Project #6
Forecasting Video Game Sales

Claudia Dai

February 2019

1 Plan your analysis

The data meets the criteria of a time series dataset.

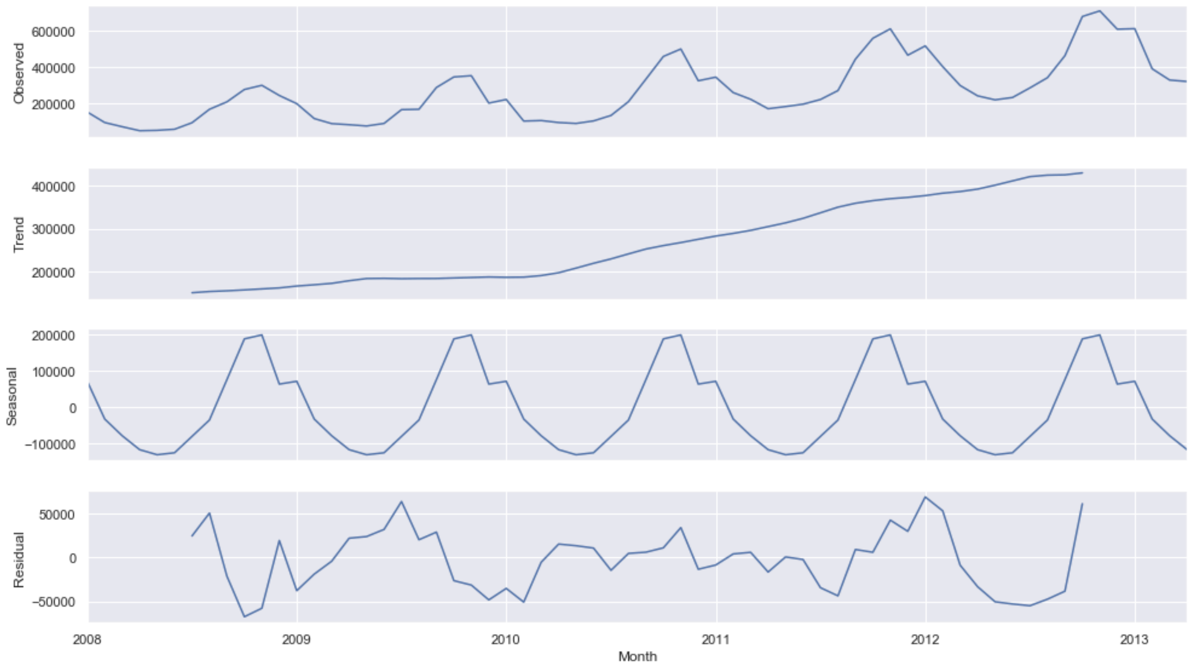
	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Month	69	69	2008-11	1	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Monthly Sales	69.000	NaN	NaN	NaN	276623.188	166398.892	51000.000	154000.000	243000.000	347000.000	711000.000

- data points are spread over continuous time interval
- data points are sequential
- two consecutive measurements are of equal intervals
- each time unit has at most one data point

	Sales
Month	
2008-01-01 00:00:00	154000
2008-02-01 00:00:00	96000
2008-03-01 00:00:00	73000
2008-04-01 00:00:00	51000
2008-05-01 00:00:00	53000

We need to predict the sales for the next 4 months, so we will use a holdout sample from the last four given months, i.e. June 2013 until Sep 2013.

2 Determine trend, seasonal, and error components (/250)



We have the scenario: Trend-Linear (see decomposition graph), Seasonal-Constant (sales amount doesn't change, constant every November).

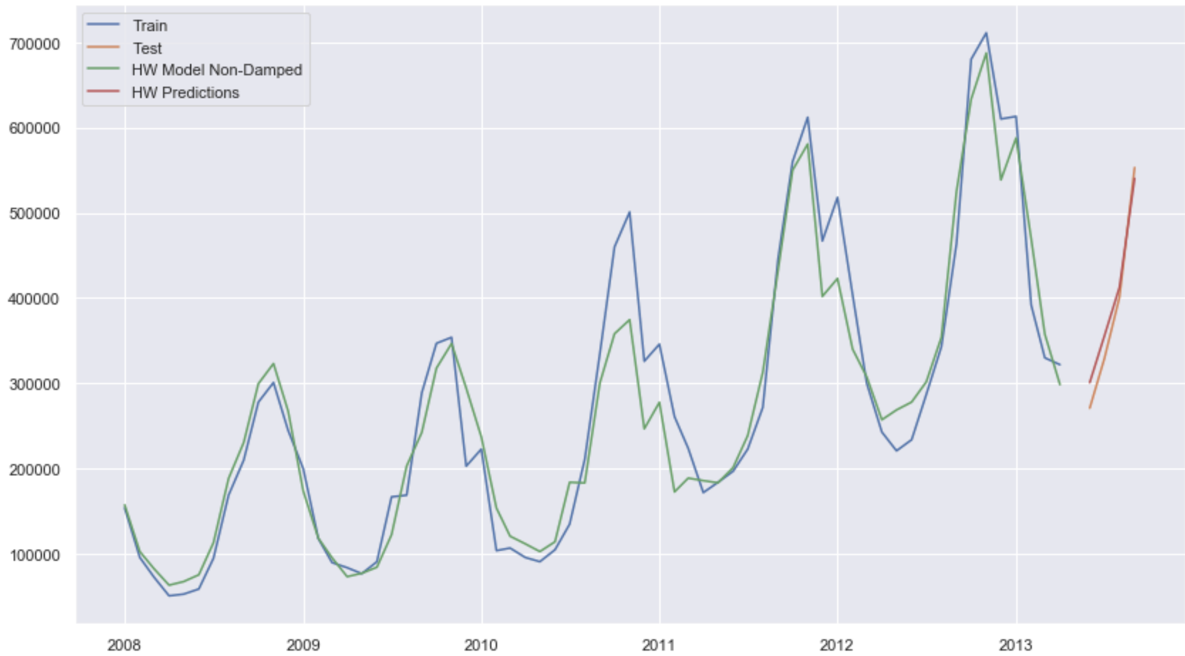
- Trend goes up linear, i.e. Additive
- Seasonality remains constant, i.e. Additive (see table excerpt below)
- Error has no trend but has fluctuations, i.e. Multiplicative

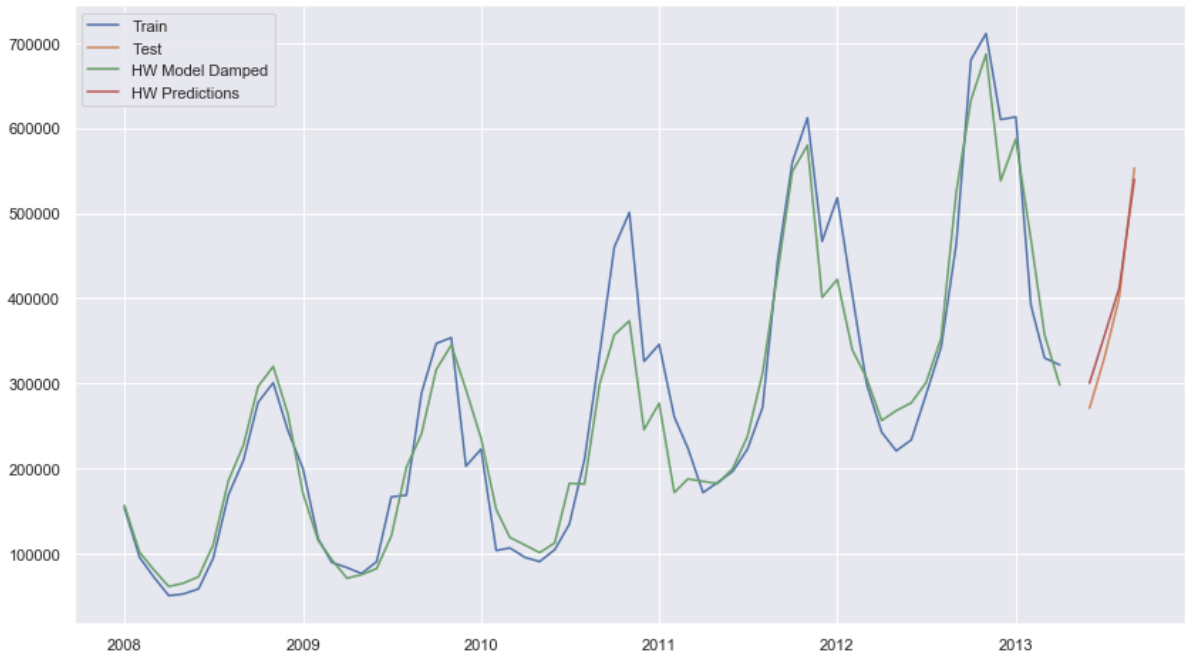
	Sales
Month	
2008-11-01	199068.750
2009-11-01	199068.750
2010-11-01	199068.750
2011-11-01	199068.750
2012-11-01	199068.750

3 Build your models (/500)

3.1 ETS model

Our ETS model terms are **ETS(M,A,A)**, which was derived from analyzing the decomposition in the previous section. We have a yearly seasonal period, i.e. 12 months, and ran a damped and a non-damped Holt-Winters model for comparison.





We look at a number of performance metrics, including:

- MSE (mean squared error)
- RMSE (root mean spared error)
- MDAE (median absolute error)
- MAE (mean absolute error)
- MSLE (mean squared logarithmic error)
- MAPE (mean absolute percentage error)
- R2 (R-squared)
- MASE (mean absolute scaled error)

Out of all of those, two key components to look at are the RMSE, which shows the in-sample standard deviation, and the MASE, which can be used to compare forecasts of different models.

Holt-Winters Non-Damped:

	0	1	2	3	4	5	6
Metric	MSE	RMSE	MDAE	MAE	MSLE	MAPE	R2
Value	474691762.828	21787.422	19580.530	20219.567	0.005	6.083	0.957

Holt-Winters Damped:

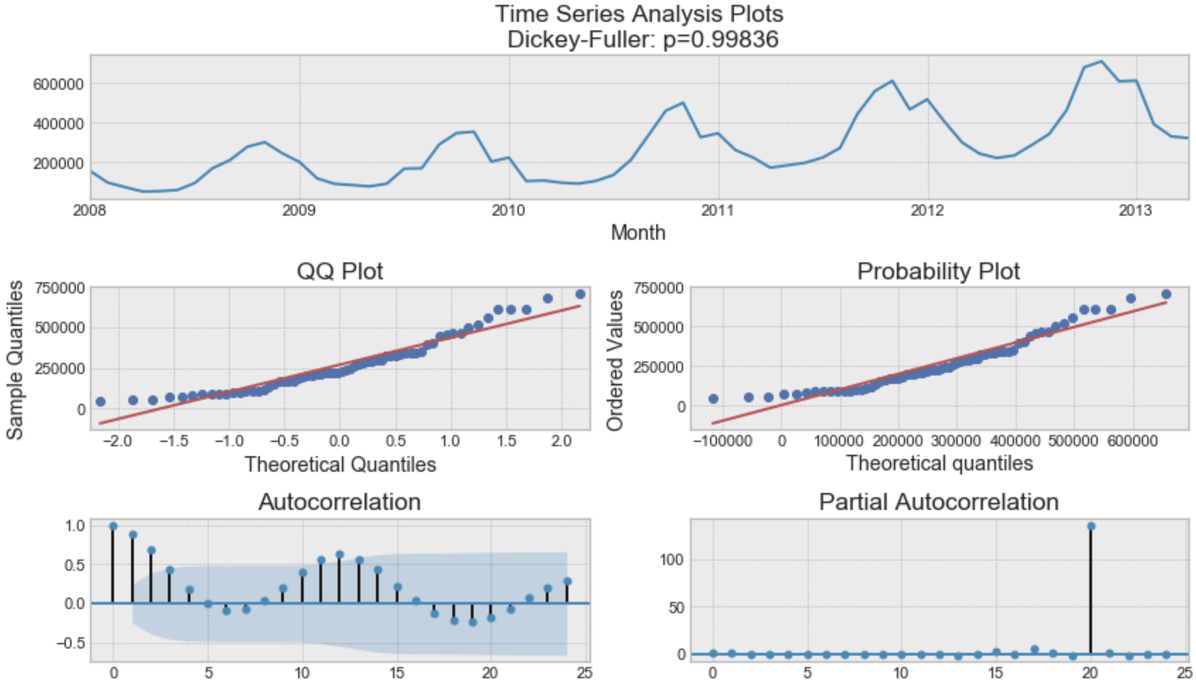
	0	1	2	3	4	5	6
Metric	MSE	RMSE	MDAE	MAE	MSLE	MAPE	R2
Value	451190217.594	21241.239	19652.580	19787.888	0.004	5.918	0.960

The damped model has a lower MSE, RMSE, MAE and MAPE. The non-damped model has a smaller MDAE, and MSLE and R2 are about the same for both models. The MASE for non-damped is 0.3388, while the MASE for the damped model is 0.3316. The MASE result can be reviewed in the Python notebook.

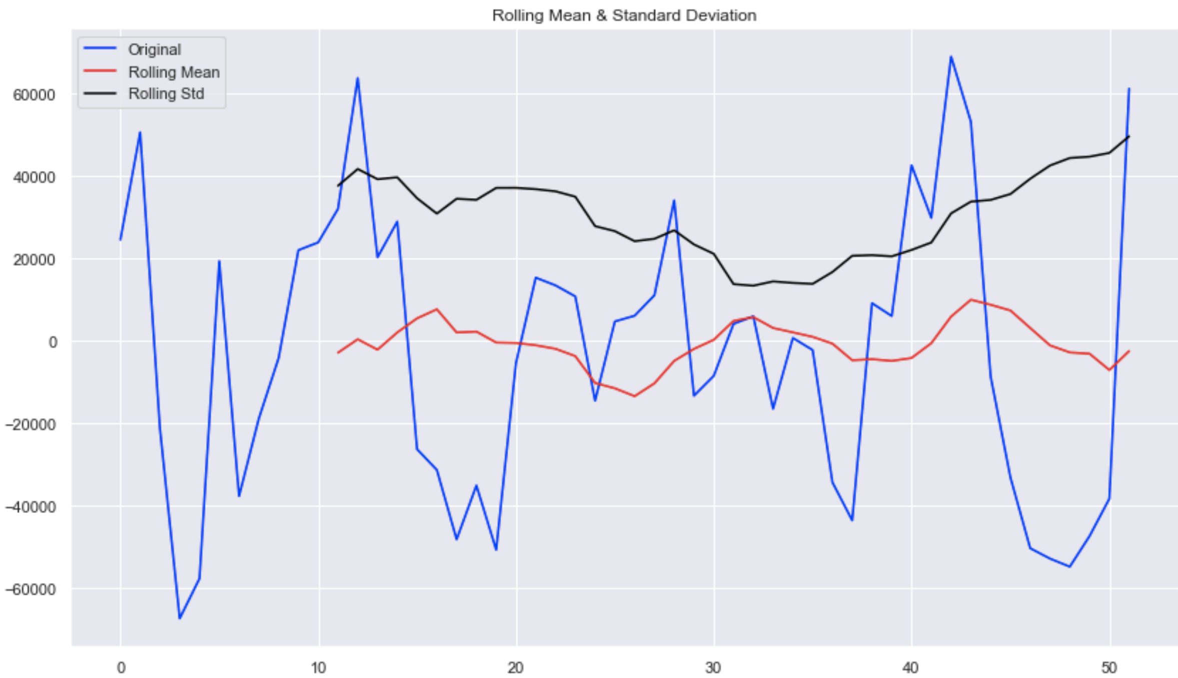
Between these two, we would choose the **damped** model for higher accuracy. Based on the RMSE, we can see that our variance is about 22000 units around the mean. The MASE of the damped model shows a good forecast at 0.33, with its value falling below the generic 1.0, the commonly accepted MASE threshold for model accuracy.

Bonus: After running a grid search for hyperparameters, the search yielded the same parameters for the ETS model as we had manually analysed. However, the search preferred a non-damped model. Please see the code notebook for details.

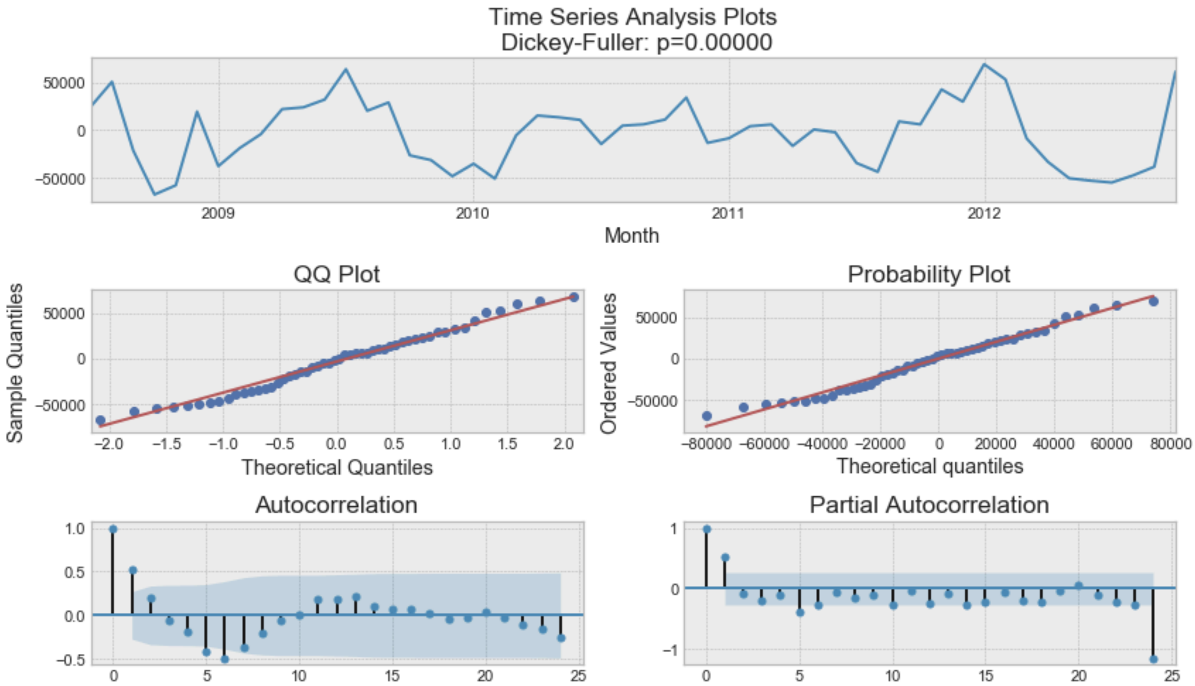
3.2 (S)ARIMA model



Our original dataset is not stationary. After differencing the seasonality and trend, our residual from the decomposition passes the Dickey-Fuller test. The test statistic is significantly lower than the 1% critical value. So our residual data is very close to stationary.




```
Results of Dickey-Fuller Test:
Test Statistic                 -5.647
p-value                         0.000
#Lags Used                      5.000
Number of Observations Used    46.000
Critical Value (1%)            -3.581
Critical Value (10%)           -2.602
Critical Value (5%)            -2.927
dtype: float64
```



Our ACF and PACF plots also don't show as much autocorrelation anymore as with the original data. Yay!

From the ACF and PACF plots we can collect valuable information for our (S)ARIMA hyperparameter optimization grid search.

- p is most probably 5 since it is the last significant lag on the PACF, after which, most others are not significant.
- q should be somewhere around 6 as well as seen on the ACF
- P none of the seasonal lags in PACF seem to be significant, hence 0, but maybe it could be 2...
- Q none of the seasonal lags in ACF seem to be significant, hence 0.

Our best model turns out to be **SARIMAX(1, 1, 1)x(0, 1, 0, 12).**

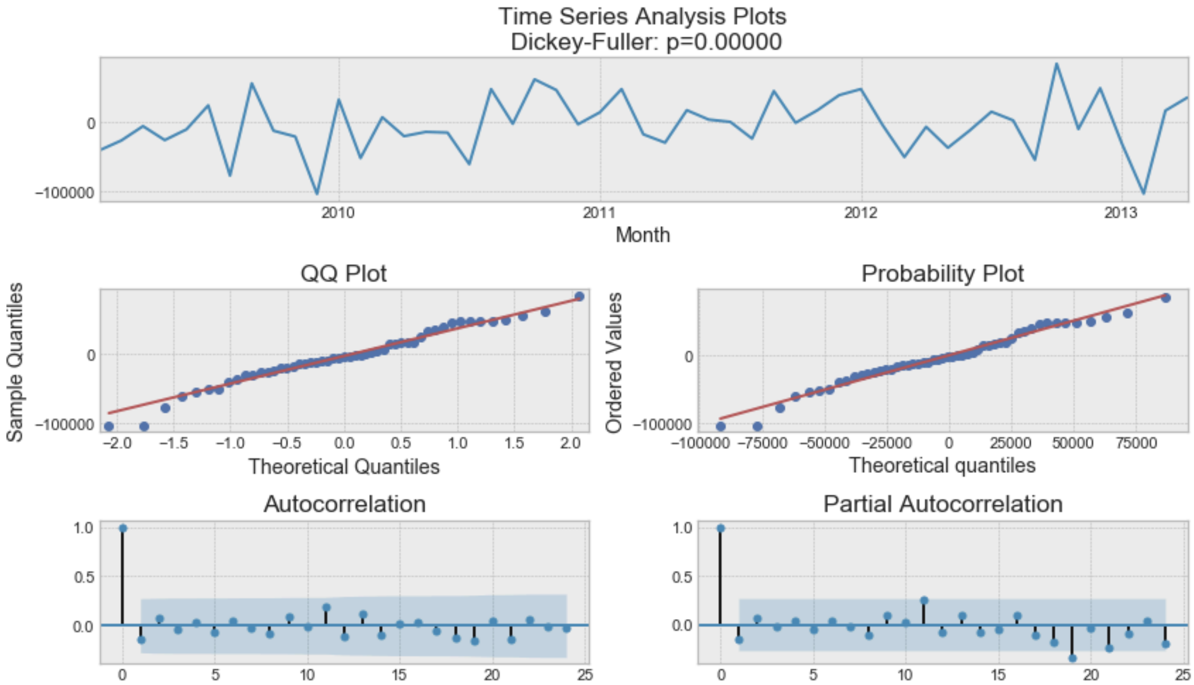
Statespace Model Results

Dep. Variable:	Sales	**No. Observations:**	64
Model:	SARIMAX(1, 1, 1)x(0, 1, 0, 12)	**Log Likelihood**	-613.014
Date:	Thu, 27 Dec 2018	**AIC**	1232.027
Time:	18:46:49	**BIC**	1237.823
Sample:	01-01-2008	**HQIC**	1234.242
	- 04-01-2013		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	0.7140	0.126	5.647	0.000	0.466	0.962
ma.L1	-0.9939	0.224	-4.432	0.000	-1.433	-0.554
sigma2	1.801e+09	1.25e-10	1.44e+19	0.000	1.8e+09	1.8e+09

Ljung-Box (Q):	34.69	**Jarque-Bera (JB):**	0.83
Prob(Q):	0.71	**Prob(JB):**	0.66
Heteroskedasticity (H):	1.22	**Skew:**	-0.31
Prob(H) (two-sided):	0.69	**Kurtosis:**	3.13

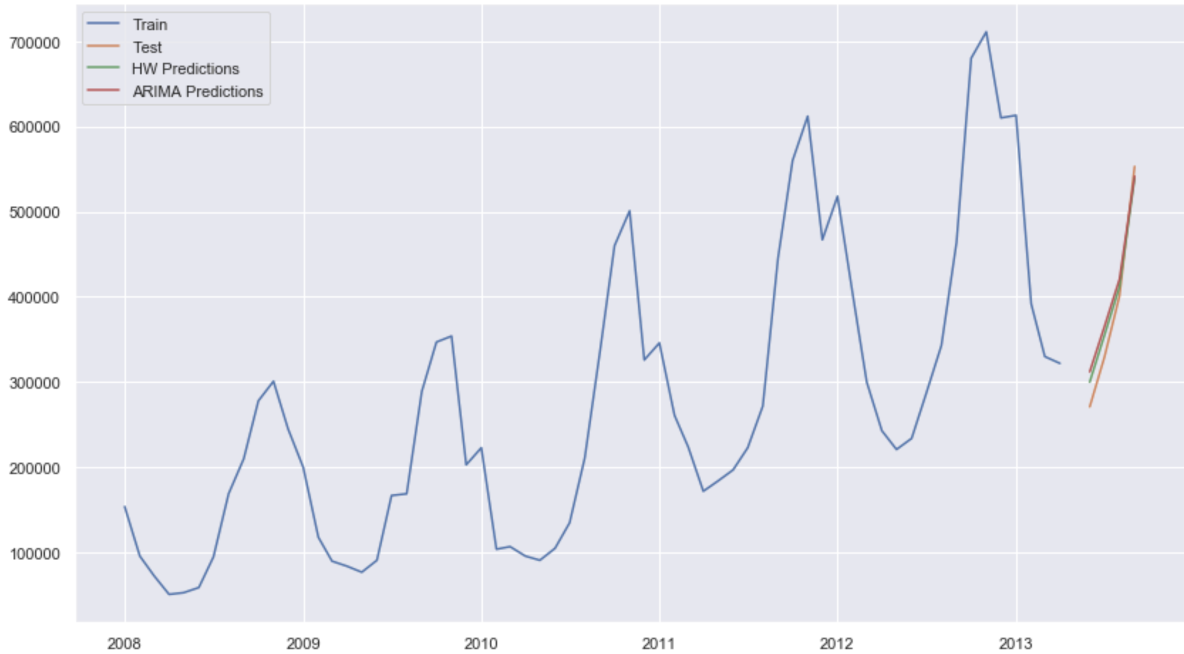
When we inspect the residuals, we see that there is no autocorrelation. Again, yay!



	0	1	2	3	4	5	6
Metric	MSE	RMSE	MDAE	MAE	MSLE	MAPE	R2
Value	874444559.740	29571.009	27784.567	27035.469	0.008	8.266	0.922

The MASE of the (S)ARIMA model is 0.4530, which, again, can be reviewed in the Python notebook.

4 Model comparison



If we compare the error measurements for both models, we see that the damped Holt-Winters model yields lower errors than the (S)ARIMA model.

Holt-Winters Damped:

	0	1	2	3	4	5	6
Metric	MSE	RMSE	MDAE	MAE	MSLE	MAPE	R2
Value	451190217.594	21241.239	19652.580	19787.888	0.004	5.918	0.960

$$MASE = 0.3316$$

(S)ARIMA:

	0	1	2	3	4	5	6
Metric	MSE	RMSE	MDAE	MAE	MSLE	MAPE	R2
Value	874444559.740	29571.009	27784.567	27035.469	0.008	8.266	0.922

$$MASE = 0.4530$$

The damped model has a lower MSE, RMSE, MDAE, MAE, MSLE, MAPE, and MASE, and a higher R2. Between the two, we would choose the **damped** model for higher accuracy. Based on the RMSE, as mentioned before, we can see that our variance is about 22000 units around the mean. The MASE of the damped model shows a good forecast at 0.33, with its value falling below the generic 1.0, the commonly accepted MASE threshold for model accuracy. It is slightly lower than ARIMA's MASE of 0.4530.

5 Forecast

The predictions for the upcoming four months are:

Damped:
```
2013-10-01    750740.217
2013-11-01    779390.732
2013-12-01    670872.884
2014-01-01    670807.384
Freq: MS, dtype: float64
```

ARIMA:
```
2013-10-01    757467.575
2013-11-01    788400.640
2013-12-01    687352.848
2014-01-01    690318.725
Freq: MS, dtype: float64
```