TIMES SERIES ANALYSIS - PROJECT

FITTING A MODEL TO IPHONE SALES IN INDIA IN MILLIONS

(FROM 2007 Q2 - 2018 Q1)

MAY 17, 2021

BY **KUSH PATEL** AND **OGUMOSHABE DISAN**

TIME SERIES ANALYSIS VT - 2021

UPPSALA UNIVERSITET

Introduction

India's production of iphones has grown faster and the iphone sales have boomed in the last couple of years. In fact,the iphone sales statistics in India by data.world indicated a drastic increase in iphone sales in millions from 2007 Q2 to 2018 Q1.

India's technology has continued to be ranked among the best in producing durable, fancy and sophisticated electronics including smart phones, computers, digital watches, televisions, biometric machines among others.
Therefore, we have found it important to analyse some of the past data of India's sales of Iphones. The purpose of this project is to try and fit an ARIMA or SARIMA model to India's Iphone sales in millions between 2007 Q2 to 2018 Q1 as provided by data.world and thoroughly predict future values with the developed model.

Data Analysis

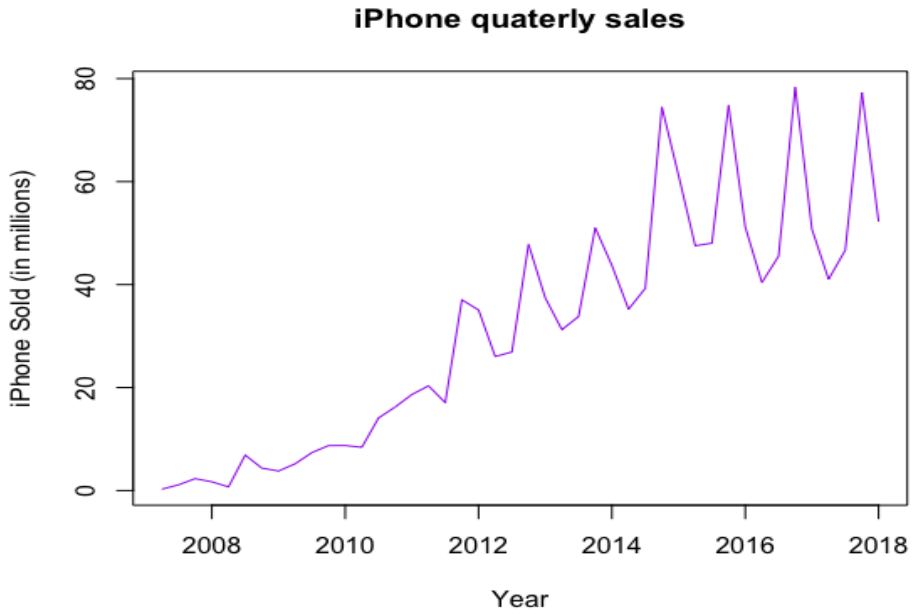

Figure 1: iPhone quarterly sales in millions.

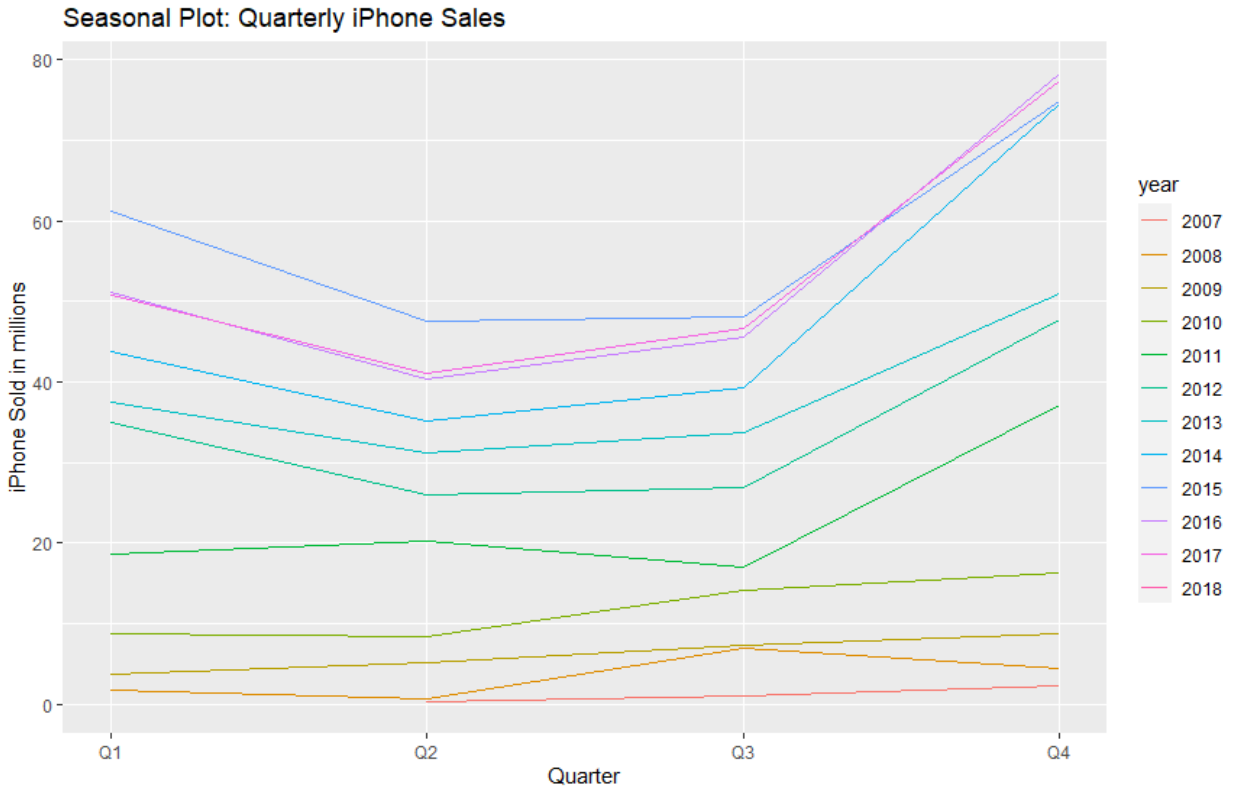

Figure 2: Seasonal plots of Quarterly iPhone sales in millions

Figure 2 shows the quarterly sales relationship with different years. As we can see here, the iPhone sales drastically increased in the fourth quarter because of the result of the new iPhone release every september. And it is lowest at the end of the third quarter because people are anticipating a new iPhone in the market so they do not want to purchase an iphone which will become the second generation after a few months.

In figure 3, We give the ACF and PACF of the series of iPhone sales and we notice ACF decaying slowly and the PACF cut sharply after the fifth lag. Usually for seasonal patterns either ACF or PACF should cut off after fourth lag but in this case it does after fifth lag, it may be because on the top of season there might be one step correlation. Since ACF is slowly reducing and PACF is cutting off, it is recommended to try seasonal AR 1 models.

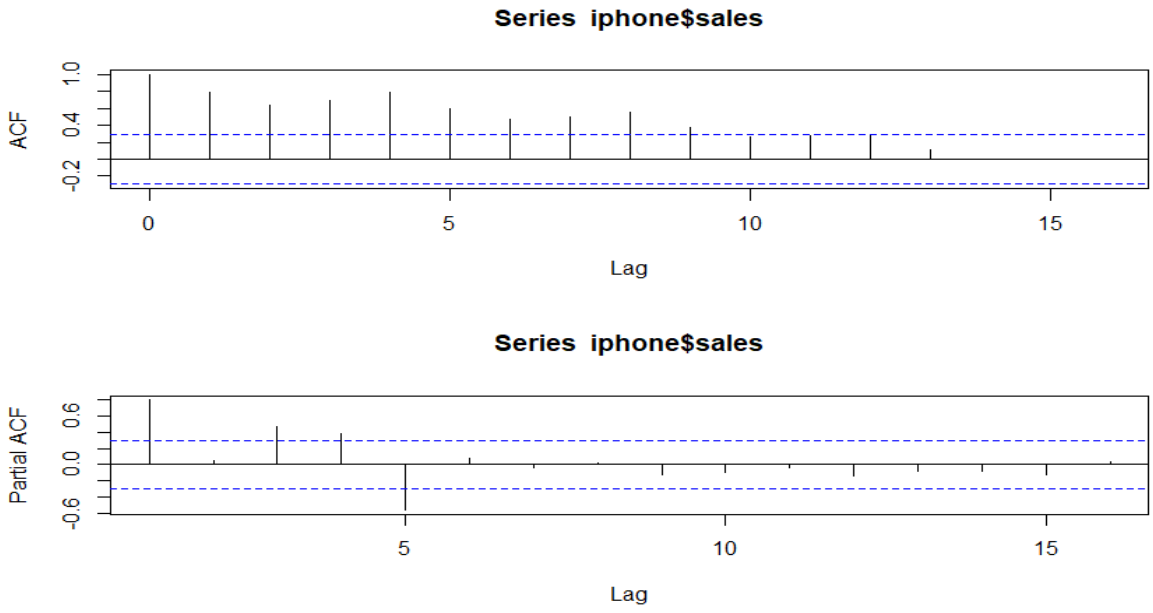

Figure 3: ACF and PCF of iPhone quarterly sales in millions.

After checking the lowest AIC, causality, invertibility, and coefficient significantly different then zero, it was determined that ARIMA(1,0,0)(1,0,0)[4] was the best fit model for the data.
In figure 4, we give ACF and PACF of residuals of iPhone sales and we notice that the ACF and the PACF are within the dotted bound, so there is no more seasonal structure left.

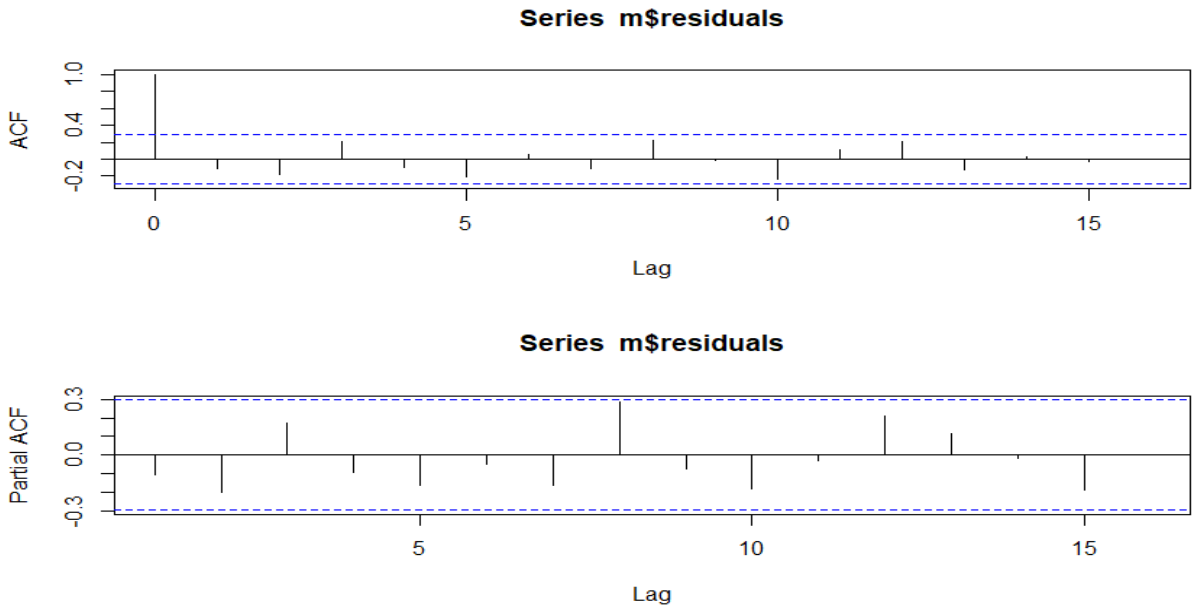

Figure 4: ACF and PACF of residuals of iPhone quarterly sales in millions

In figure 5, We plot the histogram of residues and the normal Q-Q plot and we notice the histogram fairly shows normality and follows normal distribution. The Q-Q plot also shows an increasing linear trend as clearly shown by the red line. However we discover no clear patterns of the residuals or any autocorrelation in between and this suggests a valid model.

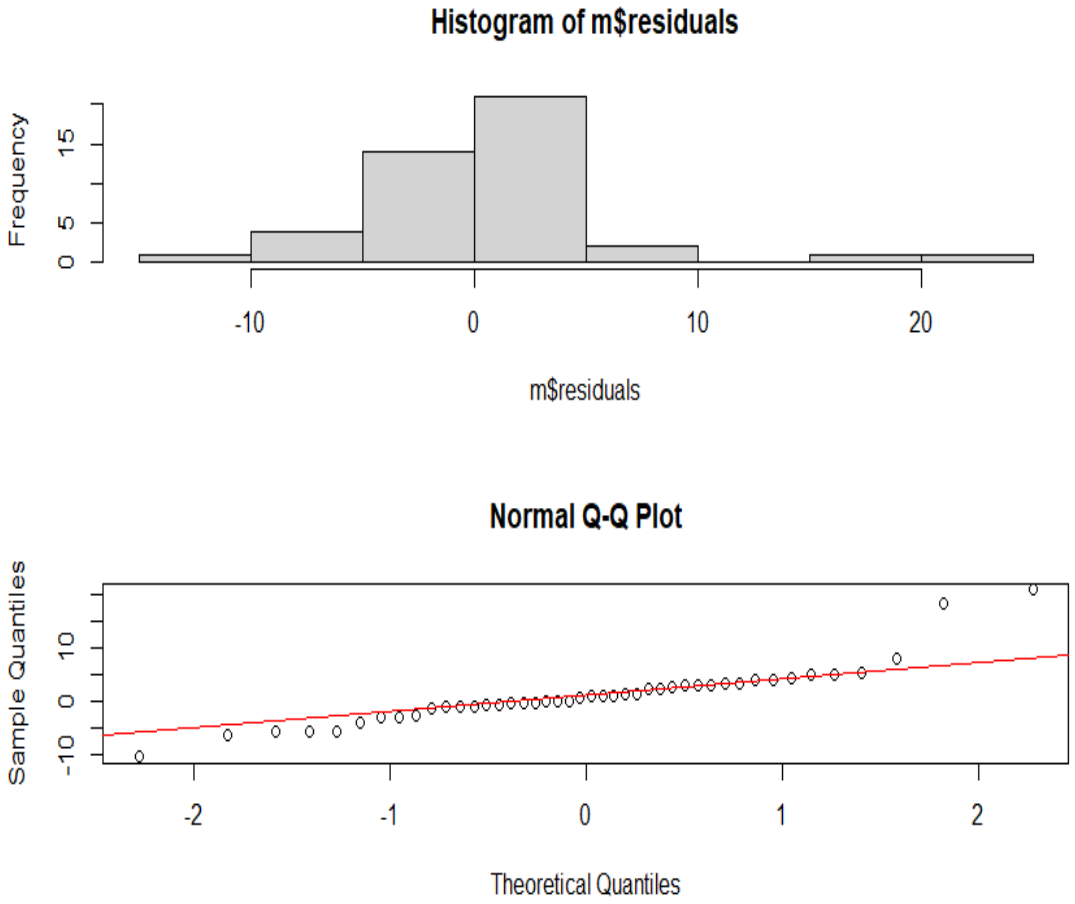

Figure 5: Histogram and normal Q-Q plot of residuals of iPhone quarterly sales in millions

In figure 6, displays residuals from ARIMA(1,0,0)(1,0,0)[4] and show that it is a good model to use since we notice normal fitting on the histogram. It is noticed that there is no visible structural pattern in residuals graph and again ACF is within dotted lines.

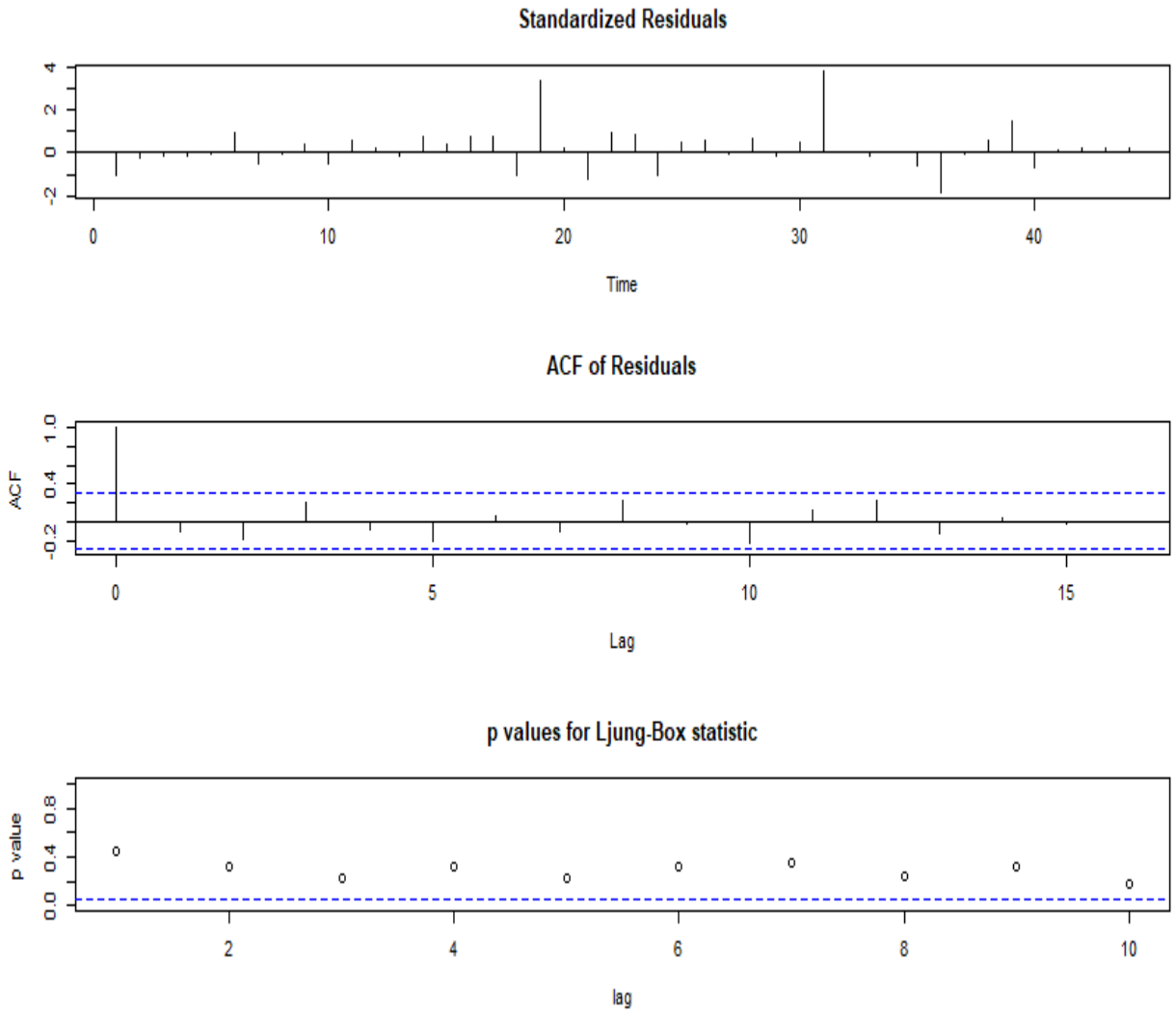

Figure 6: Standardized residuals, ACF and p values for Ljung-Box statistics.

Ljung-Box statistics test if there is any autocorrelation within the residuals. If p limit is within 0.05 then we reject the null hypothesis of no autocorrelation. But if the graph values are above 0.05 mark, we can say that there is no autocorrelation within the residuals. The model passes all the diagnostics, so we can accept this model.

Seasonal patterns of the model can also be seen by plotting a periodogram. With the use of danielle smoothing window (default in R), it was determined that smoothing with span of 4 was best fit because there is no unexpected peak and there is no visible plateau which might indicate oversmoothing.
In figure 7, there is an expected peak at 0.25 because of 4 quarters patterns. The peak at zero is because of trend and at 0.5 is the resonance peak that follows from 0.25.

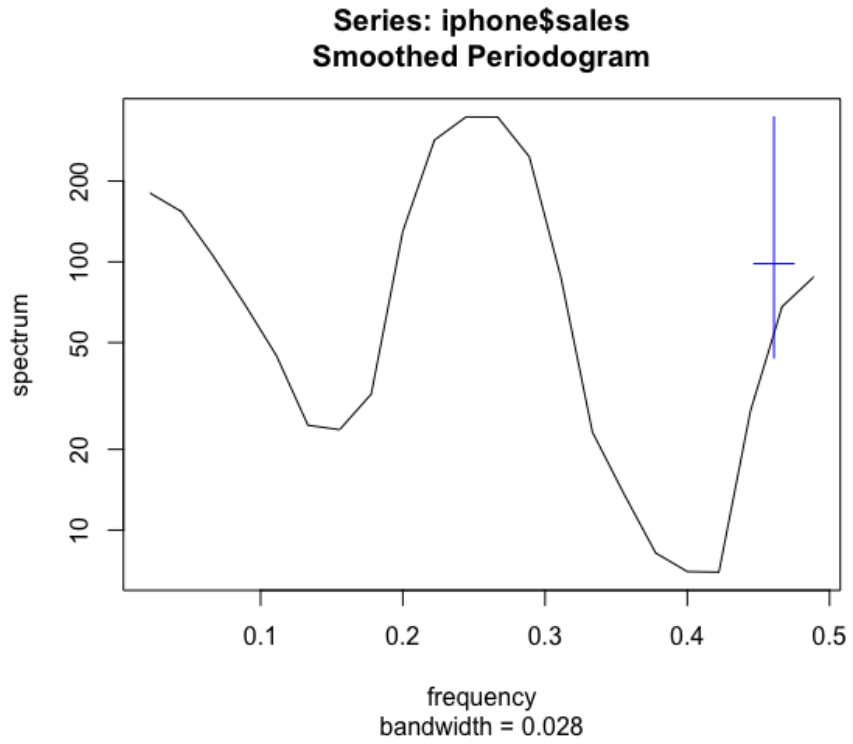

Figure 7: Smoothed periodogram of iPhone quarterly sales in millions.

Forecasts

Figure 8 below shows forecasts, we try to examine how well the model can forecast future values. The plot shows fitted values and ten step ahead-forecast from the ARIMA(1,0,0)(1,0,0)[4] which gives the 80% prediction interval (dark blue) and 95% prediction intervals (light blue) which particularly brings volatility and well predictions. As seen in the graph, the seasonal pattern still exists and the iphone sales will fairly stabilize in next fews quarters similar to past quarters.

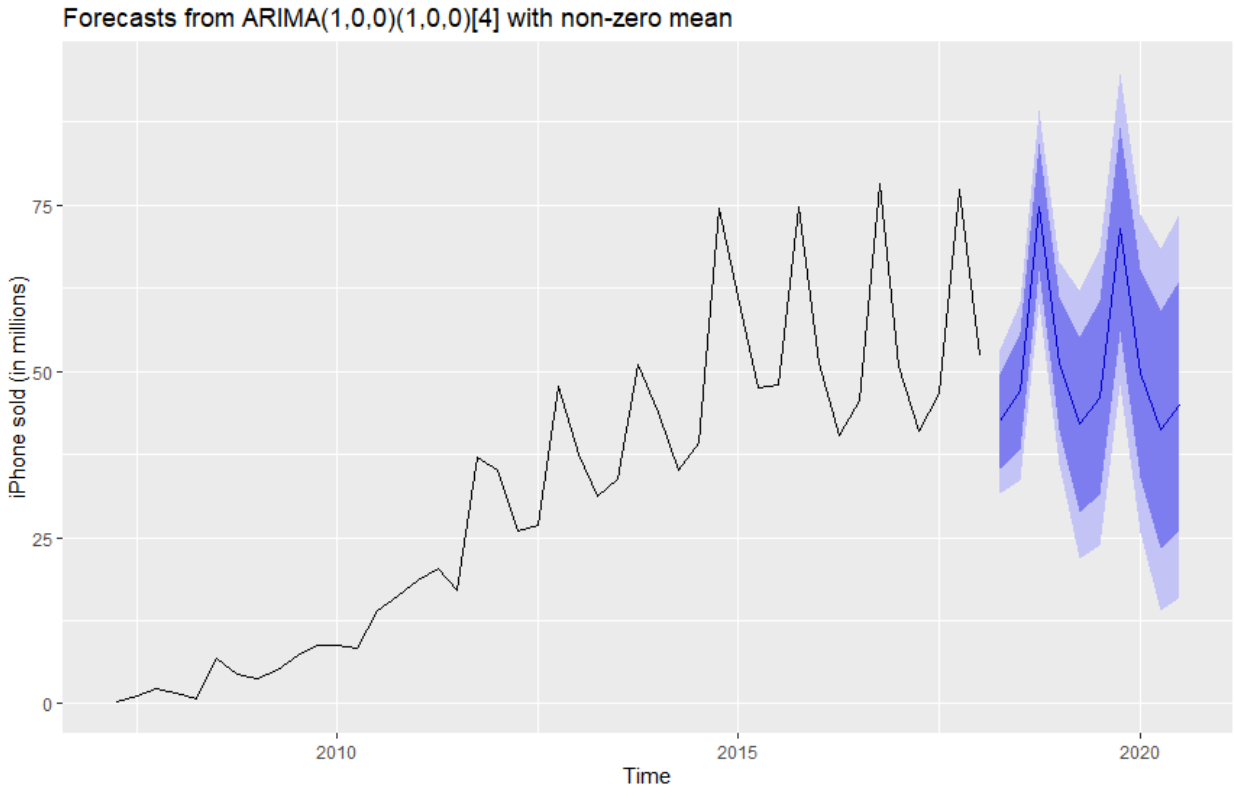

Figure 8: Forecasts of iPhone quarterly sales in millions

Forecast method: ARIMA(1,0,0)(1,0,0)[4] with non-zero mean

Model Information:

Call:
arima(x = Y, order = c(1, 0, 0), seasonal = list(order = c(1, 0, 0), period = 4))

Coefficients:

	ar1	sar1	intercept
	0.7362	0.9189	29.1032
s.e.	0.0901	0.0459	21.7412

sigma^2 estimated as 30.31: log likelihood = -141.91, aic = 291.83

Error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	1.17605	5.50583	3.556587	-49.00184	68.05845	0.5602358	-0.1088898

Forecasts:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2018 Q2	42.42282	35.36682	49.47882	31.63159	53.21405
2018 Q3	46.99215	38.23003	55.75427	33.59165	60.39265
2018 Q4	74.68761	65.12709	84.24812	60.06606	89.30915
2019 Q1	51.28598	41.31938	61.25258	36.04338	66.52857
2019 Q2	42.03722	28.90115	55.17329	21.94733	62.12710
2019 Q3	46.05284	31.48412	60.62155	23.77191	68.33377
2019 Q4	71.36758	56.07829	86.65687	47.98463	94.75053
2020 Q1	49.76432	34.09828	65.43035	25.80519	73.72345
2020 Q2	41.19245	23.38220	59.00270	13.95402	68.43087
2020 Q3	44.82861	25.95763	63.69960	15.96793	73.68929

Bibliography

1. iPhone Sales Data set by Rafael Pereira
   https://data.world/rflprr/iphone-sales