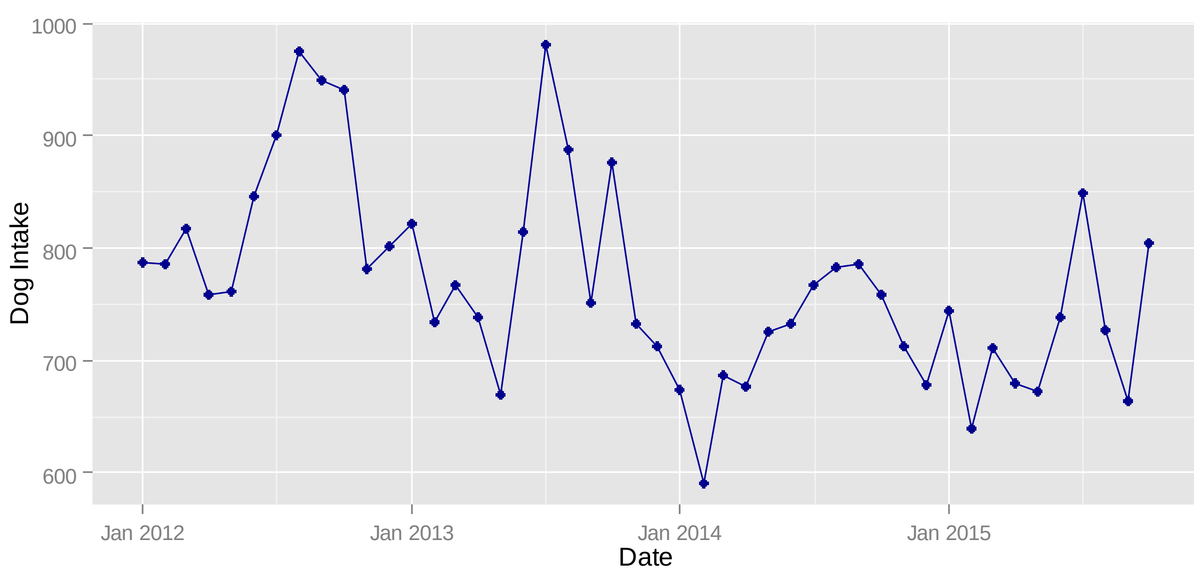
Time Series: A Study on Surrender and Adoption

While being an actuary certainly has it glory moments – studying for exams, studying for VEE credits, studying for Online Courses – it was not always the path I envisioned for myself. I had always wanted to be a veterinarian, for the most part based on the idea that I would get to play with animals for hours on end. Along with this dream job, I always hoped to own a rescue center for abandoned animals (still a dream, just a more distant, less likely one). I have volunteered at shelters, watched hours on end of The Dog Whisperer, dreamt of being the Skywalker to Cesar Millan's Obi-Wan. My dad currently owns seven dogs, all rescued from the local animal shelter and all rescued around the same time of year. This started me wondering if there is a pattern to the numbers of surrenders and adoptions animal shelters see – what a great time for a time series project!

The first thing I did was pick a shelter. Living and working in Philadelphia, it was fantastic to see that the Animal Care and Control Team of Philadelphia (AACT) had numbers on surrenders and adoptions of dogs and cats (and miscellaneous) by month for the past few years. Working on a project that can actually have some positive life-impact is always a plus (by-the-by, I am now volunteering for the shelter as reviewing the data was so depressing, it truly gave me no choice).

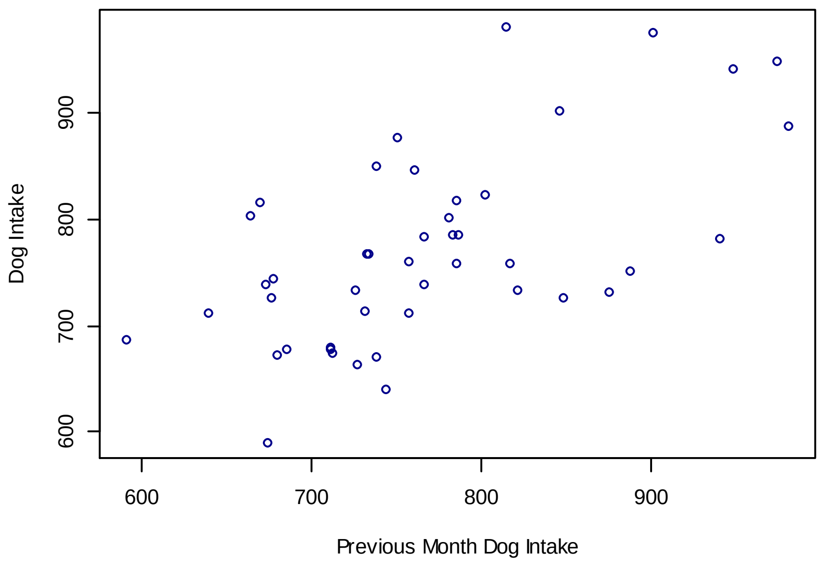
Dog Intake

I began by looking at the number of dogs the AACT had taken in starting in 2012. The graph of the data is below, Dog Intake by Date. The data is from January 2012 through October 2015, as more recent data was unavailable.

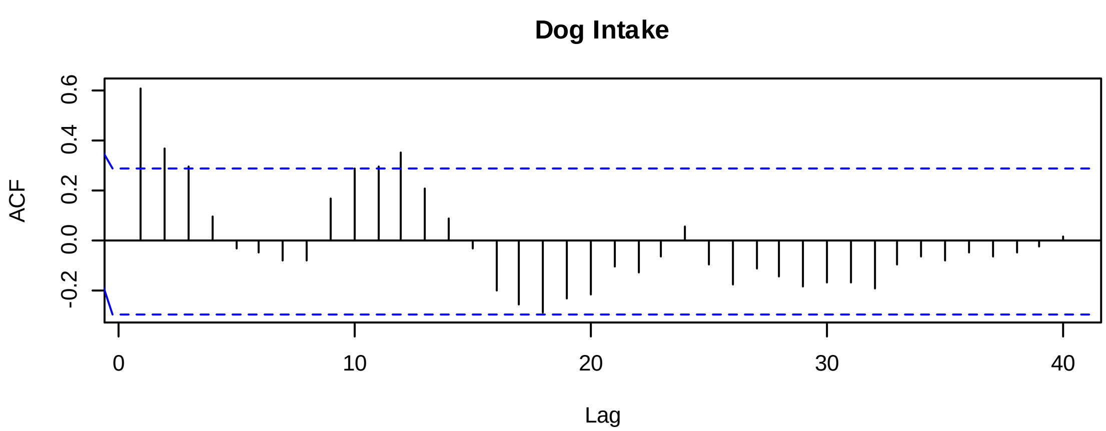


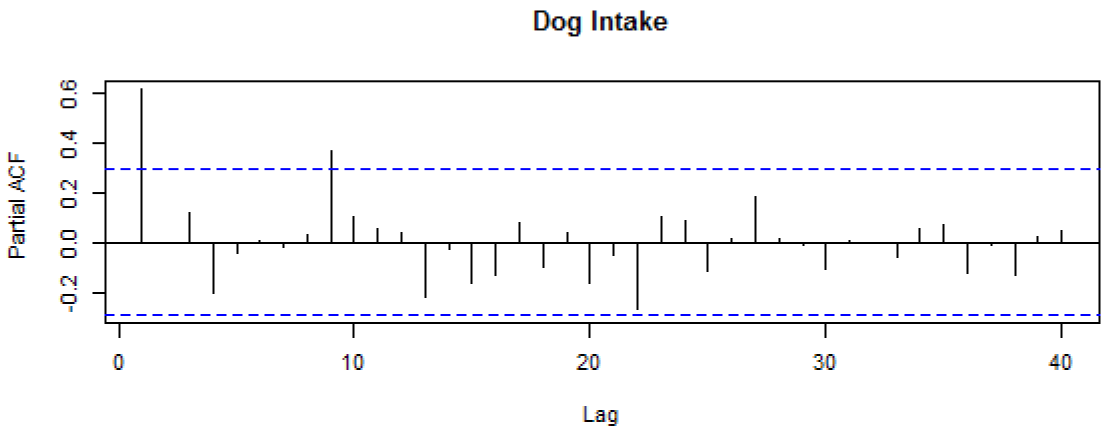
This time series doesn't give the appearance of seasonality; at first glance, it offers the impression of erratic behavior. However, upon further inspection, you can clearly see that it peaks annually around July or August and hits its low around February. This suggests that the shelter sees more dogs surrendered in the warmer months than the colder months. There are many possible explanations for this pattern; people are outside less in the winter and are therefore less likely to come across abandoned or stray dogs, dogs are more likely to breed in the spring which gives rise to more puppies in summer months, or, saddest of all, adorable Christmas presents evolve from the cutest thing ever to destructor of shoes and carpets and are “returned”.

In order to confirm my suspicion of seasonality, I looked at the correlation, plotting the number of dogs surrendered by month against the previous month’s surrenders; the correlation is 0.61091. You can clearly see a positive trend in the plot when you look at this graph of dogs surrendered versus its lag.



However, the data is not seasonal. I performed the Dickey-Fuller test to test the significance of the correlations of the Dog Intake series; the p-value of 0.03613 suggests a stationary series, so I did not need to take any differences. I verified this with the Box-Ljung test, resulting in a p-value of 8.256×10^{-7}. As can be seen in the ACF graph below, the autocorrelation early on is significant, but damped oscillatory behavior is strong in this series; correlations taper to zero by the end of the series. Because of this, I assume it is autoregressive. Taking into account its Partial ACF, I made the assumption of an AR(1) series.





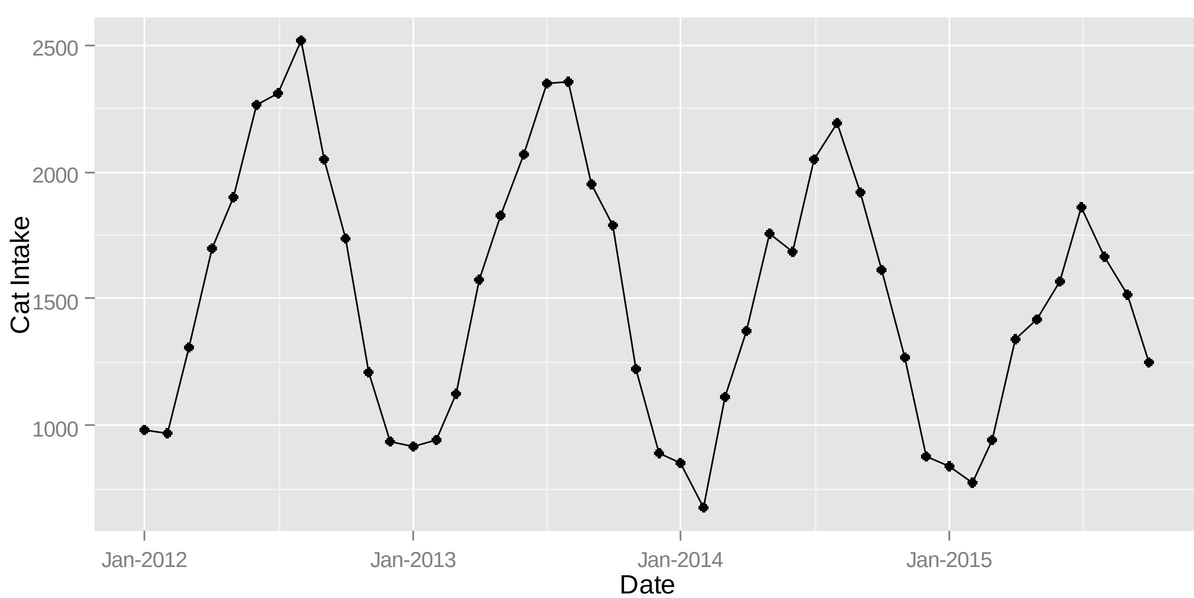
Parameter Estimation

Working with the assumption that my Dog Intake Series is AR(1), I ran some tests using Yule-Walker, Ordinary Least Squares, and Maximum-Likelihood Estimate. The parameter ϕ results returned were 0.6088, 0.6109, and 0.5987, respectively. The standard error of the estimates is $\sqrt{\frac{1-0.5987^2}{46}} = 0.135$ where 46 is *n*, the number of months used. This suggests the estimates are comparable.

I ran possibilities of AR(2), MA(1), and ARMA(1,0,1) in R, but the AIC was lowest for AR(1) and therefore allowed me to conclude that the Dog Intake Series is AR(1) with $Y_t = 0.5987Y_{t-1} + 771.0201$.

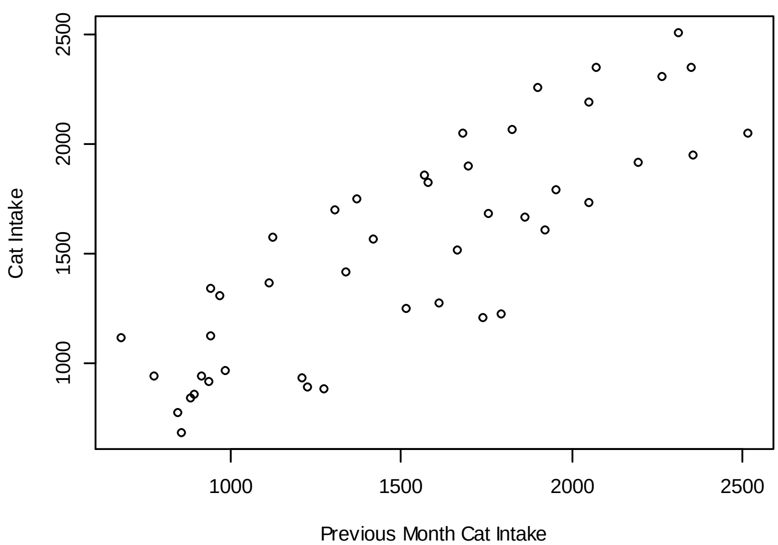
Cat Intake

I looked also at the number of cats the AACT had received since January 2012 so as not to discriminate. These numbers produced a much more clear-cut graph than the Dog Intake numbers, which I was not expecting.

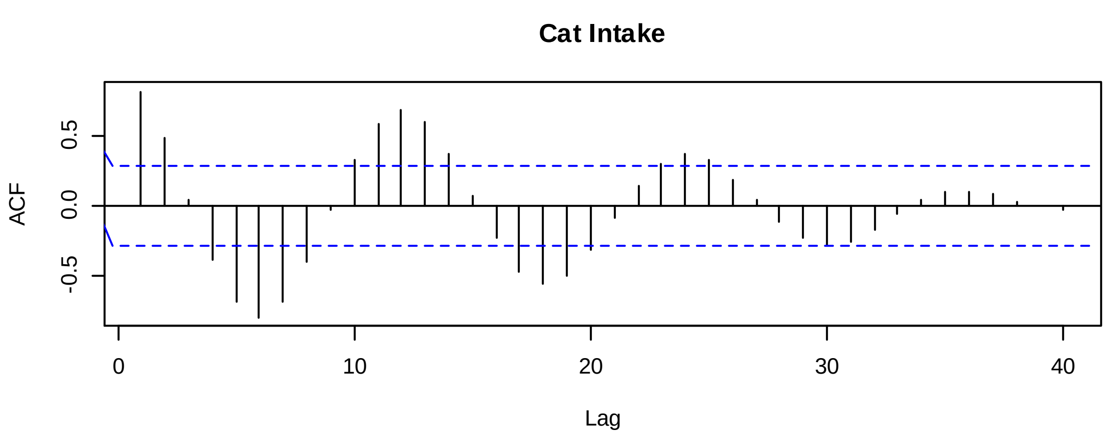


This time series displays a very clear pattern of seasonality. Unlike the Dog Intake graph, you can easily make out the peaks and valleys of these observations and there is no doubt about its seasonality. Annually, January or February can expect to see the least amount of cats surrendered to the AACT, while the rescue organization sees its peak amount of cat surrenders in July. While there is quite visibly a seasonal pattern, there is a decline in the amount of cat surrenders annually, suggesting negative trend. This can be seen when comparing January to January, or July to July intake.

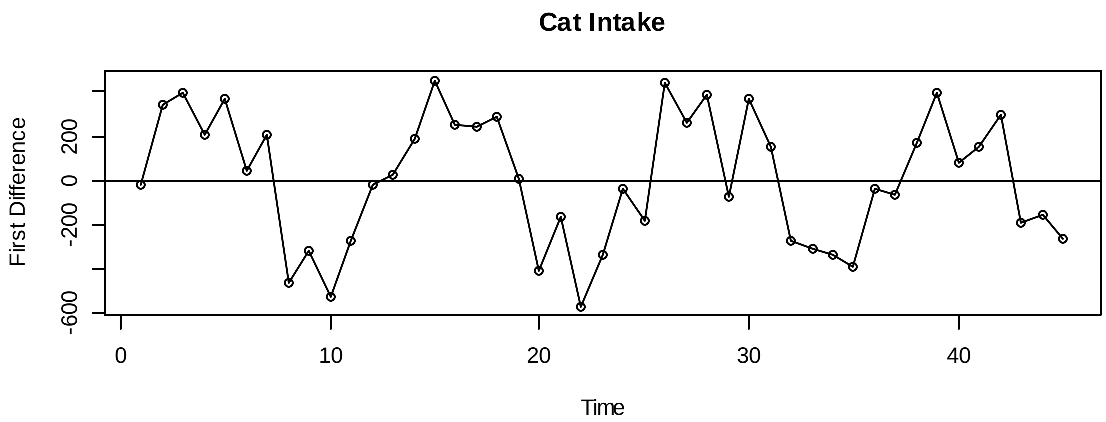
Below you can see the scatter plot that shows how strongly correlated the observations are. This graphs the number of cats surrendered against its lag, or the number of cats surrendered in the previous month. It can be expected to see a significantly stronger correlation in the Cat Intake scatter plot than the Dog Intake scatter plot due to the cleaner seasonality of the Cat observations. It is no surprise that the correlation here is 0.8298 versus the Dog Intake correlation of 0.61091.

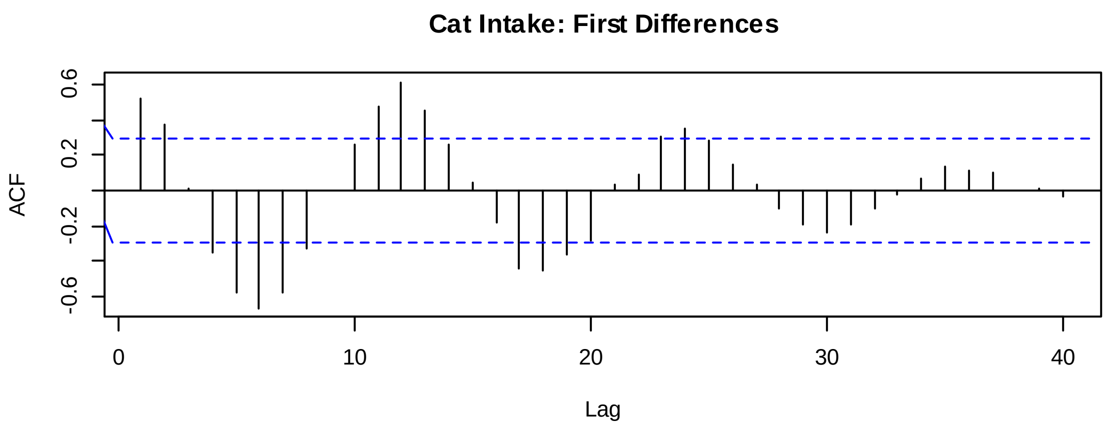


The seasonality of the data, if not already clear from the first graph of Cat Intake vs Date, can again be seen in the ACF below.

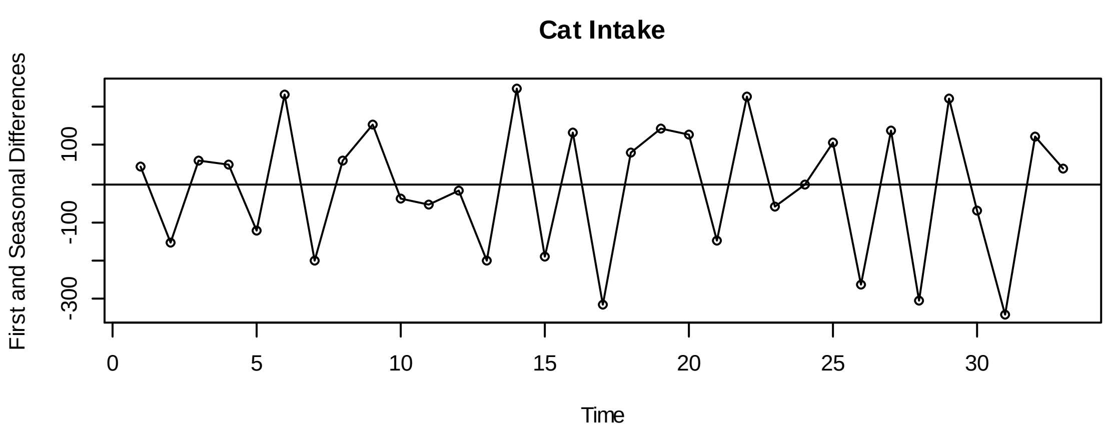


To create a stationary model, I took First Differences. Noticeably, you can see that seasonality is still strongly present, though any downward trend has disappeared.

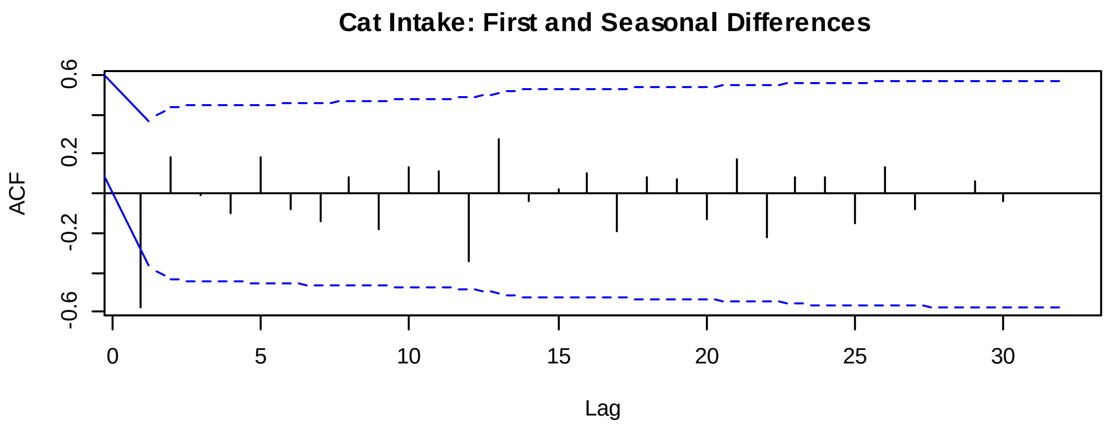




Because seasonality remained (seen in both the First Differences vs Time graph and the First Differences ACF graph above), I took first and seasonal differences in an attempt at parsimonious modeling, shown in the graph below.



This seemed to reduce or remove the seasonality, so I ran another ACF graph.

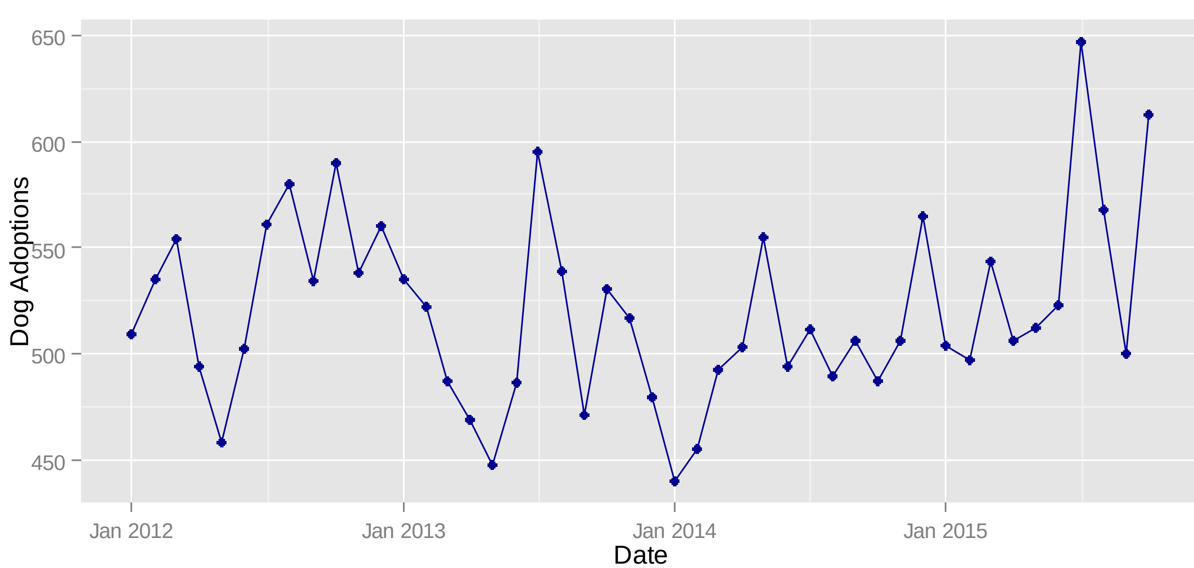


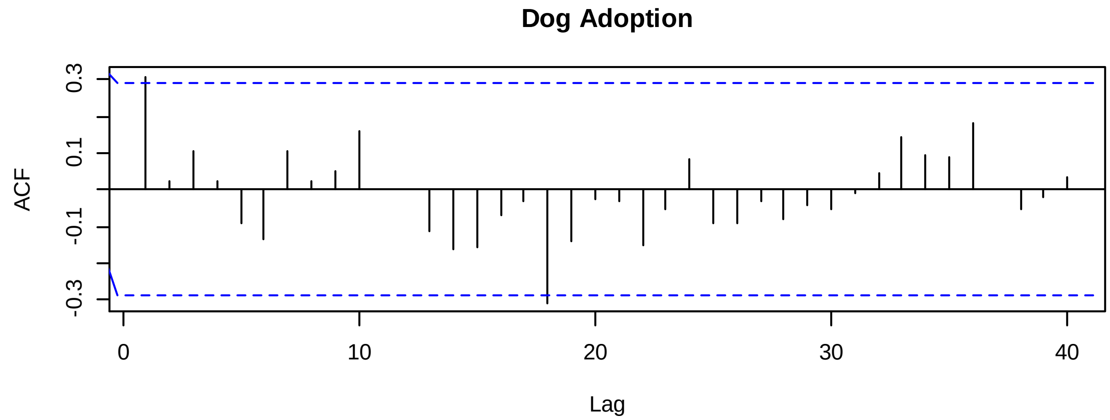
Parameter Estimation

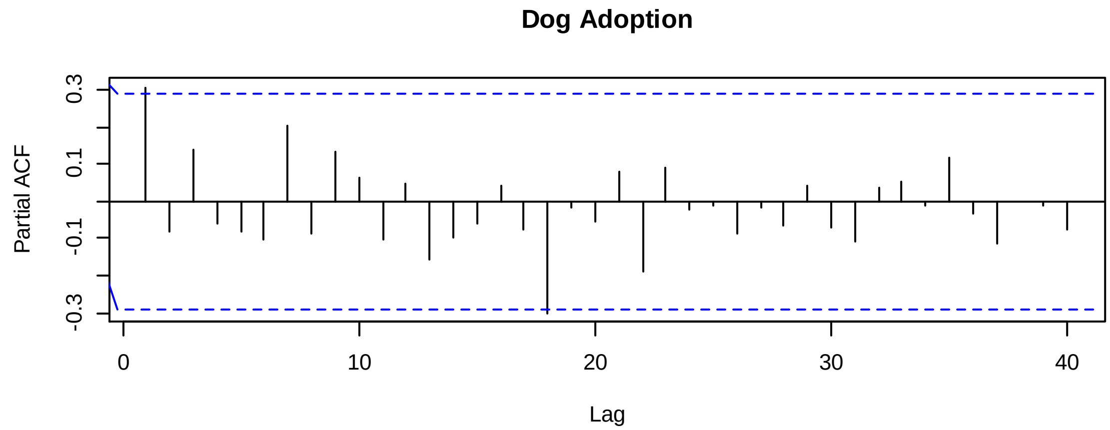
This ACF graph is significantly better, showing only lag 1 exceeding two standard deviations, which confirms that the majority of the autocorrelation no longer remains. Due to it being seasonal and because I took first difference, I started with the possibility of ARIMA(1,1,0)x$(1,1,0)_{12}$. In fitting the Cat Intake Series, I tried ARIMA(1,1,1)x$(1,1,1)_{12}$ and ARIMA(0,1,1)x$(0,1,1)_{12}$ as well, with less than favorable results. In the end, my initial assumption of ARIMA(1,1,0)x$(1,1,0)_{12}$ appears to be the best fit with ϕ=-0.5636 and Θ=-0.5068 and an intercept of 1439.6314.

Dog Adoptions

It turns out dogs are not surrendered to the AACT as cyclically as I thought, so it would be unwise to assume that dogs are adopted in a cyclical manner. The graph below shows Dog Adoptions versus Date, and as expected, it appears to have no seasonality and is perhaps even stationary. Looking at ACF and PACF graphs will help determine if this is true.





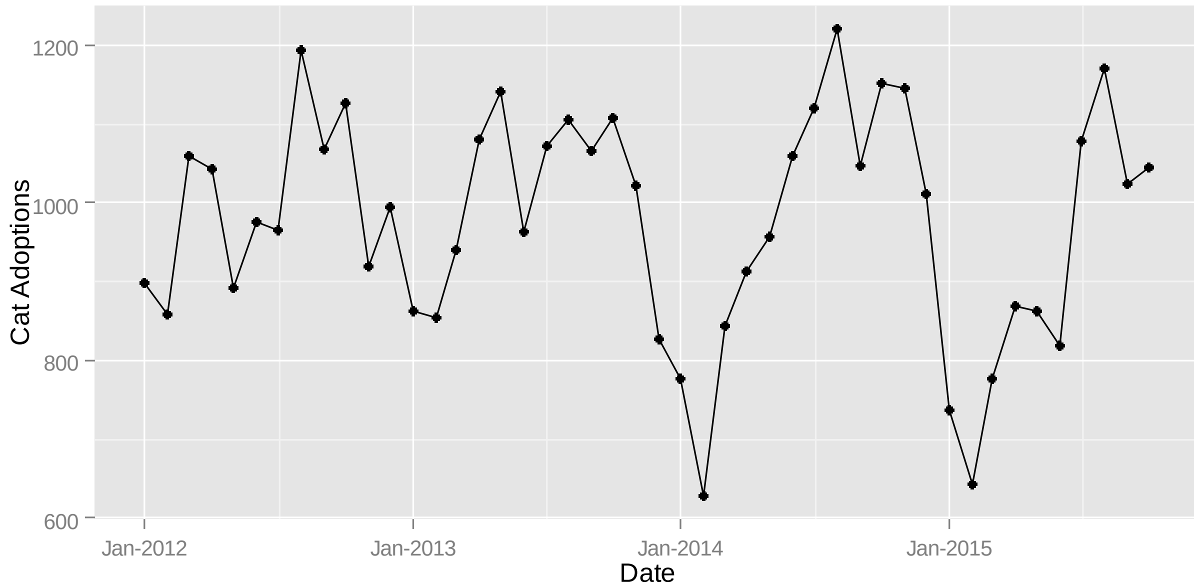


Parameter Estimation

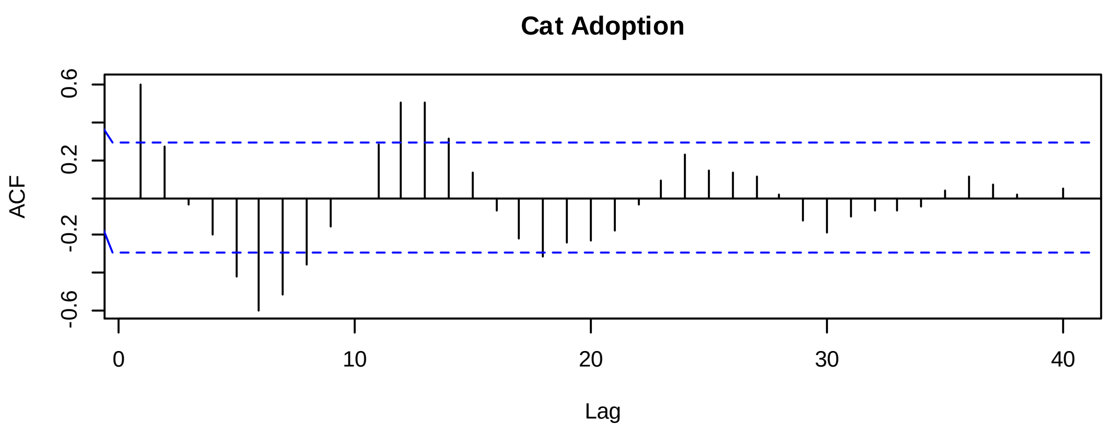
Based off of the ACF and PACF, I suspect it is an AR(1). Now, assuming that my Dog Adoption Series is AR(1), I ran some tests using Yule-Walker, Ordinary Least Squares, and Maximum-Likelihood Estimate. The parameter ϕ results returned were 0.3094, 0.3451, and 0.3366, respectively. When I ran possible models, such as ARIMA(0,0,1), the standard error was high and the coefficient was different. However, running AR(1), I got a coefficient ϕ of 0.3369. This leads me to believe my initial assumption is correct and the Dog Adoption Series is AR(1) with $Y_t = 0.3369Y_{t-1} + 520.6.08$.

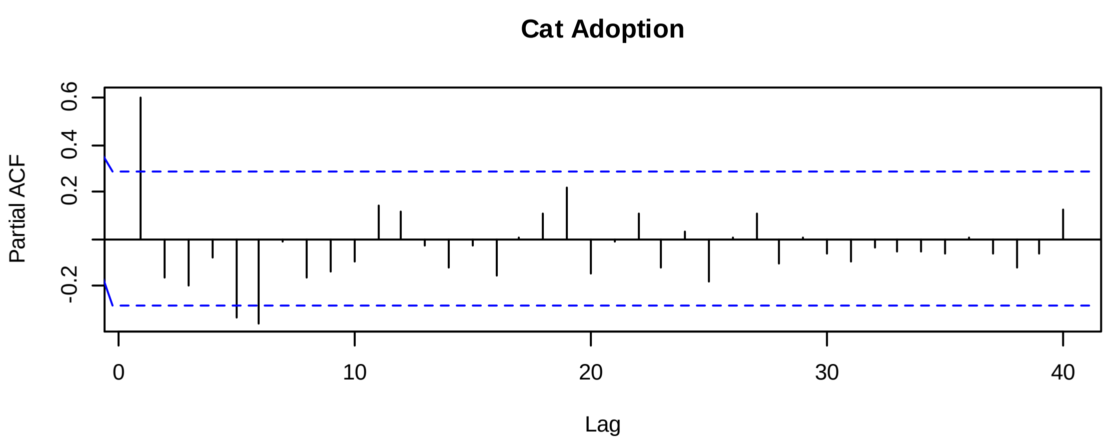
Cat Adoptions

Last, but not least, we must see how Cat Adoptions look! The data, shown in the below graph of Cat Adoptions versus Date, shows a mightily sporadic graph. This, unfortuantely, leads me to believe that cats are not adopted out in a similarly cyclical pattern as they are taken in. This easily accounts for the overcrowding in shelters and the desperate need for cat fosters (hint hint: foster a cat or two). Looking at the graph, I assumed a stationary and unseasonal series, but need to look at the ACF and PACF to confirm.



The ACF and PACF graphs are below. The ACF shows a sinusoidal graph that has significant early lag, but tapers to zero. This leads me to believe it is an AR(2) model. The fact that the PACF does not seem to taper to zero strengthens my assumption.





My Cat Adoption series does not appear to have any seasonality nor trend; in fact, it appears stationary.

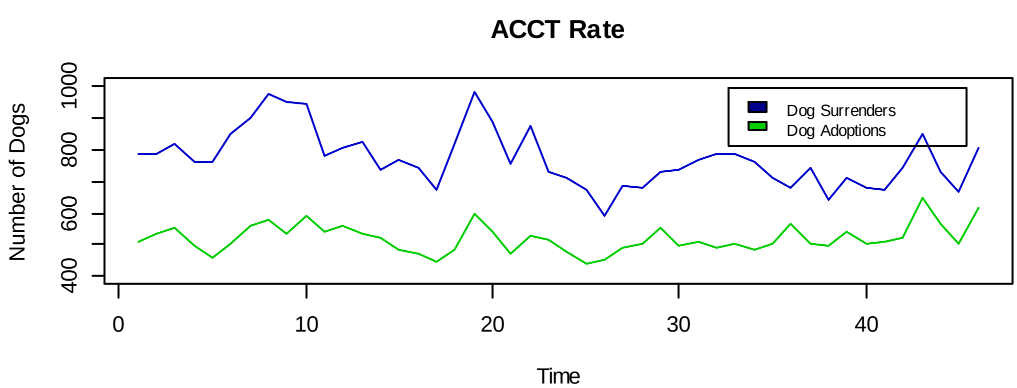
Parameter Estimation

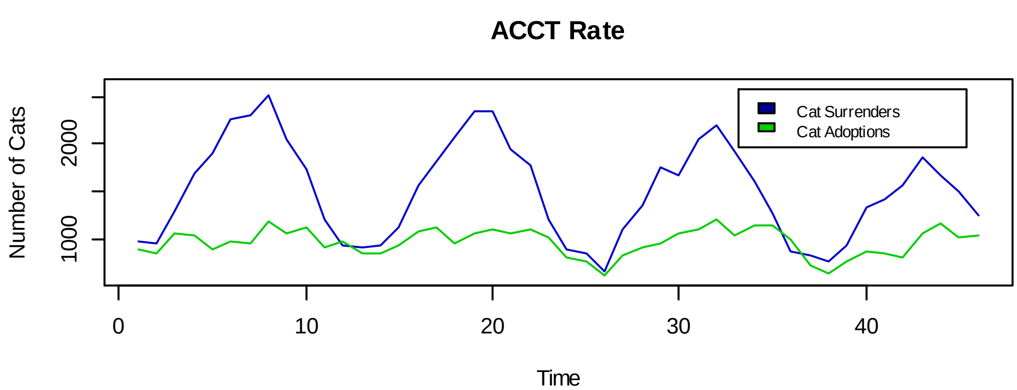
With the assumption that my Cat Adoption Series is AR(2), I ran some tests. I used assumptions of AR(2), AR(1), and ARIMA(2,0,1). The best fit was the ARIMA(2,0,1). This makes sense because the series did not require any differencing or seasonal differencing, and did not quite fit the look of AR(2). I am therefore left with $Y_t = 1.5048Y_{t-1} - 0.6870Y_{t-2} + e_t + e_{t-1} + 977.4585$ as the Cat Adoption equation.

Summary

It's all well and good to analyze the numbers separately; in this case, it showed the cyclicality of cat surrenders, which can be beneficial to the ACCT in how they run their adoption center. However, in plotting the surrender data with the adoption data, one gains a better understanding of the big picture and can see how the surrenders and adoptions work in tandem – for you cannot continue adopting animals out if you are not receiving animals to being with.

In this case, while I settled on AR(1) for both Dog Surrenders and Adoptions and could therefore expect similar patterns, it was somewhat shocking to see how similarily the data matched up, only with different intercepts. This suggests that the shelter sees quick turnaround for the dogs they receive which is highly encouraging. It is said that a dog that does not get adopted within two weeks is significantly less likely to be adopted and is more likely to be put down.





The cat data was similarly shocking, as it contained a much different pattern than that of the dog. The data suggests that cats are adopted at an almost steady rate, that you can expect an average of 1000 cats per month to be adopted, no matter the time of year. This implies that cats sit for longer amounts of time in the shelter before they are adopted because the valleys and peaks of Surrenders do not match that of the Adoptions. This creates a housing problem in shelters and overcrowding leads to more animals put down, unfortuantely. Perhaps with this data, the shelter can better prepare for long term cat tenants.

Because I have volunteered in shelters before and occasionally find myself on dog rescue sites, it came as no surprise that the adoption numbers are lower than the surrender numbers. I found it interesting that cat surrender is as seasonal as shown, while dog surrender has no seasonality. It seems as though everyone decides to get rid of their cats in the summertime, or maybe outdoor cats are being brought in with the assumption that they are strays; perhaps it is a mixture. It makes logical sense to see no cyclicality in adoptions of dogs or cats; I thought perhaps in the winter more dogs would be adopted, as in the days of yore dogs were used to keep masters warm. I believe that with this information, or a more thorough study with more years of data, shelters can better equip themselves to prepare for the onslaught of cat surrenders in the summer months and take advantage of the “slower” winter months, thereby rescuing a larger number of animals. For example, increasing the amount of volunteers, holding animal adoption events in local parks, or requesting more foster families are all ways to increase the adoption numbers in warmer months. Most importantly, I have determined it is still vitally important to adopt a pet from your local shelter – maybe two.