

Análisis de series temporales de operaciones canceladas

Lic. Malena Álvarez Brito, Ing. Florencia Florio

octubre 2020

Resumen: Este es el resumen. (aproximadamente 250 palabras).

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

1 Introducción

Introducción. Explicar de dónde viene el dataset. Qué significa cada valor. Periodicidad de la serie (frecuencia de muestreo). Qué se entiende por operación "electiva". Qué tipo de instituciones son? públicas y privadas? O sólo públicas?

2 Hipótesis y objetivos

Hipótesis y objetivos.

Objetivo: analizar la serie de tiempo de operaciones canceladas para descubrir patrones en su comportamiento, ver si existe correlación con ocupación de las camas de internación, hacer forecasting de porcentaje de cancelación. Todo esto podría contribuir a una mejor asignación de recursos del sistema de salud y toma de decisión de gestión sanitaria.

3 Material y métodos

Material y métodos.

R, RStudio. Librerías: tidyverse, zoo, xts, forecast, seasonal.

4 Resultados

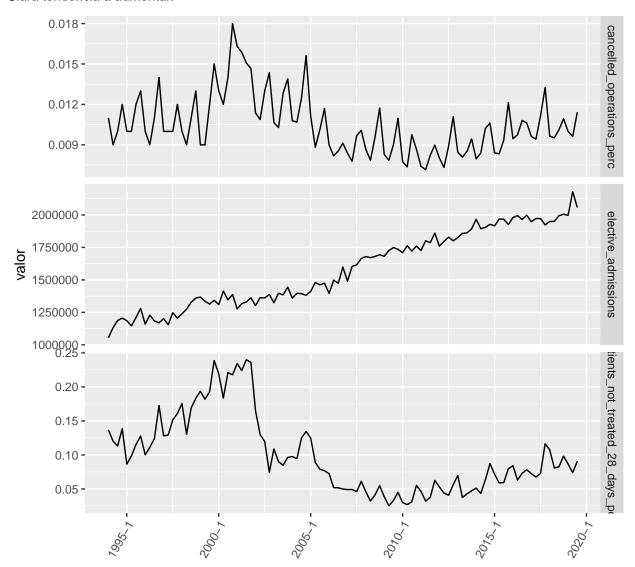
Gráfico temporal de todas las series en el dataset de operaciones canceladas.

4.0.1 Respecto al porcentaje de cancelaciones

Fuerte estacionalidad (sube, baja, sube, baja). Algo pasó en 2001 que aumentó el porcentaje de cancelaciones (casi el doble que en otros trimestres). Además, entre el 2000 y el 2010 parecería haber una tendencia a disminuir. Luego, en 2010 parece haber un cambio de tendencia y empieza a aumentar ligeramente.

4.0.2 Respecto al total de opraciones

Clara tendencia a aumentar.



Viendo el filtro moving average también se evidencia el aumento en 2001.

Operaciones canceladas

0.018 0.015 0.012 Data
4-MA

2005

Año

Warning: Removed 4 row(s) containing missing values (geom_path).

4.0.3 Análisis a partir de los gráficos de estacionalidad

2000

1995

En el siguiente gráfico se ve que el pico fue en realidad en el último trimestre de 2000 pero que todo el 2000 fue alto. También el 2001 fue alto, aunque luego fue bajando.

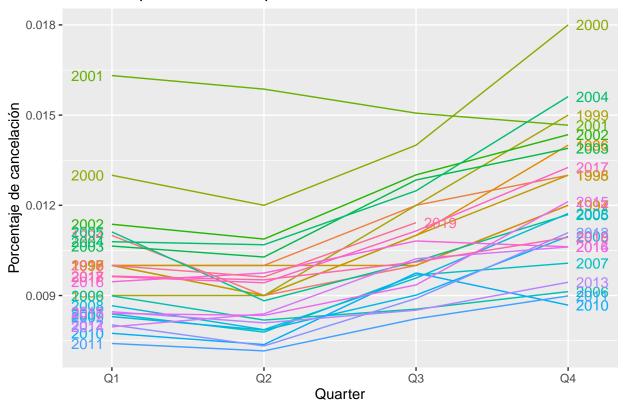
2010

2015

2020

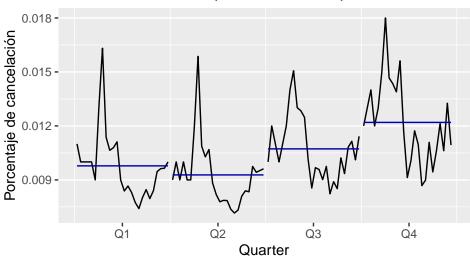
Yo veo, además, que el primer y segundo trimestres suelen ser más bajos que el tercero y el cuarto. El 2010 fue distinto porque el Q4 fue bastante menor al Q3.

Seasonal plot: cancelled operations



En el siguiente gráfico se ve que los primeros dos trimestres son, en promedio, más bajos que el tercero y el cuarto.





Autocorrelación

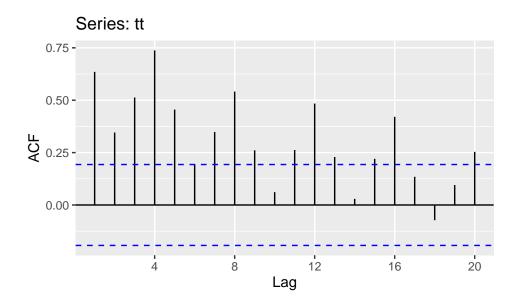
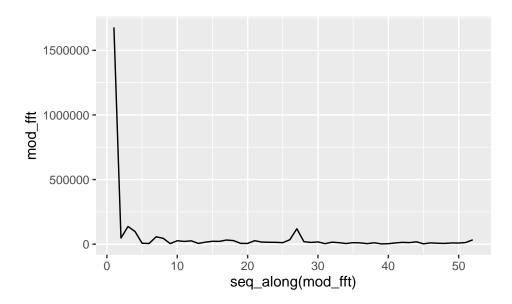
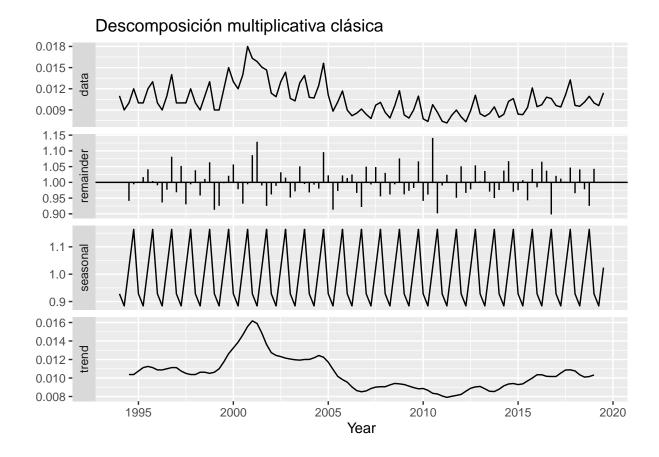


Gráfico de la fft de porcentaje de operaciones canceladas. Se distinguen 2 picos bien marcados (a parte del valor del índice 1...este sería el valor de continua???)



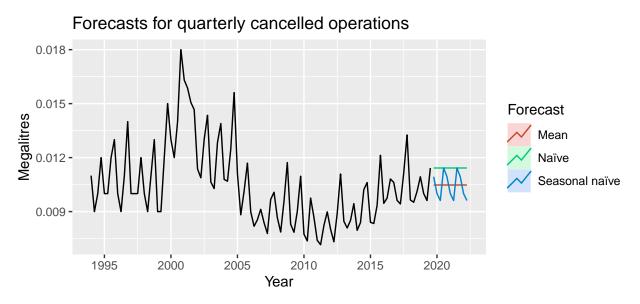
valor	posicion
1677577.0	1
136952.8	3
119513.9	27

Descomposición de la señal. El cuarto gráfico es como el de moving average. Después podemos ver con cuál nos quedamos.



4.1 Forecasting

Forecast simple (preliminar).

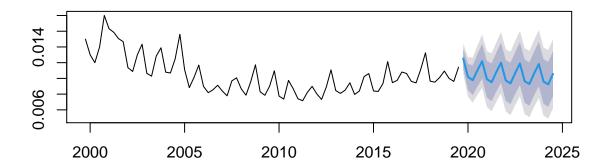


4.1.1 Auto ARIMA forecast

Se aplica un ARIMA automático con estacionalidad.

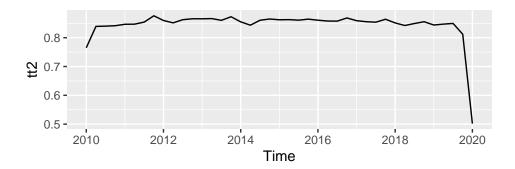
```
## Series: tt
## ARIMA(1,0,0)(0,1,1)[4]
##
##
   Coefficients:
##
            ar1
                    sma1
                 -0.8829
##
         0.8637
## s.e. 0.0546
                  0.0618
## sigma^2 estimated as 9.364e-07: log likelihood=544.63
## AIC=-1083.25
                  AICc=-1083
                               BIC=-1075.47
```

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

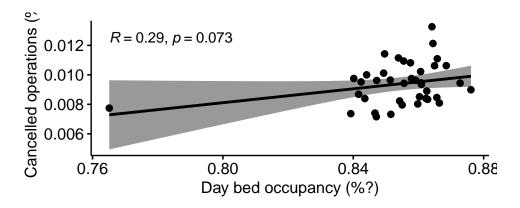


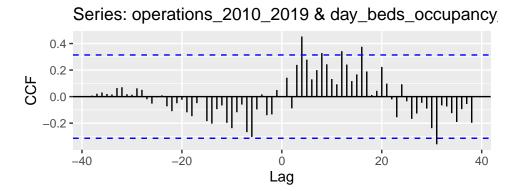
4.2 Correlación con ocupación de camas

4.2.0.1 Correlación con ocupación de camas por el día. Mirando el gráfico de dispersión, no parece haber una relación tan clara entre ambas.

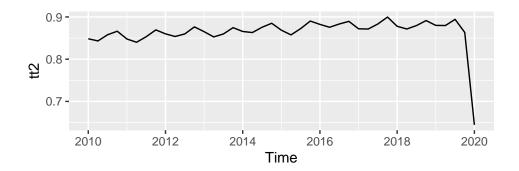


`geom_smooth()` using formula 'y ~ x'

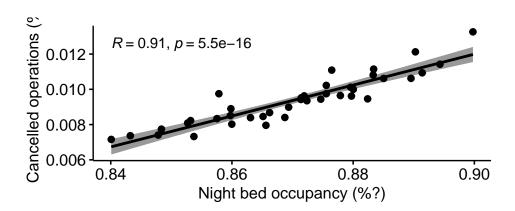




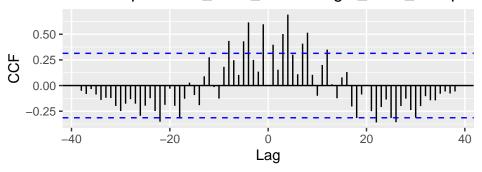
4.2.0.2 Correlación con camas día+noche. En el gráfico de dispersión se ve una relación entre ambas. En decir, cuando hay mucha ocupación de camas nocturnas, hay mayor porcentaje de cancelación de operaciones. Esta relación (aparentemente lineal) no implica causalidad.



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



Series: operations_2010_2019 & night_beds_occupan-



5 Discusión y conclusiones

6 Bibliografía