Cases of Mumps In New York City

Isis Ramirez

4/16/2019

Background and Key Questions

Mumps is a highly contagious viral infection that presents itself with fever, muscle aches, weakness, and headaches. There is no cure, but a vaccine was introduced in 1967 that reduces the chances of contracting the disease.

Since mumps can affect the health of various individuals, it is worthwhile to explore any patterns presented in mumps outbreaks. These revelations can help us learn on how to protect ourselves more effectively.

We will be looking at a time series dataset on the monthly count of mumps cases in New York City from 1928-1972.

The following questions will be addressed:

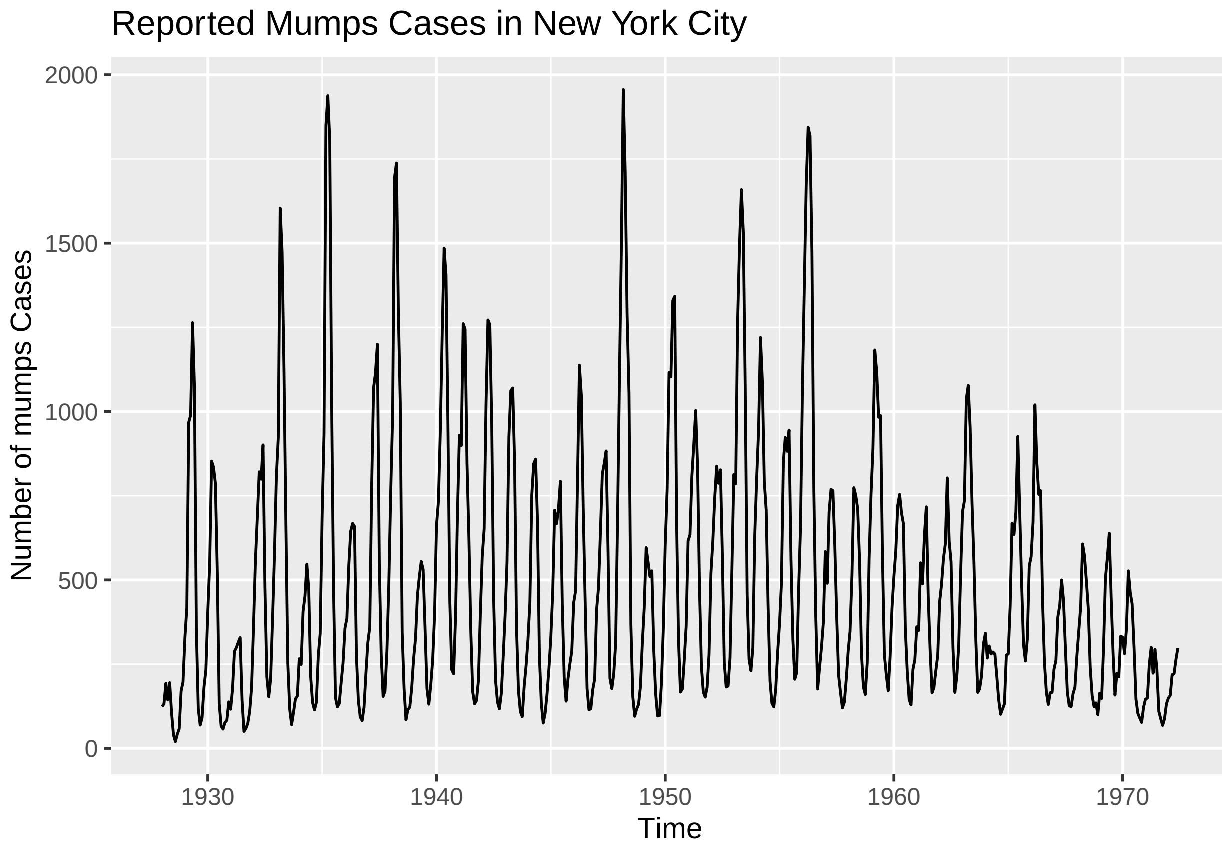
- Is there a temporal pattern to mumps occurrences?
- Is the vaccine effective at preventing mumps?

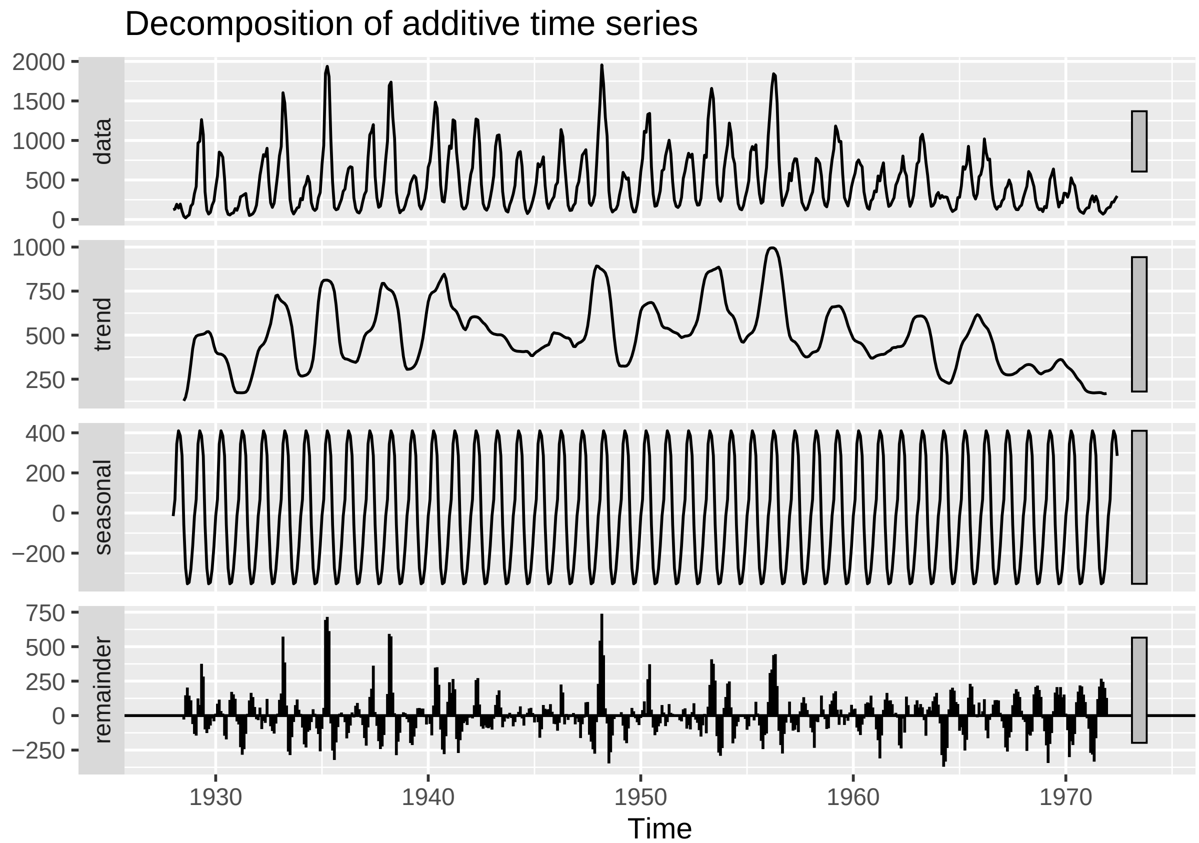
In order to answer these questions, we will focus on exploring the seasonality of the data and developing a model using trigonometric functions and exogenous variables.

Data: Definition, Acquisition and Exploration

Monthly data of reported cases of mumps in New York City from 1928-1972 was acquired from the Time Series Data Library.

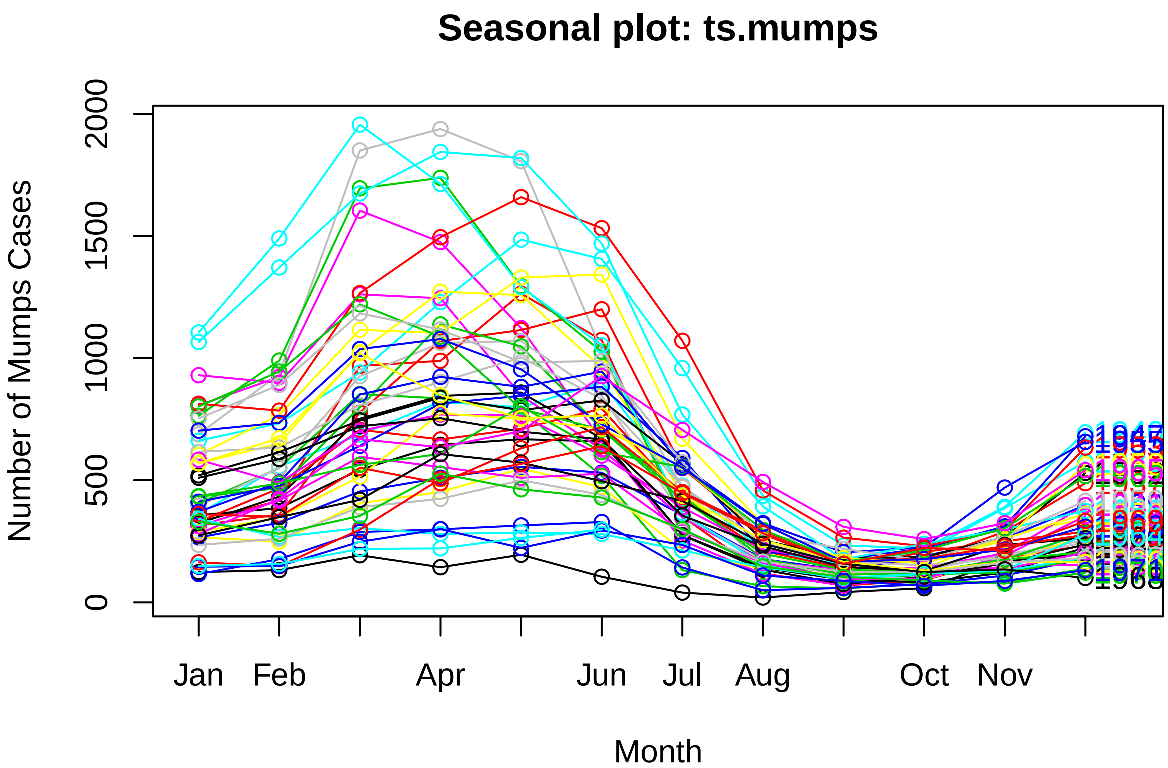
Below is a time series plot of the data, which shows decreasing numbers in the later years and periodic oscillations across the entire series. These oscillations are also evident in the seasonal component of the decomposed data.


Reported Mumps Cases in New York City
Number of mumps Cases
2000
1500
1000
500
0
1930
1940
1950
1960
1970
Time

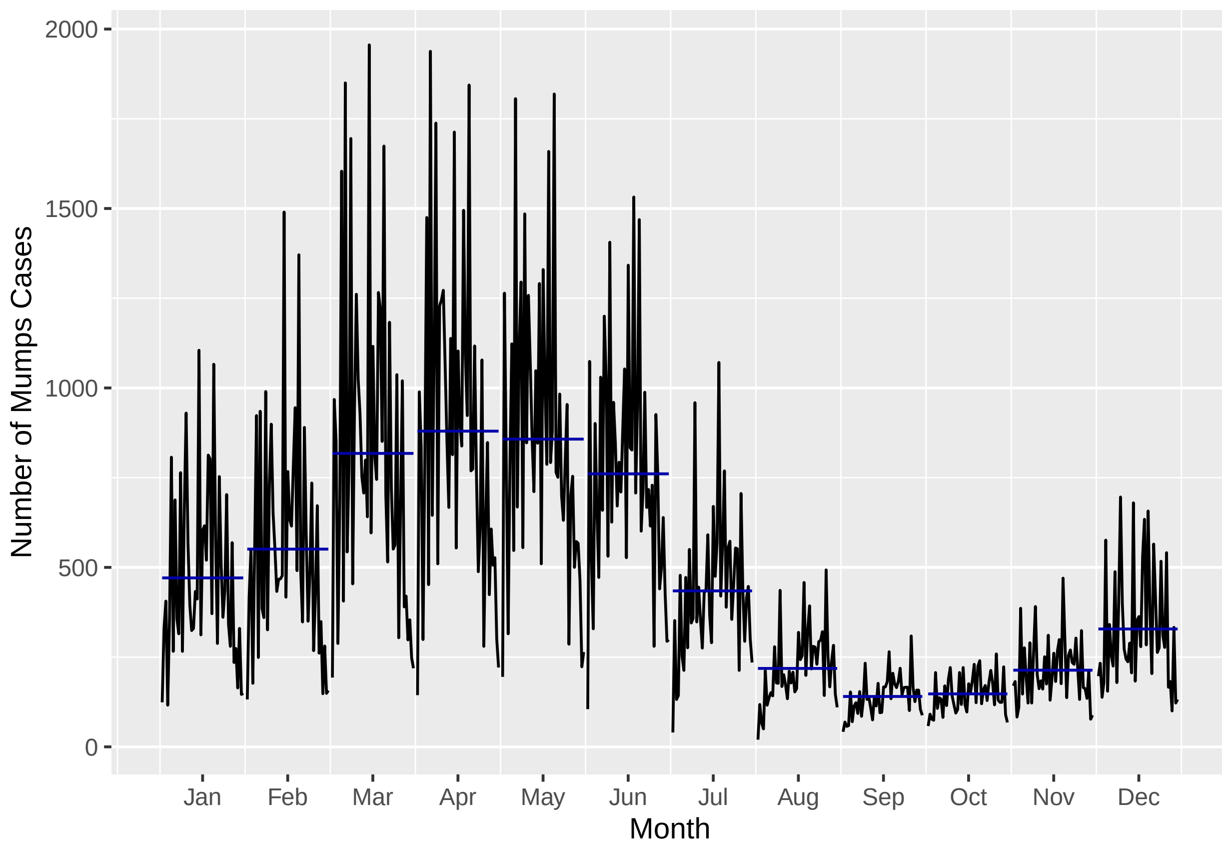


Seasonality

A seasonal plot reveals a yearly reocurring pattern. The number of cases reported are at the highest during late winter and spring months.



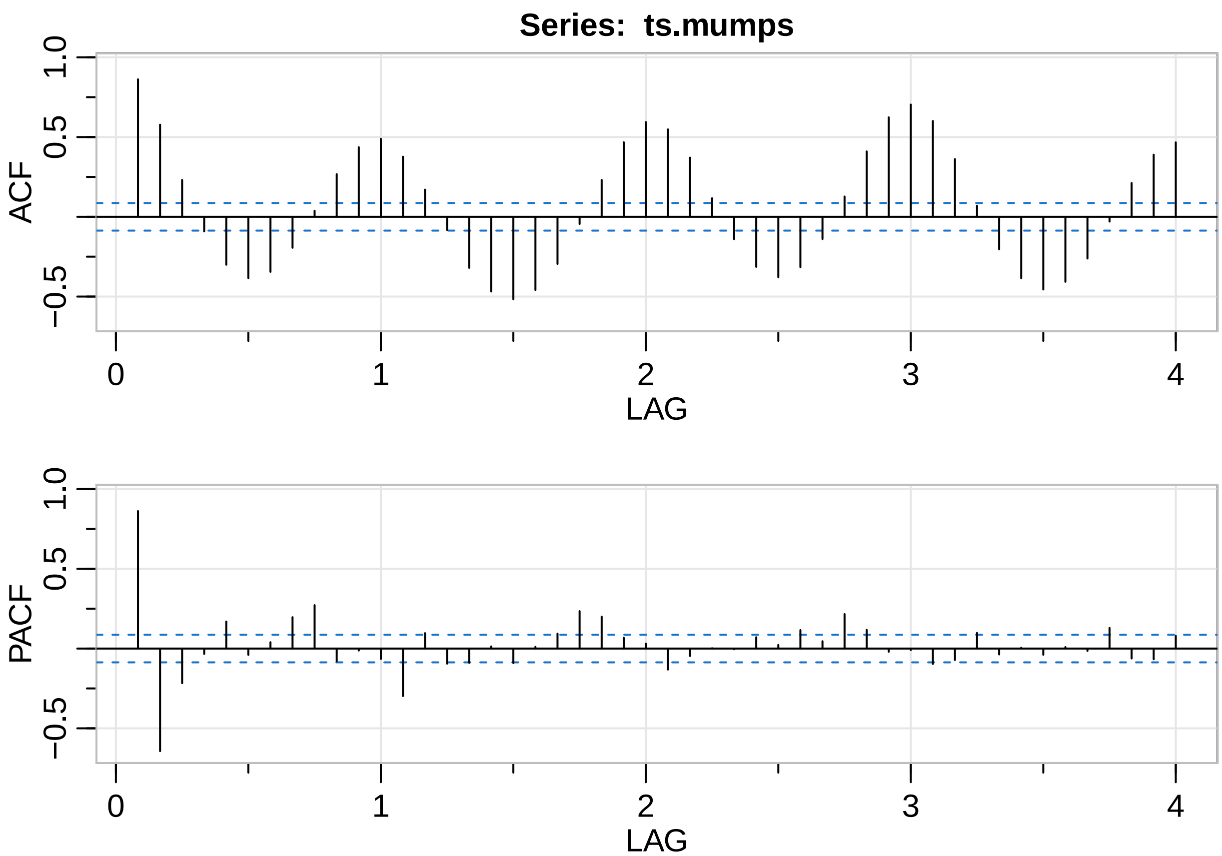
Even in current times, this temporal pattern of higher incidences in later winter and early spring is confirmed by the CDC. This knowledge is valuable, especially for individuals that may not be eligible to get vaccinated, such as those with weakened immune systems, pregnant women, or those with history of seizures. With this information, people can now be more cautious and vigilant during months of higher occurences of mumps.



The subseries plot shows a decrease in reported cases of mumps across all months in the later years. Further investigation will be done on whether this decrease is a result of a vaccine being introduced to the public in early 1967.

Finding a Model

The ACF shows the seasonal component. We previously saw an annual pattern indicating a period of 12 months. We also see a sinusoidal in the ACF, but we do not see any decay. A seasonal arima model will be using with exogenous variables. Due to the periodic behaviour seen in the ACF, trigonometric functions will be used to estimate coefficients.



In order to determine the effect of the vaccine, we will use the following formulas to detect a change in amplitude. Indicators are used to determine whether an incidence occurred before or after the vaccine was introduced. A pre-vaccine indicator (I_{pre}) value of 1 implies an incidence that occurred before the vaccine and a pre-vaccine indicator value of 0 implies an incidence that did not occur before the vaccine; the post-vaccine indicator (I_{post}) works similarly. Note that the A and U variables correspond to incidences occuring before the vaccine while the B and V variables correspond to incidences occurring after the vaccine was introduced.

$$x_t = A[I_{pre}cos(\frac{2\pi t}{12} + \phi_{pre})] + B[I_{post}cos(\frac{2\pi t}{12} + \phi_{post})]$$

The exogenous matrix will consists of the terms seen in the trignometric function below, which is a transformation of the previous function:

$$x_t = [U_1(I_{pre}cos(\frac{2\pi t}{12})) + U_2(I_{pre}sin(\frac{2\pi t}{12}))] + [V_1(I_{post}cos(\frac{2\pi t}{12})) + V_2(I_{post}sin(\frac{2\pi t}{12}))]$$

Note the following:

$$U_1 = Acos(\phi_{pre})$$

$$U_2 = -Asin(\phi_{pre})$$

$$V_1 = Bcos(\phi_{post})$$

$$V_2 = -Bsin(\phi_{post})$$

Fitting a Model

Several SARIMA models with exogenous variables were used to fit the data. Below are the results for a SARIMA(2,0,0)x(1,0,1)[12]. SARIMA models of higher orders returned similar results.

Table 1: SARIMA(2,0,0)x(1,0,1)[12] Parameter Estimates

	Estimate	SE	p.value
ar1	1.21	0.04	0.00
ar2	-0.36	0.04	0.00
sar1	0.99	0.01	0.00
sma1	-0.89	0.03	0.00
intercept	436.98	133.10	0.00
U1	-165.50	68.81	0.02
U2	345.43	85.03	0.00
V3	-88.38	84.51	0.30
V4	156.43	95.96	0.10

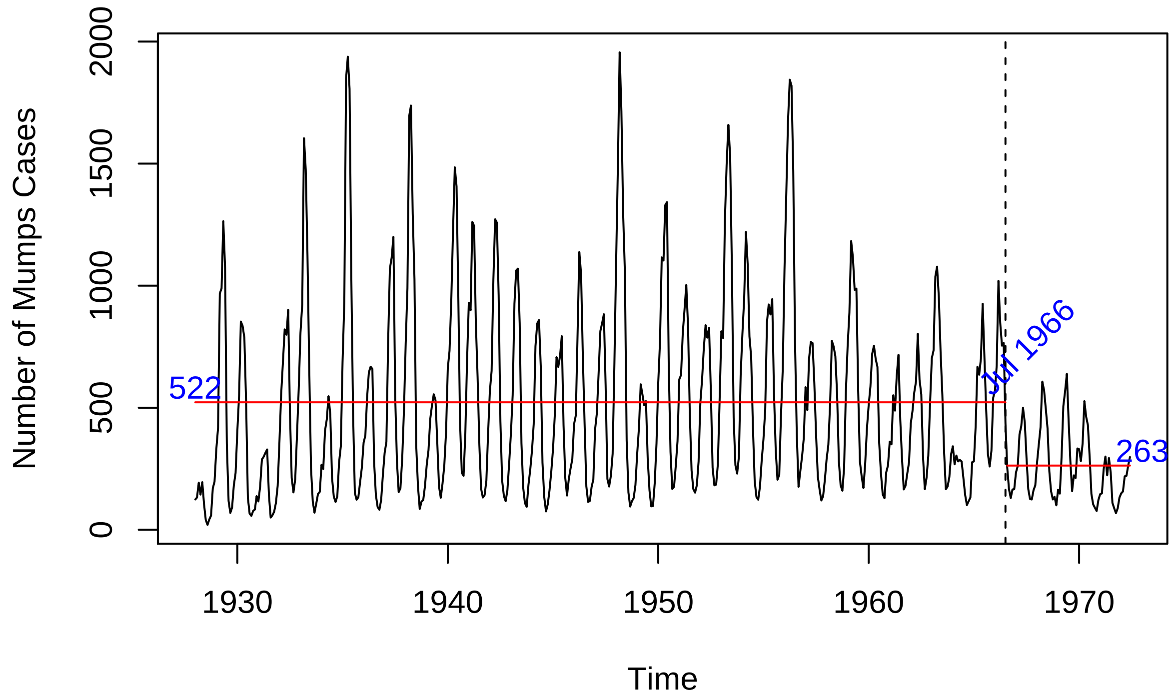
After extracting the appropriate coefficients, the amplitudes for pre-vaccination and post-vaccination periods were calculated. We see that the post-vaccine amplitude is smaller than the pre-vaccine amplitude. This suggests that the vaccine did indeed help reduce the numbers of mumps cases.

Table 2: Amplitudes Pre/Post Vaccine

	Amplitudes
Pre-Vaccine (A)	383.03
Post-Vaccine (B)	179.67

Change in Mean

The effect of the vaccine was further validated by using the changepoint package to detect a change in mean. Below we see that there is a change in mean in the data beginning in July 1966, which is before the vaccine was introduced to the public. We previously saw a seasonal pattern of decreasing occurrences in the summer, which may have affected the detection of the changing point time.

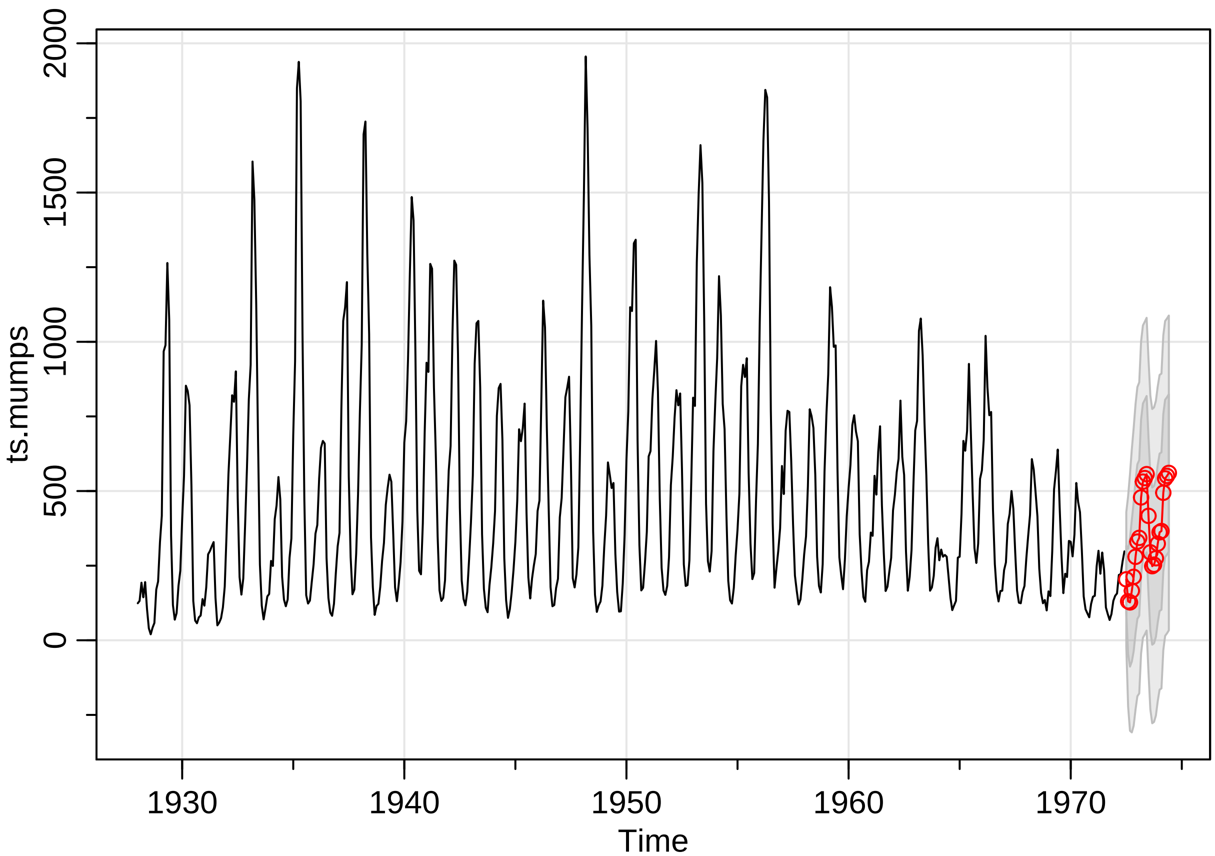


Interestingly, in the months of November and December of 1966, a sample of people were vaccinated as part of clinical trials before the wave of rising occurrences typically seen in late winter and spring. The exact number of people that participated in these trials was not retrieved. The vaccine was immediately licensed in distributed in March 1967 following these clinical trials. The overlap of the change and reduction in the mean following the clinical trials and the introduction of the vaccine to the public suggests that the vaccine did indeed reduce the occurrences of mumps.

Note that the estimated mean of the process is 437, which lies between the two different means.

Forecasting 1973-1974

Using the previously fitted SARIMA(2,0,0)x(1,0,1)[12] model, predictions for the 24 months following the series are calculated. Below we can see the plotted forecasts up to June 1974 with error bounds. The predictions mimic the seasonal behavior seen in the time series. We also see a smaller amplitude in the predictions compared to earlier years. The process converges to a mean of 437, which is higher than the mean of 263 calculated for the post-vaccine era by the changepoint package. Predictions could be improved to reflect the reducing effect of the vaccine.



Summary Conclusions

There is an evident seasonal pattern in the data that reveals higher numbers of mumps occurrences in late winter and early spring months. Additionally, we saw numbers decreasing in the later years included in the time series. After fitting a SARIMA(2,0,0)x(1,0,1)[12] with exogenous variables, a decrease in amplitude was determined in observations following the introduction of a vaccine. A reduction in the mean was also ascertained following clinical trials and the distribution of the vaccine to the public, which indicates that the vaccine does prevent mumps.

The SARIMA model used is able to capture the seasonal behavior of the data, but additional post-vaccine data would be helpful in capturing the decrease in occurrences following the introduction of a vaccine for forecasts.