Name: Alexis R. Feraer

Course: Time Series

Session: Fall 2014

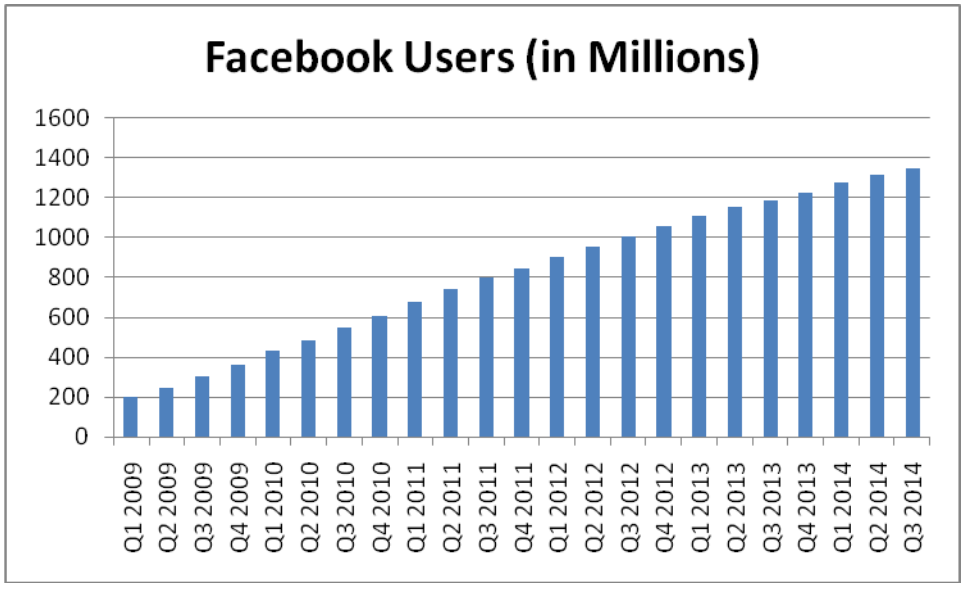
FACEBOOK USERS

INTRODUCTION

Due to technological advancements and increased dependence on the internet in the past decade, social media has quickly become a part of our everyday lives. Since its launch in 2004, Facebook has been among the most popular social networks with an active user count of 1.35 billion as of the 3rd quarter of 2014. For my student project, I will try to model the time series of active Facebook users around the world.

DATA

To qualify as an active user, one must have logged into Facebook during the past 30 days. The data is obtained from the following site:
http://www.statista.com/statistics/264810/number-of-monthly-active-facebook-users-worldwide/
The data is measured on a quarterly basis from 2009 to 2014. Based on the graph, it is evident that there is an upward trend on the observed time period.

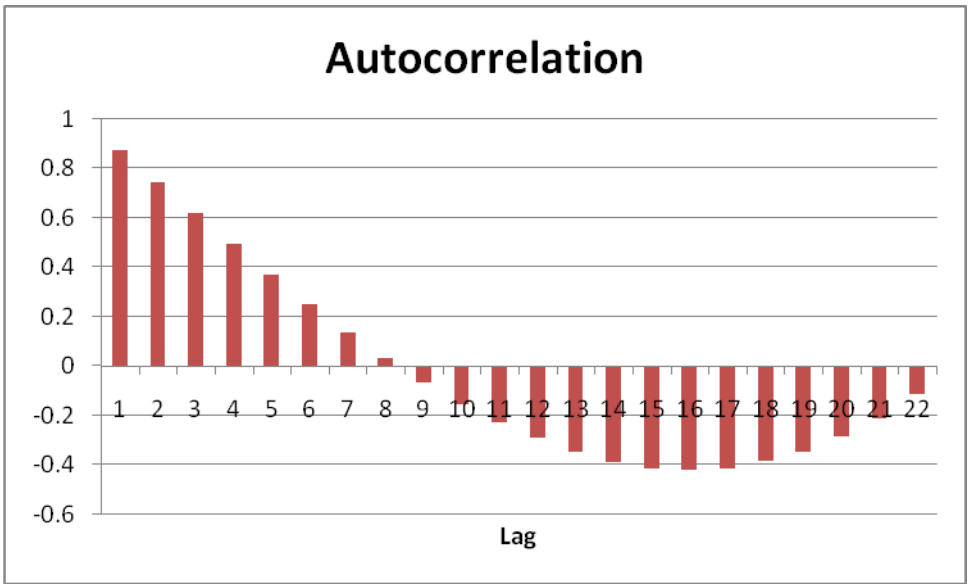


MODEL SPECIFICATION

To assess the model to be used, we calculate the sample autocorrelation function with the formula:

$$r_k = \frac{\sum_{t=k+1}^{n}(Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$

where k is from 1 to 22 (see Excel file for actual computations). Looking on the correlogram plot, the autocorrelations do not die down quickly but rather decrease exponentially as the number of lags k increases. This implies that AR(p) models could be suitable.



MODEL FITTING AND DIAGNOSTICS

To find the parameters for our AR(p) models, we use Excel's Regression Add-in tool. We start with AR(1) and the results are as follows:

AR(1)

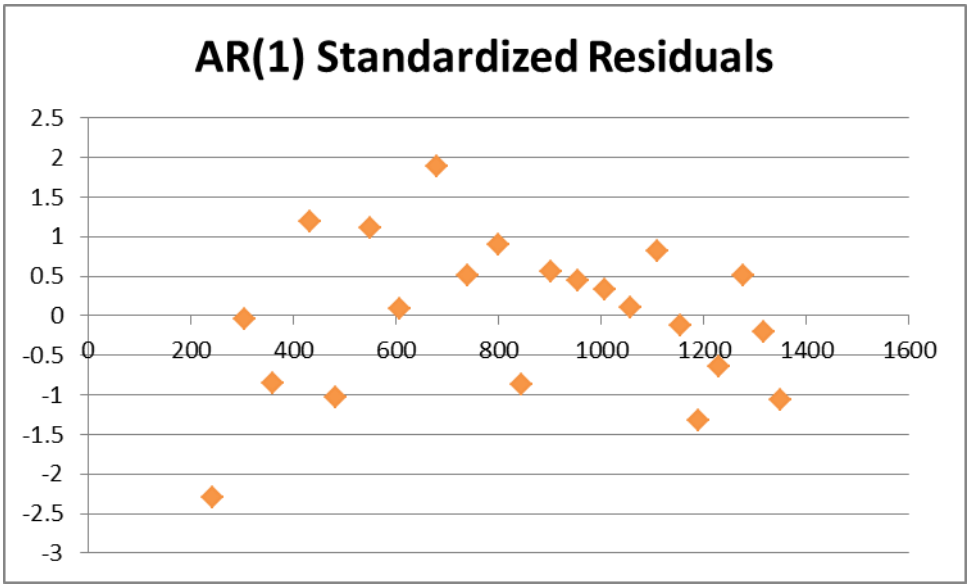
Regression Statistics	
Multiple R	0.999723
R Square	0.999447
Adjusted R Square	0.999419
Standard Error	8.414105
Observations	22

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	2556783	2556783	36114.2	4.76E-34
Residual	20	1415.943	70.79717		
Total	21	2558199			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	68.26086	4.46276	15.29566	1.68E-12
x	0.979995	0.005157	190.0374	4.76E-34

The fitted AR(1) model is $Y_t = 68.26086 + 0.979995Y_{t-1}$. The R square is 0.999447, meaning 99.94% of the variations of this time series is explained by the AR(1) model. Moreover, $|\Phi_1| = 0.979995$, which is <1, implies the model is stationary. The resulting standardized residuals plot also suggests no significant pattern, with only 1 point beyond 2 standard deviations.



AR(2)

Regression Statistics	
Multiple R	0.99978
R Square	0.999561
Adjusted R Square	0.999512
Standard Error	7.288346
Observations	21

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	2176549	1088275	20487.1	6.07E-31
Residual	18	956.1597	53.11998		
Total	20	2177505			

	Coefficients	*Standard Error*	*t Stat*	*P-value*
Intercept	77.73853	14.01969	5.544952	2.9E-05
x1	0.911134	0.20041	4.546352	0.00025
x2	0.062496	0.196871	0.317444	0.754559

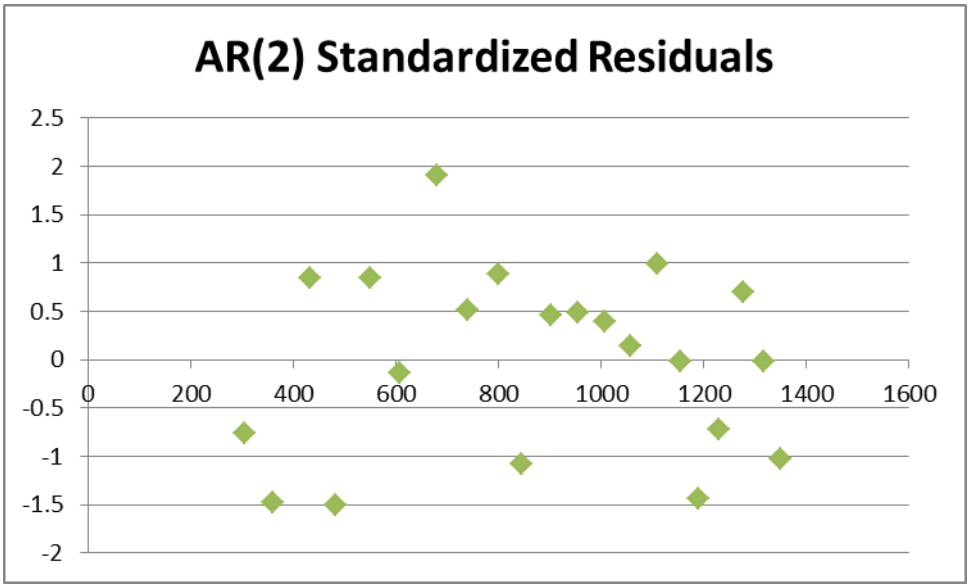
The fitted AR(2) model is $Y_t = 77.73853 + 0.911134Y_{t-1} + 0.062496\ Y_{t-2.}$ The R square is 0.999561, meaning 99.96% of the variations of this time series is explained by the AR(2) model. The model is also stationary as shown by:

a) $\Phi_1 + \Phi_2 = 0.97363 < 1$

b) $\Phi_2 - \Phi_1 = -0.84864 < 1$

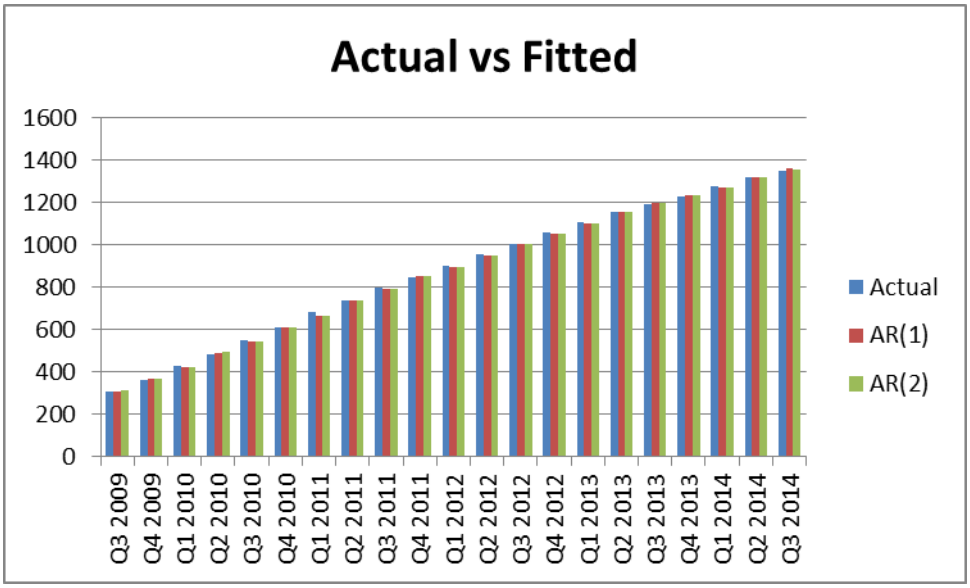
c) $|\Phi_2| = 0.062496 < 1$

The resulting standardized residuals plot also suggests no significant pattern, with all points within 2 standard deviations.



Actual vs. Fitted

Shown below is a graph of the actual data vs the AR(1) and AR(2) model:



Based on the figure above and the R square statistic, both models provide excellent estimates for our observed series. However, we can argue that although AR(2) has the higher R square, its added value is actually very minimal (0.000114) compared to the complexity added by the second Φ term. By the principle of parsimony, we therefore keep the model as simple as possible, and stick with AR(1).

CONCLUSION

After examining the above statistics and by the principle of parsimony, we can conclude that the AR(1) model of $\mathbf{Y_t = 68.26086 + 0.979995Y_{t-1}}$ is the most suitable model for determining the number of active Facebook users in the world.