



# Análisis de series temporales de operaciones canceladas en Inglaterra

Lic. Malena Álvarez Brito, Ing. Florencia Florio

octubre 2020

**Resumen:** En la presente investigación aplicamos distintas técnicas con el fin de evaluar patrones de comportamiento en las cancelaciones de operaciones en Gran Bretaña. La información de la ocupación de camas nos parecía un dato muy relevante para agregar al análisis así como también para entender el comportamiento de los datos. En su desarrollo hemos aplicado técnicas de forecasting con el fin de predecir el porcentaje de cancelación de operaciones lo que ayudaría principalmente a la correcta planificación de los recursos en el sistema de salud británico.

Palabras clave: series de tiempo, sistema de salud, forecasting

## 1 Introducción

Los quirófanos representan un costo considerable para los centros de salud dado que requieren una gran cantidad de recursos disponibles para garantizar su funcionamiento. Por este motivo, la subutilización de los quirófanos por parte de los centros de salud, va en detrimento de las finanzas del centro. En la presente investigación hemos trabajado con datos publicados por el servicio de salud de Inglaterra acerca de la cancelación a último minuto de operaciones electivas en hospitales públicos.

## 2 Hipótesis y objetivos

El objetivo del presente trabajo es analizar la serie de tiempo de operaciones electivas canceladas para descubrir patrones en su comportamiento y realizar, de ser posible, un forecasting de valores futuros en el porcentaje de cancelaciones. También se busca analizar qué tanto contribuye el porcentaje de ocupación de camas trimestrales en los hospitales públicos al porcentaje de cancelación de operaciones electivas. Se cree que la variable ocupación de camas contribuye al aumento de las cancelaciones.

El resultado de este tipo de análisis podría contribuir a una mejor asignación de recursos del sistema de salud y toma de decisión de gestión sanitaria.

### 3 Material y métodos

Los datos utilizados corresponden a aquellos publicados por el servicio de salud de Inglaterra acerca de las operaciones electivas en centros de salud públicos <sup>1</sup>. El dataset consta de datos trimestrales sobre:

- Cantidad de operaciones electivas programadas (no incluye procedimientos ambulatorios menores).
- Cantidad de cancelaciones de operaciones a último minuto por parte del hospital debido a motivos no clínicos.
- Cantidad de pacientes no tratados en un lapso 28 días posteriores a la cancelación (este punto no fue incorporado al análisis).

Ejemplos de motivos de cancelaciones (son no clínicos):

- Cirujano no disponible
- Anestesista no disponible
- Otro personal no disponible
- Cama de sala no disponible
- Cama de cuidados intensivos no disponible
- Caso de emergencia que necesita quirófano
- Falla en el equipamiento médico
- Error administrativo

Dado que el dataset no aporta estadísticas sobre la prevalencia de los motivos de cancelación, se utiliza otro data set provisto por el servicio de salud de Inglaterra referido al porcentaje de ocupación de las camas de internación en sus hospitales públicos <sup>2</sup> para analizar qué tanto aporta esta variable a la cancelación. El data set de ocupación se encuentra diferenciado por camas para internación diurna o nocturna.

El análisis de series temporales tiene utilidad en el estudio de la relación entre variables que cambian con el tiempo y que pueden llegar a influenciarse entre sí. Los datos se pueden comportar de diferentes maneras a través del tiempo: puede que se presente una tendencia, estacionalidad o simplemente no presenten una forma definida.

Se utilizaron los siguientes conceptos:

- Descomposición de series temporales
- Análisis en el dominio de la frecuencia
- Autocorrelación de una serie temporal
- Forecasting
- Correlación entre series de tiempo

La investigación fue realizada con R versión 4.0.2 en el IDE RStudio. Las principales librerías utilizadas fueron: tidyverse, zoo, xts, forecast, seasonal. Para el presente informe se utilizó RMarkdown.

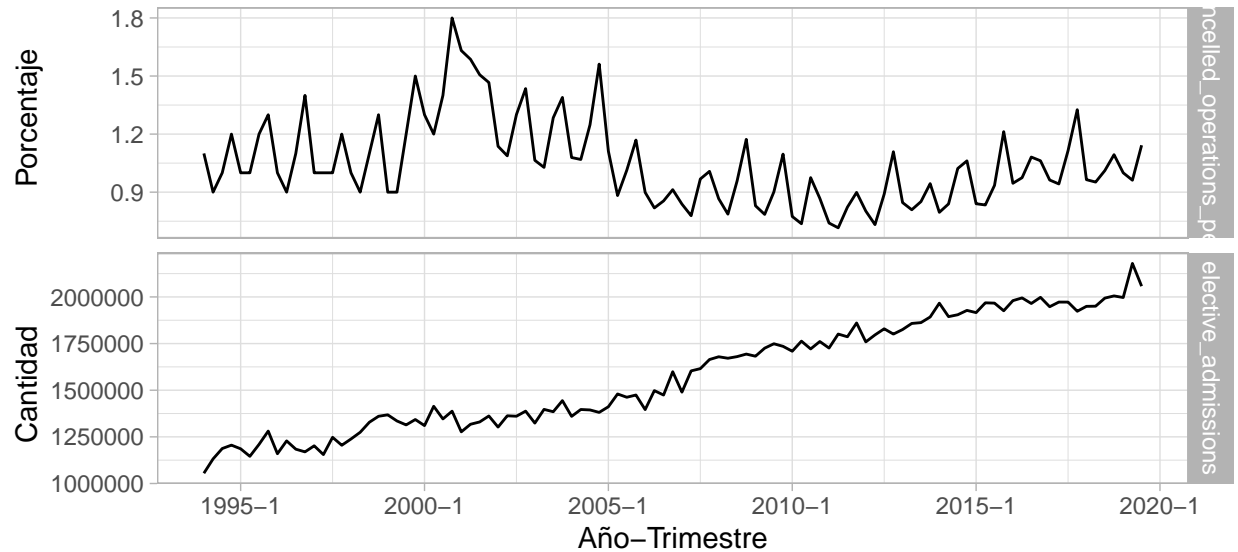
### 4 Resultados

Se analiza la cancelación de operaciones como porcentaje del total de operaciones para quitarle la tendencia ascendente que tiene esta última. Se observa que dicho porcentaje tiene un pico en 2001 (casi el doble que en otros trimestres) y depresión entre 2010 y 2012. En la descomposición multiplicativa, la

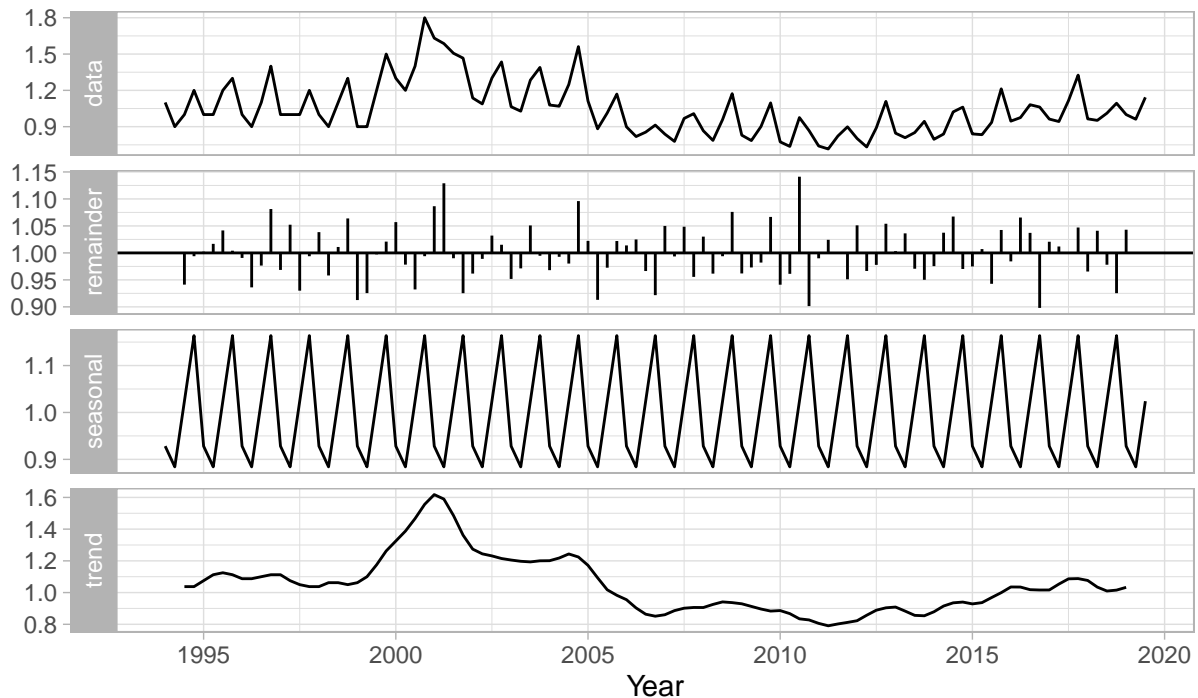
<sup>1</sup><https://www.england.nhs.uk/statistics/statistical-work-areas/cancelled-elective-operations/>

<sup>2</sup><https://www.england.nhs.uk/statistics/statistical-work-areas/bed-availability-and-occupancy/>

curva suavizada resultante muestra dichas oscilaciones no estacionales. También, se evidencia la fuerte estacionalidad de los datos.



### Descomposición multiplicativa clásica



Comparando los ciclos anuales, se observa que el pico ocurrió específicamente en el último trimestre de 2000 pero que todo el 2000 fue atípicamente alto. También el 2001 fue alto, aunque luego fue bajando. Tanto en este gráfico como en el de subestacionalidad se ve que los primeros dos trimestres son, en promedio, más bajos que el tercero y el cuarto.

Gráfico estacional

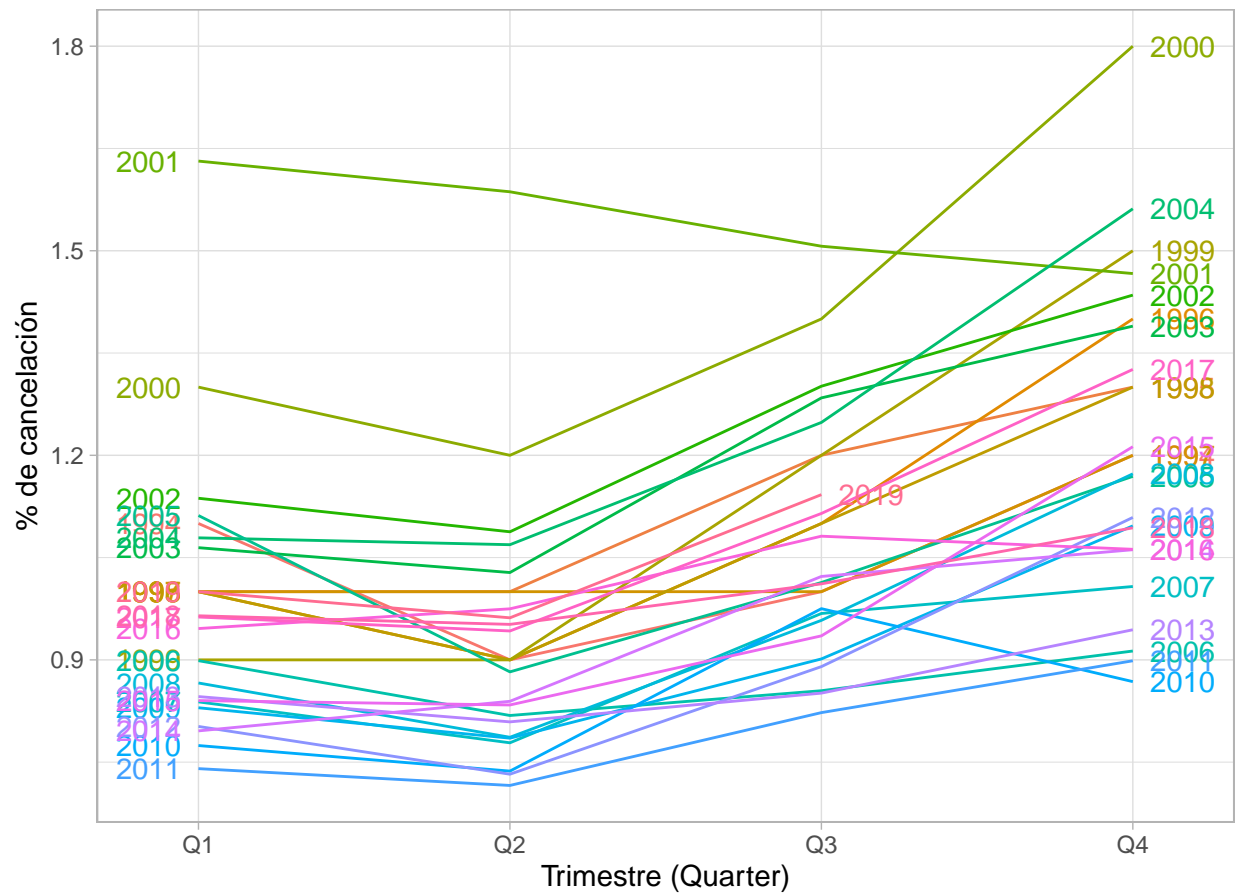
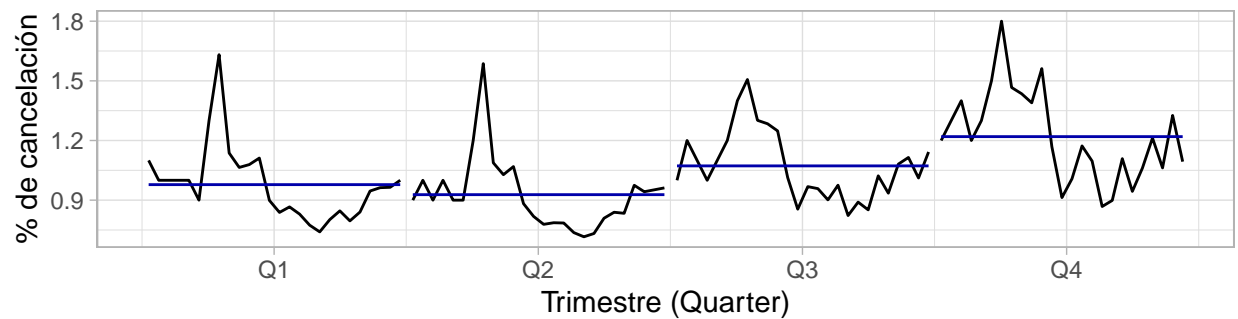


Gráfico de subestacionalidad



Respecto a la autocorrelación, se ve que el pico más alto está en 4. Los picos tienden a tener una separación de 4 (observaciones) trimestres y van disminuyendo en su amplitud. Forman picos debido a la estacionalidad. El hecho de que la correlación sea positiva en casi todos los lags se debe a que la serie tiene una tendencia positiva.

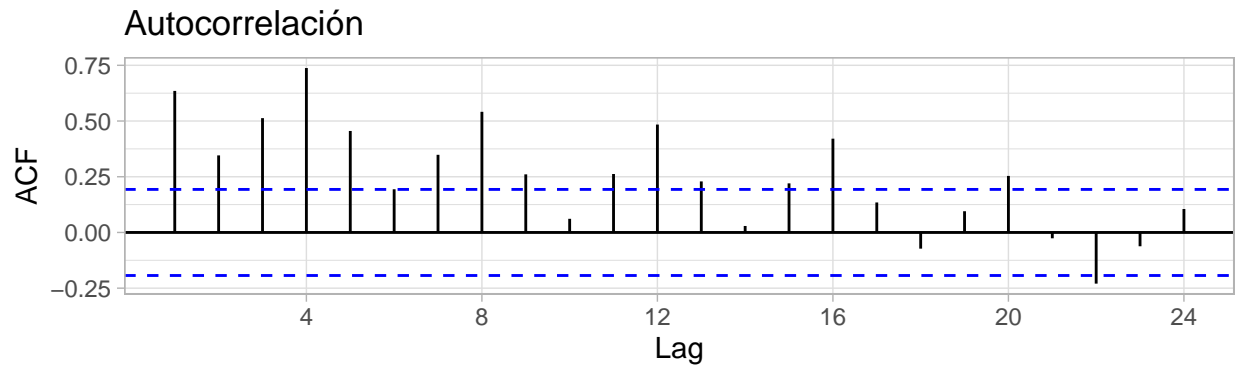
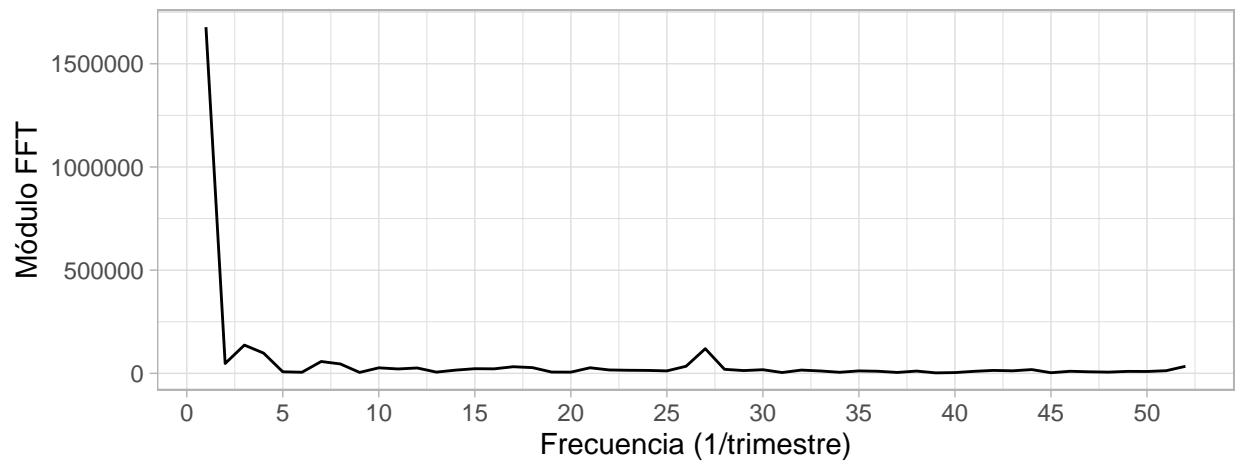


Gráfico de la fft de porcentaje de operaciones canceladas. Se distinguen 2 picos bien marcados (a parte del valor del índice 1).



Valor	Frecuencia
1677577	1
136953	3
119514	27

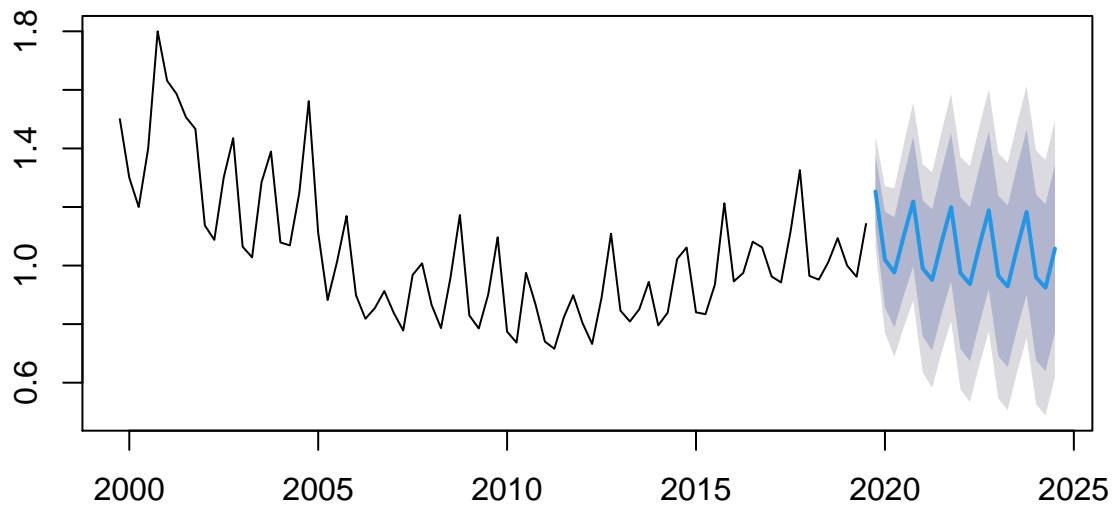
## 4.1 Forecasting

Para el forecasting se aplica un ARIMA automático con estacionalidad.

```
## Series: tt
## ARIMA(1,0,0)(0,1,1)[4]
##
## Coefficients:
##      ar1      sma1
##    0.8637  -0.8829
## s.e. 0.0546  0.0618
##
## sigma^2 estimated as 0.009364: log likelihood=88.72
```

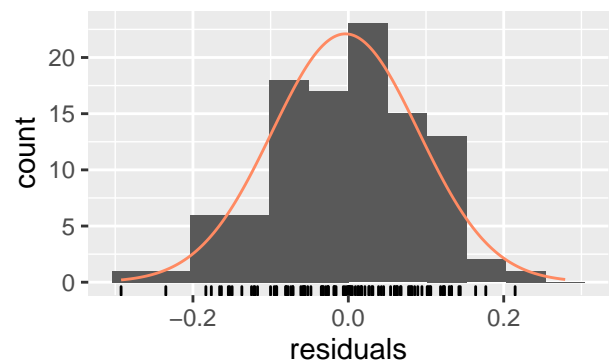
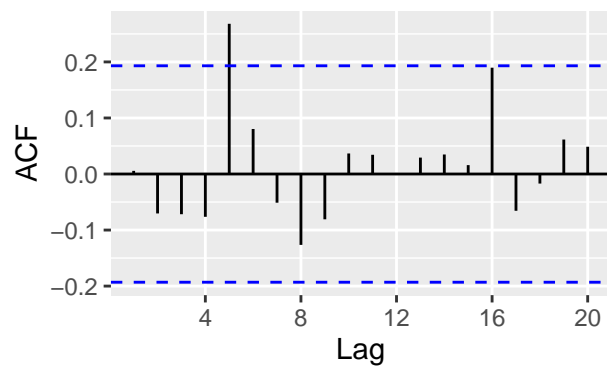
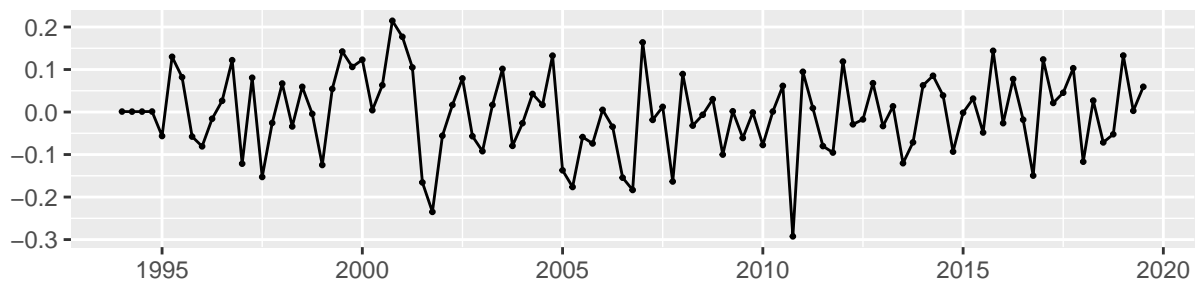
## AIC=-171.43 AICc=-171.18 BIC=-163.65

### Forecasts from ARIMA(1,0,0)(0,1,1)[4]



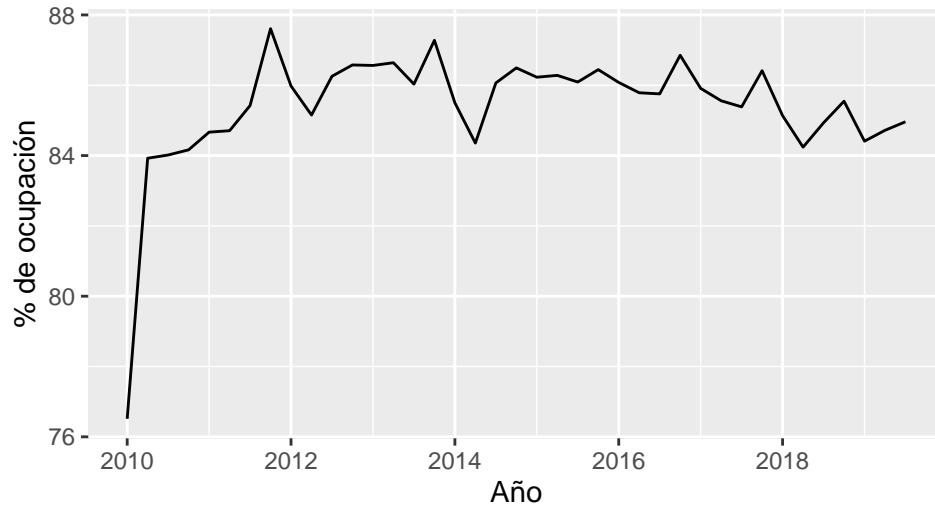
Residuos del forecast:

### Residuals from ARIMA(1,0,0)(0,1,1)[4]



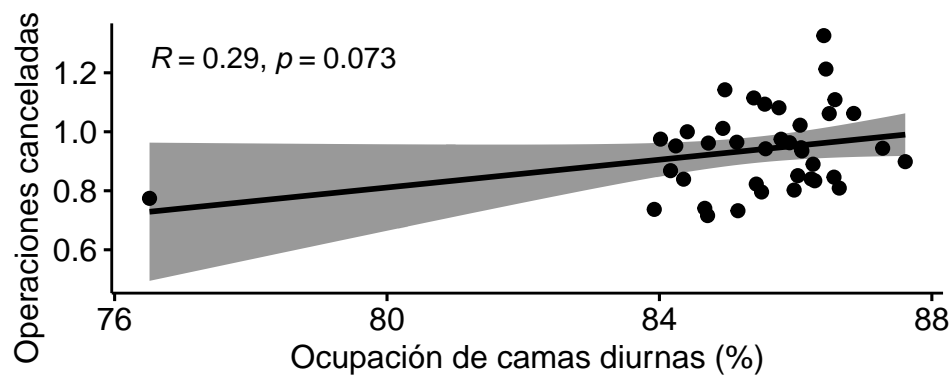
## 4.2 correlación con ocupación de camas

### 4.2.1 Ocupación de camas (Diurnas)



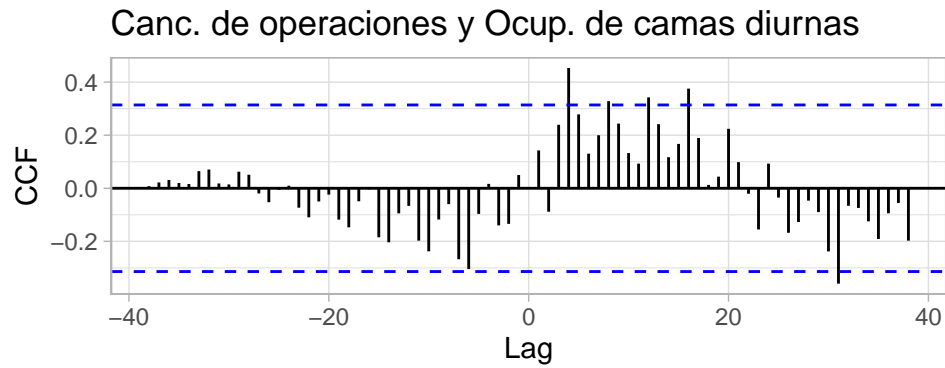
Correlación lineal con Ocupación de camas (Diurnas).

```
## `geom_smooth()` using formula 'y ~ x'
```



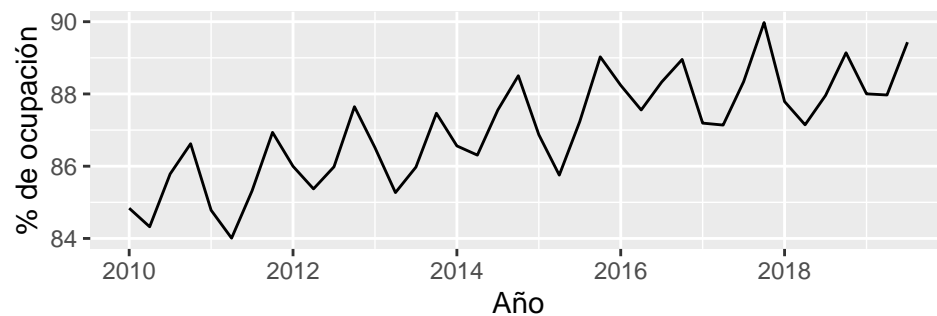
Correlación temporal con Ocupación de camas (Diurnas):

Un aumento (o disminución) en el % de ocupación de las camas diurnas por sobre la media, tiene correlación con un aumento (o disminución) en el % de cancelación de operaciones 1 año (4 muestras) después.



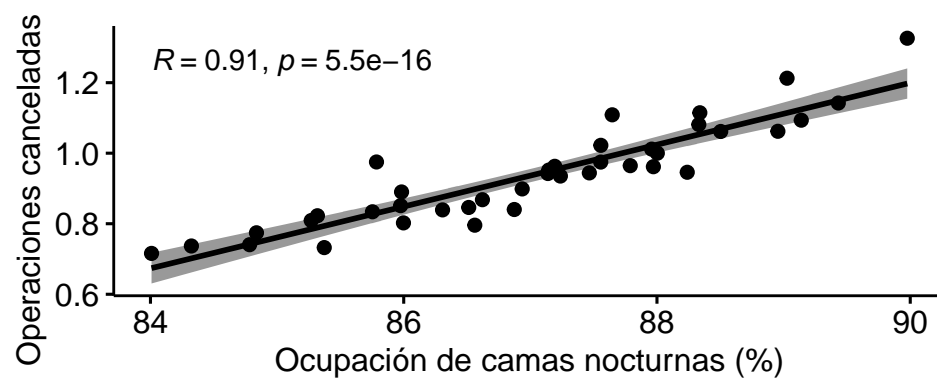
#### 4.2.2 Ocupación de camas (Nocturnas)

La correlación temporal entre las señales no muestra un claro anticipo de una a la otra.



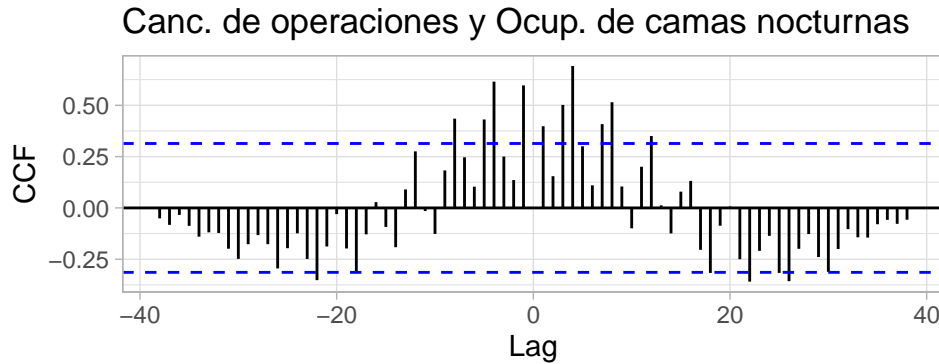
Correlación lineal con Ocupación de camas (Nocturnas).

```
## `geom_smooth()` using formula 'y ~ x'
```



Correlación temporal con Ocupación de camas (Nocturnas).





## 5 Discusión y conclusiones

La principal característica encontrada en la serie temporal de porcentaje de cancelación de operaciones electivas es su marcada estacionalidad. Tiene un periodo anual, dentro del cual se evidencia un mayor porcentaje de cancelación en la segunda mitad del año respecto a la primera mitad. El pico anual se suele alcanzar en el último trimestre del año. Este fenómeno podría estar explicado por un aumento en la ocupación de las camas que se da en la época invernal.

También parecería existir un ciclo, no asociado al calendario, más amplio. Además, existe una elevación atípica alrededor del año 2001 y una depresión alrededor del 2011. Globalmente, la serie de cancelaciones presenta una ligera tendencia a aumentar.

El modelo de forecasting resultó satisfactorio ya que los residuos se aproximan a una distribución normal y prácticamente no muestran autocorrelación, es decir que se parecen al ruido blanco (aleatorio). Existió solamente un pico significativo en la autocorrelación de los residuos. Además, estos se ven menos “aleatorios” alrededor de las irregularidades de la serie en los años 2001 y 2011. Estos últimos hallazgos en los residuos del modelo dan cuenta de que existe una pequeña parte de la información de la serie que no logró ser capturada por el modelo.

Respecto a la correlación con la ocupación de camas diurnas, se observa que no existe una clara relación lineal pero sí temporal. Sin embargo, al incorporar la información de las camas de noche sí se observa una fuerte relación lineal, pero no una clara relación temporal con las cancelaciones de operaciones.

Se concluye que el conjunto de técnicas aplicadas fue lo suficientemente sensible como para detectar patrones interesantes en los datos así como también verificar la hipótesis de la correlación entre la ocupación de camas y el aumento de la cancelación de operaciones.

Como mejoras al proyecto, se sugiere realizar alguna transformación sobre la serie previo a aplicar el modelo ARIMA para que este pueda capturar por completo sus cualidades.

Sería interesante, además, incorporar variables internas de la administración sanitaria, así como también de tipo contextuales del país en estudio.

Futuramente, se intentará replicar este estudio en datos de la República Argentina con el fin de contribuir a mejorar la planificación de los recursos en la gestión del sistema de salud.

## 6 Bibliografía

Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. <https://otexts.com/fpp2/>. Accedido el 22 de octubre de 2020.

Theophano Mitsa (2010) *Temporal Data Mining*, Chapman and Hall/CRC