

# Colgate vs Crest

Daniel Corral Ruiz

23-11-2020

## Contents

Objetivos informe . . . . .	1
Importación de datos y desarrollo del modelo . . . . .	1
Carga de datos . . . . .	1
Análisis exploratorio (EDA) . . . . .	2
Creación del modelo . . . . .	3
ARIMAX . . . . .	5
Predicción para el año 1963 . . . . .	6
Modelo de función de transferencia entre cuotas . . . . .	7
Conclusiones del informe . . . . .	8

## Objetivos informe

Consideramos la lucha por la cuota de mercado entre la crema dental Colgate de Colgate-Palmolive y la crema dental Crest de Procter y Gamble. El objetivo es predecir las 16 semanas del año 1963.

Procter y Gamble introdujeron Crest en el mercado estadounidense en 1956. Antes de ese momento, Colgate disfrutaba de un liderazgo de mercado con una participación de mercado cercana al 50%. Durante los siguientes 4 años, Colgate siguió siendo un competidor dominante y Crest solo logró una participación de mercado relativamente modesta pero estable del 15%. Sin embargo, el 1 de agosto de 1960, el Consejo de Terapéutica Dental de la American Dental Association (ADA) aprobó a Crest como una “ayuda importante en cualquier programa de higiene dental”.

El análisis incluye para las dos series la estimación de un modelo ARIMA, la detección de atípicos/outlier y un modelo de intervención. Por último un modelo de función de transferencia entre las dos cuotas.

## Importación de datos y desarrollo del modelo

En el fichero “data.xlsx” se encuentran las cuotas del mercado de Colgate y Crest semanales desde 1958 hasta abril de 1963.

## Carga de datos

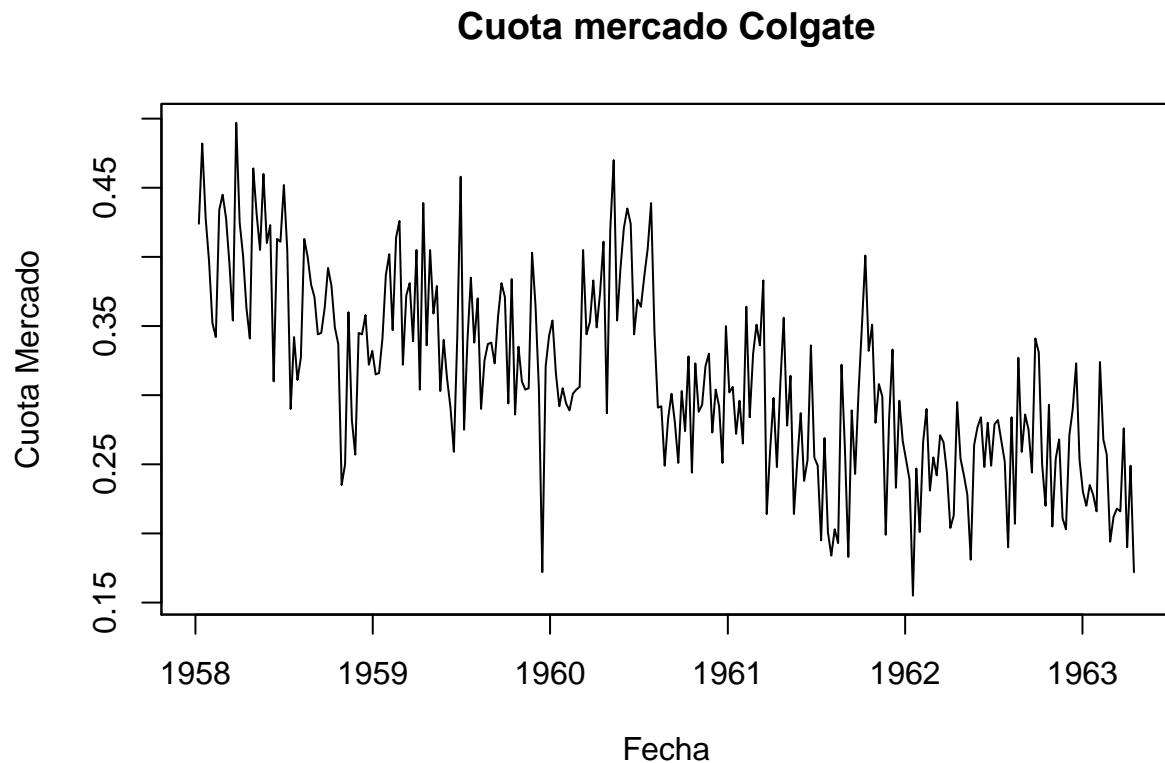
Comenzamos cargando el dataset mediante archivo “data.xlsx”, además de la transformación de los formatos.

```
## # A tibble: 10 x 4
##   Year Week Crest Colgate
##   <dbl> <dbl> <dbl>   <dbl>
## 1 1958     1 0.108   0.424
## 2 1958     2 0.166   0.482
## 3 1958     3 0.126   0.428
## 4 1958     4 0.115   0.397
## 5 1958     5 0.119   0.352
```

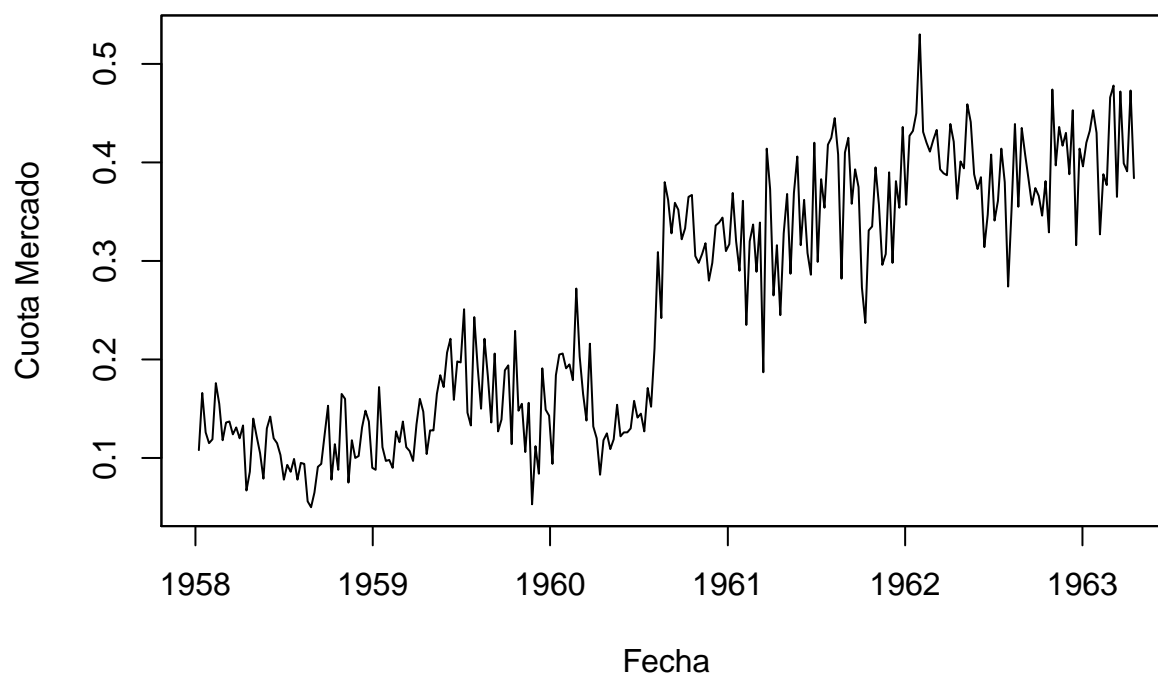
##	6	1958	6	0.176	0.342
##	7	1958	7	0.155	0.434
##	8	1958	8	0.118	0.445
##	9	1958	9	0.136	0.428
##	10	1958	10	0.137	0.395

## Análisis exploratorio (EDA)

Antes de comenzar con la realización del modelo predictivo necesitaremos conocer los datos con los que vamos a trabajar. Para ello deberemos conocer la existencia de valores nulos, así como análisis visual de los datos. No existen observaciones con valores nulos o NA, por lo que no tendremos problemas en ese aspecto a la hora de tratar con los datos. Los gráficos siguientes muestran las cuotas de mercados para ambas marcas (Colgate y Crest) entre 1958 y 1963. Podemos observar la tendencia decreciente de Colgate, al contrario que la marca Crest, en especial el aumento de la cuota de mercado en el momento que el Consejo de Terapéutica Dental de la American Dental Association aprobó a Crest como una ayuda importante al higiene dental. Se aprecia a mediados de agosto de 1960 un importante crecimiento en la cuota de Crest.



## Cuota mercado Crest



Podemos resumir: la serie temporal de la empresa Crest presenta tendencia creciente, dando un salto importante en su cuota de mercado. Además, la cuota no se estabiliza con respecto a las medias, por lo que observamos que se trata de un evento escalón.

## Creación del modelo

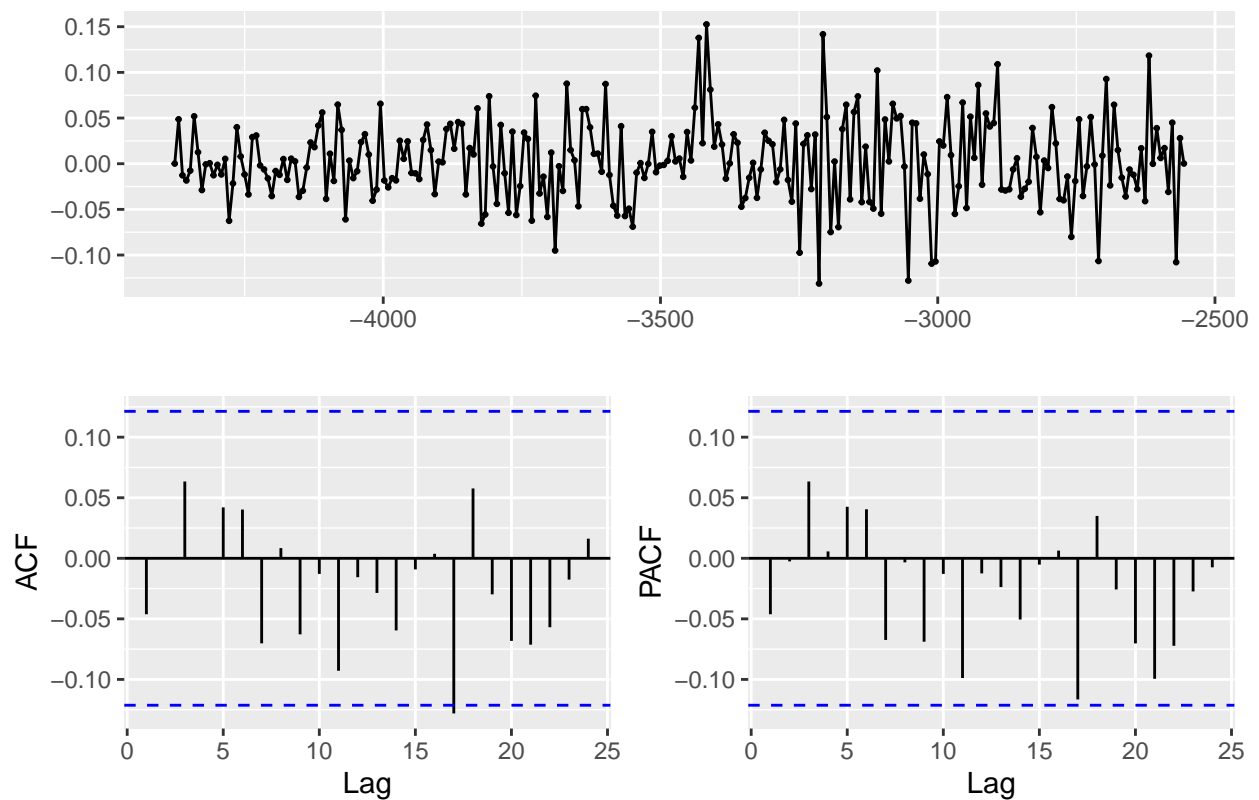
Para la realización del modelo deberemos de dividir los datos en dos conjuntos: - El primero de ellos es el conjunto “train”. Este primer conjunto se utilizará para entrenar al modelo de predicción. En nuestro caso obtendrá los datos para cada una de las empresas desde 1958 hasta diciembre de 1962. - El segundo conjunto es el “test”. Este último conjunto obtendrá los datos para predecir las 16 semanas del año 1963, por lo tanto obtendrá los datos para la estimación de este año.

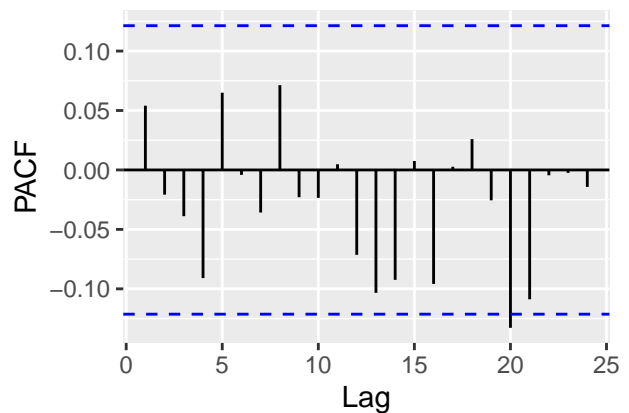
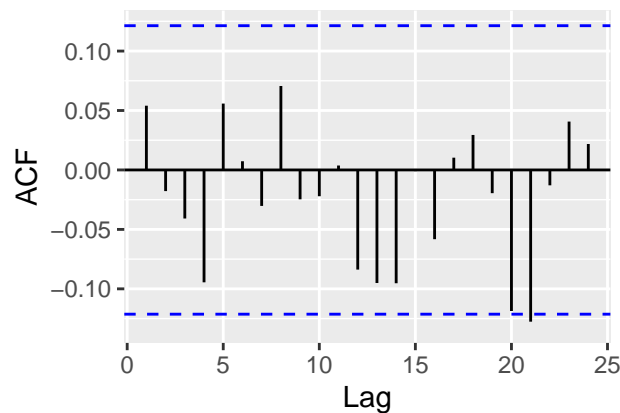
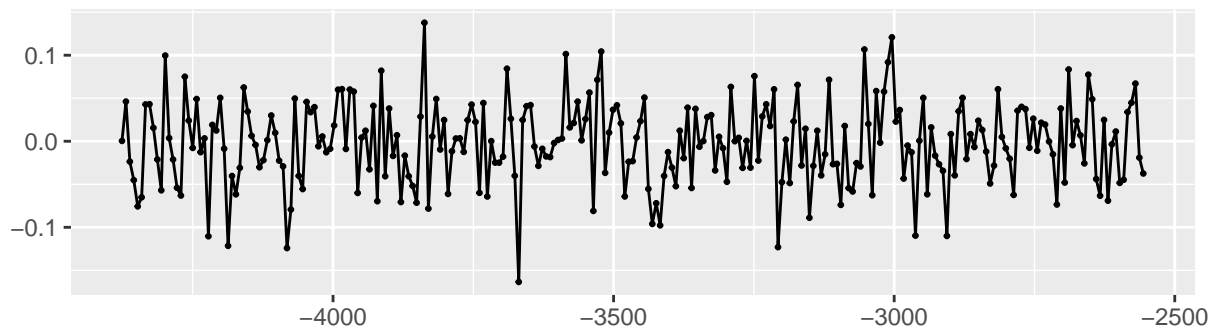
```
## Series: cresttrain
## ARIMA(0,1,1)
##
## Coefficients:
##      ma1
##      -0.6494
## s.e.    0.0447
##
## sigma^2 estimated as 0.002046:  log likelihood=436.25
## AIC=-868.51  AICc=-868.46  BIC=-861.38
##
## Training set error measures:
##              ME      RMSE      MAE      MPE      MAPE      MASE
## Training set 0.003007244 0.04505787 0.03432215 -2.913157 17.20469 0.1396754
##              ACF1
```

```
## Training set -0.04618293
## Series: colgatetrain
## 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
```

Una vez realizado los modelos arimas para ambas empresas, obtenemos como para ambas empresas el modelo ARIMA más adecuado es un ARIMA (0,1,1). Ambas sin características de serie estacional.

En las siguientes gráficas se muestran los outliers, con el fin de identificar outliers aditivos e innovativos. En el caso de Crest se identifican outliers aditivos en las observaciones 135, 136 y 138. Estos dos últimos podrían ser efectos de la primera observación.





```
##           [,1]      [,2]      [,3]
## ind      135.000000 136.000000 138.000000
## lambda2   3.934052  4.389722  4.020833
## [1] "No A0 detected"
## [1] "No I0 detected"
## [1] "No I0 detected"
```

## ARIMAX

Estudiaremos los modelos ARIMAX como modelos de intervención. Al identificar el momento en el que existe el escalón ocasionado por el ADA, estudiamos estos puntos, causantes del outlier aditivo.

```
##
## Call:
## arimax(x = cresttrain, order = c(0, 1, 1), xreg = data.frame(out136 = 1 * (seq(cresttrain) ==
##      136), out138 = 1 * (seq(cresttrain) == 138)), method = "ML", xtransf = data.frame(ADA.Int = 1 *
##      (seq(cresttrain) >= 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
```

Obtenemos un segundo modelo para Crest con un AIC de -886.62.

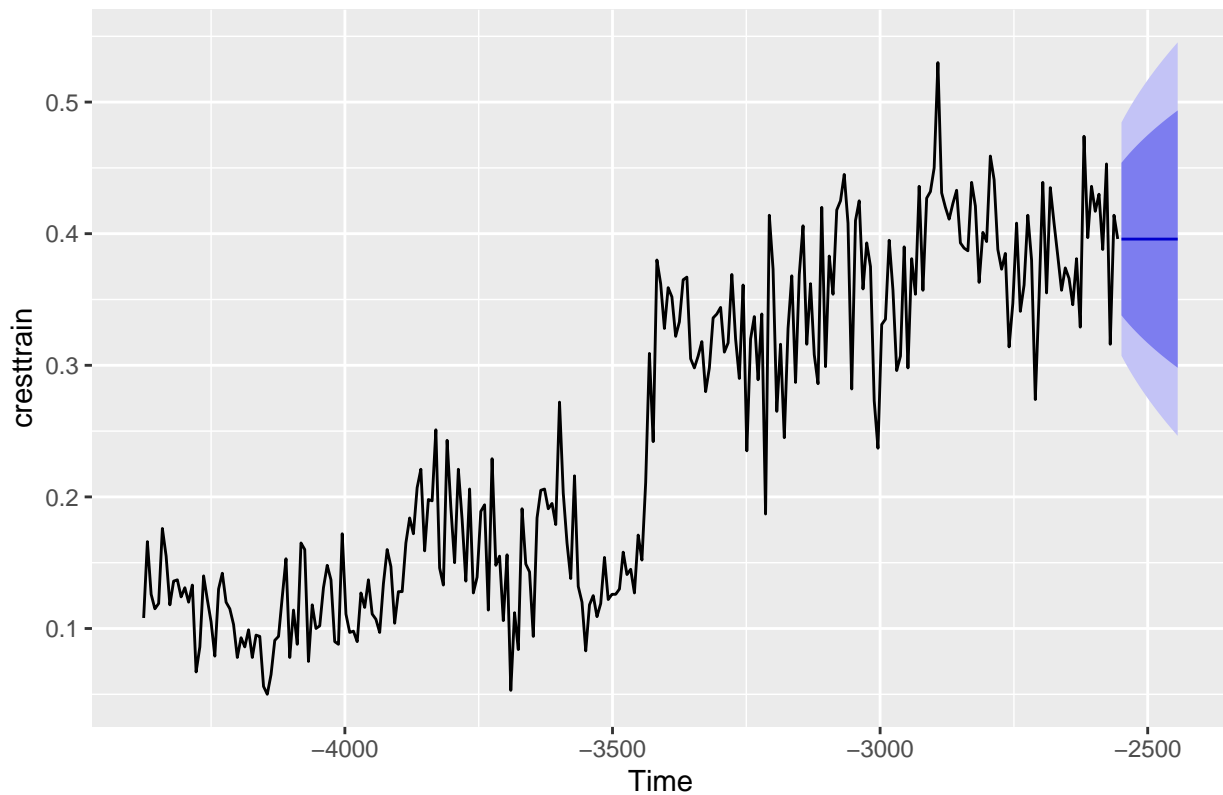
```
##
## Call:
## arimax(x = colgatetrain, order = c(0, 1, 1), xreg = data.frame(out136 = 1 *
##   (seq(colgatetrain) == 136), out138 = 1 * (seq(colgatetrain) == 138)), method = "ML",
##   xtransf = data.frame(ADA.Int = 1 * (seq(colgatetrain) >= 135)), transfer = list(c(0,
##     0)))
##
## Coefficients:
##      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
```

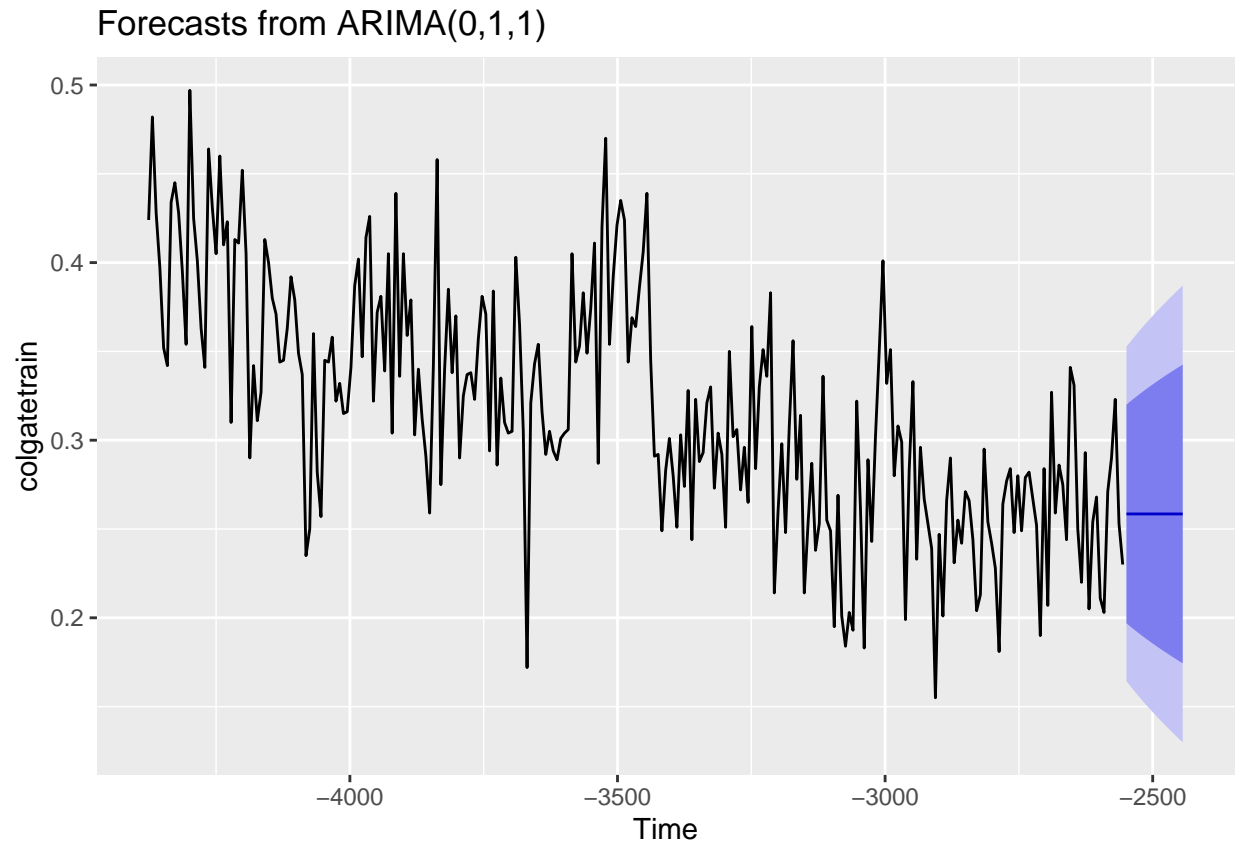
Obtenemos un segundo modelo para Colgate con un AIC de -845.81.

## Predicción para el año 1963

Continuamos con la predicción de ambas compañías para el año 1963. EL primer gráfico muestra la predicción para la empresa Crest, mientras que el segundo gráfico para la empresa Colgate.

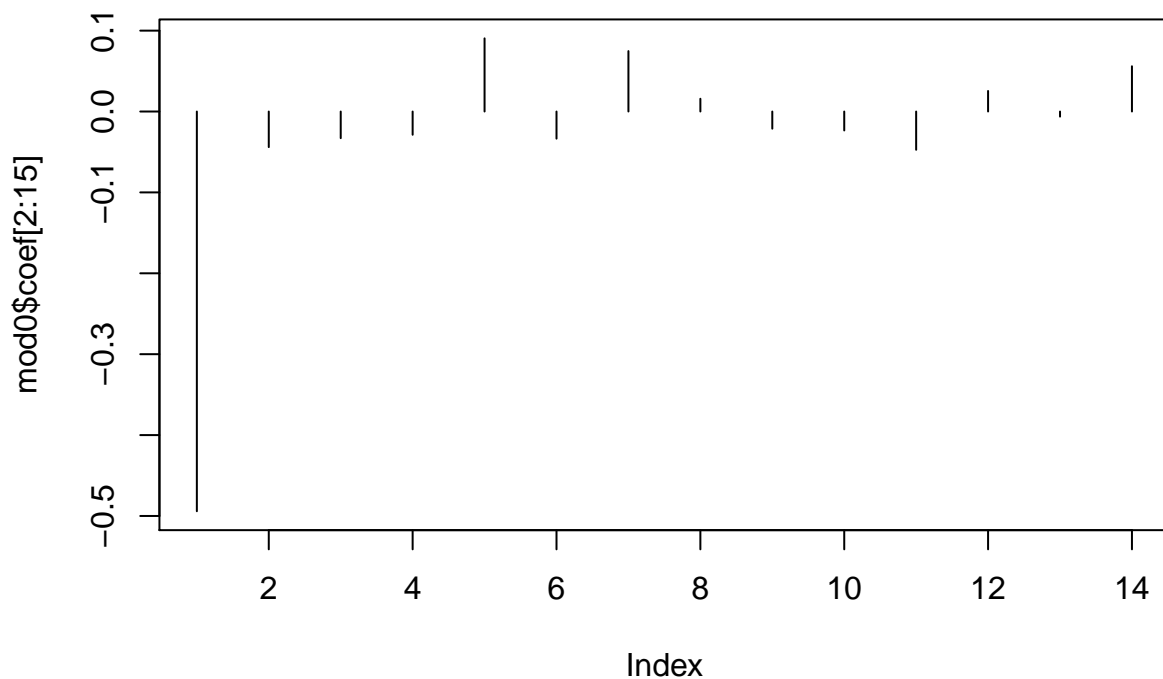
Forecasts from ARIMA(0,1,1)





### Modelo de función de transferencia entre cuotas

Realizamos la función de transferencia con la observación 135 como punto de corte, ya que es la observación anterior al anuncio del ADA. Podemos observar en la gráfica el efecto de los 15 retardos.



```
##
## Call:
## arimax(x = dcolgate, order = c(0, 1, 1), include.mean = TRUE, fixed = c(NA,
##      NA, 0, 0, NA), method = "ML", xtransf = dcrest, transfer = list(c(1, 2)))
##
## Coefficients:
##          ma1  xcrest.Close-AR1  xcrest.Close-MA0  xcrest.Close-MA1
##      -0.9999          0.2459              0              0
## s.e.   0.0096          2.9924              0              0
##      xcrest.Close-MA2
##          -0.0206
## s.e.         0.1787
##
## sigma^2 estimated as 0.003445:  log likelihood = 361.28,  aic = -716.56
```

## Conclusiones del informe

- La noticia del ADA influyo en ambas empresas, dando un impulso en la cuota de mercado de Crest y viceversa para Colgate.
- Observamos una relación inversa entre el impacto del EDA en ambas empresas, gracias a la función de transferencias al tener coeficientes negativos.