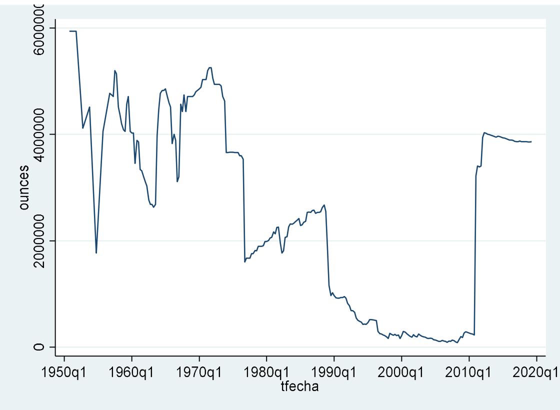
Tarea

1. Realizar un análisis de una serie trimestral empleando SARIMA

Se usará la serie de Kaggle “Country Gold Reserves Annual, quarterly, & monthly” de la pagina https://www.kaggle.com/datasets/eliasdabbas/gold-reserves-by-country-quarterly para Mexico.

   a. Análisis previo (hay que diferenciar)

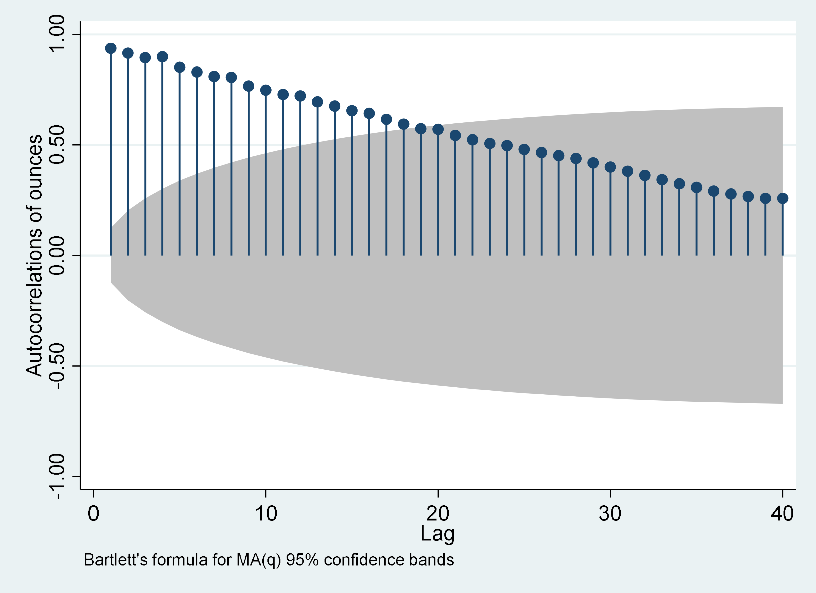
Graficamos la serie



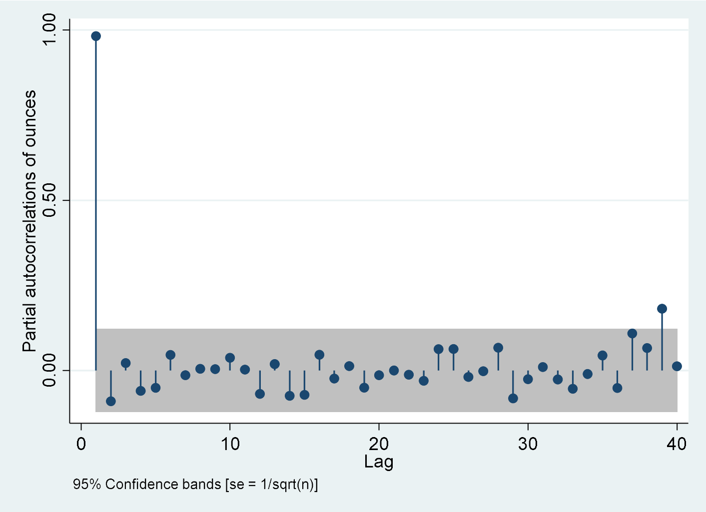

```
summarize ounces
```

Variable	Obs	Mean	Std. Dev.	Min	Max
ounces	257	2473928	1758374	81694.98	5943000

Graficando la autocorrelación simple



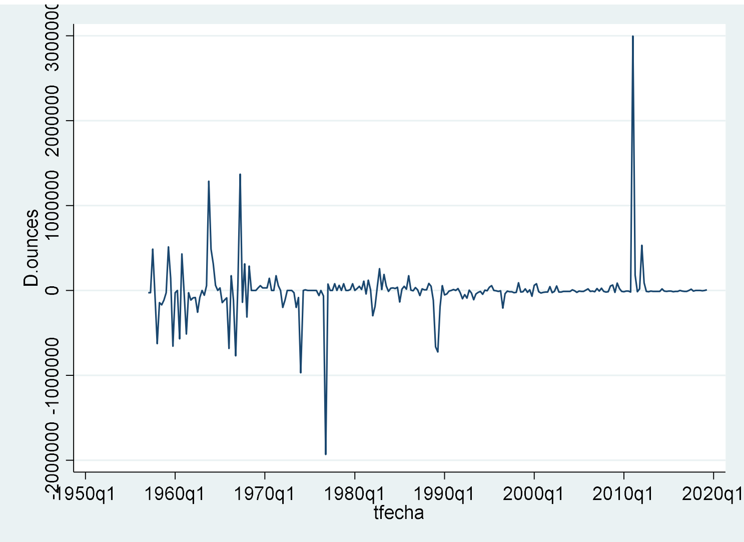
Graficando la autocorrelación parcial



Verificamos si la serie es estacionaria para diferenciar, como el valor critico es mayor al calculado, el test indica que hay raíz unitaria por lo que no es estacionaria.

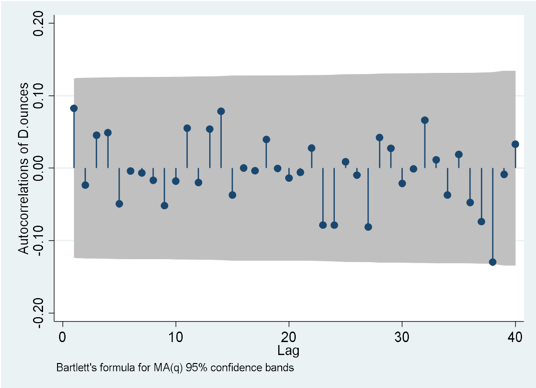
```
Dickey-Fuller test for unit root                    Number of obs   =       250

                               ──────── Interpolated Dickey-Fuller ────────
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
──────────────────────────────────────────────────────────────────────────────
 Z(t)             -1.596            -3.460            -2.880            -2.570
──────────────────────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.4856
```

Al diferenciar tenemos:

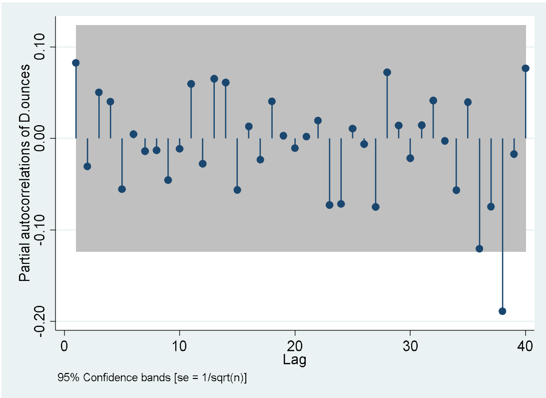


b. Test de raíz unitaria, test de estacionalidad, correlogramas (AC y PAC)

Graficando la autocorrelación simple



Graficando la autocorrelación parcial



Volvemos a revisar si es estacionaria y comprobamos el valor critico es menor al calculado, la prueba indica que no hay raíz unitaria por lo que es estacionaria.

```
Dickey-Fuller test for unit root                   Number of obs  =      249

                               ——————— Interpolated Dickey-Fuller ———————
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
─────────────────────────────────────────────────────────────────────────────
 Z(t)            -14.466            -3.460            -2.880            -2.570
─────────────────────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.0000
```

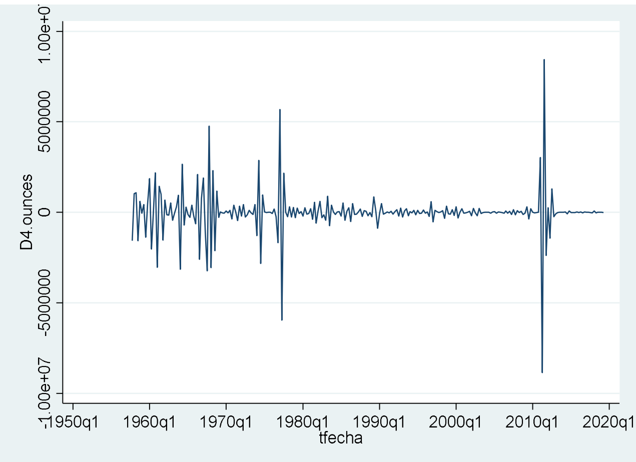
```
Phillips-Perron test for unit root                 Number of obs   =      249
                                                   Newey-West lags =        4

                               ——————— Interpolated Dickey-Fuller ———————
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
─────────────────────────────────────────────────────────────────────────────
 Z(rho)         -230.649           -20.297           -13.998           -11.199
 Z(t)            -14.478            -3.460            -2.880            -2.570
─────────────────────────────────────────────────────────────────────────────
MacKinnon approximate p-value for Z(t) = 0.0000
```

Con kpss vemos que en todos el calculado es menor al crítico por lo que aceptamos Ho es No estacionario

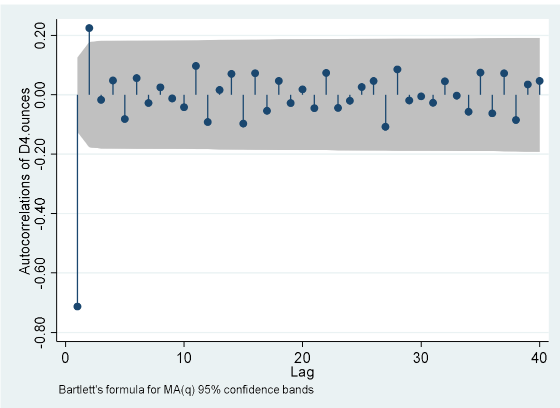
```
KPSS test for D.ounces

Maxlag = 5 chosen by Schwert criterion
Autocovariances weighted by Bartlett kernel

Critical values for H0: D.ounces is trend stationary

10%: 0.119   5% : 0.146   2.5%: 0.176   1% : 0.216

Lag order    Test statistic
    0            .0768
    1            .0714
    2            .0711
    3            .0698
    4             .068
    5             .068
```

c. Estimación de modelo SARIMA

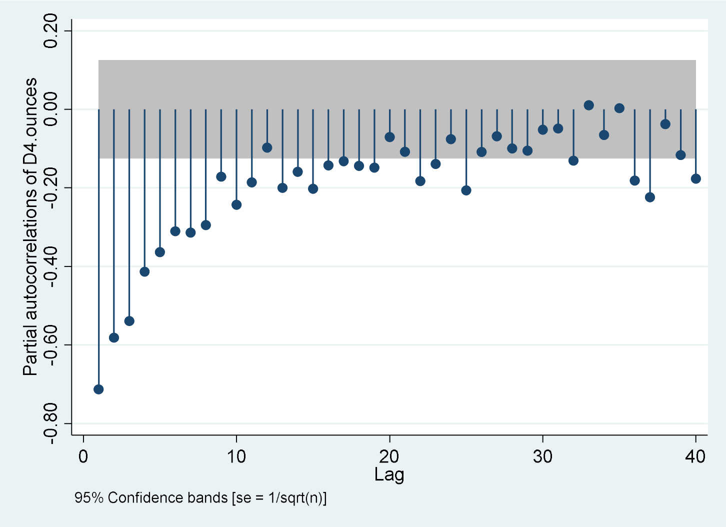
Al ser una serie trimestral graficamos con 4 rezagos



Graficando la autocorrelación simple



Graficando la autocorrelación parcial



Estimando el modelo tenemos las siguientes opciones

Con un sarima (1,1,1,4)

```
ARIMA regression

Sample:  1951q4 - 2019q2, but with gaps         Number of obs     =        253
                                                Wald chi2(2)      =     187.00
Log likelihood = -3759.474                      Prob > chi2       =     0.0000

------------------------------------------------------------------------------
             |                 OPG
   S4.ounces |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
ounces       |
       _cons |   -29006.1   57807.74    -0.50   0.616    -142307.2    84294.99
-------------+----------------------------------------------------------------
ARMA4        |
          ar |
         L1. |   .9418268   .0794862    11.85   0.000     .7860368    1.097617
             |
          ma |
         L1. |  -.9999889   155.7668    -0.01   0.995    -306.2973    304.2974
-------------+----------------------------------------------------------------
      /sigma |     681693   5.31e+07     0.01   0.495            0    1.05e+08
------------------------------------------------------------------------------
Note: The test of the variance against zero is one sided, and the two-sided
      confidence interval is truncated at zero.


Akaike's information criterion and Bayesian information criterion

-----------------------------------------------------------------------------
       Model |          N   ll(null)  ll(model)      df         AIC        BIC
-------------+---------------------------------------------------------------
           . |        253          .  -3759.474       4    7526.947   7541.081
-----------------------------------------------------------------------------
Note: BIC uses N = number of observations. See [R] BIC note.
```

Con un sarima(1,1,2,4)

```
ARIMA regression

Sample:  1951q4 - 2019q2, but with gaps         Number of obs     =        253
                                                Wald chi2(3)      =       0.10
Log likelihood = -3761.175                      Prob > chi2       =     0.9923
```

S4.ounces	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf. Interval]	
ounces						
_cons	-20065.92	48905.37	-0.41	0.682	-115918.7	75786.85
ARMA4						
ar						
L1.	.0254688	3.437931	0.01	0.994	-6.712753	6.763691
ma						
L1.	-.0339756	3.415305	-0.01	0.992	-6.727851	6.6599
L2.	-.0205673	.0781546	-0.26	0.792	-.1737474	.1326129
/sigma	692059.4	15978.61	43.31	0.000	660741.9	723376.9

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	253	.	-3761.175	5	7532.351	7550.017

Note: BIC uses N = number of observations. See [R] BIC note.

Con un sarima (2,1,1,4)

```
ARIMA regression

Sample:  1951q4 - 2019q2, but with gaps         Number of obs     =        253
                                                Wald chi2(4)      =   25986.70
Log likelihood = -3752.333                      Prob > chi2       =     0.0000
```

S4.ounces	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf. Interval]	
ounces						
_cons	-19916.52	45163.51	-0.44	0.659	-108435.4	68602.33
ARMA4						
ar						
L1.	-.031502	.0416416	-0.76	0.449	-.113118	.0501141
L2.	-.9873568	.0290745	-33.96	0.000	-1.044342	-.9303719
ma						
L1.	-.0232064	1.49115	-0.02	0.988	-2.945808	2.899395
L2.	.999974	127.7715	0.01	0.994	-249.4275	251.4275
/sigma	651640.7	4.16e+07	0.02	0.494	0	8.22e+07

Note: The test of the variance against zero is one sided, and the two-sided confidence interval is truncated at zero.

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
.	253	.	-3752.333	6	7516.666	7537.867

Note: BIC uses N = number of observations. See [R] BIC note.

De estos tres elegimos el sarima (1,1,2,4)

d. Verificación de supuestos (estacionariedad principalmente)

Probamos estacionalidad

```
Number of observations  :  246
Deterministic variables : Seasonal dummies
Optimal lag selection method: Modified AIC
Lags tested: 5
Augmented by lags : 0

                Stat    1% critical  5% critical  10% critical
---------------------------------------------------------------------------------
t[0]           -7.110     -3.424       -2.845       -2.548
t[Pi]          -8.007     -3.423       -2.845       -2.548

F[Pi/2]        54.224      8.861        6.628        5.592
F[All seas]    58.393      7.666        5.944        5.132
F[All]         56.065      7.001        5.556        4.869

Testing regression:
```

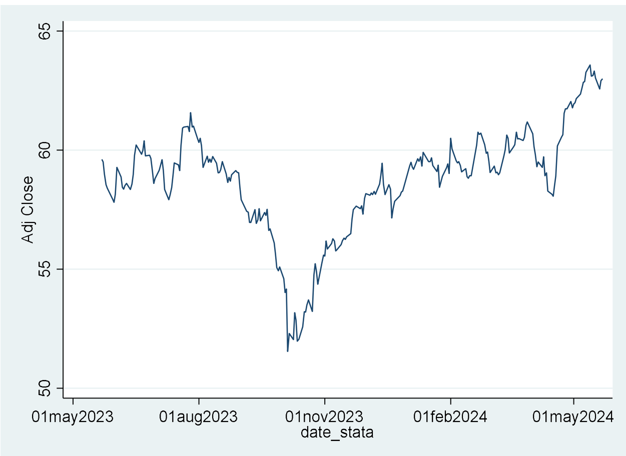
Revisamos t[Pi] y F[all], vemos que son mayores en valor absoluto por lo que sugiere no estacionaridad en los datos.

2. Realizar un análisis de un precio de una acción de forma diaria empleando ARCH, GARCH, etc.

Analizamos a coca cola (KO)

a. Análisis previo (hay que diferenciar)

Graficamos la serie



Vemos un resumen de los datos

```
. summarize adjclose
```

Variable	Obs	Mean	Std. Dev.	Min	Max
adjclose	253	58.6689	2.28728	51.55385	63.58

Hacemos la prueba de raíz unitaria y vemos que el valor critico es mayor al calculado, el test indica que hay raíz unitaria por lo que no es estacionaria.

Dickey-Fuller test for unit root   Number of obs = 198

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(t)	-0.772	-3.477	-2.883	-2.573

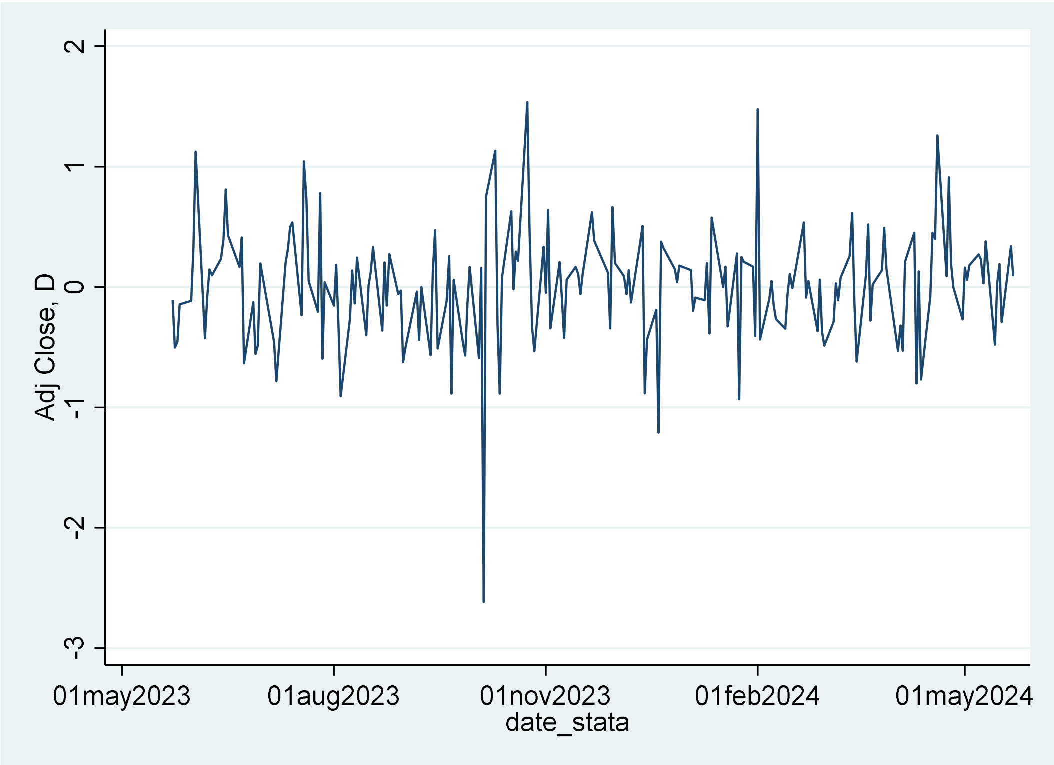
MacKinnon approximate p-value for Z(t) = 0.8273

Phillips-Perron test for unit root   Number of obs = 198
Newey-West lags = 4

	Test Statistic	Interpolated Dickey-Fuller 1% Critical Value	5% Critical Value	10% Critical Value
Z(rho)	-2.274	-20.127	-13.896	-11.131
Z(t)	-0.768	-3.477	-2.883	-2.573

MacKinnon approximate p-value for Z(t) = 0.8283

Diferenciamos para hacerla estacionaria



Corroboramos nuevamente que sea estacionaria y vemos que si es estacionaria. Al ver que el valor critico es menor al valor calculado en los test.

```
Dickey-Fuller test for unit root                   Number of obs   =       145

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(t)            -12.023            -3.495            -2.887            -2.577
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000
```

```
Phillips-Perron test for unit root                 Number of obs   =       145
                                                   Newey-West lags =         4

                               ---------- Interpolated Dickey-Fuller ---------
                  Test         1% Critical       5% Critical      10% Critical
               Statistic           Value             Value             Value
------------------------------------------------------------------------------
 Z(rho)         -145.225           -19.950           -13.790           -11.060
 Z(t)            -12.021            -3.495            -2.887            -2.577
------------------------------------------------------------------------------
MacKinnon approximate p-value for Z(t) = 0.0000
```

b. Test ARCH con n rezagos

```
. reg adjclose

      Source |       SS           df       MS      Number of obs   =       253
-------------+----------------------------------   F(0, 252)       =      0.00
       Model |           0         0           .   Prob > F        =         .
    Residual |  1318.37604       252  5.23165096   R-squared       =    0.0000
-------------+----------------------------------   Adj R-squared   =    0.0000
       Total |  1318.37604       252  5.23165096   Root MSE        =    2.2873

------------------------------------------------------------------------------
    adjclose |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       _cons |    58.6689   .1438001   407.99   0.000     58.38569     58.9521
------------------------------------------------------------------------------
Number of gaps in sample:  54
LM test for autoregressive conditional heteroskedasticity (ARCH)
---------------------------------------------------------------------------
    lags(p)  |          chi2               df                 Prob > chi2
-------------+-------------------------------------------------------------
       1     |        173.980               1                   0.0000
---------------------------------------------------------------------------
             H0: no ARCH effects       vs.  H1: ARCH(p) disturbance
```

c. Estimación de modelo

Estimado un arch

```
ARCH family regression

Sample: 22may2023 - 22may2024, but with gaps      Number of obs    =       253
Distribution: Gaussian                            Wald chi2(.)     =         .
Log likelihood = -441.8821                        Prob > chi2      =         .

------------------------------------------------------------------------------
             |                 OPG
    adjclose |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
adjclose     |
       _cons |   59.22553   .0473809  1249.99   0.000     59.13266    59.31839
-------------+----------------------------------------------------------------
ARCH         |
        arch |
         L1. |   .9307497   .1212165     7.68   0.000     .6931698     1.16833
             |
       _cons |   .1788525   .0401575     4.45   0.000     .1001453    .2575598
------------------------------------------------------------------------------
```

Estimando un garch

```
ARCH family regression

Sample: 22may2023 - 22may2024, but with gaps      Number of obs    =       253
Distribution: Gaussian                            Wald chi2(.)     =         .
Log likelihood = -439.6973                        Prob > chi2      =         .

------------------------------------------------------------------------------
             |                 OPG
    adjclose |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
adjclose     |
       _cons |   59.24767   .0480749  1232.40   0.000     59.15344    59.34189
-------------+----------------------------------------------------------------
ARCH         |
        arch |
         L1. |    .954494   .1185837     8.05   0.000     .7220742    1.186914
             |
       garch |
         L1. |  -.0285958   .0151416    -1.89   0.059    -.0582727    .0010812
             |
       _cons |   .2219635   .0634891     3.50   0.000     .0975271    .3463999
------------------------------------------------------------------------------
```

El modelo GARCH es el que mejor modela la serie.

d. Verificación de supuestos

Verificamos autocorrelación

Dado que el valor p es menor que el nivel de significancia, se rechaza la hipótesis nula de que los residuos son ruido blanco.

Portmanteau test for white noise

Portmanteau (Q) statistic	=	2260.0228
Prob > chi2(40)	=	0.0000

Verificación de la No Heterocedasticidad en los Residuos

Como el p-valor es menor al 5% no hay heterocedastsidad, los residuos son homocedasticos.

Time-series regression

Sample: 22may2023 - 22may2024, but with gaps	Number of obs	=	253
Distribution: Gaussian	Wald chi2(.)	=	.
Log likelihood = -567.8134	Prob > chi2	=	.

resid	Coef.	OPG Std. Err.	z	P>|z|	[95% Conf.	Interval]
_cons	-.5787732	.1594846	-3.63	0.000	-.8913573	-.2661891
/SIGMA2	5.210973	.4229635	12.32	0.000	4.381979	6.039966

Verificación de Normalidad de los Residuos

Como el p-valor es menor al 5% aceptamos que no los residuos siguen una distribución normal

. sktest resid

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	joint adj chi2(2)	joint Prob>chi2
resid	253	0.0000	0.0122	21.67	0.0000