

# Colgate vs Crest

Antonio Pascual Hernández

20/11/2020

## 1. Resumen

Este informe explica la utilidad de la metodología de las series temporales y los modelos ARIMA en un caso de negocio real de competencia entre Procter and Gamble y Colgate cuando el primero introdujo su marca Crest en 1956. En ese momento, Colgate lideraba el mercado de cuidado bucal con una participación cercana a 50%, cuando el 01 de Agosto de 1960, el respaldo de la American Dental Association (ADA) ocasionó un impacto sobre la participación del mercado, teniendo efectos casi permanentes sobre el rendimiento de las ventas de Crestel.

## 2. Modelo Descriptivo

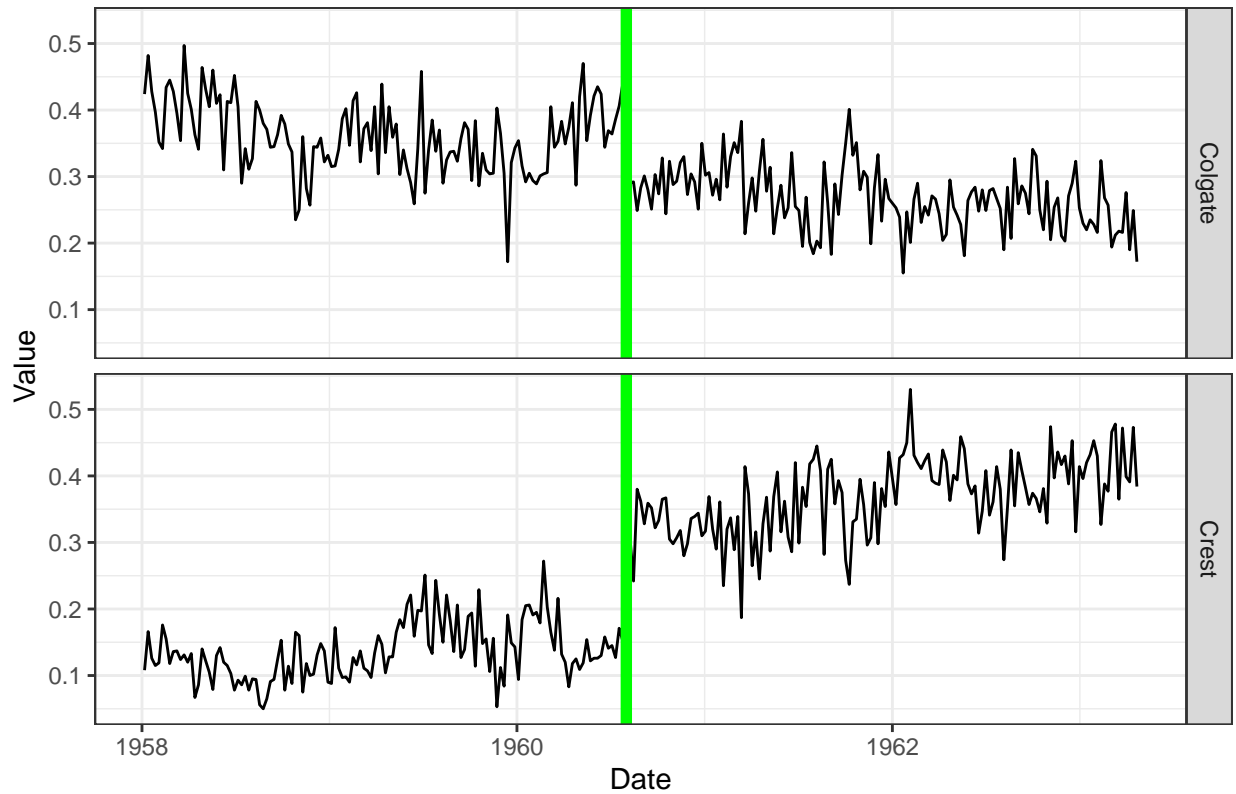
```
xColgate <- xts(colgate, order.by = fechas)
xColgate <- to.weekly(xColgate)
zColgate <- as.zoo(xColgate$xColgate.Close)
xCrest <- xts(crest, order.by = fechas)
xCrest <- to.weekly(xCrest)
zCrest <- as.zoo(xCrest$xCrest.Close)
names(zColgate) <- 'Marketshare'
names(zCrest) <- 'Marketshare'
```

Para la selección de un modelo predictivo adecuado para la información, primero se analizará el comportamiento de las series temporales de las cuotas de mercado de Crest y Colgate-Palmolive.

## 2.1. Análisis Exploratorio de los datos

El dataset testá formado por 276 observaciones semanales (desde Junio de 1958 hasta Abril de 1963).

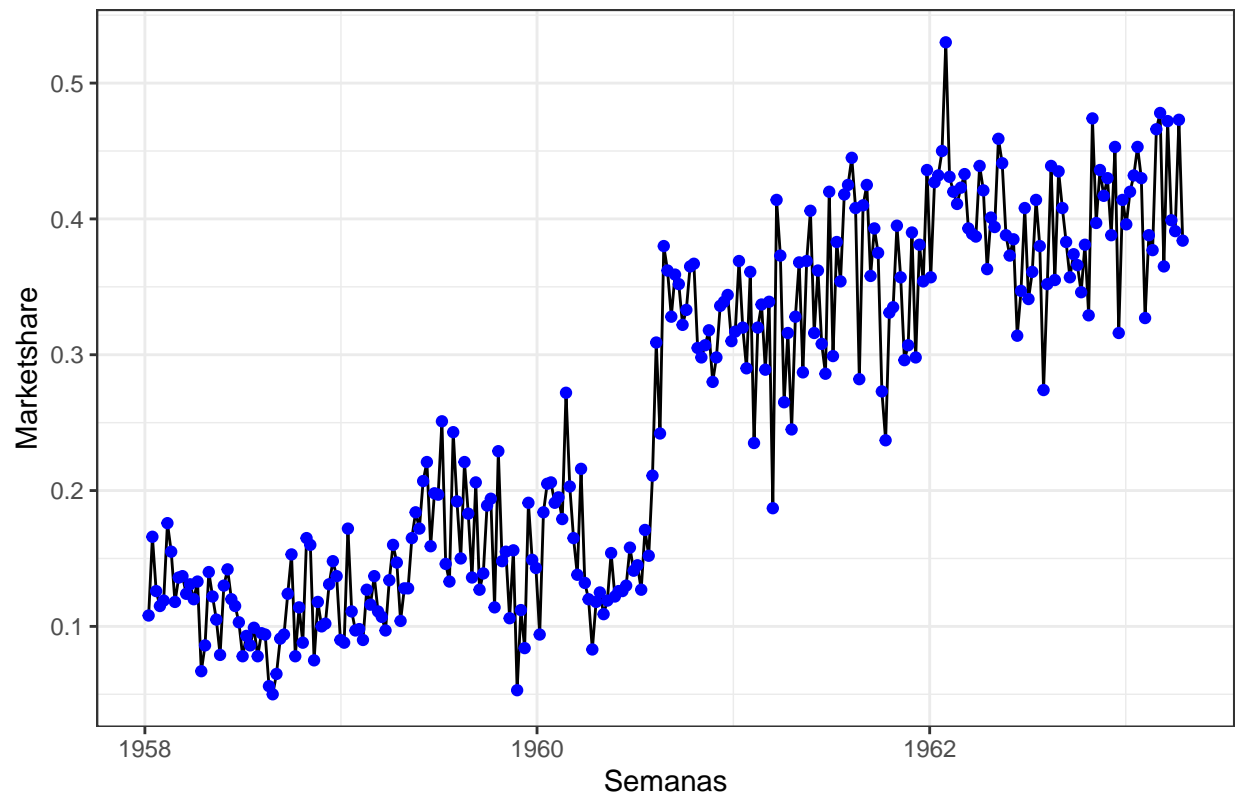
Market Share Toothpaste 1958 – 1963



Las ventas de cremas dentríficas no es estacional dado que no dependen de alguna época específica del año.

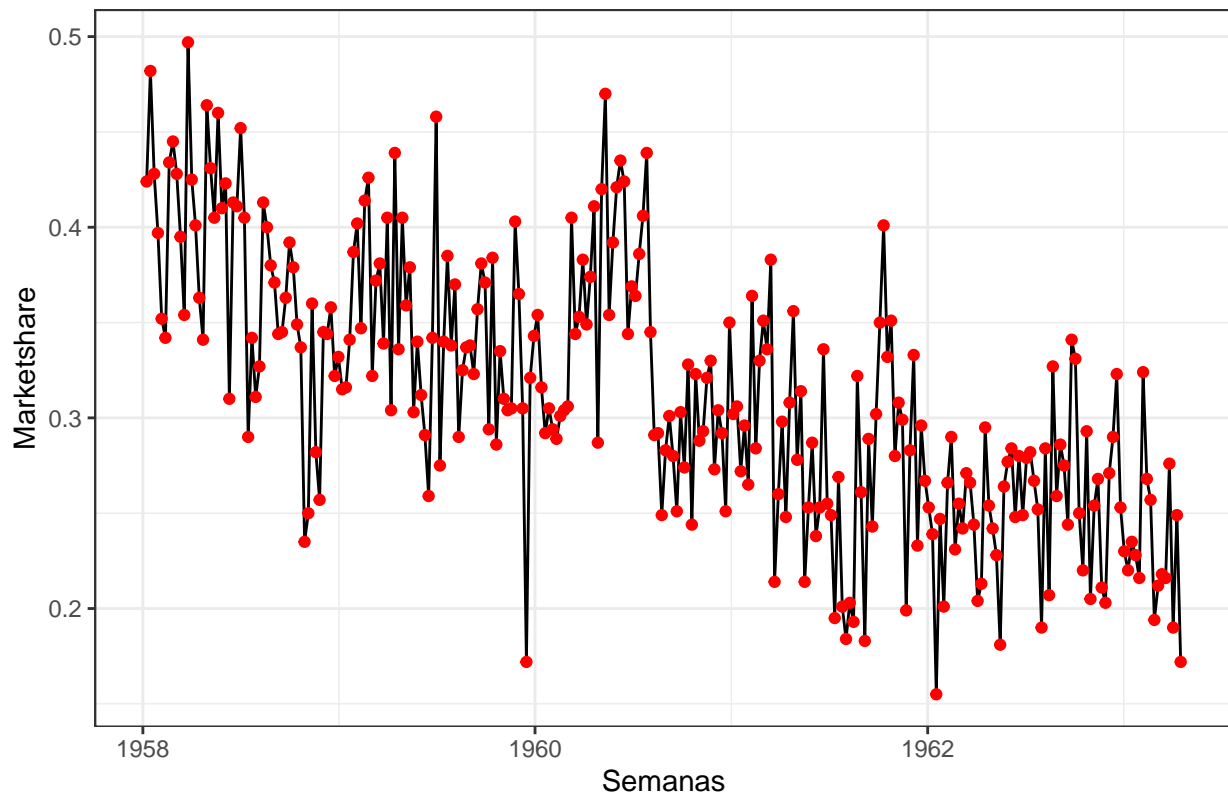
```
autoplot(zCrest) +  
  geom_point(color = "blue") +  
  labs(x = "Semanas", y = "Marketshare", title = "Marketshare Semanal de Crest") +  
  theme_bw()
```

Marketshare Semanal de Crest



```
autoplot(zColgate) +  
  geom_point(color = "red") +  
  labs(x = "Semanas", y = "Marketshare", title = "Marketshare Semanal de Colgate") +  
  theme_bw()
```

### Marketshare Semanal de Colgate



Entre las principales características se encuentran: la serie temporal de Crest tiene tendencia creciente, y da un salto en 1960, donde da un impulso pero no se estabiliza con respecto a las medias, lo que denota que es un evento escalón. Otras características que tiene es que no es estacionaria ni estacional. En el caso de Colgate-Palmolive, también tiene el efecto del escalón en 1960 pero en sentido decreciente y no es estacionaria ni estacional.

### 3. Desarrollo del Modelo.

```
trainCrest <- window(zCrest, start = index(zCrest[1]), end = dCutoff)
trainColgate <- window(zColgate, start = index(zColgate[1]), end = dCutoff)
```

Se separa la información en dos subconjuntos:

Train: Incluirá la información de Junio del 1958 hasta Diciembre del 1962 y se utilizará como entrenamiento para el modelo predictivo. Test: La información de 1963 será utilizada para revisar el modelo ARIMA, y predecir las 16 semanas de la manera eficiente.

```
Crest.arima <- auto.arima(trainCrest)
summary(Crest.arima)
```

```
## Series: trainCrest
## ARIMA(0,1,1)
##
## Coefficients:
```

```
##          ma1
##        -0.6494
## s.e.    0.0447
##
## sigma^2 estimated as 0.002046:  log likelihood=436.25
## AIC=-868.51   AICc=-868.46   BIC=-861.38
##
## Training set error measures:
##              ME          RMSE          MAE          MPE          MAPE          MASE
## Training set 0.003007244 0.04505787 0.03432215 -2.913157 17.20469 0.1396754
##              ACF1
## Training set -0.04618293
```

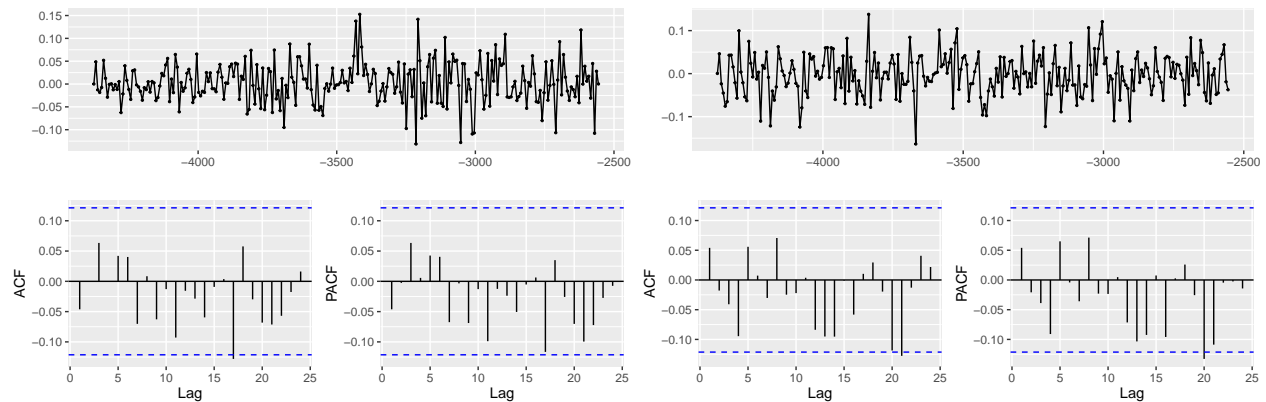
El modelo ARIMA más adecuado para la serie de Crest es de orden 0,1,1 que no tiene características de ser estacional. En cuanto a la parte regular se tiene 0 como valor de p (número de términos autorregresivos), d es 1 que indica que el modelo es estacionario y q es 1 que significa el número de retardos que necesita el modelo.

```
Colgate.arima <- auto.arima(trainColgate)
summary(Colgate.arima)
```

```
## Series: trainColgate
## ARIMA(0,1,1)
##
## Coefficients:
##          ma1
##        -0.7600
## s.e.    0.0463
##
## sigma^2 estimated as 0.002309:  log likelihood=420.38
## AIC=-836.76   AICc=-836.71   BIC=-829.64
##
## Training set error measures:
##              ME          RMSE          MAE          MPE          MAPE          MASE
## Training set -0.002752692 0.04786537 0.03775483 -3.107405 12.80497 0.1189681
##              ACF1
## Training set 0.05403327
```

El modelo de la serie Colgate-Palmolive también es un ARIMA(0,1,1), que no tiene características de parte estacional, como se demostró en el análisis exploratorio de datos.

Debido a que la serie temporal no es estacionaria, tenemos que convertirla mediante la diferencia de orden d. Como se observa en los correlogramas ACF el modelo es eficiente por que lo que queda de la diferenciación es ruido blanco.



La prueba Box-Ljung contrasta la hipótesis nula de que los residuos tienen correlación, se busca rechazar esta hipótesis con la finalidad de identificar que el error residual no está relacionado y se puede considerar como ruido blanco. Con un p-value es de 0.6505 para el caso de Crest y de 0.5704 de Colgate se rechaza la hipótesis nula para ambas empresas.

```
##
## Box-Ljung test
##
## data: Crest.arima$residuals
## X-squared = 13.303, df = 16, p-value = 0.6505

##
## Box-Ljung test
##
## data: Colgate.arima$residuals
## X-squared = 14.381, df = 16, p-value = 0.5704
```

## Detección de Outliers

El siguiente paso es identificar si se tiene outliers aditivos (que afectan la serie temporal) e innovativos (que afectan el error).

```
##           [,1]      [,2]      [,3]
## ind      135.000000 136.000000 138.000000
## lambda2    3.934052  4.389722  4.020833

## [1] "No IO detected"
```

Para el caso de Crest se identifica la intervención de ADA como un outlier aditivo (Observación 135). También se encontraron otros outliers aditivos pero podrían ser efectos de la observación 135. No se han identificado outliers innovativos. En el caso de Colgate no se detectaron outliers aditivos ni innovativos.

```
## [1] "No AO detected"

## [1] "No IO detected"
```

## Modelo de Intervención

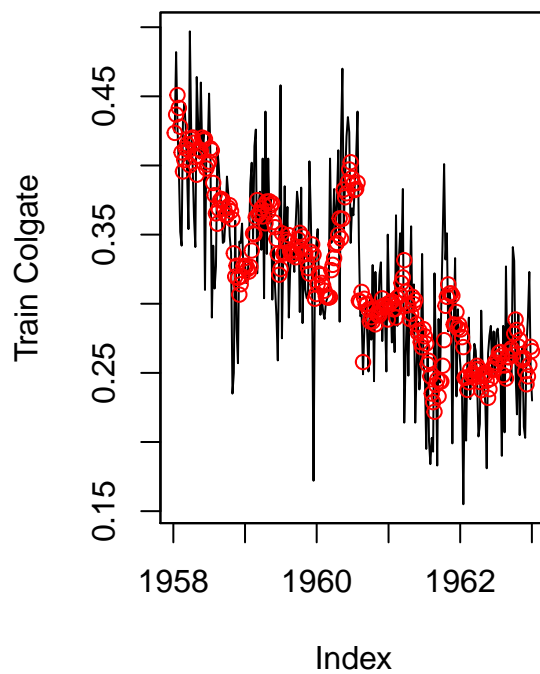
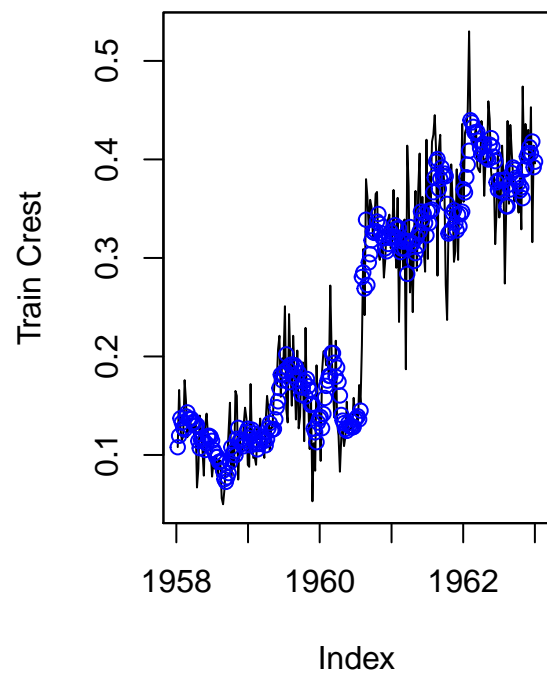
El primer paso será usar una función de transferencia simple (siguiendo la propiedad de parsimonia) de orden 0,0 y se identifica el momento en que ADA en la función de transferencia del tipo escalón debido a que es un outlier aditivo:

```
##
## Call:
## arimax(x = trainCrest, order = c(0, 1, 1), xreg = data.frame(out136 = 1 * (seq(trainCrest) ==
##      136), out138 = 1 * (seq(trainCrest) == 138)), method = "ML", xtransf = data.frame(ADA.Int = 1 *
##      (seq(trainCrest) >= 135)), transfer = list(c(0, 0)))
##
## Coefficients:
##          ma1  out136  out138  ADA.Int-MA0
##      -0.7446  0.0224  0.0768         0.1338
## s.e.    0.0492  0.0431  0.0414         0.0326
##
## sigma^2 estimated as 0.00187:  log likelihood = 447.31,  aic = -886.62
```

El modelo obtiene un AIC de -886 y calcula el coeficiente de la diferenciación de la media -0.7446. Se ha estimado como 0.13338 para la intersección de la incidencia de ADA en la serie temporal. Este valor sirve para estimar los efectos de la intervención.

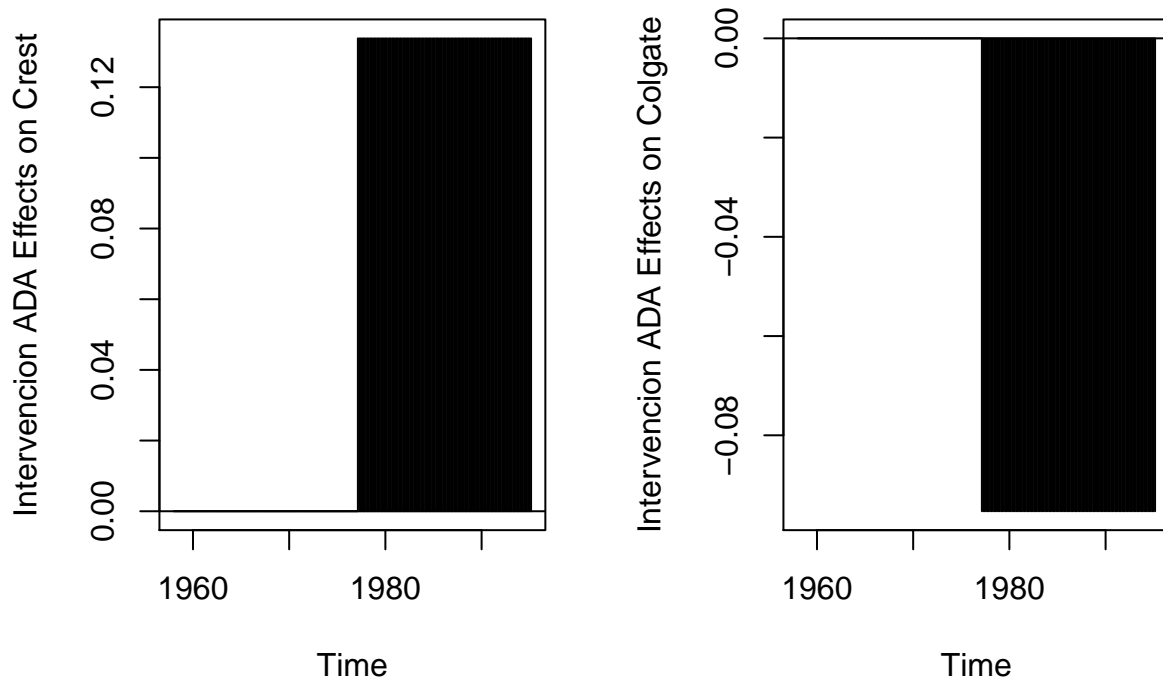
```
##
## Call:
## arimax(x = trainColgate, order = c(0, 1, 1), xreg = data.frame(out136 = 1 *
##      (seq(trainColgate) == 136), out138 = 1 * (seq(trainColgate) == 138)), method = "ML",
##      xtransf = data.frame(ADA.Int = 1 * (seq(trainColgate) >= 135)), transfer = list(c(0,
##      0)))
##
## Coefficients:
##          ma1  out136  out138  ADA.Int-MA0
##      -0.8049 -0.0092 -0.0475        -0.0953
## s.e.    0.0435  0.0463  0.0454         0.0292
##
## sigma^2 estimated as 0.002186:  log likelihood = 426.91,  aic = -845.81
```

El modelo de Colgate- Palmolive obtiene un AIC de -845 y calcula el coeficiente de la diferenciación de la media -0.8049. Se ha estimado como -0.0953 para la intersección de la incidencia de ADA en la serie temporal. Este valor sirve para estimar los efectos de la intervención.



Los modelos se ajustan a las series temporales adecuadamente. Sobre los efectos de ADA se tiene los siguientes graficos:



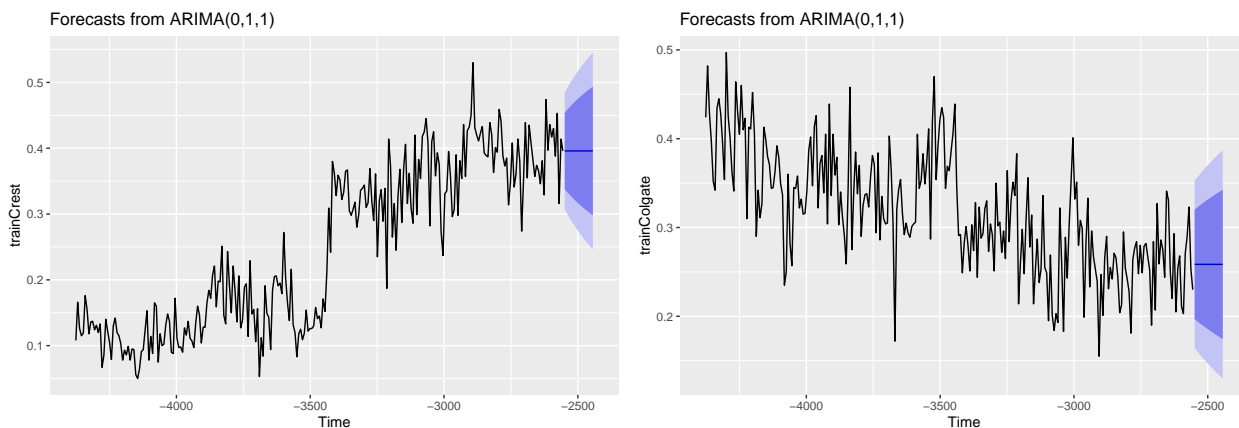


La intervención de ADA agrega un 0.13 a la serie temporal de Crest, en cambio para Colgate, significa la reducción del 0.09.

Según la prueba de Wald (función `coeftest`) la intervención de ADA es significativa y se comprueba en ambos modelos al calcular un coeficiente `ADA.Int-MA0`. Sin embargo, los outliers 136 y 138 no son significativos.

## Predicción

Sobre la predicción de las 16 semanas, se utiliza el modelo ARIMA. La predicción del ARIMA se representa por una línea debido a que solo usamos un retardo y no tiene estacionalidad.

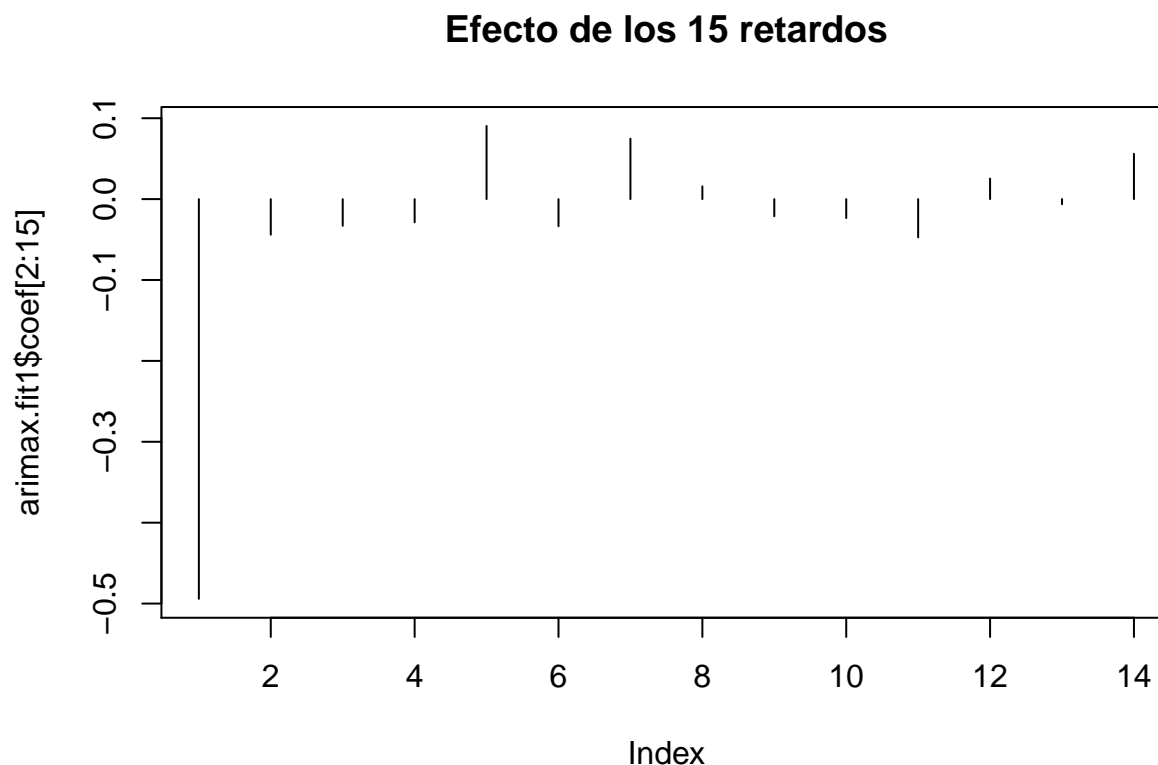


## Relación entre Crest y Colgate e Impacto de ADA

Dado que la intervención de ADA demuestra ser muy significativa, ahora queda comprobar esta afirmación con una función de transferencia de la siguiente manera: se considera la observación 135 como punto de corte para realizar diferencias a partir de ellas, y en la serie temporal de colgate-palmolive se le asigna como función de transferencia el efecto de la serie temporal de crest.

La función de transferencia inicial es de orden 15 para el numerador.

El gráfico del efecto de los 15 retardos muestra lo siguiente:



Con el gráfico del efecto de los 15 retardos se observa, que la función de transferencia corresponde a un ajuste de (NA,NA,0,0,NA). Tiene un decaimiento complicado ( $r=2$ ) y un retardo de 1 ( $b=1$ ).

```
##
## Call:
## arimax(x = dColgate134, order = c(0, 1, 1), include.mean = TRUE, fixed = c(NA,
##      NA, 0, 0, NA), method = "ML", xtransf = dCrest134, transfer = list(c(1,
##      2)))
##
## Coefficients:
##      ma1  Marketshare-AR1  Marketshare-MA0  Marketshare-MA1
##      -0.9999          0.2459              0              0
## s.e.    0.0096          2.9924              0              0
##      Marketshare-MA2
##      -0.0206
## s.e.    0.1787
##
```

```
## sigma^2 estimated as 0.003445: log likelihood = 361.28, aic = -716.56
```

El modelo resultante tiene un AIC de -716.56. Se observa basado en los coeficientes que el impacto de Crest sobre Colgate es negativo.

#### 4. Conclusiones

La venta de cremas dentríficas no es estacional ni estacionaria.

La intervención del Consejo de Terapéutica Dental de la American Dental Association (ADA) influyó tanto en el rendimiento de las ventas de la marca Crest como también en las ventas de Colgate-Palmolive.

Sobre la relación dinámica entre las ventas de Crest y Colgate, se demostró mediante la función de transferencia que es negativa, que el incremento de las ventas en Crest produjo un efecto negativo en el rendimiento de Colgate.