

## OBJETIVO

Nuestro objetivo en este trabajo es realizar una predicción para la base de datos “*Primas\_mapfre*” que contiene 4 variables de 40 observaciones cada una. Los datos se nos presentan en cuatrimestres.

Las variables son las siguientes:

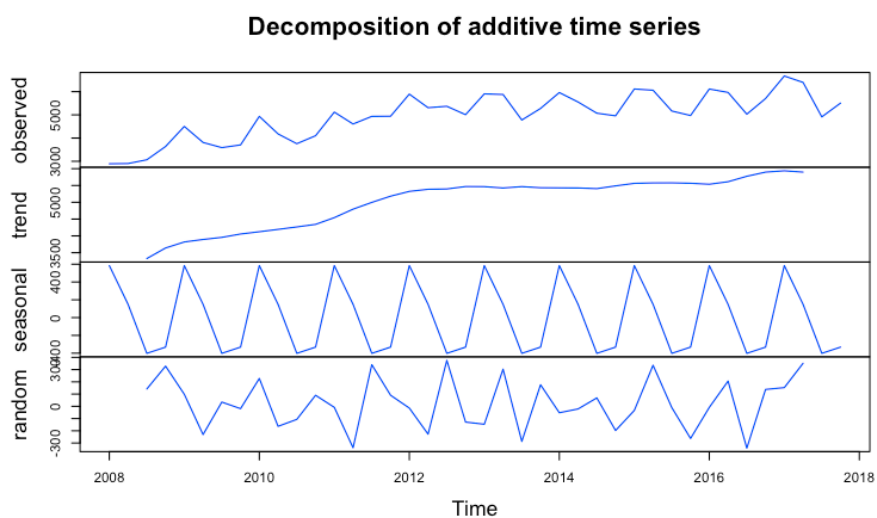
- **Trimestres:** Nos muestra los cuatrimestres desde 2008 a 2017
- **Fecha:** La fecha del dato en formato Mes/día/año
- **Primas\_vida:** Primas de los seguros de vida
- **Primas\_no\_vida:** Primas de los seguros que no son de vida.

Vamos a sumar las primas de los seguros de vida y las primas de los seguros que no son de vida, para posteriormente realizar la predicción del modelo para el 2018.

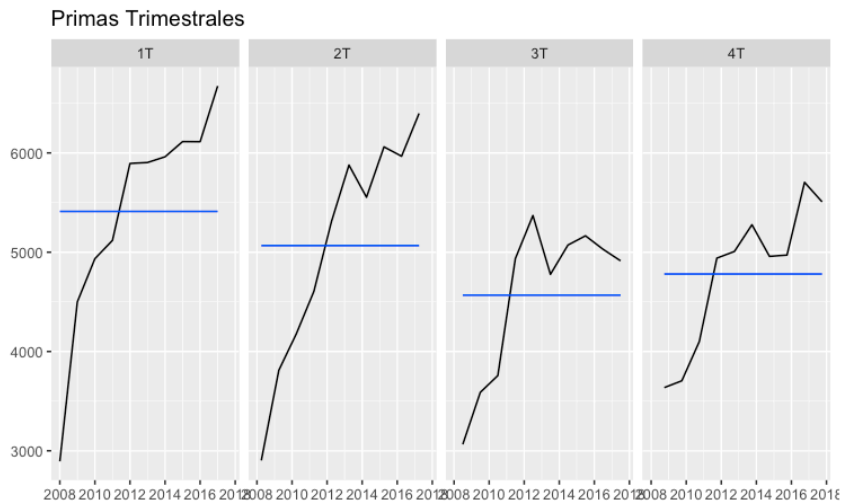
## Índice

- **Análisis exploratorio**
- **Manipulación de datos**
- **Modelo ETS**
- **Modelo ARIMA**

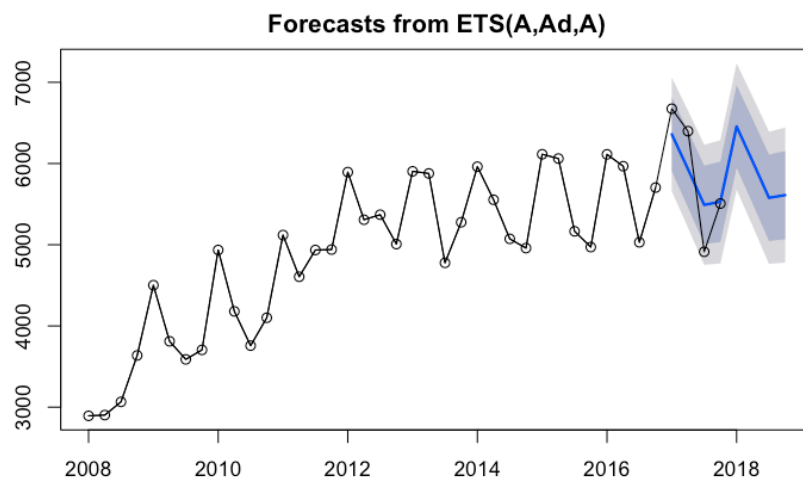
En primer lugar, realizamos una descomposición de la serie temporal:



Gracias al cual podemos observar como la tendencia es creciente positiva, así como que existe estacionalidad.



Mediante el siguiente gráfico podemos observar la distribución de las primas por cuatrimestres desde 2008 a 2018; en los cuatro cuatrimestres ha habido una subida de las primas de los seguros siendo esta más alta en el primero y en el segundo.



Realizamos la predicción para 2018 con el **modelo ETS**, siendo el resultado un modelo **"A, Ad, A"** (*A = Additive, Ad = Additive damped*).

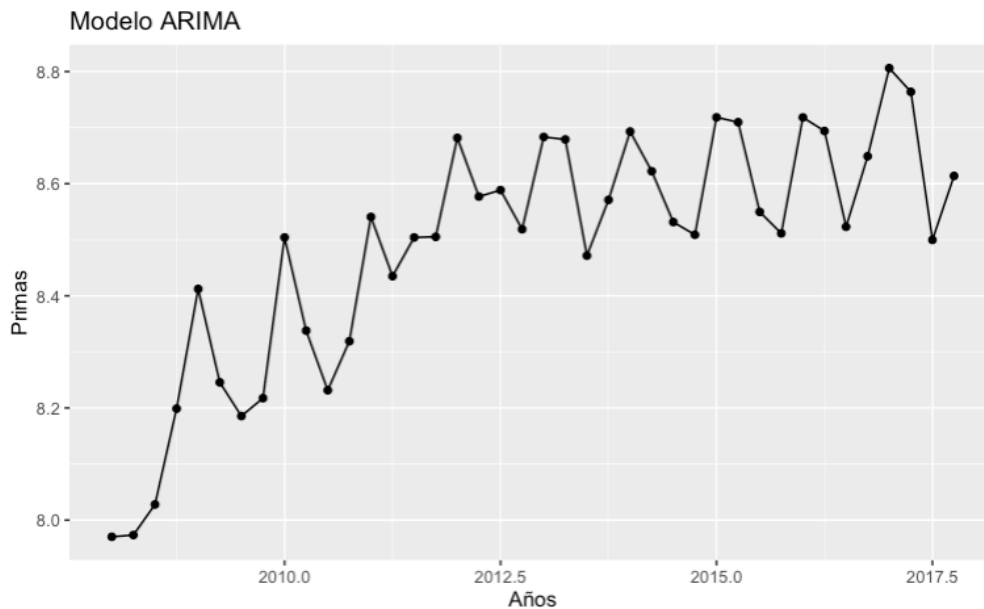
El resultado del modelo se corresponde con las siglas ETS, es decir:

**Error = A:** Tendencia lineal | **Tendencia = Ad:** Tendencia lineal amortiguada | **Estacionalidad = A:** Tendencia lineal.

La parte gris nos indica el intervalo de confianza de la predicción mientras que la línea azul es la media que calcula dentro de ese intervalo. Obtenemos un error medio cuadrático de 308.7 que posteriormente compararemos con el modelo ARIMA para ver cual realiza una mejor predicción.

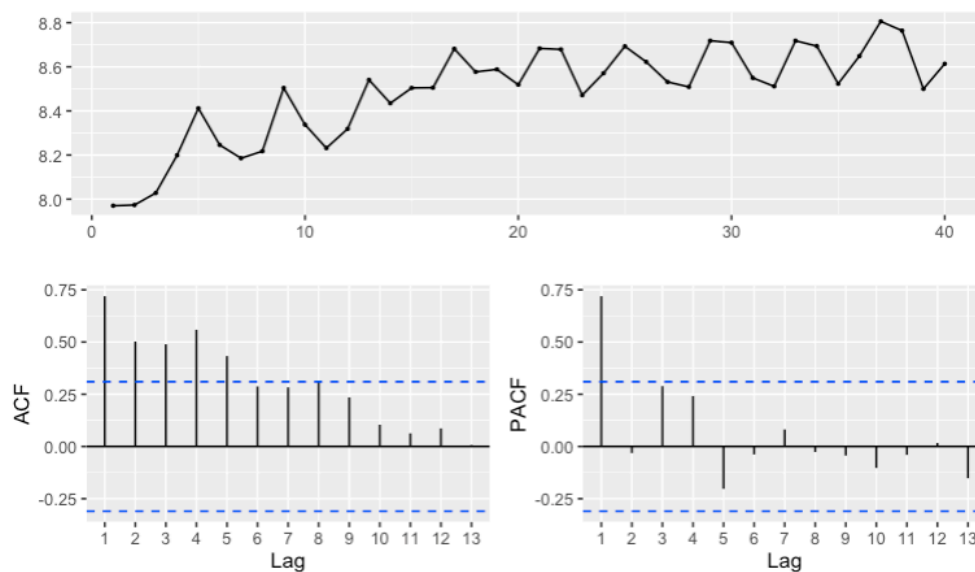
## MODELO ARIMA

Nuestra serie es no estacionaria pues no es constante ni en varianza ni en media. Por lo tanto, tenemos que realizar una transformación logarítmica.



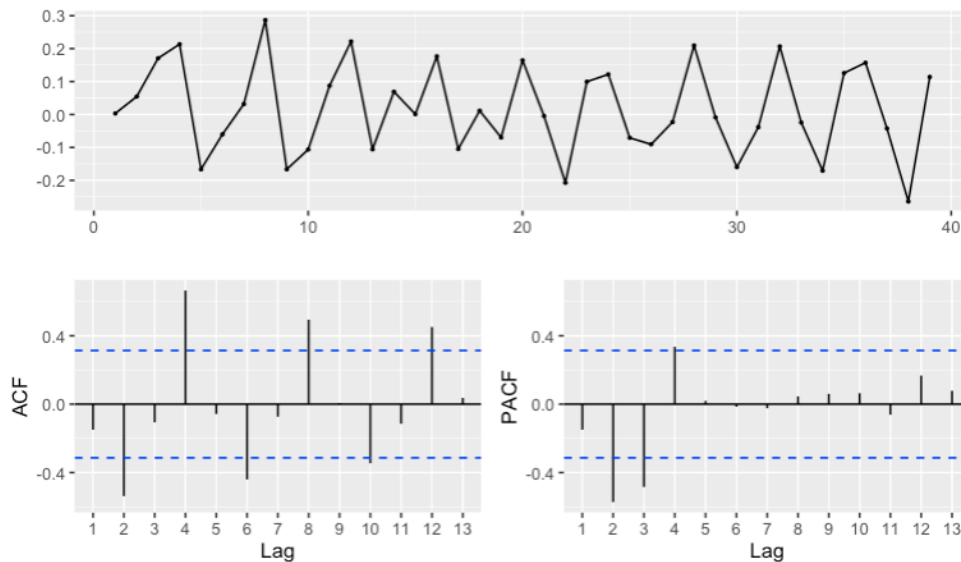
- **Serie estacionaria en varianza:**

Los ACF y los PACF son las funciones de autocorrelación que nos muestran como resultado que las correlaciones que se encuentran dentro del intervalo son estadísticamente 0, por lo que lo que ocurrió hasta el periodo 5 no nos interesa, pero si desde el 6 en adelante.



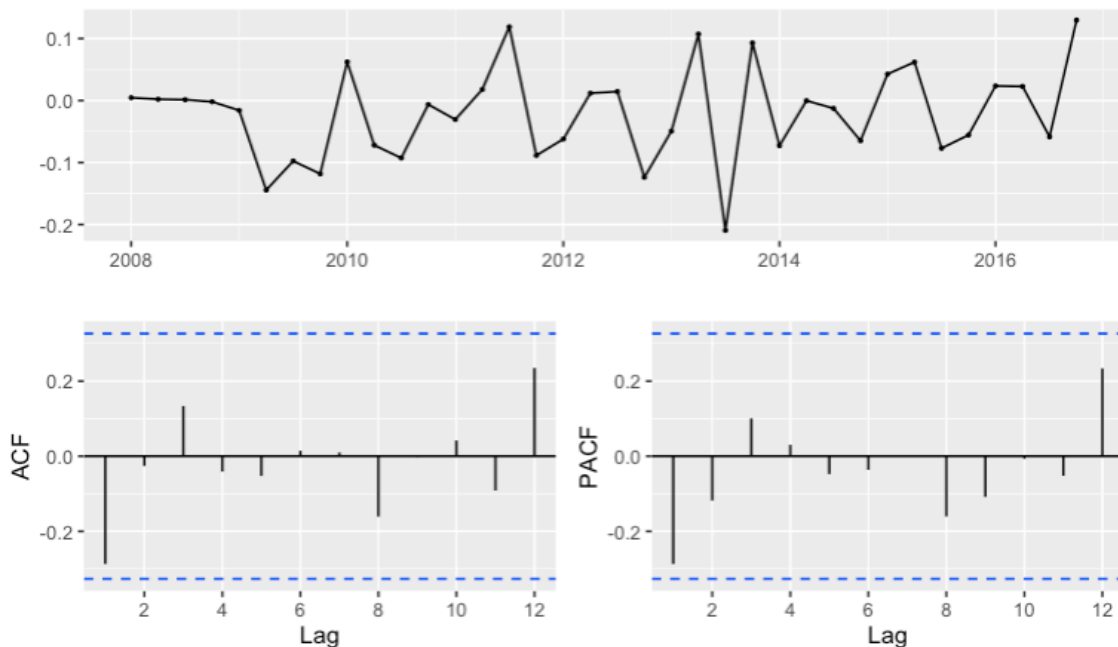
- **Serie estacionaria en media:**

Procedemos a hacer la serie estacionaria en media, realizando una diferencia del logaritmo para posteriormente poder realizar el modelo ARIMA.



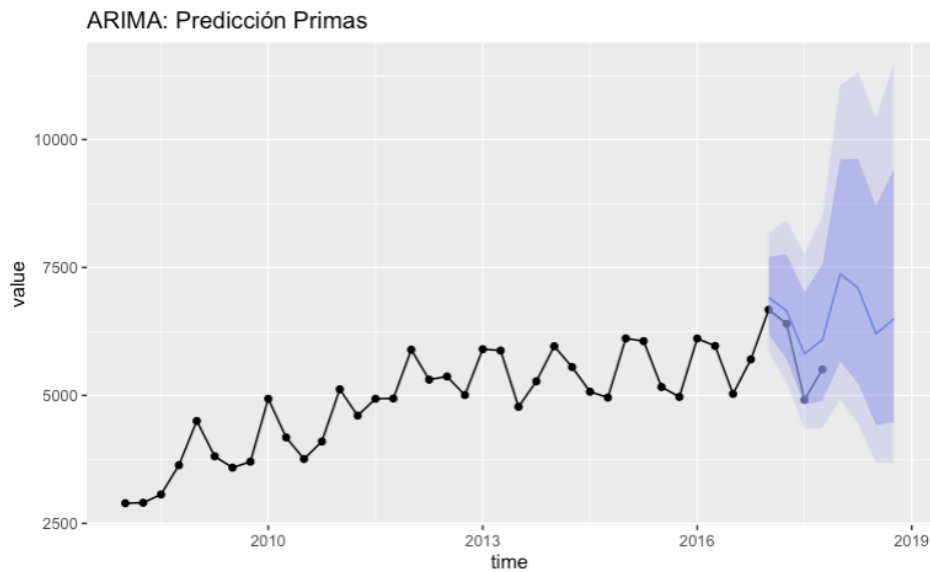
- **Modelo ARIMA**

Mediante ARIMA obtenemos el modelo sin los residuos en ACF y PACF.



Realizamos el *test box – Ljung*, por el cual se acepta la hipótesis nula porque nuestro nivel de significación es superior a 0.05. Es decir, a un nivel de confianza de 95% se acepta la hipótesis nula.

## PREDICCIÓN MODELO ARIMA



Obtenemos un error medio cuadrático de 388.9

## CONCLUSIÓN:

Comparando el modelo ARIMA con el modelo ETS realizado anteriormente podemos determinar que en este caso realiza una mejor predicción el modelo ETS que ARIMA ya que el primero tiene un error medio cuadrático menor que el segundo