Introducción

La base utilizada en este proyecto contiene información de cotización sobre el precio de las acciones de Google del 23 de marzo del 2017 al 24 de marzo de 2022, está información es pública debido a que la empresa cotiza en la bolsa de valores. La serie de tiempo a utilizar corresponde a los precios de cierre. La frecuencia con la que se registraron los datos fue diaria.

La base se obtuvo de la página kaggle en el siguiente enlace: https://www.kaggle.com/datasets/varpit94/google-stock-data?resource=download

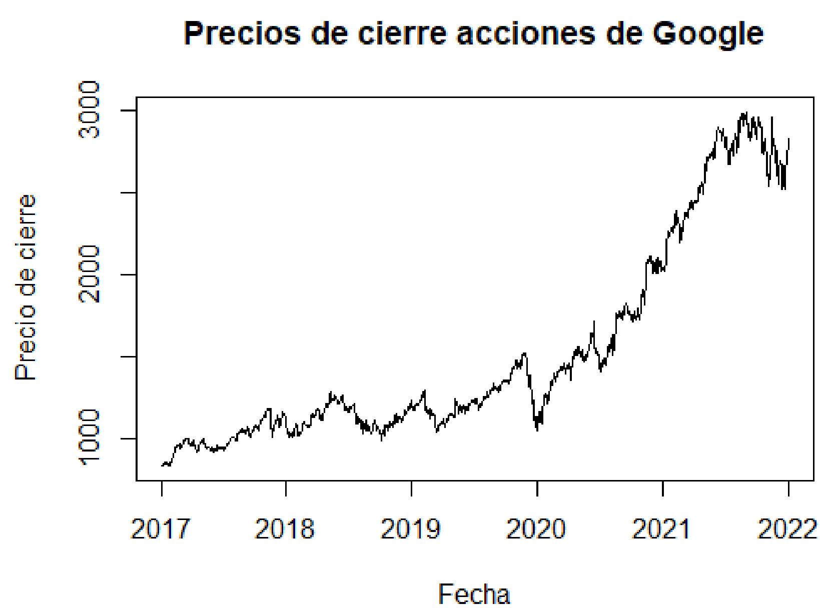
El objetivo de este proyecto es ajustar un modelo ARMA-GARCH a la serie de los log rendimientos de los precios de cierre de las acciones de Google.

Hechos estilizados

Antes de trabajar con la serie vamos a comprobar la presencia de algunas de las características de una serie financiera.

- **No es estacionaria**

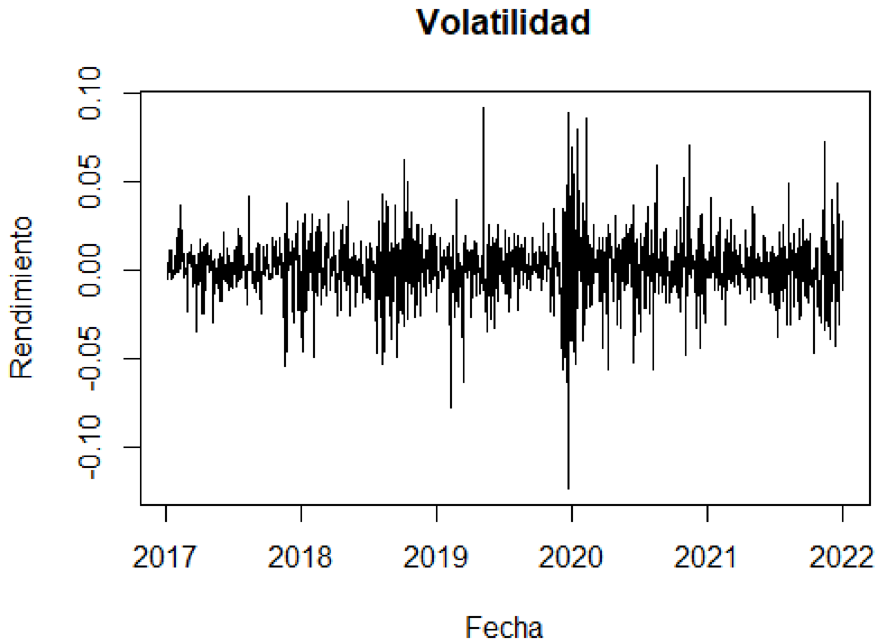
Gráfica 1.



De la gráfica 1 se observa que la serie no es estacionaria, de 2017 a 2020 oscila alrededor de 1000, luego de 2020 a mediados de 2021 tiene un comportamiento creciente y luego parece estabilizarse nuevamente alrededor de 2700.

- **Agrupamiento volatilidad**

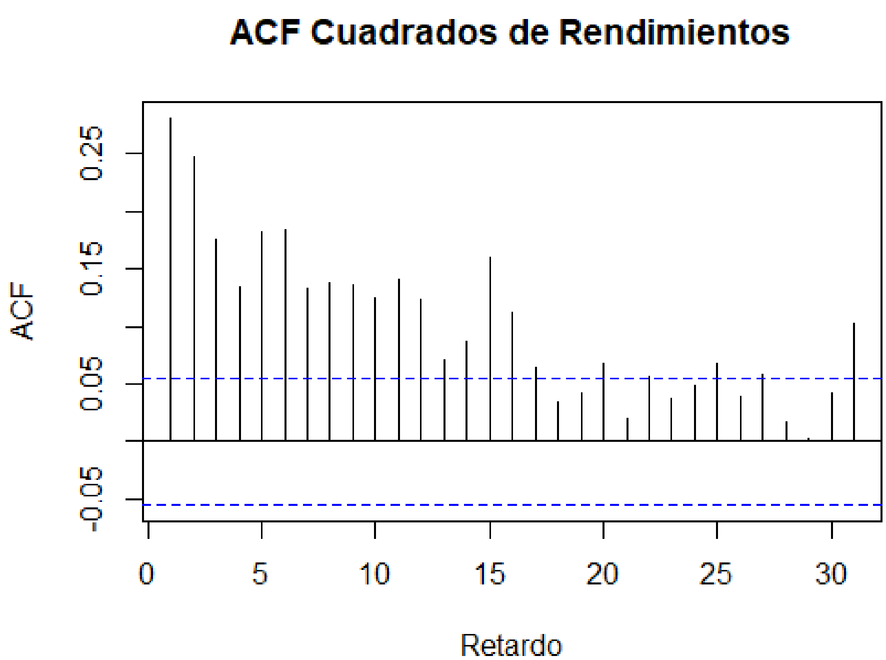
Gráfica 2.



De la gráfica 2 se observa que existen periodos donde se presenta alta volatilidad seguidos de periodos con una volatilidad más estable. Por ejemplo, a inicios del año 2020 la volatilidad se disparó y luego se redujo considerablemente.

- **Correlación en cuadrado de los rendimientos**

Gráfica 3.



De la gráfica 3 se observa que los cuadrados de los rendimientos parecen estar correlacionados.

Antes de ajustar un modelo a nuestra serie vamos a verificar la presencia del efecto ARCH, pues en caso de no estar presente sería contradictorio intentar ajustar un modelo GARCH.

Prueba del efecto ARCH

Se realizará la prueba ARCH para los log rendimientos tomando en cuenta hasta el lag 5.

$$H_0: \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = \alpha_5 = 0$$

$$H_1: algún\ \alpha_i \neq 0$$

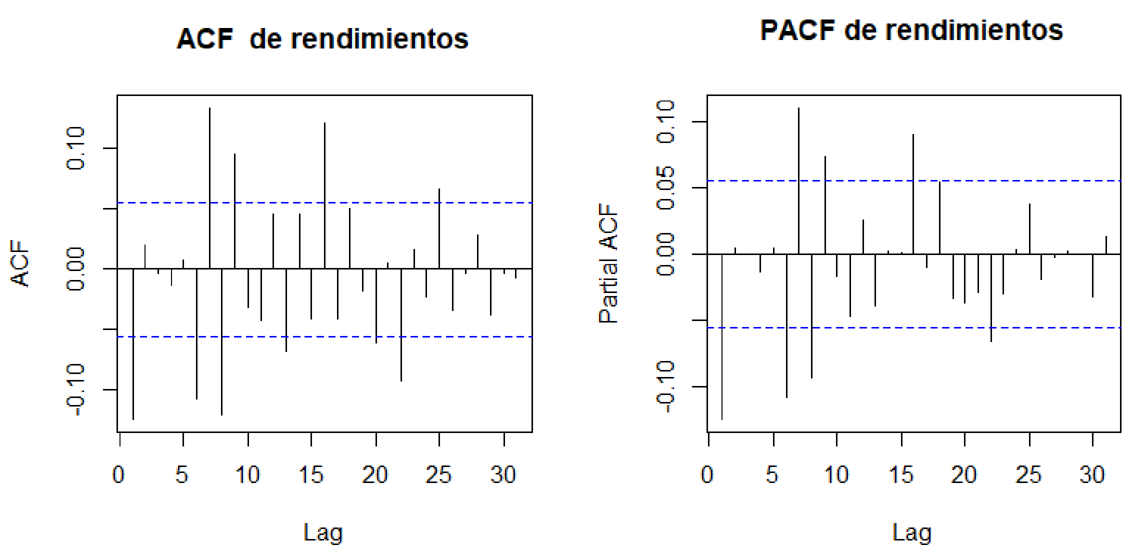
La prueba se realizó en el software R con la función "archtest" se obtuvo un p-valor de $2.2 * 10^{-16} < 0.05$ por lo cual con un nivel de significancia del 5% se rechaza la hipótesis nula, es decir, existe presencia del efecto ARCH. Por lo anterior entonces se procede a ajustar un modelo.

Ajuste del modelo

Componente ARMA

El primer paso consiste en ajusta un modelo ARMA a los log rendimientos, para ello se calcula el ACF y el PACF.

Gráfica 4.



De la gráfica 4 notamos que no parece quedar claro cual es el modelo más adecuado por lo cual se recurre a calcular el EACF. Al estimarlo en el software R se obtienen los resultados que se muestran a continuación:

AR/MA	0	1	2	3	4	5	6	7	8	9	10	11	12	13
0	x	o	o	o	o	x	x	x	x	o	o	o	x	o
1	o	o	o	o	o	o	o	o	o	o	o	o	o	o
2	o	o	o	o	o	o	o	o	o	o	o	o	o	o
3	o	x	o	o	o	o	o	o	o	o	o	o	o	o
4	x	x	o	o	o	o	o	o	o	o	o	o	o	o
5	o	x	o	x	x	o	o	o	o	o	o	o	o	o
6	x	x	o	x	o	x	o	o	o	o	o	o	o	o
7	x	x	o	x	o	x	x	o	o	o	o	o	o	o

Del resultado anterior notamos que uno de los modelos más simples que sugiere el EACF es el ARMA (0,1) por lo que se utilizará para los siguientes pasos. Hasta ahora se tiene lo siguiente:

$$r_t = ARMA(0,1) + \eta_t$$

Donde r_t son los log rendimientos.

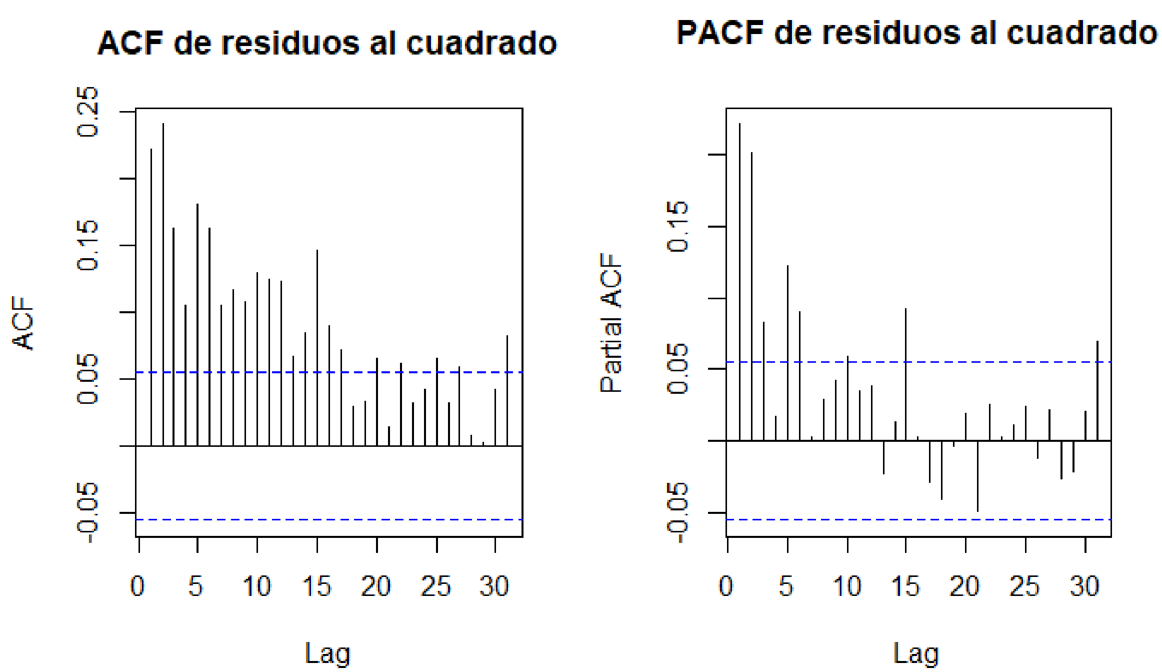
Al ajustar el modelo se calculan los residuales η_t y se les aplica nuevamente la prueba del efecto ARCH para verificar que los residuales puedan ser modelados con un GARCH. Se realizó la prueba en R y se obtuvo un p-valor menor a 0.05 por lo que los residuales si tienen presencia del efecto así que se continua con el procedimiento.

Componente GARCH

Para ajustar un modelo GARCH (p,q) a los residuales η_t es necesario trabajar con ${\eta_t}^2$ pues dado un resultado estos siguen un proceso ARMA (r,q) donde r=max{p,q}.

Nuevamente se calcula el ACF y el PACF en este caso para ${\eta_t}^2$ y se obtienen los siguientes resultados:

Gráfica 5.



De la gráfica 5 no parece quedar claro que modelo ARMA utilizar, aunque pudiese sugerir un comportamiento exponencial el ACF, y debido al PACF podríamos sugerir un AR (4) o un AR (7) se optara por calcular el EACF como se muestra a continuación:

```
AR/MA
  0 1 2 3 4 5 6 7 8 9 10 11 12 13
0 x x x x x x x x x x x  x  x  x
1 x x o o x o o o o o o  o  o  o
2 x x o x x x o o o o o  o  o  o
3 x o x o o o x o o o o  o  o  o
4 x x x x o o x o o o o  o  o  o
5 x x x x x o x o o o o  o  o  o
6 o x x x x x o o o o o  o  o  o
7 x x o x o x x o o o o  o  o  o
```

Entre los modelos sugeridos por el EACF destacamos el ARMA (r=2, q=2) y el ARMA (r=3, q=1). Aunque también sugiere el ARMA (1, 2) este modelo no es válido puesto que $r < q$.

Ajuste del modelo ARMA-GARCH

Utilizando el ARMA (r=3, q=1) nuestro único valor posible para p es de $p = 3$. Así que ajustaremos un modelo ARMA (0, 1) + GARCH (3, 1), esto lo haremos con la función "ugarchfit" en R.

Al hacer el ajuste obtenemos que los supuestos se cumplen, pero los coeficientes α_2 y α_3 no resultaron ser significativos como se muestra a continuación:

```
Optimal Parameters
------------------------------------
        Estimate  Std. Error   t value Pr(>|t|)
mu      0.001185    0.000412  2.876132 0.004026
ma1    -0.034301    0.032374 -1.059514 0.289366
omega   0.000017    0.000005  3.642858 0.000270
alpha1  0.080255    0.030518  2.629799 0.008544
alpha2  0.000000    0.048095  0.000007 0.999995
alpha3  0.000000    0.039931  0.000003 0.999998
beta1   0.863206    0.027367 31.541797 0.000000
```

Debido a lo anterior se optó por ajustar un modelo ARMA (0, 1) + GARCH (1, 1). Se obtuvieron los siguientes resultados:

Validación de supuestos

```
Weighted Ljung-Box Test on Standardized Residuals
------------------------------------
                        statistic p-value
Lag[1]                   0.003328  0.9540
Lag[2*(p+q)+(p+q)-1][2]  0.243758  0.9974
Lag[4*(p+q)+(p+q)-1][5]  1.175603  0.9172
d.o.f=1
H0 : No serial correlation
```

```
Weighted Ljung-Box Test on Standardized Squared Residuals
------------------------------------
                         statistic p-value
Lag[1]                    0.008537  0.9264
Lag[2*(p+q)+(p+q)-1][5]   0.288208  0.9849
Lag[4*(p+q)+(p+q)-1][9]   1.228692  0.9747
d.o.f=2
```

Las dos pruebas anteriores son de autocorrelación en los residuales, dado que los p-valores son mayores a 0.05 se rechazan con un nivel de significancia del 5%, es decir, no hay evidencia suficiente para rechazar la cero autocorrelación.

```
Weighted ARCH LM Tests
------------------------------------
            Statistic Shape Scale P-Value
ARCH Lag[3]   0.07637 0.500 2.000  0.7823
ARCH Lag[5]   0.52426 1.440 1.667  0.8764
ARCH Lag[7]   1.11634 2.315 1.543  0.8939
```

Esta prueba verifica la presencia del efecto ARCH en los residuales. En este caso los p-valores son mayores a 0.05, es decir, con un nivel de confianza del 5% se rechaza la hipótesis nula, no hay evidencia suficiente para rechazar la hipótesis nula, no hay presencia del efecto ARCH en los residuales.

Por los resultados anteriores se han comprobado los supuestos y se le ha dado validez al modelo.

Coeficientes del modelo

```
Optimal Parameters
------------------------------------
        Estimate  Std. Error  t value Pr(>|t|)
mu      0.001186    0.000411   2.8835 0.003933
ma1    -0.034122    0.032327  -1.0555 0.291188
omega   0.000017    0.000004   3.7412 0.000183
alpha1  0.080711    0.012797   6.3071 0.000000
beta1   0.862773    0.026061  33.1056 0.000000
```

En este caso nuestros coeficientes de la componente GARCH sí resultaron ser significativos con un nivel de significancia del 5%. Cabe mencionar que los coeficientes de la componente ARMA no resultaron significativos así que podríamos optar por no considerarlo, para fines de este proyecto se mantendrá dicha componente.

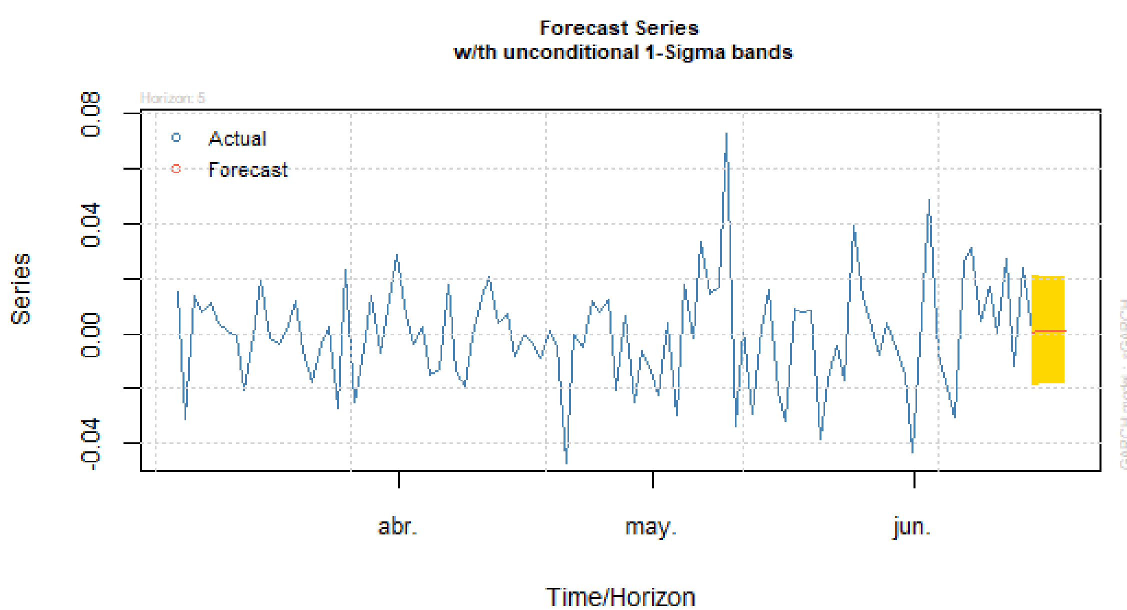
Los modelos que derivaban de utilizar un ARMA (2, 2) para el cuadrado de los residuales no cumplían los supuestos así que se omitió su análisis.

Pronóstico

Una vez que se ha ajustado el modelo lo que sigue es utilizarlo para realizar predicciones sobre los rendimientos y la volatilidad. Esto se realiza en el software R a través de la función “ugarchforecast”. Los valores estimados 5 tiempos hacia adelante para los rendimientos y la volatilidad son los siguientes:

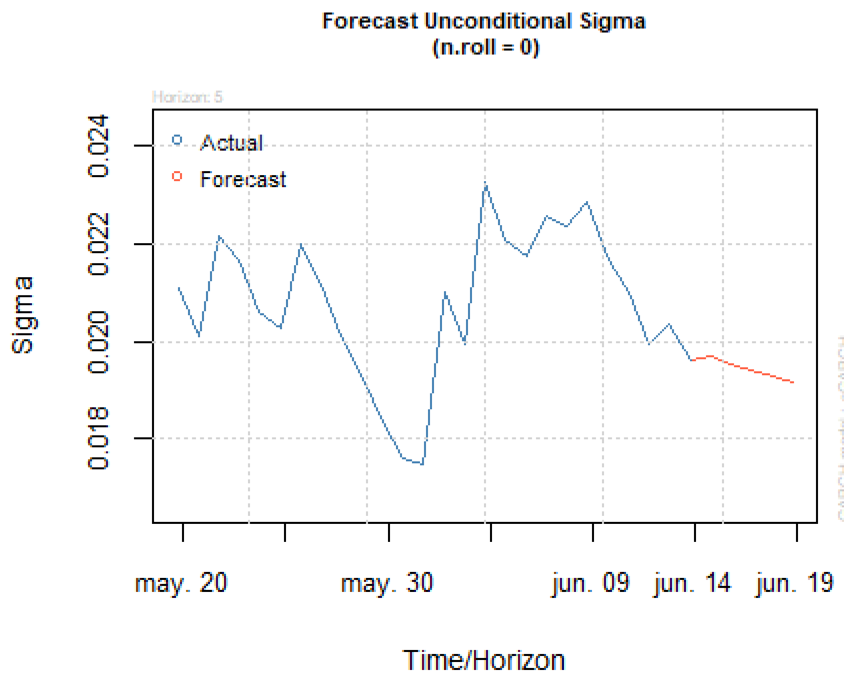
```
        Series    Sigma
T+1 0.0004362 0.01969
T+2 0.0012100 0.01955
T+3 0.0012100 0.01942
T+4 0.0012100 0.01929
T+5 0.0012100 0.01918
```

Gráfica 6.



En la gráfica 6, se observa en amarillo el intervalo de predicción para los rendimientos 5 tiempos hacia adelante.

Gráfica 7.



En la gráfica 7 se observa en rojo los valores que estiman para la volatilidad en los 5 tiempos futuros.

En este proyecto se pudo apreciar la utilidad de ajustar un modelo GARCH para una serie financiera además de su procedimiento paso a paso. Es importante destacar sus características y la importancia de estos modelos, en el contexto financiero pueden llegar a tener un gran impacto pues permiten a los inversores y a las empresas anticiparse a los cambios en el mercado, las acciones y demás instrumentos.