ARIMA Model with United States Incarceration Rates

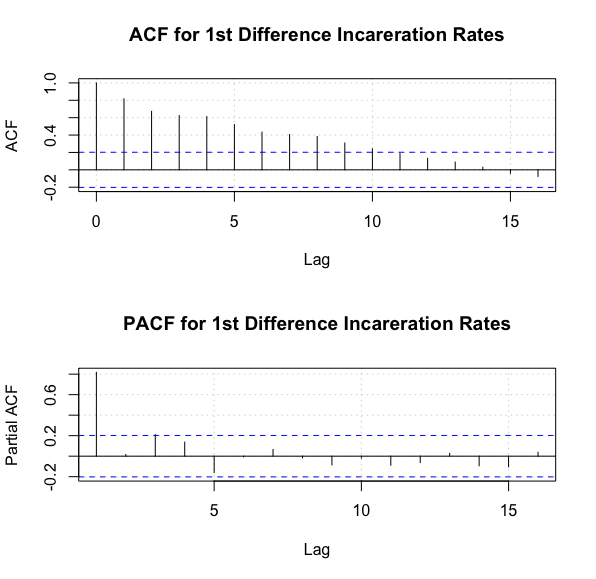
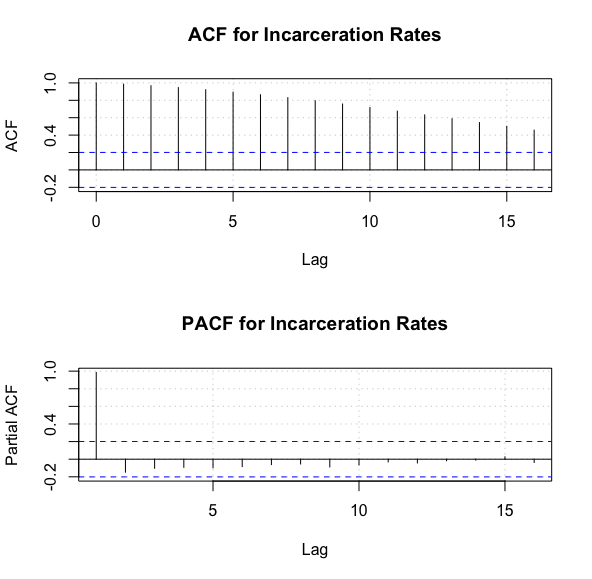
The current U.S. incarceration rate is nine times higher than Germany, eight times higher than Italy, five times higher than the U.K., and 15 times higher than Japan, leading many to question why so many Americans face prison time.[[1]](#footnote-1) However, the prison population has not always been so large. Between 1925-1976 only 0.2% of US population were in federal prisons. However, when Ronald Reagan took office in 1980, the prison population saw a dramatic increase, with more of the male population finding themselves behind bars. While Nixon declared a “War on Drugs” back in the 1970’s, Reagan took a more militant approach to individuals using drugs, invoking longer prison sentences for drug possession. [[2]](#footnote-2)

Concerns about drug use prompted Congress to pass the Comprehensive Crime Control Act of 1984 and the Anti-Drug Abuse Act of 1986, the latter of which granted $1.7 billion to fight drugs and established a mandatory minimum penalty for drug offenses. The Anti-Drug Abuse Act of 1986 increased penalties for drug use and established the National Office for Drug Control Policy. The 1994 Crime Bill provided for 100,00 new police officers, $9.7 billion in funding for prisons. The federal death penalty was expanded under this bill, and since it’s passaged the rates of men incarcerated have increased.2 These significant upward trends from these time periods show that this series is not stationary and there is a trend component in this series.

Chart, line chart

Description automatically generated

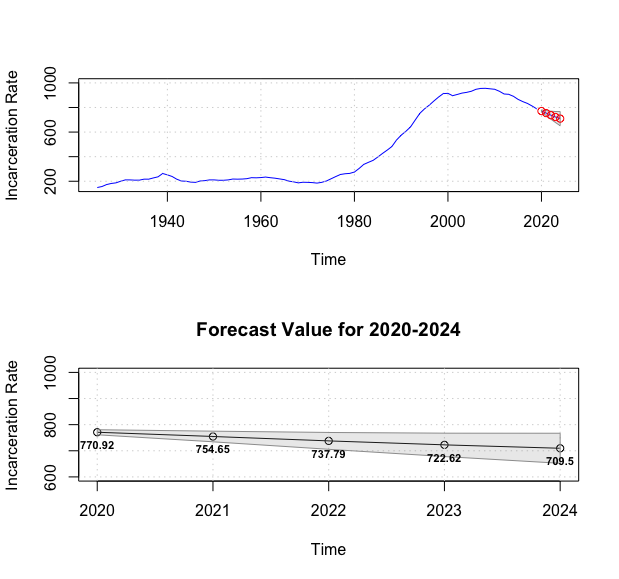
Next, we plot the ACF and PACF functions for the original series. ACF plots are bar charts of the coefficients of correlation of a time series and lags with itself. Within the ACF plot, there is a significant geometric decay as the lag increased with several autocorrelations that are non-zero. This series can therefore be classified as non-random. The ACF is also slowly declining in this model, suggesting the original series isn’t stationary. In addition, the PACF of the original series shows one significant spike that exceeds the standard error bar. This shows that performing first differencing will turn the original series into a stationary series.

After taking the first difference, the series decays more rapidly. Since there are many data points where the autocorrelation is significant, the PACF chart shows that the expected correlation due to the propagation of correlation is at lag 1 and 3. This means that all higher order autocorrelations are explained effectively by lag 1 and 3. Based on these plots, the order of the ARIMA model can be determined. In this model d=1, since a first difference was performed to transform the original series into a stationary series. 

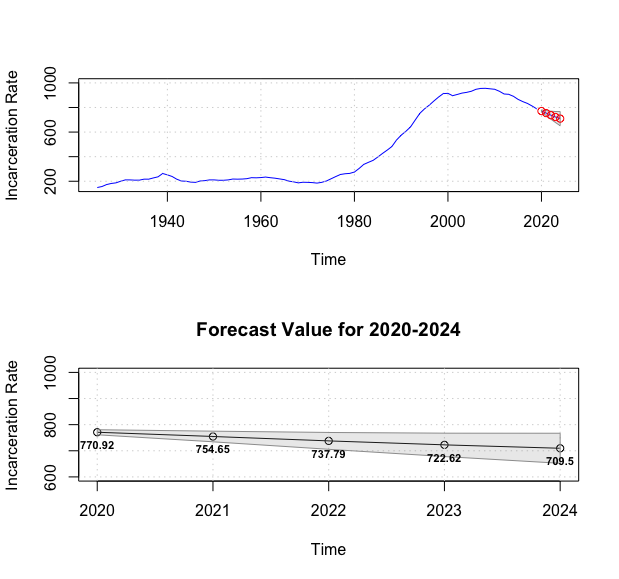
Since the ACF shows a declining pattern after performing differencing and significant spikes, the order for p=1 can be determined from the PACF, while q=0 due to the declining pattern of ACF. Therefore, a suitable ARIMA model for this series could be ARIMA(1,1,0). For more conclusive study, several alternatives were considered: ARIMA(3,1,3), ARIMA(3,1,0), ARIMA(1,1,1).

|  |  |
| --- | --- |
| ARIMA Model | AIC Value |
| ARIMA(3,1,3) | 694.4 |
| ARIMA(3,1,0) | 697.51 |
| ARIMA(1,1,1) | 700.51 |
| ARIMA(1,1,0) | 698.85 |

Based on the lowest AIC value, ARIMA(3,1,3) was chosen to use moving forward for forecasting. In addition, based on Ljung-Box p value of 0.724, the null hypothesis was accepted and indicates that the residuals are white noise. This model was used to generate the prison population moving forward 5 years ahead (2020-2024).



As more individuals have taken interest in prison reform in recent years, the increasing rates that the nation saw between the 1980’s-2010’s has reversed in recent years. This trend was occurring even before the coronavirus pandemic saw more, and the estimated incarceration rate is projected to fall from 0.077-0.0709% between 2020-2024. While recent data has not been made publicly available, it will be interesting to see how the coronavirus may have changed these forecasted results as many courts shut down during the lockdown and are now operating significantly behind.



1. https://www.prisonstudies.org/highest-to-lowest/prison-population-total [↑](#footnote-ref-1)
2. https://drugpolicy.org/issues/brief-history-drug-war [↑](#footnote-ref-2)