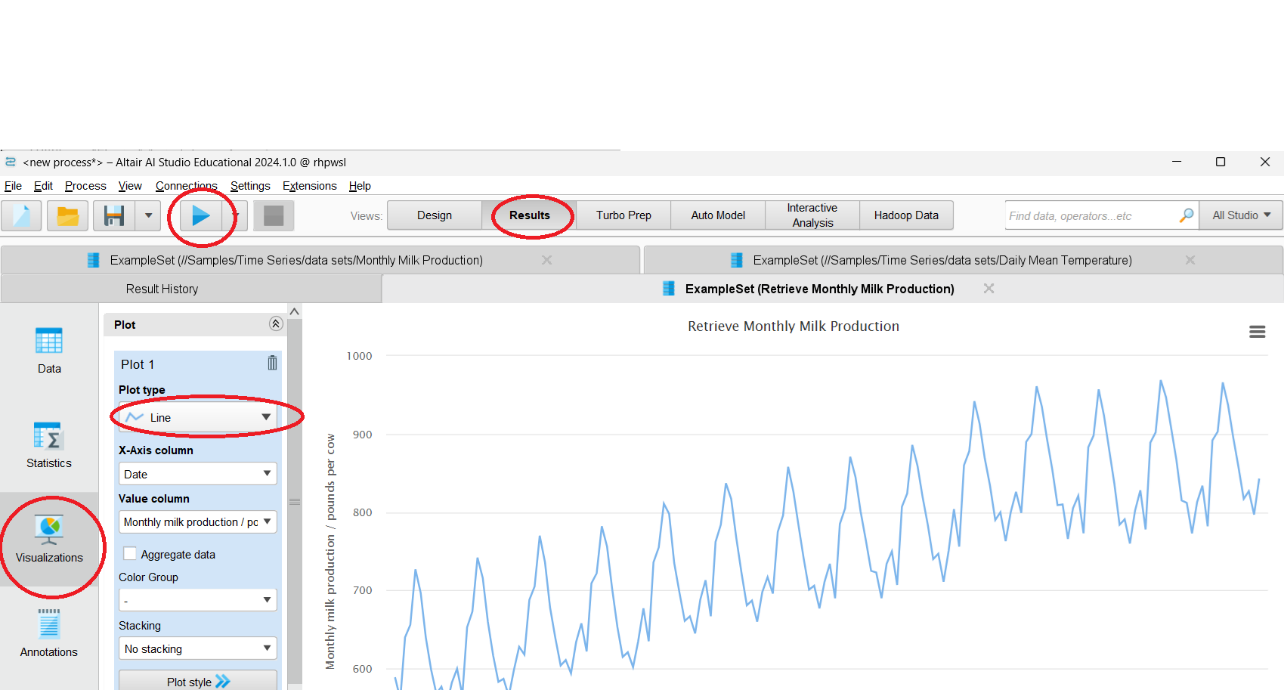


## Visualizar os dados

* Ligue o dataset à Saída de resultados premindo o botão esquerdo do rato enquanto define a ligação.



* Depois de estabelecida a ligação, faça “Run” do processo clicando no botão run .
* No separador de Resultados opte por visualizar os dados utilizando um gráfico de linhas.

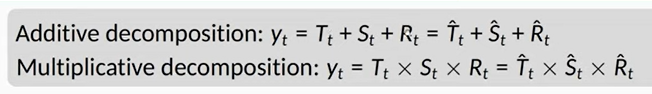


* O gráfico que visualiza apresenta a produção de leite desde dezembro de 1961 e novembro de 1975. Podemos verificar que existe uma clara tendência de crescimento na produção.

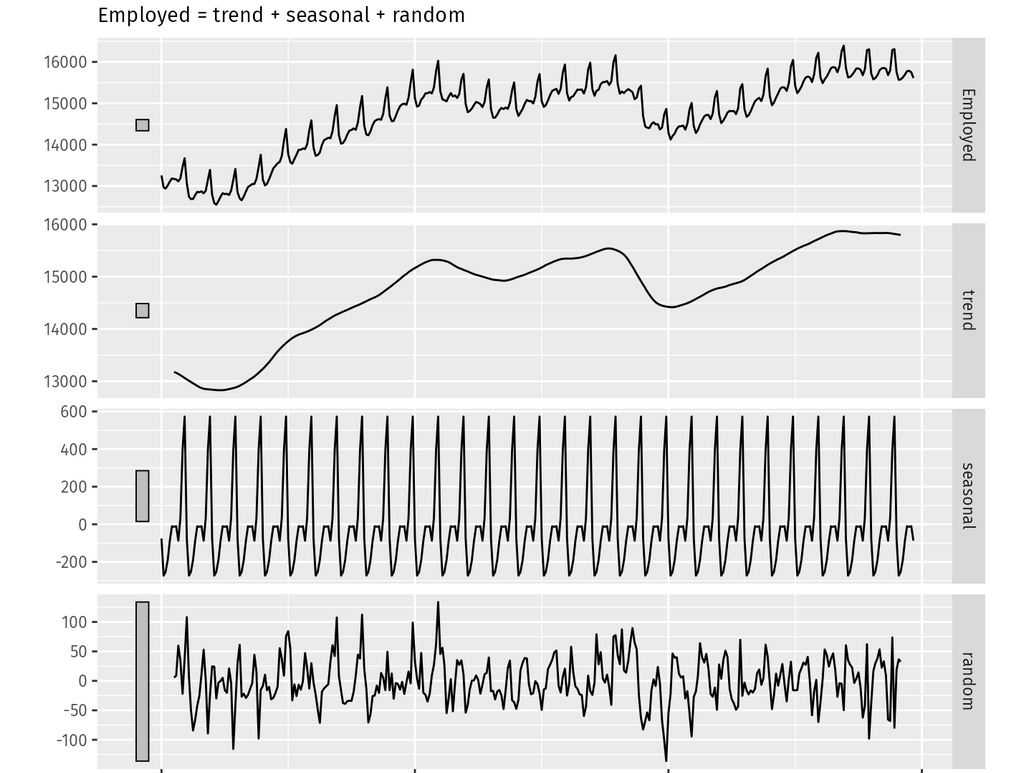
## Visualizar a linha de tendência

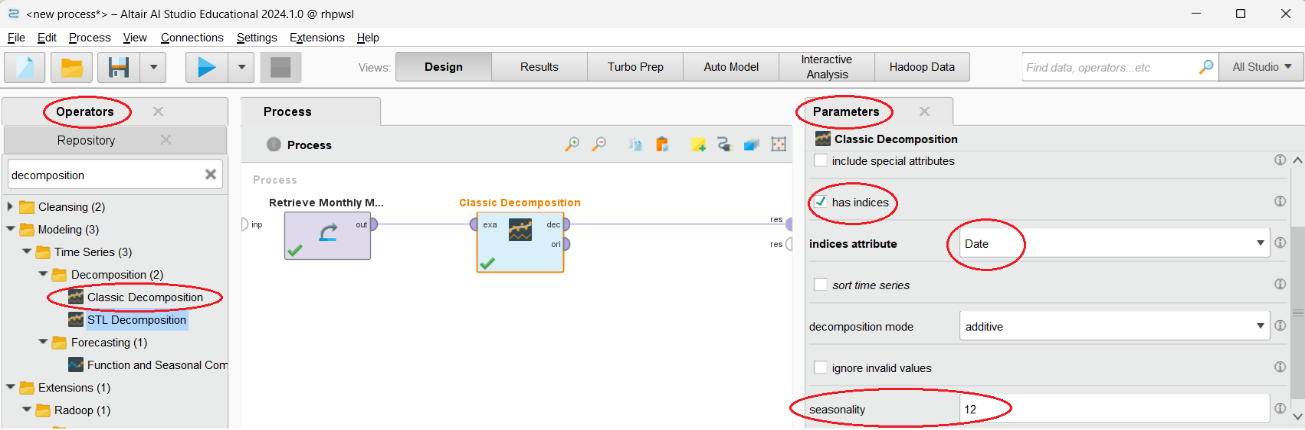
* Volte para a área de design e no separador Operators, procure por “decomposition” e encontrará o Classic Decomposition. Este operador permite decompor a série temporal em várias métricas, nomeadamente a Tendência.

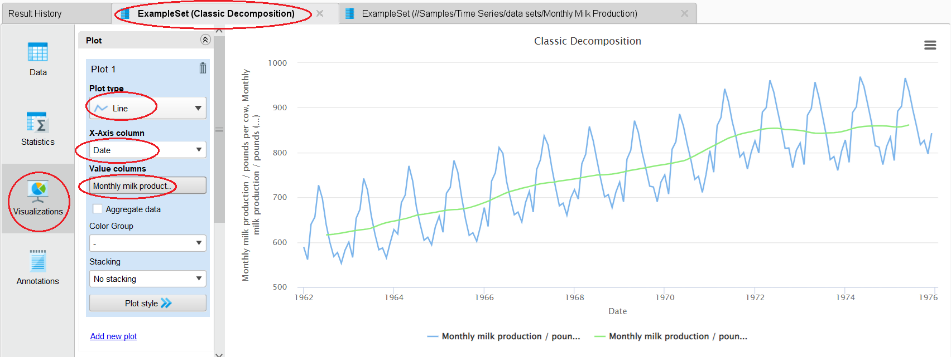
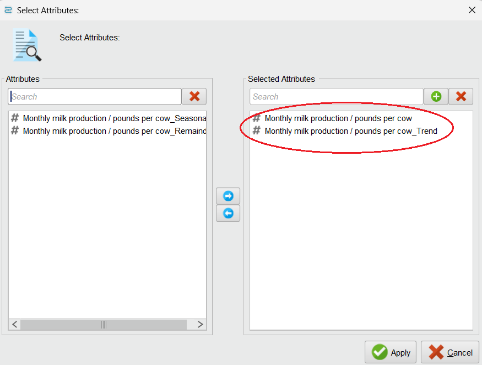
Neste exercício, vamos optar por uma decomposição aditiva.



Na decomposição clássica, assumimos que a componente sazonal é constante de ano para ano.



* Arreste-o e altere as ligações do processo conforme a figura.
* No separador Parameters, defina o campo que contem a data (Date)como índice (activar has índices) e a sazonalidade 12.
* Volte a correr o processo  e visualize os resultados. Para selecionar as linhas de valores e de tendência, no botão “Value Columns” selecione os atributos correspondentes.

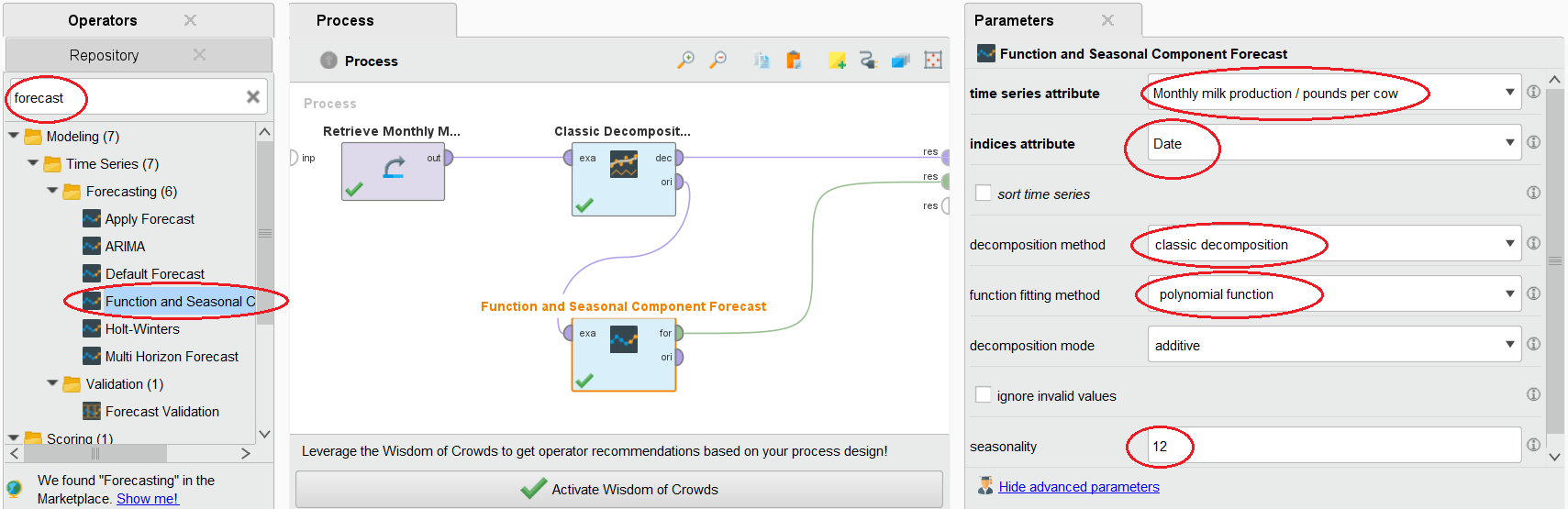
 

* O gráfico permite visualizar como a produção de leite variou ao longos dos anos, sendo clara a sazonalidade da variação e a tendência de aumento até 1972.

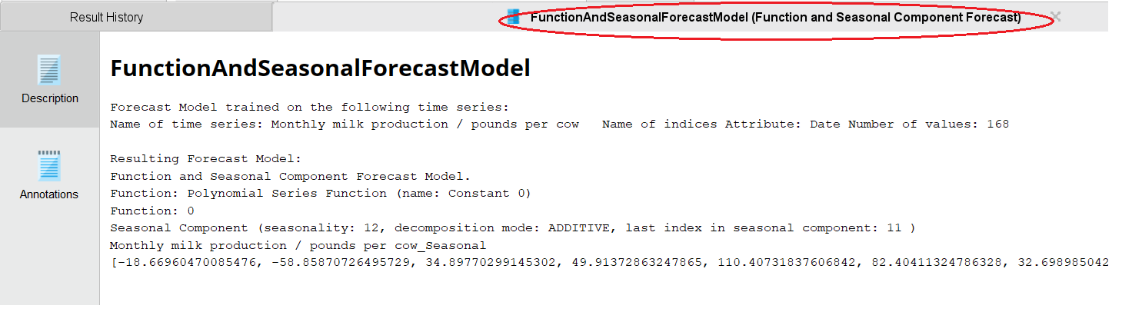
## Previsão para os anos seguintes

Sendo conhecida a sazonalidade e a tendência de variação, vamos realizar uma previsão da evolução da produção de leite.

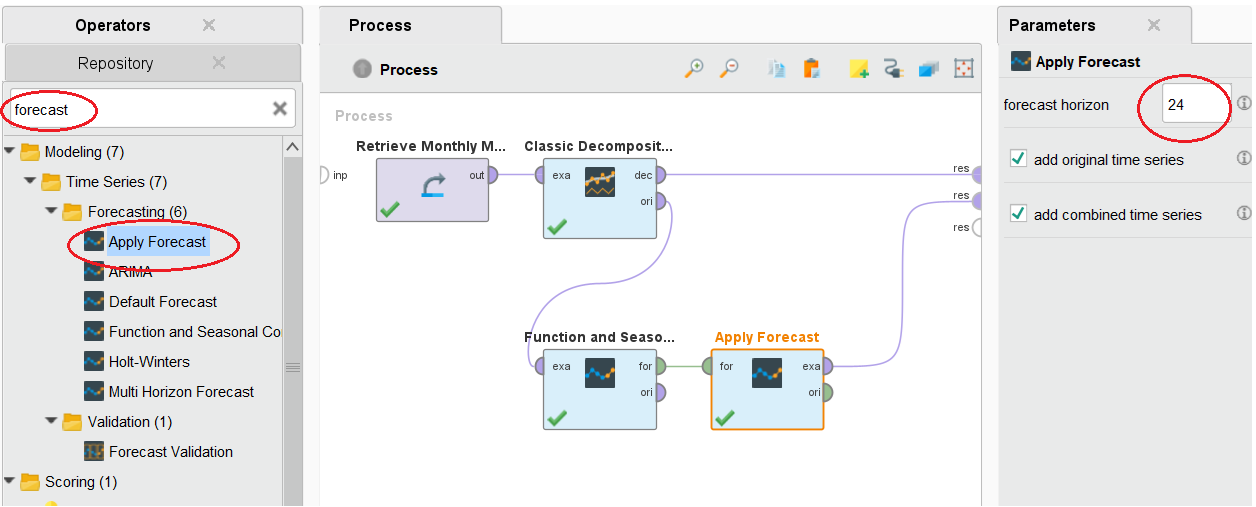
* Procurando um operador por “Forecast” e encontrará vários. Adicione ao processo o “Function And Seasonal Component Forecast”, altere as ligações e parâmetros conforme a figura.
* A saída “ORI” do operador de decomposição contem os dados originais, pelo que vamos utilizar esta saída para alimentar o nosso modelo de Forecast.
* Relativamente aos parâmetros, o atributo de entrada é o valor da produção e o índice a data. Vamos utilizar o mesmo modelo de decomposição da alínea anterior e sazonalidade de 12 meses.



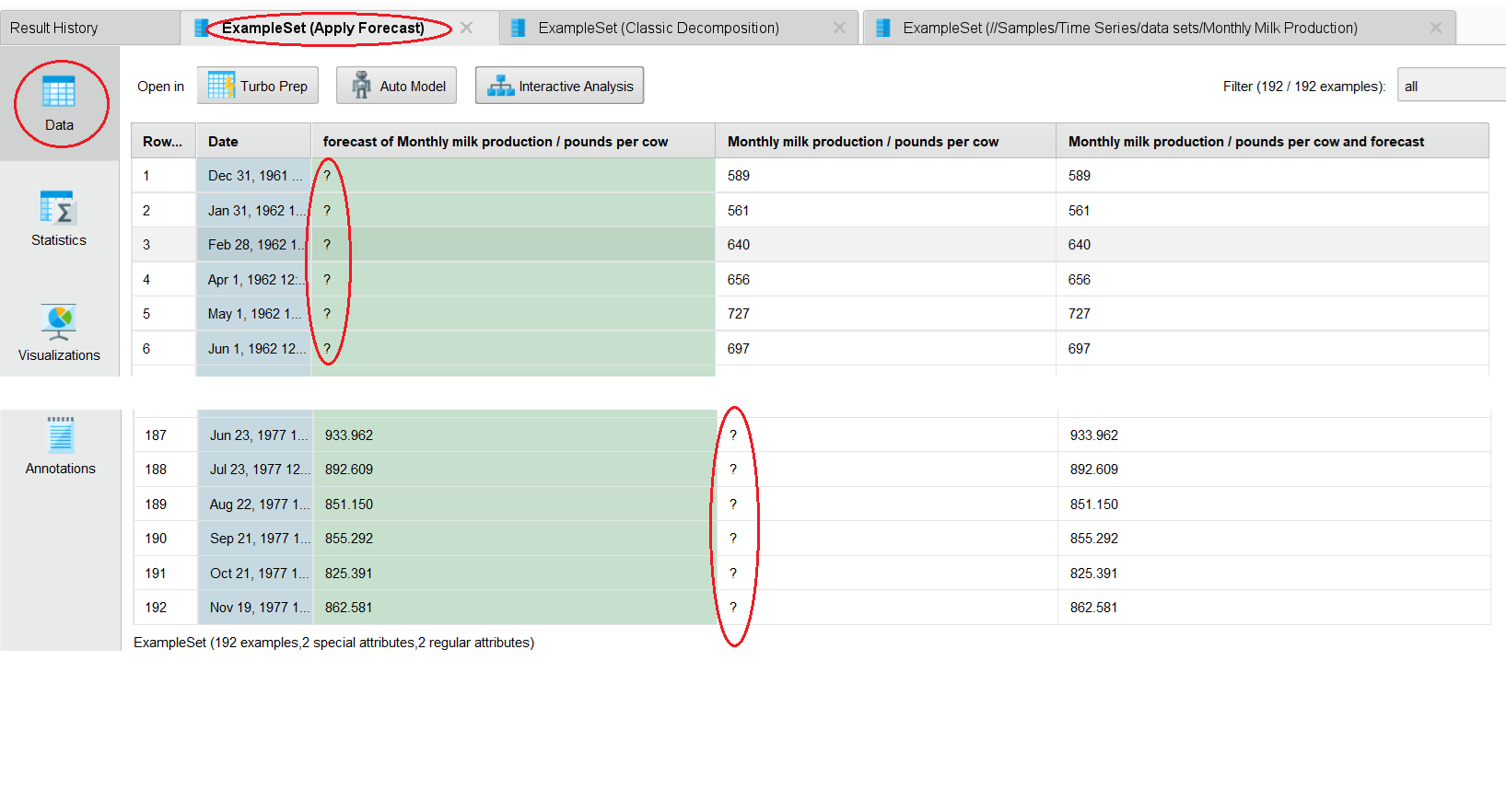
* No separador de resultados conseguimos visualizar o modelo e os respectivos parâmetros.



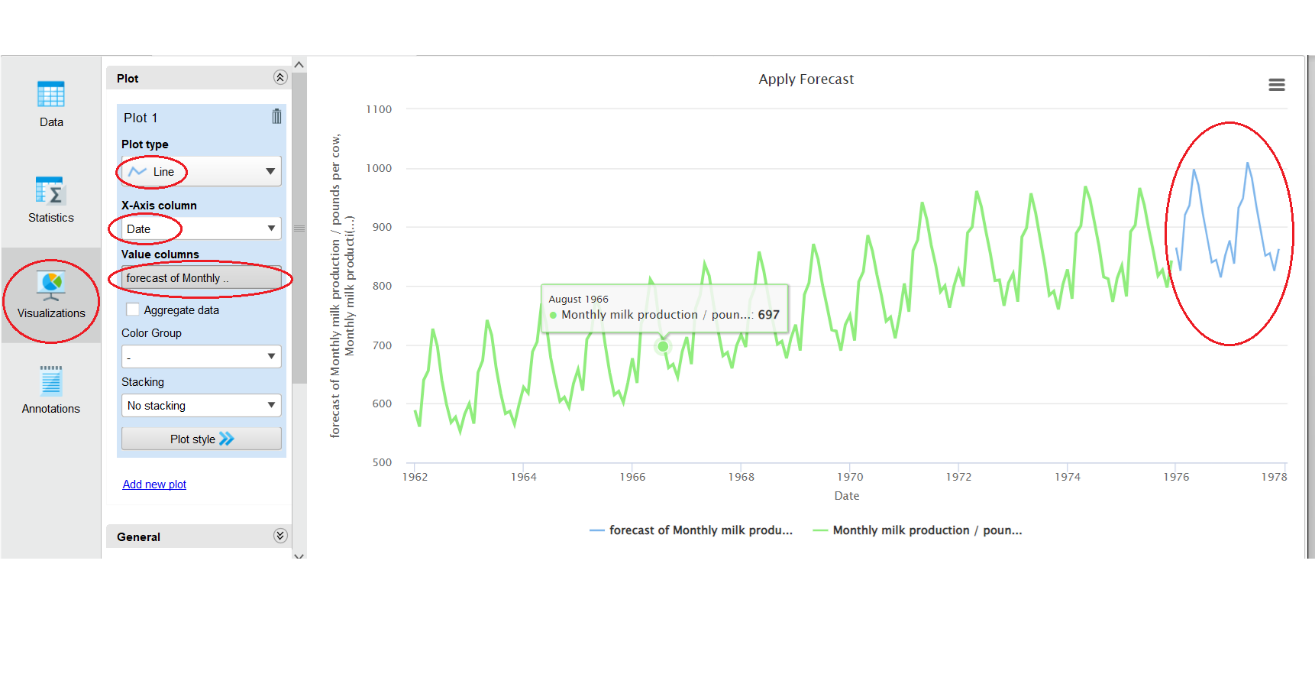
* Até aqui o processo só obtém o modelo, ainda não faz previsões.
* Procure por um operador “Aply Forecast”, adicione-o ao processo e altere as ligações conforme a figura.



* Observe que o modelo deixa de ser apresentado na saída e passa a ser entregue ao operador que aplica as previsões.
* Opte por fazer uma previsão para 24 meses.
* Visualize a previsão para os dois anos no separador de resultados, mas desta vez como uma tabela (botão Data).



* Observe que a tabela tem quatro colunas: o número de registo, a data do valor da produção, o a previsão da produção, o valor conhecido da produção e, na última, a junção dos valores conhecidos e previstos.
* Agora visualize graficamente os dados.
* Utilizando o botão “Value Columns”, selecione a terceira e a quarta coluna da tabela anterior.
* O gráfico que é apresentado, apresenta a previsão da produção de leite para os dois anos seguintes, mantendo a linha de tendência e de sazonalidade.



* Guarde o processo no repositório ou exporte-o para um ficheiro externo.

