

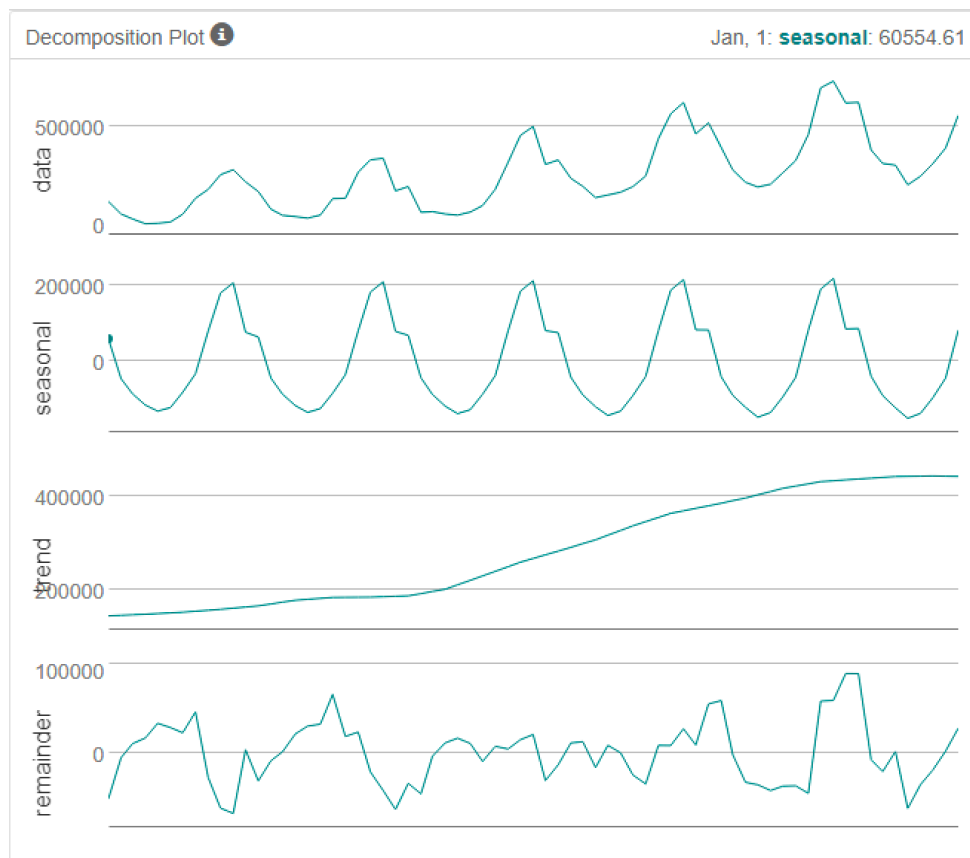
Projeto: Previsão de vendas

Passo 1: 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 “monthly_sales” de uma empresa que vende videogames, analisando os dados no arquivo chegamos a conclusão que o conjunto de dados atende as especificações da série temporal, já que a ordem das vendas importa, temos o numero de vendas de forma sequencial e em intervalos de tempo iguais, temos dados de 2008-01 ate 2013-09, portanto temos varias instancias do mesmo mes dentro do conjunto de dados.
2. Quais registros devem ser usados como amostra de retenção?
O tamanho da amostra de retenção depende de quão longa e a serie temporal e de quantos períodos de tempo queremos prever, o objetivo e prever as vendas dos próximos 4 meses, portanto pegamos os últimos 4 registros como amostra de retenção, de 2013-06 ate 2013-09.

Passo 2: Determine os componentes tendência, sazonalidade e erro

1. Qual é a tendência, a sazonalidade e o erro da série temporal? Mostre como você conseguiu determinar os componentes usando gráficos de séries temporais. Inclua esses gráficos.



A serie temporal original e dividida em três sub-series, determinando os componentes sazonal, de tendencia e de erro. A serie temporal possui as características de repetição dentro de um período de tempo de a cada 12 meses, depois podemos observar o gráfico de tendencia o qual reflete o comportamento da serie temporal em longos períodos de tempos, podemos apreciar uma tendencia crescente nesse caso, o gráfico de "remainder" estima o componente de erro da serie temporal.

Passo 3: Construa seus modelos

Analise seus gráficos, determine as medidas apropriadas para serem aplicadas aos seus modelos ARIMA e ETS e descreva os erros de ambos os modelos (limite de 500 palavras).

Responda à seguinte pergunta:

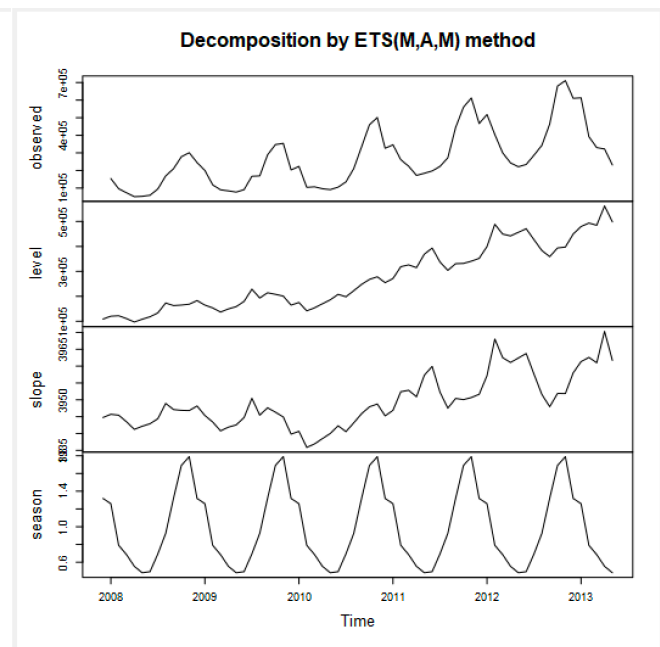
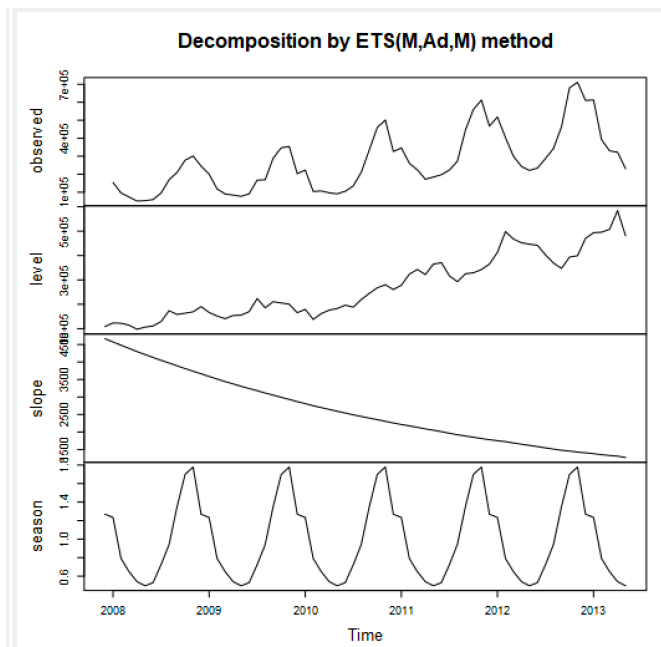
1. Quais são os termos modelo para o ETS? Explique por que você escolheu esses termos.
 - a. Descreva os erros na amostra. Use pelo menos RMSE e MASE ao examinar os resultados.

A sazonalidade e crescente ligeiramente ao longo do tempo, então aplicamos o modelo multiplicativamente. Há um comportamento crescente nos dados, então existe uma tendência. A tendência é linear, então aplicamos o modelo de forma aditiva. O erro está aumentando ou diminuindo ao longo do tempo, aplicamos o erro multiplicativamente, ja que apresenta variabilidade. Em seguida, executamos o ETS amortecido (dampened) e não amortecido (M, A, M), vamos a analisar os resultados:

Modelo ETS MAM Dampened

VS

Modelo ETS MAM



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:

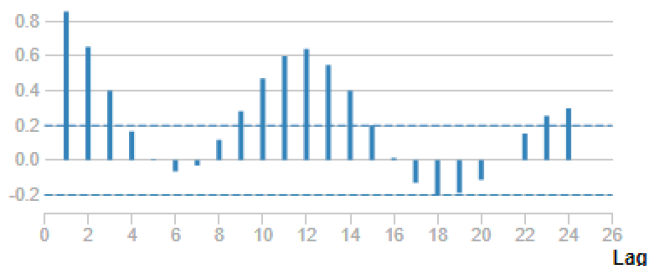
AIC	AICc	BIC
1639.7367	1652.7579	1676.7012

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 32992.73 e MASE é 0.372. 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 a sua vez, o MASE do modelo ETS amortecido é inferior ao modelo ETS não amortecido. Podemos concluir que o modelo ETS amortecido é o escolhido, dado que o AIC é um pouco inferior ao modelo ETS não amortecido.

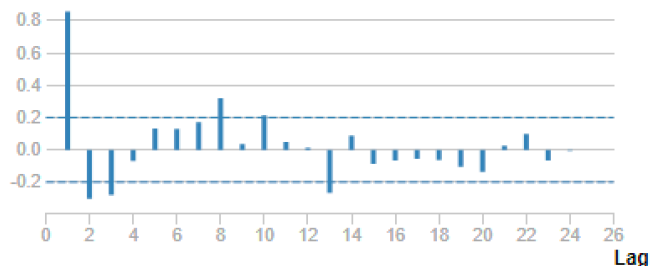
2. Quais são os termos modelo para o ARIMA? Explique por que você escolheu esses termos. Crie um gráfico com a função de correlação automática (ACF) e lotes de função de autocorrelação parcial (PACF) para as séries temporais e o componente sazonal e use esses gráficos para justificar a escolha dos termos do modelo.
 - a. Descreva os erros na amostra. Use pelo menos RMSE e MASE ao examinar os resultados.
 - b. Refaça os gráficos ACF e PACF tanto para a série temporal como para a diferença sazonal e inclua esses gráficos em sua resposta.

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, existem algumas tendências ou componentes sazonais e, portanto, 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 serie temporal.

ACF

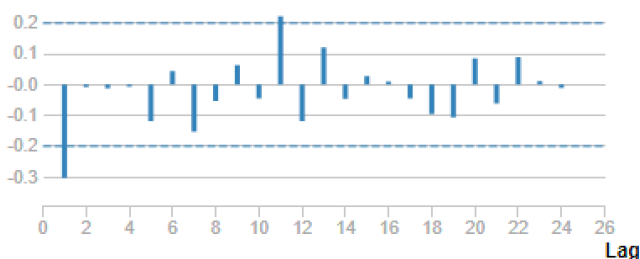


PACF

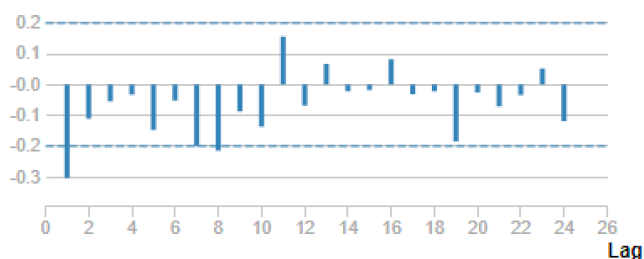


Notamos que o ACF mostra uma oscilação, indicando uma série sazonal, notamos que os picos ocorrem com defasagens de 12 meses e 24 meses. Além disso, observamos que um pico na defasagem 1 em um gráfico de ACF indica uma forte correlação entre cada valor da série e o valor anterior. Logo, precisamos ajustar a série temporal para estacionária. Séries não estacionárias podem ser corrigidas por uma transformação como a diferenciação. Consideramos a primeira diferença sazonal, podemos observar no gráfico abaixo como a série temporal foi estacionada.

ACF

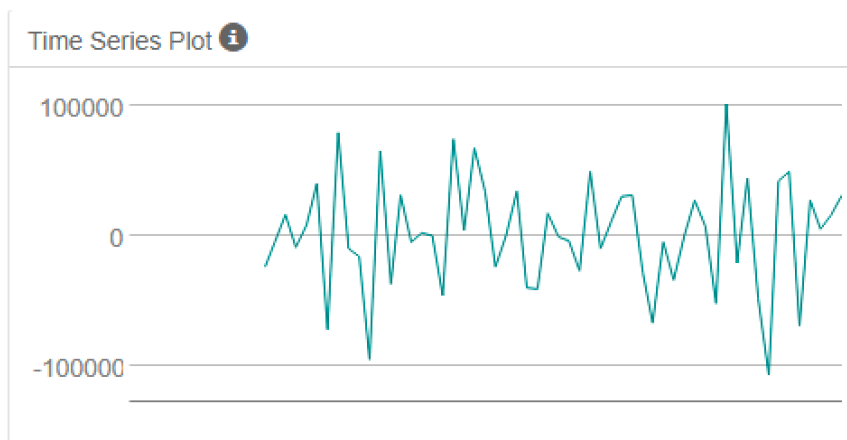


PACF



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, examinamos 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, notamos 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:

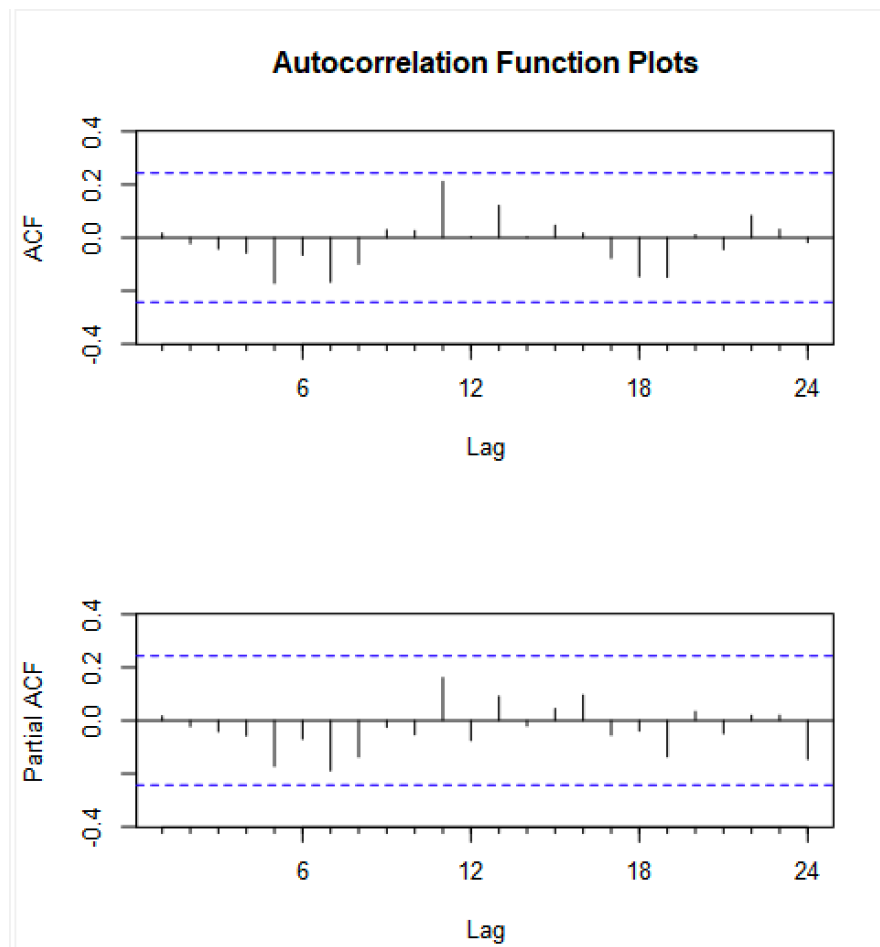
AIC	AICc	BIC
1256.5967	1256.8416	1260.4992

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

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

Abaixo, podemos apreciar o ACF e PACF:



Observamos que o modelo não apresenta correlação. Sendo assim, confirmamos que temos um modelo ARIMA estacionário pronto.

Passo 4: Previsão

1. Qual modelo você escolheu? Justifique sua resposta mostrando: medições de erro na amostra e medidas de erro de previsão contra a amostra de retenção.

Na análise dos modelos ETS anteriores, especificamente nos modelos ETS amortecido e não amortecido, apreciamos que o ETS amortecido apresentou um menor índice de AIC, portanto estabeleceremos 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	NA
ETS_MAM_DAMPENED	-41317.07	60176.47	48833.98	-8.3683	11.1421	0.8116	NA

- ARIMA Amortecido:

Actual and Forecast Values:

Actual	ARIMA_MANUAL
271000	263228.48013
329000	316228.48013
401000	372228.48013
553000	493228.48013

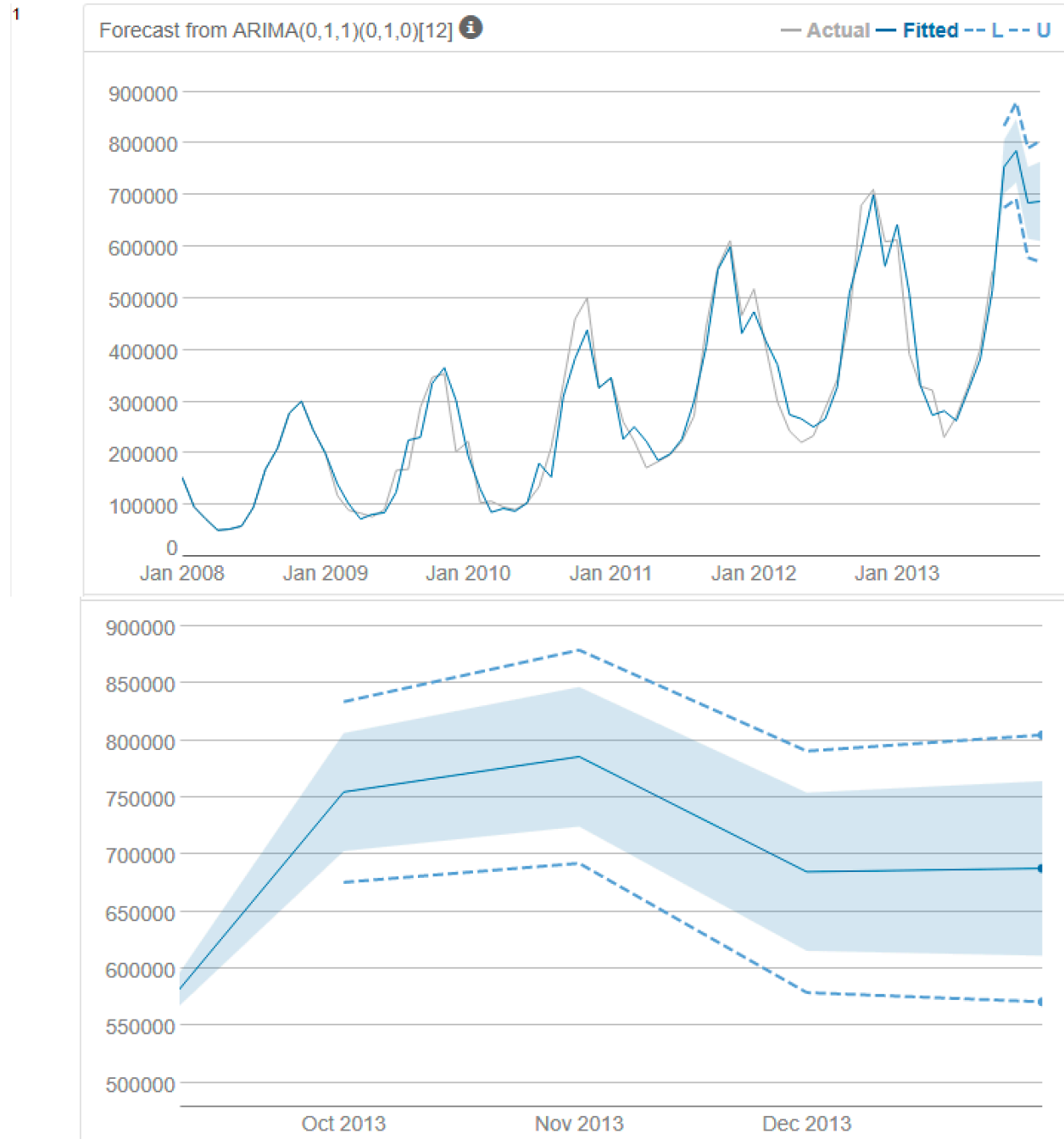
Accuracy Measures:

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

Se considerarmos as medições de erro na amostra, para o modelo ETS amortecido, RMSE é 33153.53 e MASE 0.367. 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 compararmos 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 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? Crie um gráfico com os resultados, usando intervalos de confiança de 95% e 80%.

Apresentamos, o gráfico de previsão para o período de 2013-10 to 2014-01:



Period	Sub_Period	forecast	forecast_high_95	forecast_high_80	forecast_low_80	forecast_low_95
6	10	754854.460048	834046.21595	806635.165997	703073.754099	675662.704146
6	11	785854.460048	879377.753117	847006.054462	724702.865635	692331.166979
6	12	684854.460048	790787.828211	754120.566407	615588.35369	578921.091886
7	1	687854.460048	804889.286634	764379.419903	611329.500193	570819.633462