

Previsão de vendas Companhia de Vídeo Game

Visão geral.

O departamento da cadeia de suprimentos de uma empresa de videogames foi incumbido de prever dados mensais de vendas, a fim de sincronizar a oferta com a demanda, auxiliar na tomada de decisões que ajudará a construir uma infraestrutura competitiva e medir o desempenho da empresa. O analista de cadeia foi designado para ajudar o gerente a executar os números através de uma série temporal modelo de previsão.

Planeje sua análise

1. O conjunto de dados atende aos critérios de um conjunto de dados da série temporal?

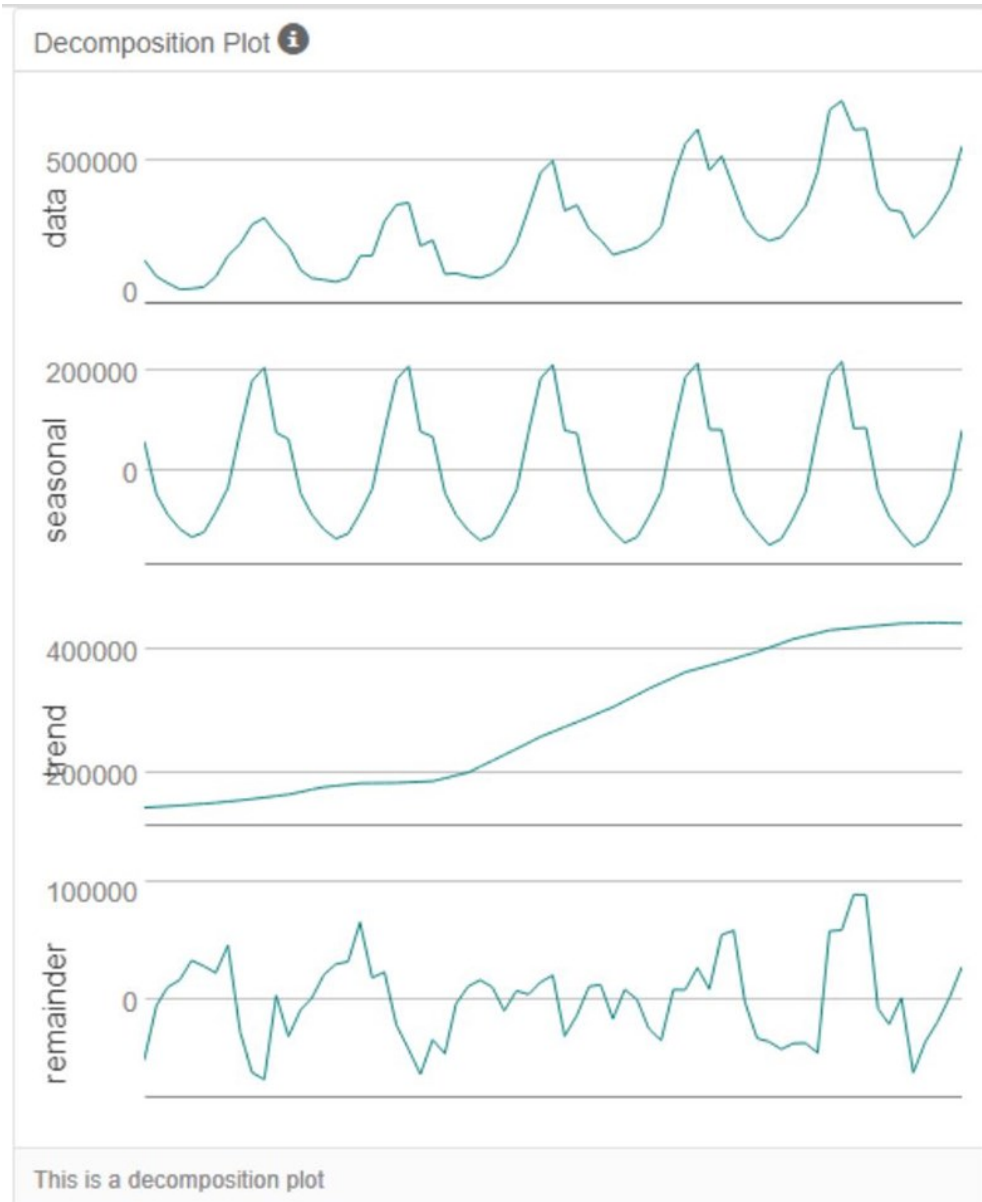
Temos os dados de vendas mensais no arquivo “monthlhy_sales” de uma empresa que vende videogames, analisando os dados no arquivo chegamos à conclusão que o conjunto de dados atende as especificações da série temporal, já que a ordem das vendas importa, temos o número de vendas de forma sequencial e em intervalos de tempo iguais, os dados estão 2008-01 até 2013-09, portanto temos apenas um único dado para cada intervalo de tempo, sendo assim, cada mês tem apenas um valor dentro do conjunto de dados.

2. Quais registros devem ser usados como amostra de retenção?

Como o objetivo é prever as vendas dos próximos 4 meses, foi pego os últimos 4 registros como amostra de retenção, que compreende 2013-06 até 2013-09. Pois o tamanho da amostra de retenção depende de quão longa é a serie temporal e de quantos períodos de tempo será previsto.

Determine os componentes tendência, sazonalidade e erro

1. Qual é a tendência, a sazonalidade e o erro da série temporal?



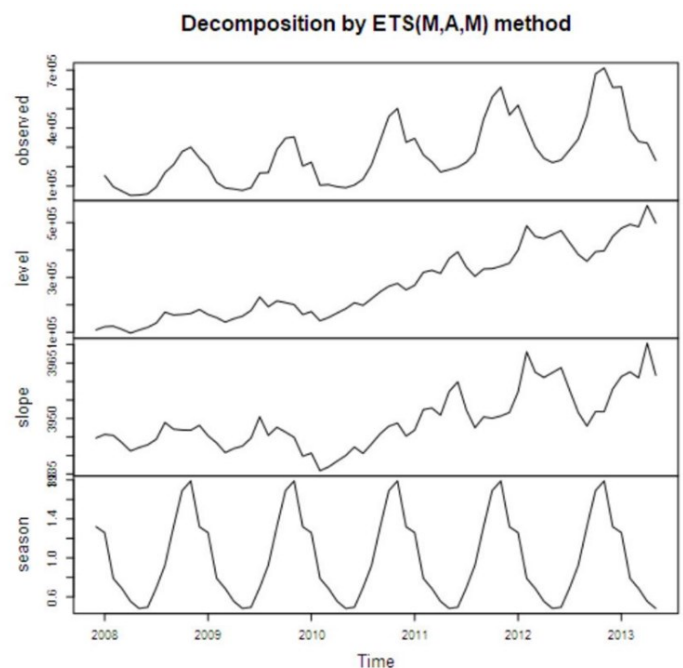
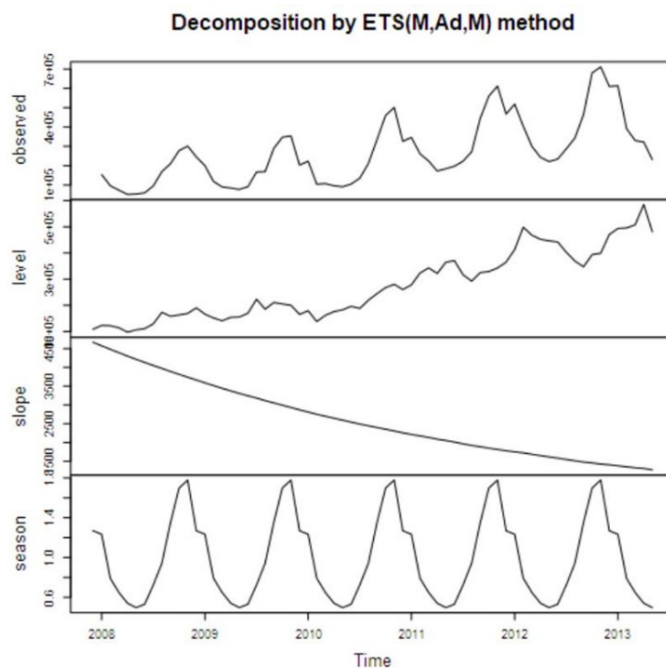
A serie temporal original é dividida em três subséries, determinados de componentes sazonal, tendência e de erro. A serie temporal possui as características de repetição dentro de um período de a cada 12 meses, depois podemos observar o gráfico de tendência o qual reflete o comportamento da série temporal em longos períodos de tempos, podemos notar uma tendência crescente neste caso, o gráfico de "remainder" estima o componente de erro da série temporal.

Construa seus modelos

1. Quais são os termos modelo para o ETS?
 - a. Descreva os erros na amostra.

A sazonalidade é ligeiramente crescente ao longo do tempo, então foi aplicado o modelo multiplicativamente. Há um comportamento crescente nos dados, então existe uma tendência, no caso essa tendência é linear, então aplicamos o modelo de forma aditiva.

O erro está aumentando ou diminuindo ao longo do tempo, aplicamos o erro multiplicativamente, já que apresenta variabilidade. Em seguida, executamos o ETS amortecido (dampened) e não amortecido (M, A, M), como mostra os resultados abaixo.



Method:
ETS(M,Ad,M)

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
5597.130809	33153.5267713	25194.3638912	0.1087234	10.3793021	0.3675478	0.0456277

Information criteria:

AIC	AICc	BIC
1639.465	1654.3346	1678.604

Method:
ETS(M,A,M)

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
2818.2731122	32992.7261011	25546.503798	-0.3778444	10.9094683	0.372685	0.0661496

Information criteria:

AIC	AICc	BIC
1639.7367	1652.7579	1676.7012

Consideramos duas medidas de precisão, RMSE e MASE, para julgar se é um bom modelo de previsão. Para o modelo ETS não amortecido, RMSE é de 32992.73 e MASE é 0.372; já para o modelo ETS amortecido, o RMSE é 33153,53 e MASE 0,367.

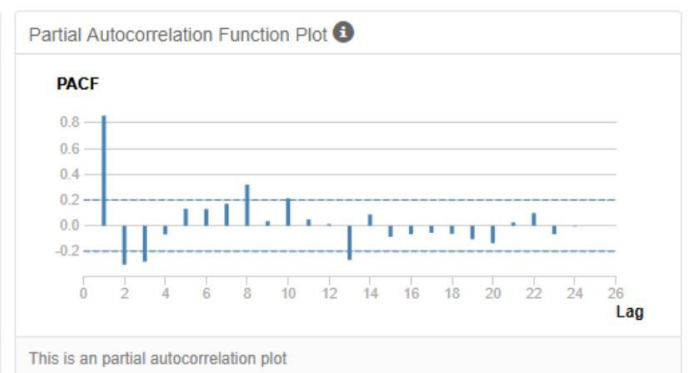
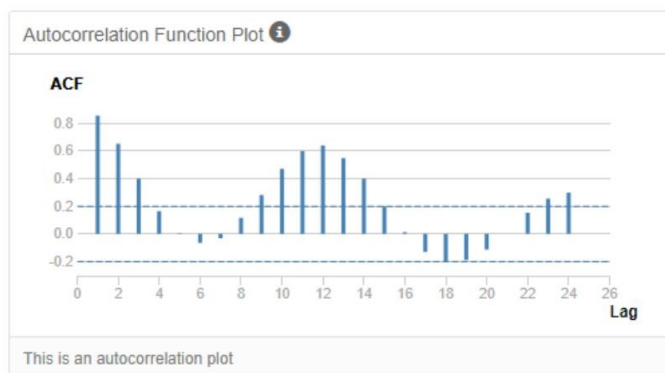
O RMSE do modelo ETS amortecido é maior que o modelo ETS não amortecido, e por sua vez, o MASE do modelo ETS amortecido é inferior ao modelo ETS não amortecido. Podemos concluir que o modelo ETS amortecido é a melhor escolha, dado que o AIC é um pouco inferior ao modelo ETS não amortecido.

2. Quais são os termos modelo para o ARIMA?
- Descreva os erros na amostra. Use pelo menos RMSE e MASE ao examinar os resultados.

O modelo ARIMA exige que a série seja estacionária. Os dados de vendas mensais não são estacionários. O gráfico da série temporal mostra uma tendência ascendente e sazonalidade. Além disso, há algumas tendências ou componentes sazonais, isso implica que suas propriedades estatísticas não são constantes ao longo do tempo.

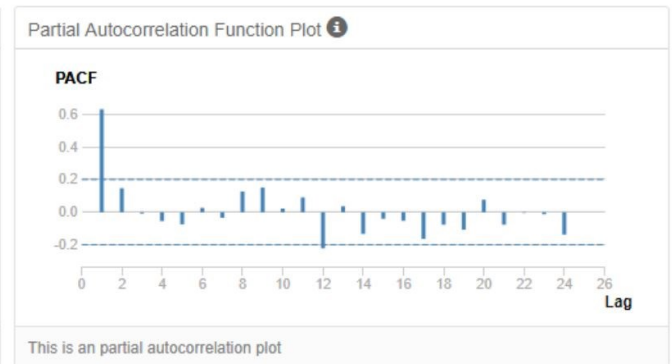
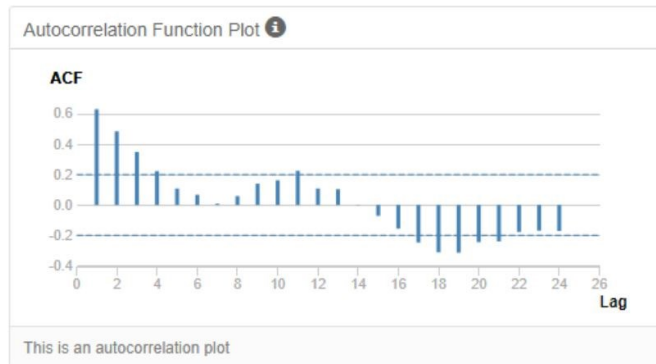
Os gráficos de ACF e PACF ajudam a determinar a existência de autocorrelação ao longo da série temporal.

- ACF e PACF da série temporal.



No gráfico ACF, podemos notar um decaimento lento da autocorrelação em direção ao 0, o que aumenta nos ciclos sazonais. Como a autocorrelação é alta, é necessário realizar a diferença sazonal da série temporal.

- ACF e PACF da diferença sazonal.

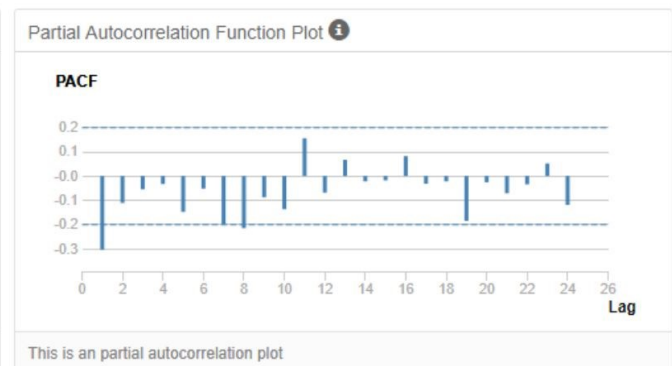
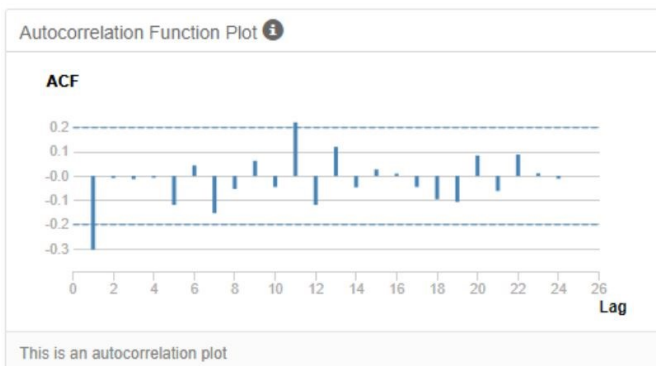


Ao plotar o ACF e PACF da diferença sazonal podemos notar que os resultados comparados com os gráficos iniciais sem a diferenciação são similares, distinguindo-se apenas com uma correlação menor.

Para eliminar essa correlação, é necessário realizar novamente a diferenciação sazonal.

Nos gráficos ACF e PACF acima podemos perceber que a correlação entre os lags sazonais foi eliminada, lag 24 ficou invertido com o 12, porém ainda há uma alta correlação entre os lags sequenciais. Por isso foi necessário aplicar uma nova diferenciação para tornar a série estacionária, mas dessa vez uma diferenciação não-sazonal. Sendo a primeira diferenciação sazonal ($D=1$) e a segunda não-sazonal ($d=1$).

- ACF e PACF da primeira diferença sazonal.



Ao plotar o gráfico ACF podemos ver não há mais correlação entre os lags, o que indica que não há necessidade de nova diferenciação. Contudo ainda há lags

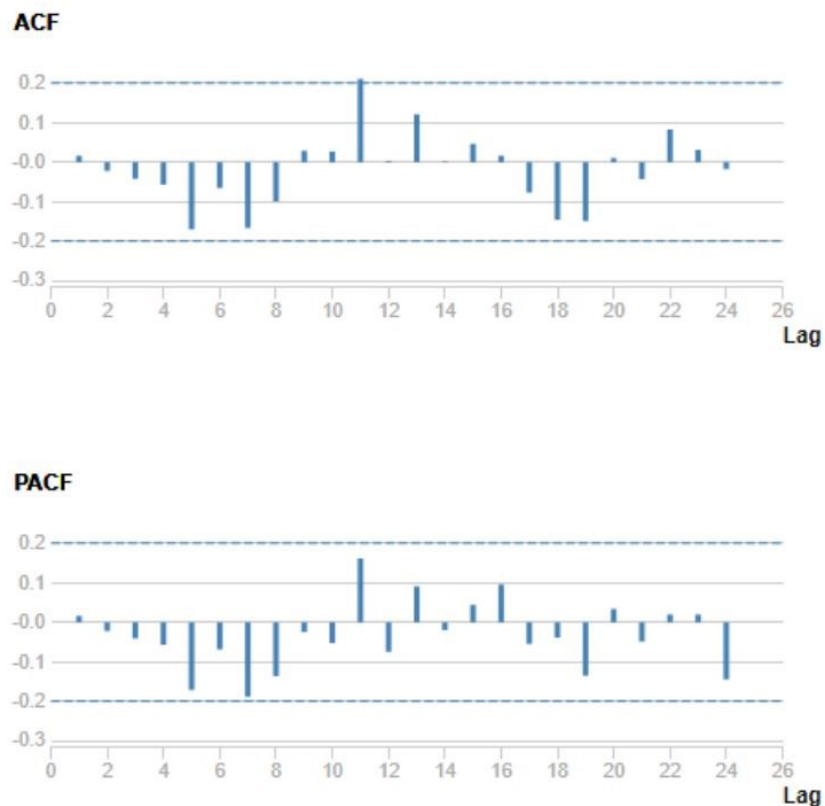
significativos (como o primeiro lag), nesse caso devemos avaliar a aplicação de termos AR ou MA.

A correlação restante poderá ser tratada utilizando os termos autogressivos (AR) e a média móvel (MA), os termos de diferenciação serão $d(1)$ e $D(1)$.

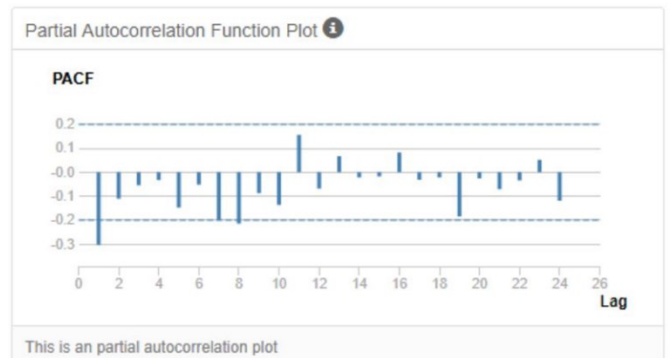
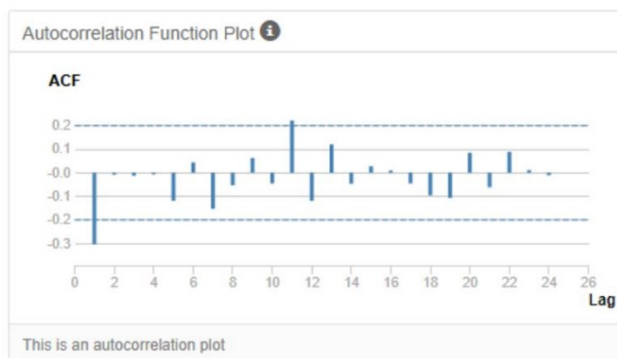
O gráfico ACF mostra uma forte correlação negativa no período 1 confirmado no PACF; isso indica um modelo MA(1), pois há um período significativo. Os períodos sazonais 12, 24 e etc., tanto no ACF quanto no PACF não apresentam correlação significativa, devido a isso não há necessidade dos termos sazonais de autoregressão (AR) ou de média móvel (MA).

- ACF e PACF após a aplicação do modelo ARIMA

O modelo de ARIMA é $(0, 1, 1) (0, 1, 0)$ [12]. Nota-se que nos gráficos ACF e PACF do modelo de ARIMA não há correlação significativa entre os períodos, indicando que não há a necessidade de termos adicionais em AR e MA.



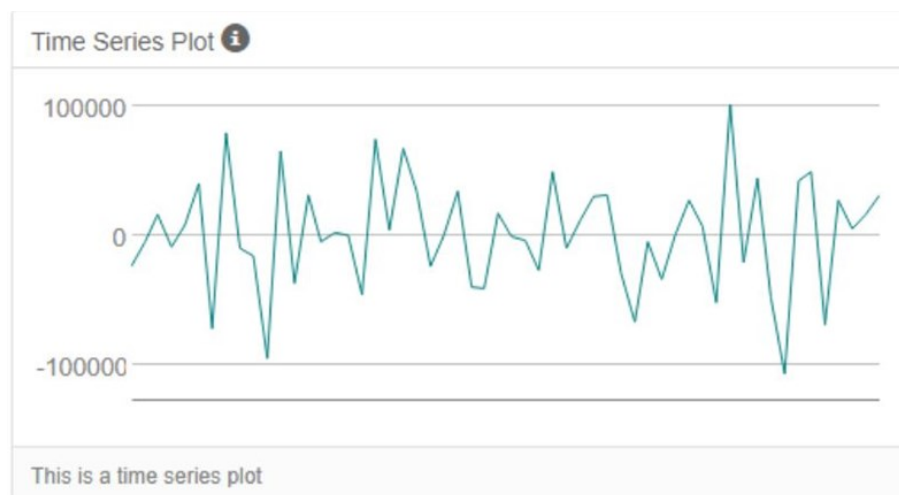
Os gráficos ACF e PACF tanto para a série temporal como para a diferença sazonal.



Observando os gráficos de função de autocorrelação (ACF) e autocorrelação parcial (PACF) da primeira diferença sazonal, podemos identificar os números de termos de AR e / ou MA necessários.

Para series não sazonais, foi examinado as defasagens anteriores e observamos um pico no valor de defasagem 1 da série ACF, o que indica termos de AM não sazonais.

Para termos sazonais, nota-se que não há mais picos ocorrendo em intervalos de 12 meses e 24 meses. Então, o modelo que se ajusta é ARIMA (0, 1, 1) (0, 1, 0) [12].



Information Criteria:

AIC	AICc	BIC
1256.5967	1256.8416	1260.4992

In-sample error measures:

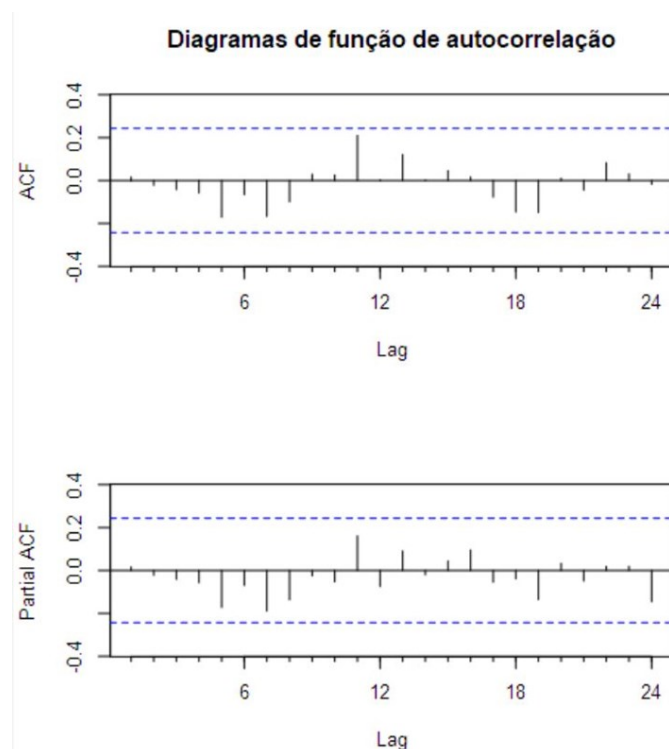
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-356.2665104	36761.5281724	24993.041976	-1.8021372	9.824411	0.3646109	0.0164145

Ljung-Box test of the model residuals:

Chi-squared = 16.4458, df = 23, p-value = 0.83553

Os erros de previsão e as medidas de precisão do modelo, RMSE e MASE são 36761,52 e 0,364, respectivamente.

Abaixo, podemos ver os gráficos ACF e PACF:



Observa-se que o modelo não apresenta correlação. Sendo assim, confirmamos que o modelo ARIMA estacionário está pronto.

Previsão

1. Qual modelo você escolheu?

Na análise dos modelos ETS anteriores, especificamente nos modelos ETS amortecido e não amortecido, foi visto que o ETS amortecido apresentou um menor índice de AIC, portanto estabeleceu-se uma comparação entre o modelo ETS amortecido e o modelo ARIMA.

Medições de erro na amostra:

- ETS Amortecido:

Method:

ETS(M,Ad,M)

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
5597.130809	33153.5267713	25194.3638912	0.1087234	10.3793021	0.3675478	0.0456277

Information criteria:

AIC	AICc	BIC
1639.465	1654.3346	1678.604

- ARIMA

Information Criteria:

AIC	AICc	BIC
1256.5967	1256.8416	1260.4992

In-sample error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
-356.2665104	36761.5281724	24993.041976	-1.8021372	9.824411	0.3646109	0.0164145

Medidas de erro de previsão contra a amostra de retenção:

- ETS amortecido

Actual and Forecast Values:

Actual	ETS_MAM_DAMPENED
271000	255966.17855
329000	350001.90227
401000	456886.11249
553000	656414.09775

Accuracy Measures:

Model	ME	RMSE	MAE	MPE	MAPE	MASE
ETS_MAM_DAMPENED	-41317.07	60176.47	48833.98	-8.3683	11.1421	0.8116

- ARIMA amortecido

Actual and Forecast Values:

Actual	ARIMA_MANUAL
271000	263228.48013
329000	316228.48013
401000	372228.48013
553000	493228.48013

Accuracy Measures:

Model	ME	RMSE	MAE	MPE	MAPE	MASE
ARIMA_MANUAL	27271.52	33999.79	27271.52	6.1833	6.1833	0.4532

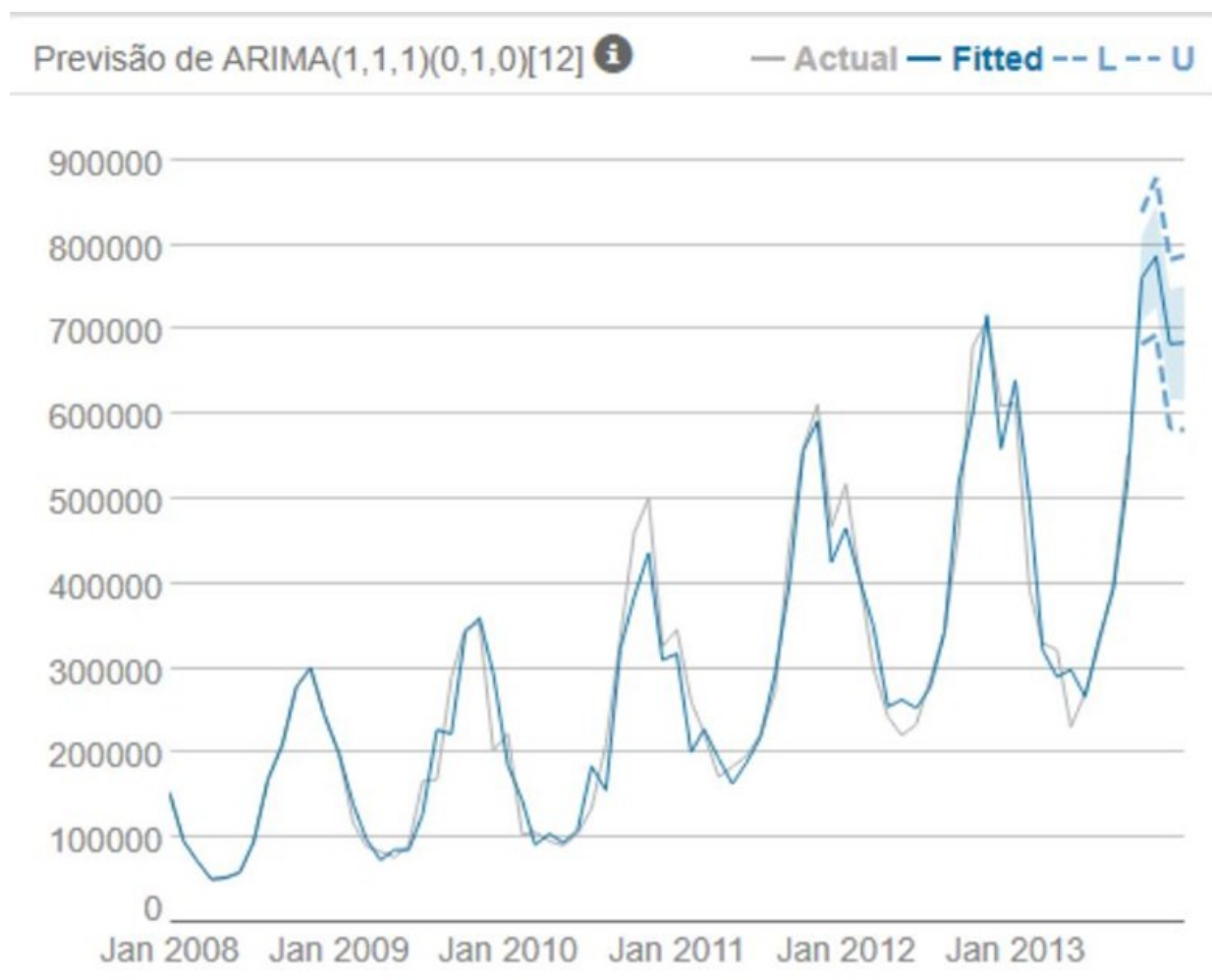
Se considerarmos as medições de erro na amostra, para o modelo ETS amortecido, RMSE é 33153.53 e MASE de 0.367; já para o modelo ARIMA, o RMSE é 36761.53 e o MASE é 0.364, o RMSE do modelo ETS amortecido é inferior ao modelo ARIMA, mas se comparamos o MASE dos dois modelos, a diferença é pequena.

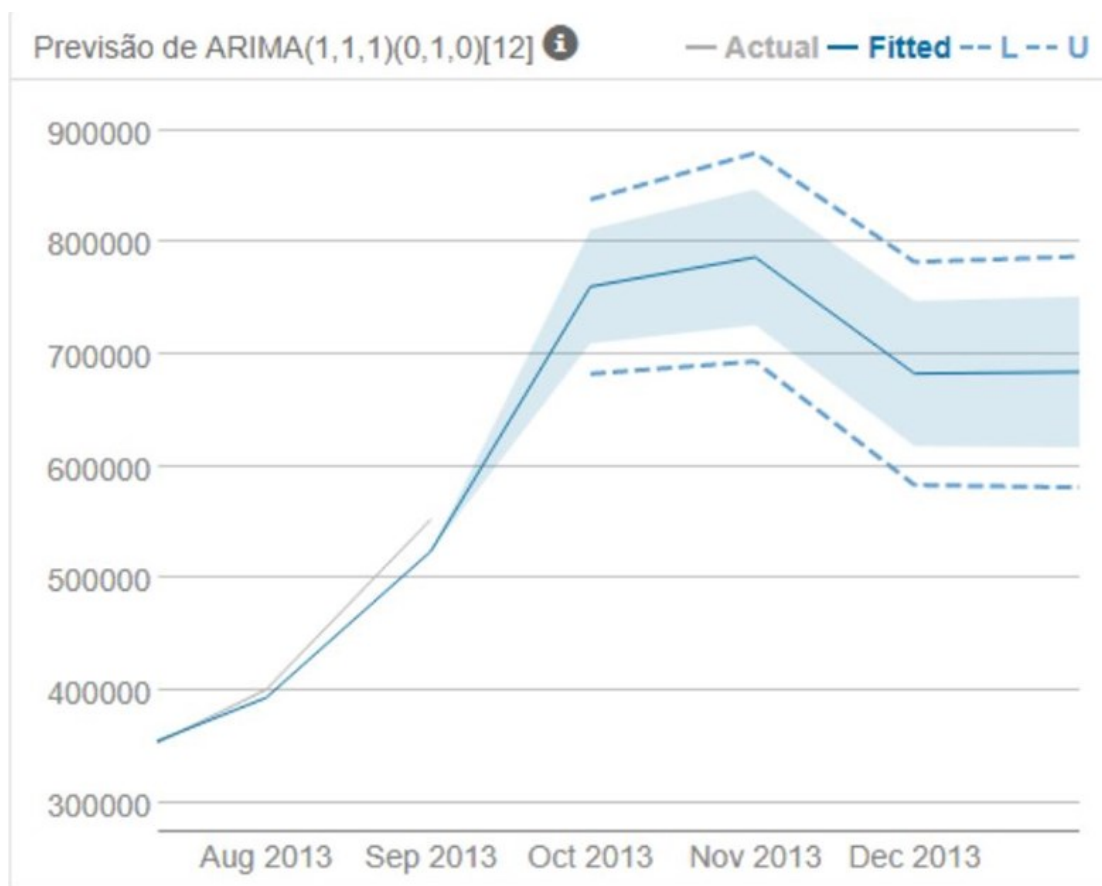
No entanto, se considerarmos as previsões de vendas dos modelos contra a amostra de holdout, o modelo ARIMA é melhor, já que o RMSE e MASE do modelo ARIMA são inferiores ao modelo ETS amortecido. Para o modelo ARIMA, o RMSE é

33999,79 e o MASE de 0,453; para o modelo ETS amortecido, o RMSE é 60176,47 e o MASE é 0,81. Portanto, o modelo ARIMA é o escolhido.

2. Qual é a previsão para os próximos quatro períodos?

Abaixo segue o gráfico de previsão para o período de 2013-10 até 2014-01 e os resultados nos intervalos de confiança de 95% e 80%.

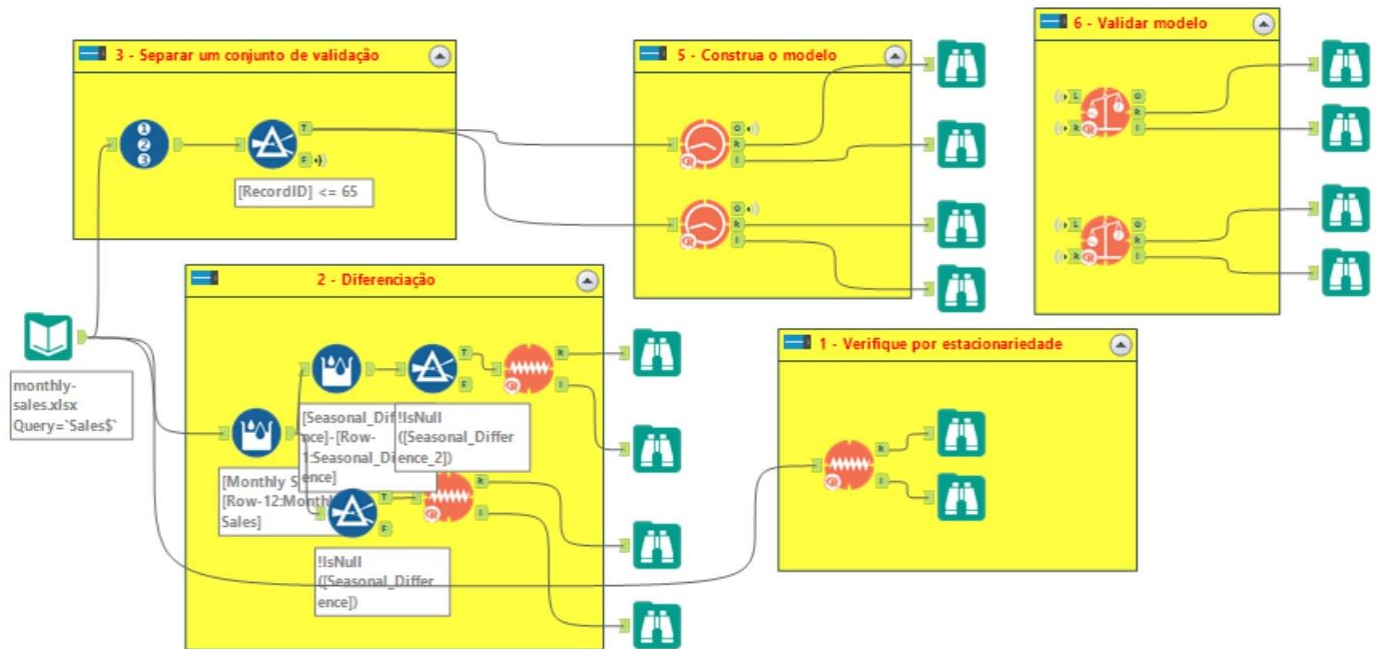




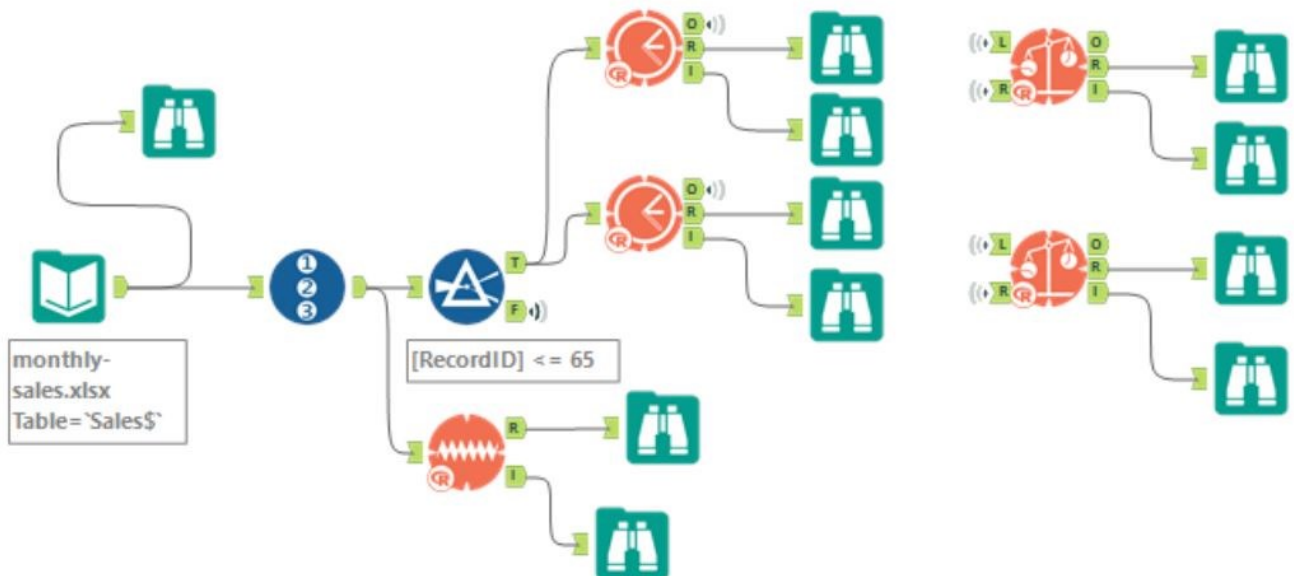
Period	Sub_Period	previsão	previsão_high_95	previsão_high_80	previsão_low_80	previsão_low_95
2013	10	754854.46	834046.2159	806635.166	703073.7541	675662.7041
2013	11	785854.46	879377.7531	847006.0545	724702.8656	692331.167
2013	12	684854.46	790787.8282	754120.5664	615588.3537	578921.0919
2014	1	687854.46	804889.2866	764379.4199	611329.5002	570819.6335

Alteryx Workflow

1- Modelo ARIMA



2- Modelo ETS



3- Forecasting

