Proyección de afiliados por provincia

Dirección Actuarial

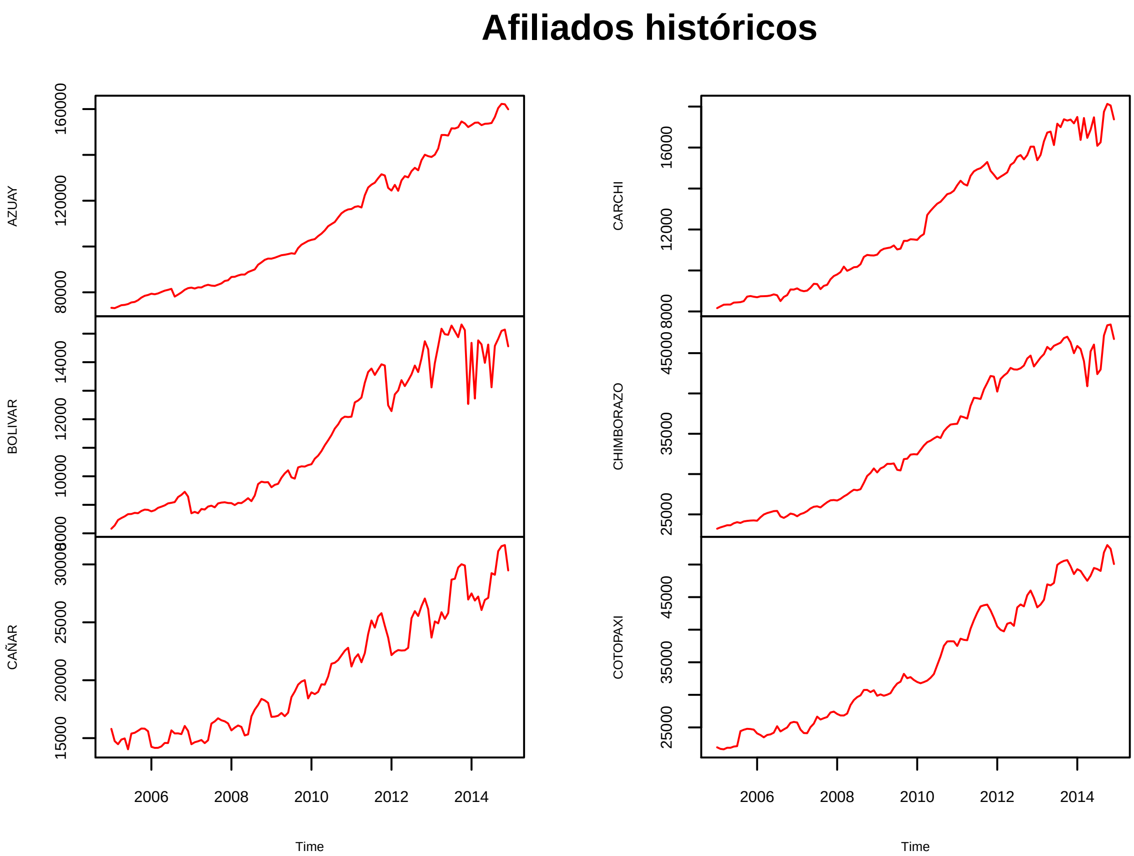
Febrero 2016

1. Información

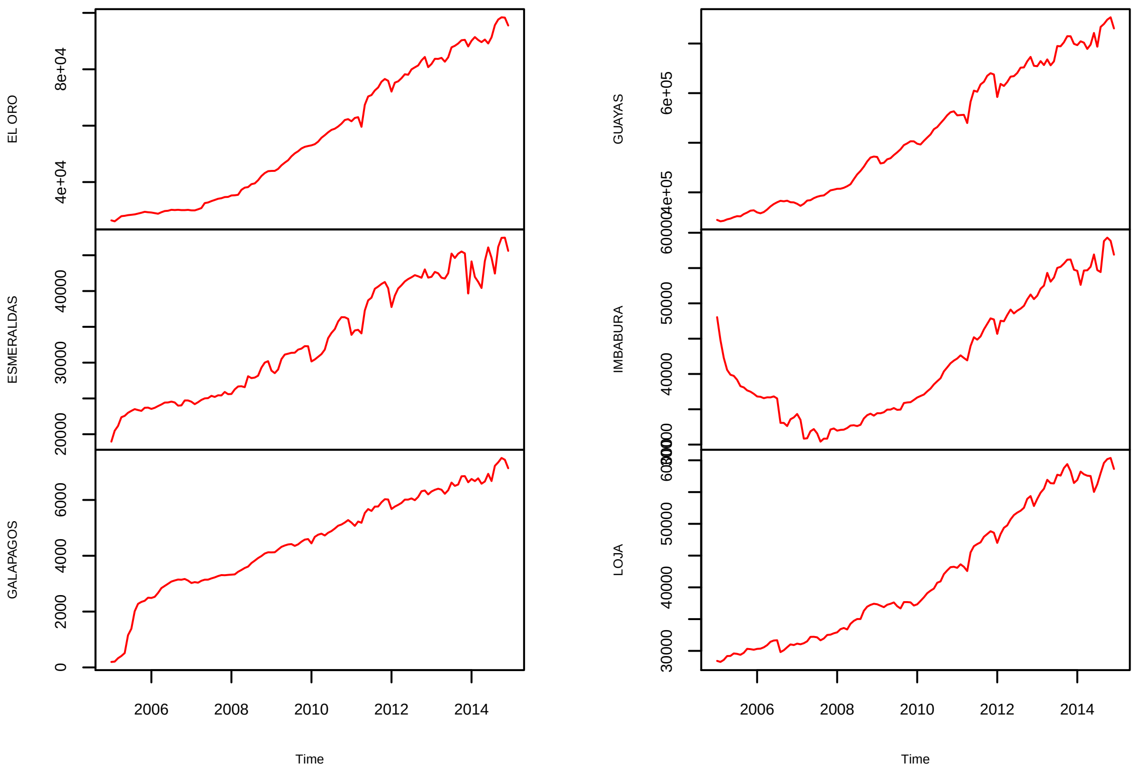
La información proporcionada para el análisis corresponde al total de afiliados por provincia en el periodo Enero - 2005 y Diciembre 2015.

2. Metodología

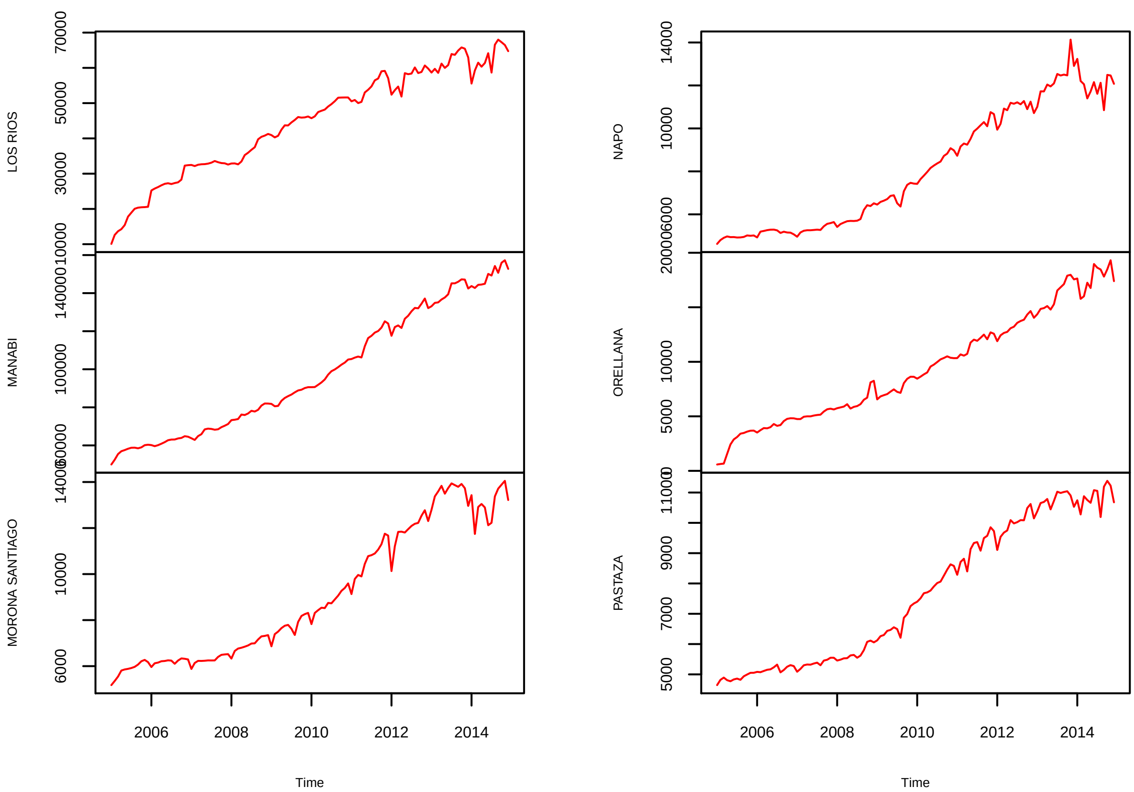
El siguiente paso consiste en ajustar un modelo de series de temporales a la sucesión estimada de afiliados por provincia utilizando la metodología Box-Jenkins que permita obtener predicciones.

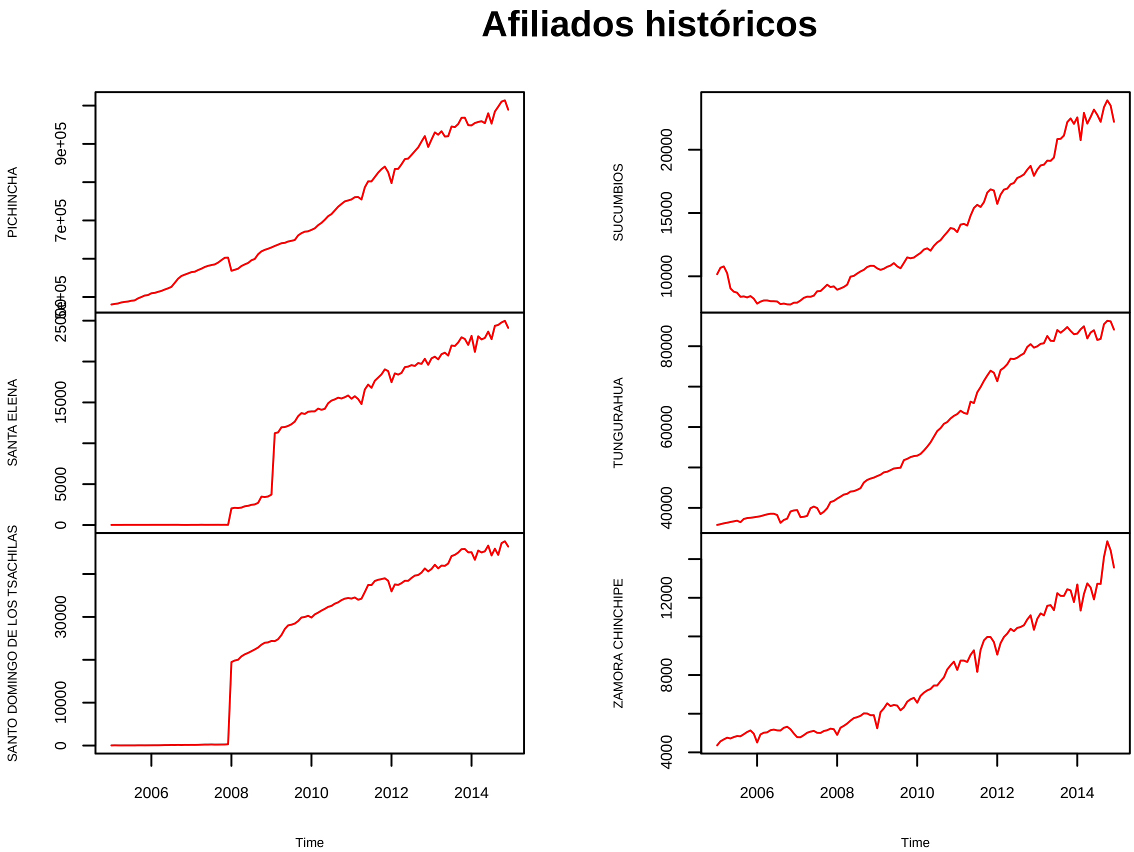


Afiliados históricos



Afiliados históricos





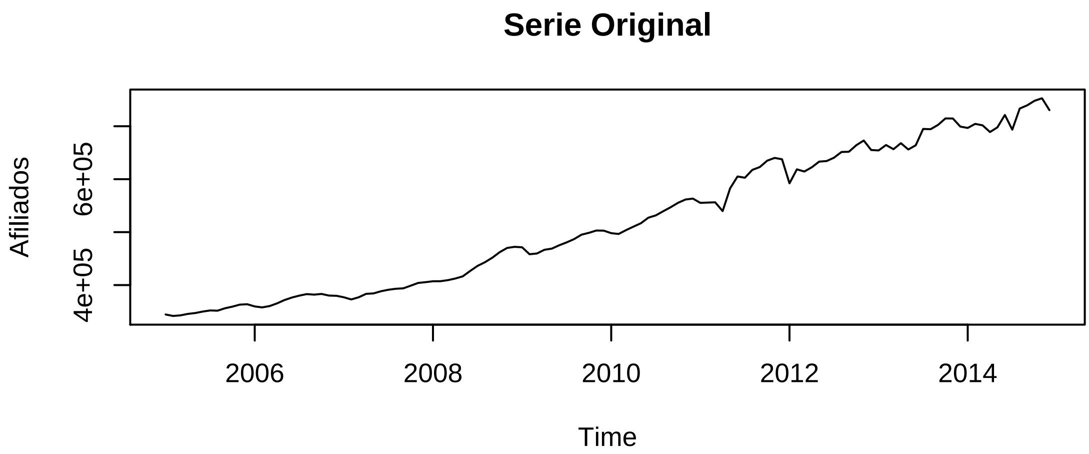
Al observar las gráficas de las series se evidencia que las mismas no son estacionarias pues presentan una tendencia estocástica.

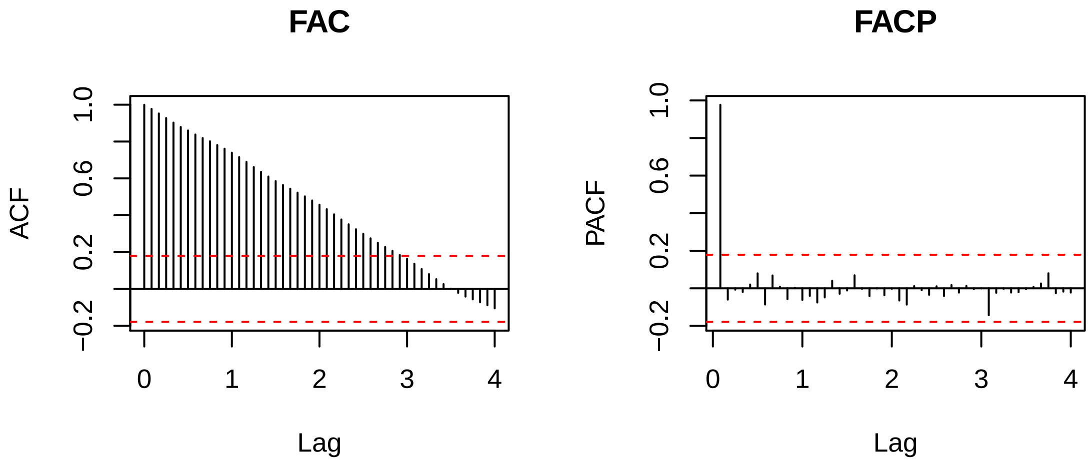
La no estacionariedad se debe a la existencia de raíces unitarias en las series de afiliados, generalmente para obtener una serie estacionaria en media se procede a diferenciar la serie y seguido aplicar una prueba estadística que nos garantice la estacionariedad en varianza.

Para el presente análisis implementaremos la prueba de estacionariedad de Kwiatkowski-Phillips-Schmidt-Shin (KPSS), la cual a diferencia de la prueba de raíz unitaria de Dickey Fuller, nos proporciona la prueba directa de la hipótesis nula de estacionariedad frente a la hipótesis alternativa de existencia de una raíz unitaria. La hipótesis nula y alternativa de la prueba KPSS es la siguiente:

$$\begin{cases} H_0\text{: La serie es estacionaria} \\ H_1\text{: Existen raíces unitarias} \end{cases}$$

2.1. Ajuste de afiliados Guayas

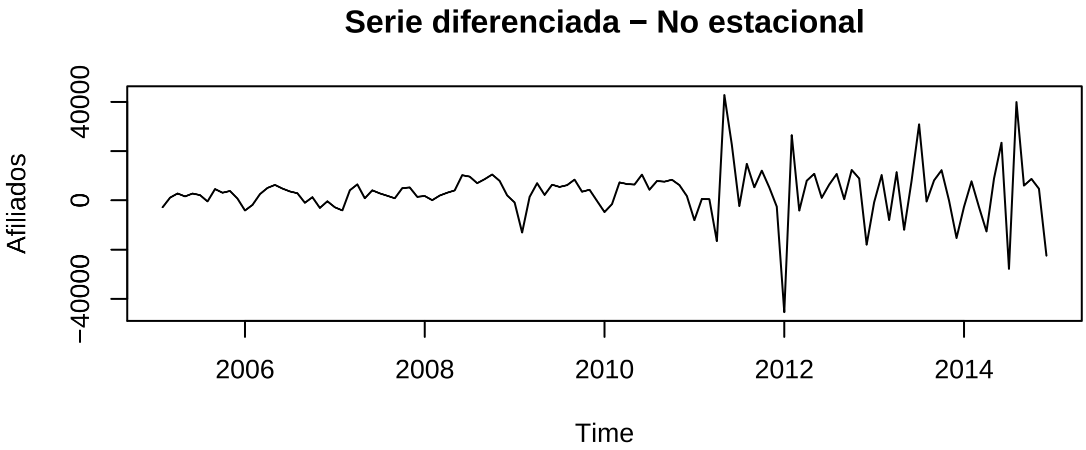


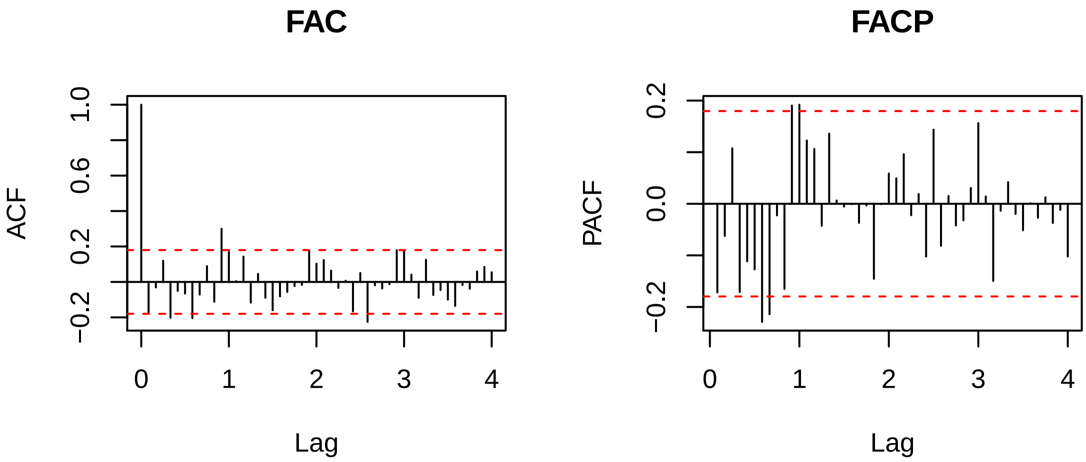


Los correlogramas muestran un decaimiento lento de la función ACF y un pico en el retardo 1 de la función PACF, esto muestra un comportamiento no estacionario y la necesidad de diferenciar la serie.

```
##
##  KPSS Test for Level Stationarity
##
## data:  data[, 10]
## KPSS Level = 4.0757, Truncation lag parameter = 2, p-value = 0.01
```

Dado que el p-value de la prueba KPSS es menor al nivel de significancia 0.05, se rechaza la hipótesis nula. Esta prueba corrobora la existencia de raíces unitarias lo cual ocasionan la no estacionariedad.

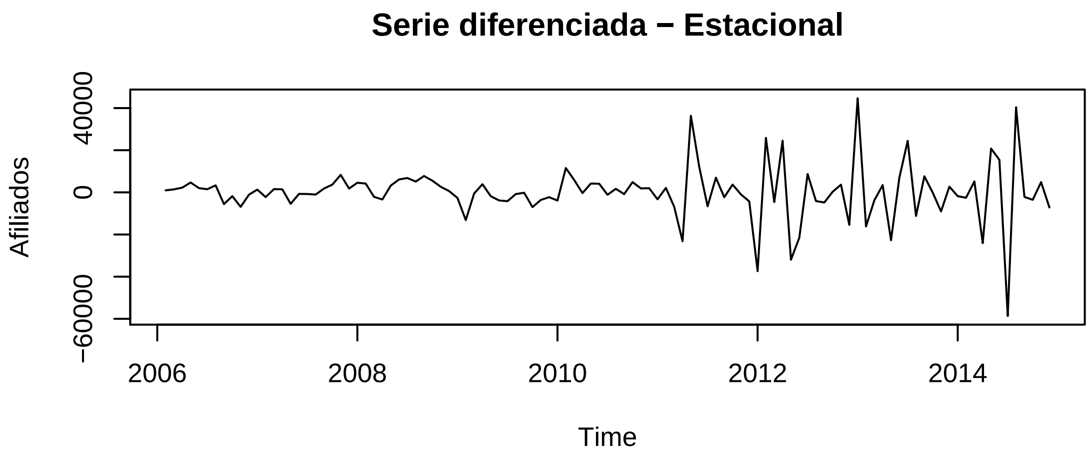


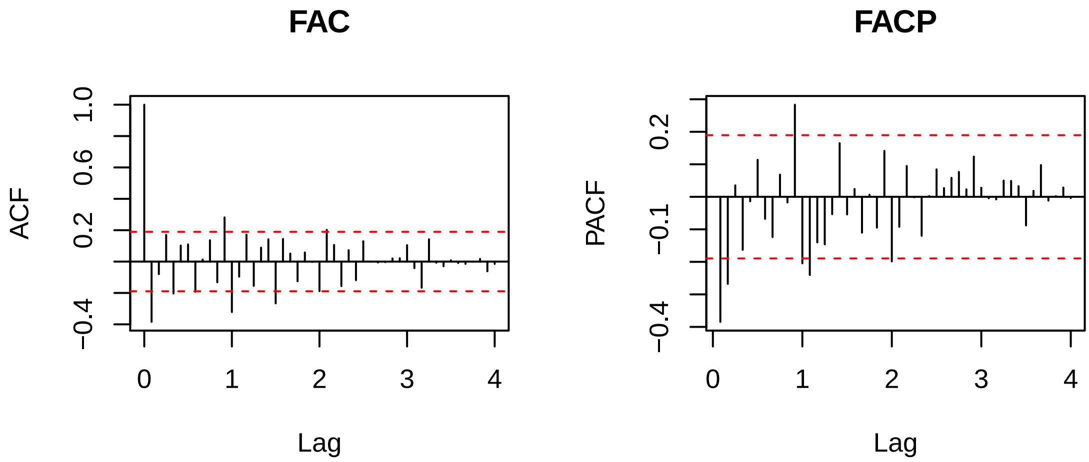



```
##
##  KPSS Test for Level Stationarity
##
## data:  diff(data[, 10])
## KPSS Level = 0.094912, Truncation lag parameter = 2, p-value = 0.1
```

Observamos que el p-value de la prueba KPSS es superior al nivel de significancia 0.05, por lo cual aceptamos la hipótesis nula, es decir, la serie es estacionaria.

Observamos que la función PACF de la serie diferenciada presenta picos en los retardos 12, 24, 36, 48 con decaimiento lento, lo cual nos conduce a la necesidad de diferenciación estacional de los datos.

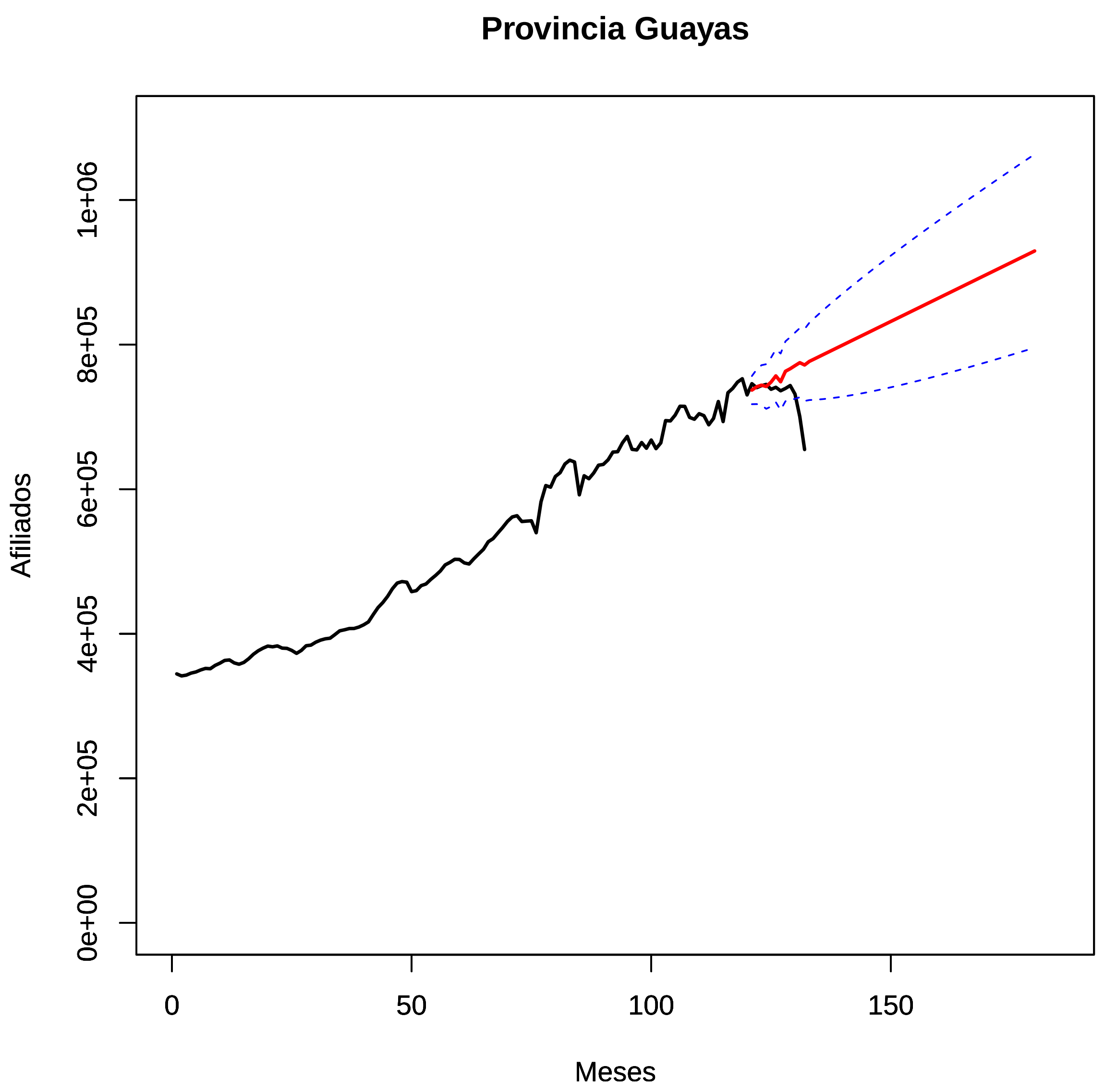




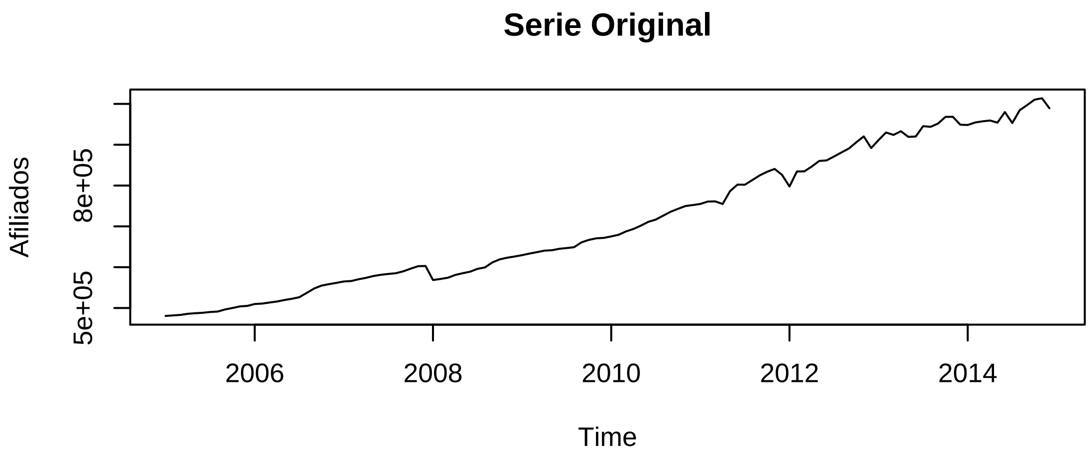
La función ACF presenta un pico en el retardo 12 y valores menores para los retardos 24, 36, 48. Por lo tanto, se sugieren los modelos MA estacional de orden 1 (Q=1). De igual manera los retardos no estacionales $h = 1, 2, \ldots, 11$ de la función PACF decaen, indicando que se debe ajustar un modelo para el comportamiento no estacional de la serie, en principio tomamos $q = 1$.

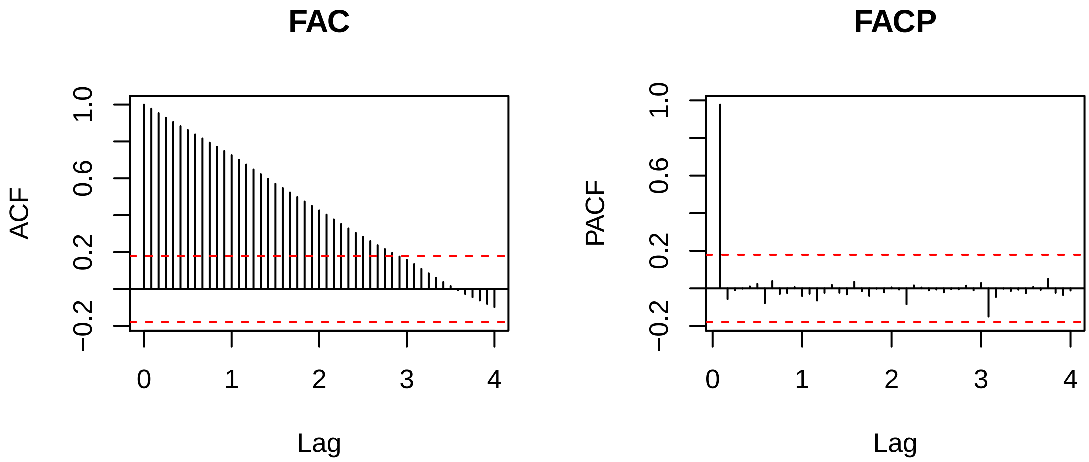
Una vez identificado el modelo a ajustarse sobre la serie correspondiente al número de afiliados de la provincia de Guayas se procede al cálculo del mismo.

```
## Series: data[, 10]
## ARIMA(0,1,1)(0,0,1)[12] with drift
##
## Coefficients:
##           ma1    sma1      drift
##       -0.2913  0.2949  3249.9255
## s.e.   0.0962  0.1032   821.4847
##
## sigma^2 estimated as 98937829:  log likelihood=-1264.84
## AIC=2537.68   AICc=2538.03   BIC=2548.79
```



2.2. Ajuste de afiliados Pichincha

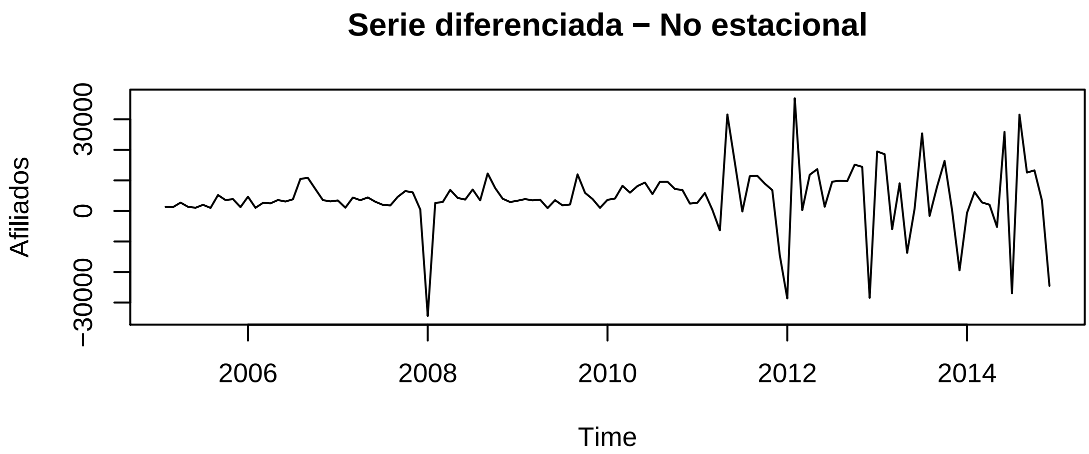


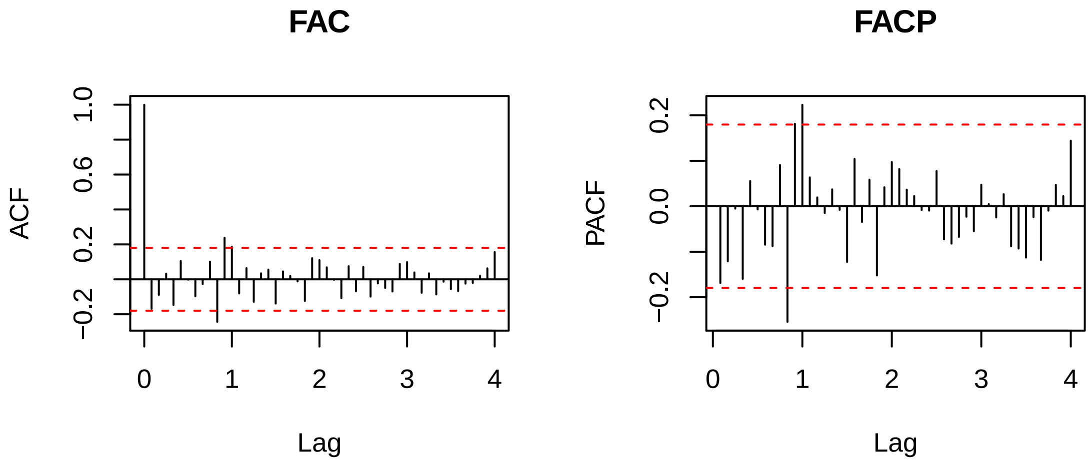


Los correlogramas muestran un decaimiento lento de la función ACF y un pico en el retardo 1 de la función PACF, lo cual implica que la serie tiene un comportamiento no estacionario y surge la necesidad de diferenciar la misma.

```
##
##  KPSS Test for Level Stationarity
##
## data:  data[, 19]
## KPSS Level = 4.0451, Truncation lag parameter = 2, p-value = 0.01
```

La prueba KPSS corrobora el hecho que la serie es no estacionaria, pues el p-value es menor al nivel de significancia 0.05 lo cual implica la existencia de raíces unitarias.

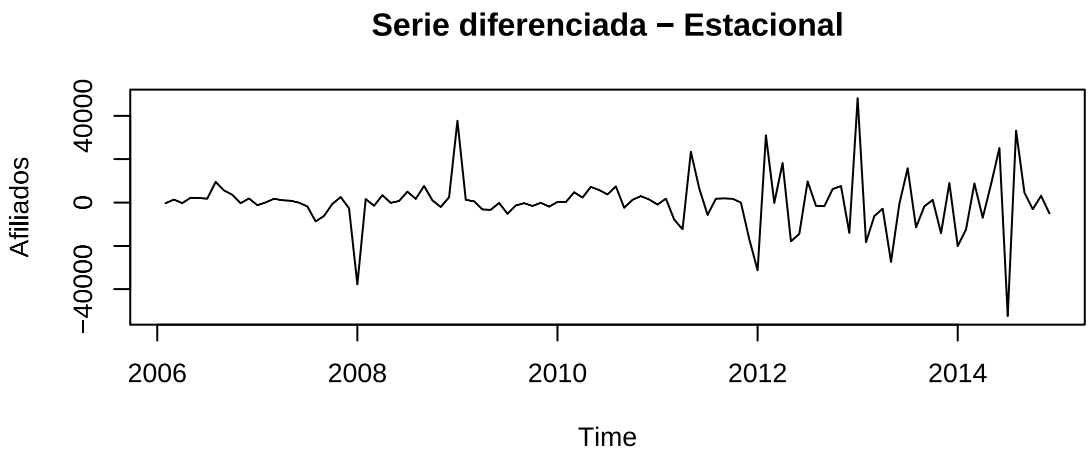


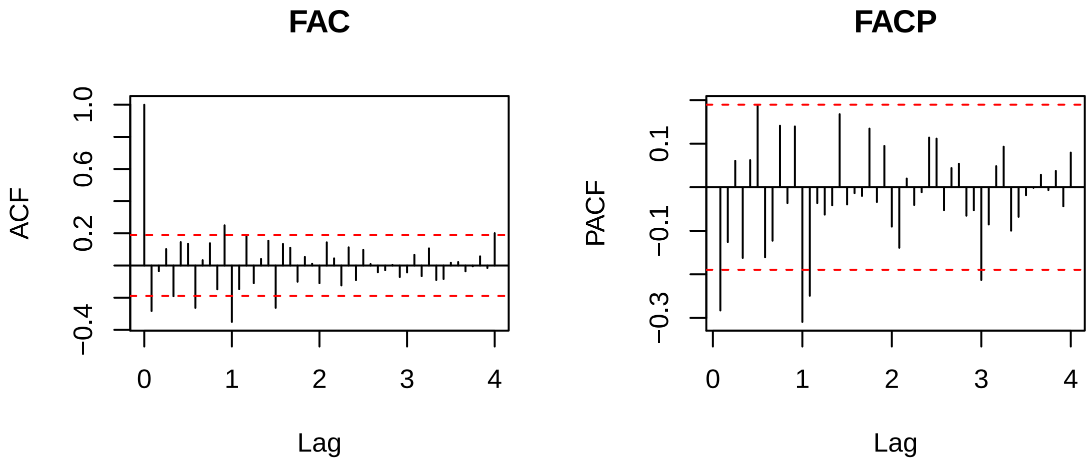



```
##
##  KPSS Test for Level Stationarity
##
## data:  diff(data[, 19])
## KPSS Level = 0.14489, Truncation lag parameter = 2, p-value = 0.1
```

Observamos que el p-value de la prueba KPSS es superior al nivel de significancia 0.05, por lo cual aceptamos la hipótesis nula, es decir, la serie es estacionaria.

Observamos que la función PACF de la serie diferenciada presenta picos en los retardos 12, 24, 36, 48 con decaimiento lento, lo cual nos conduce a la necesidad de diferenciación estacional de los datos.

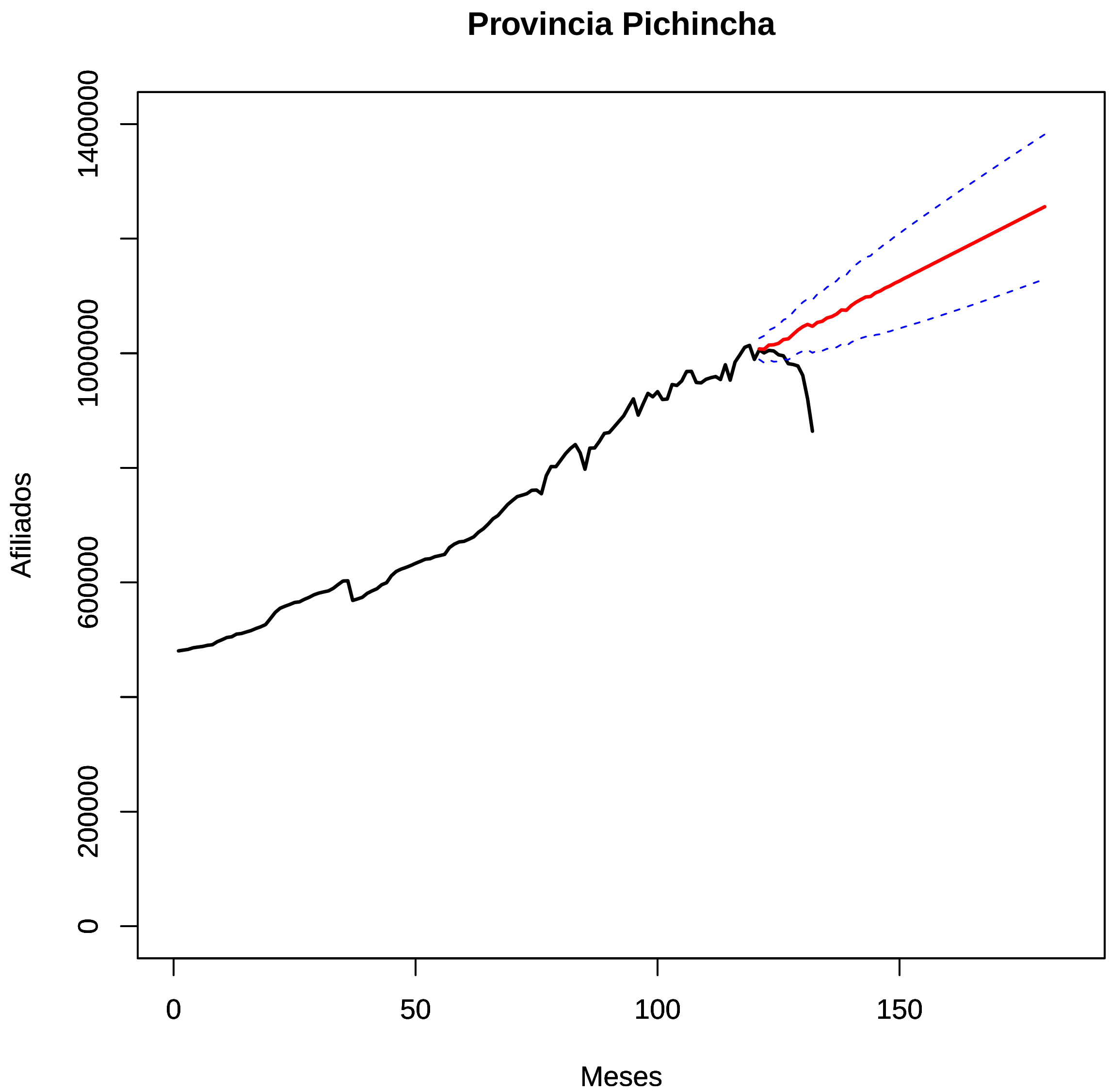




La función ACF presenta dos picos los retardo 12 y 13, mientras valores menores para los retardos 24, 25, 36, 37. Por lo tanto, se sugieren los modelos MA estacional de orden 2 (Q=2). De igual manera los retardos no estacionales $h = 1, 2, \ldots, 11$ de la función PACF decaen, indicando que se debe ajustar un modelo para el comportamiento no estacional de la serie, en principio tomamos $p = 1$ y $q = 2$.

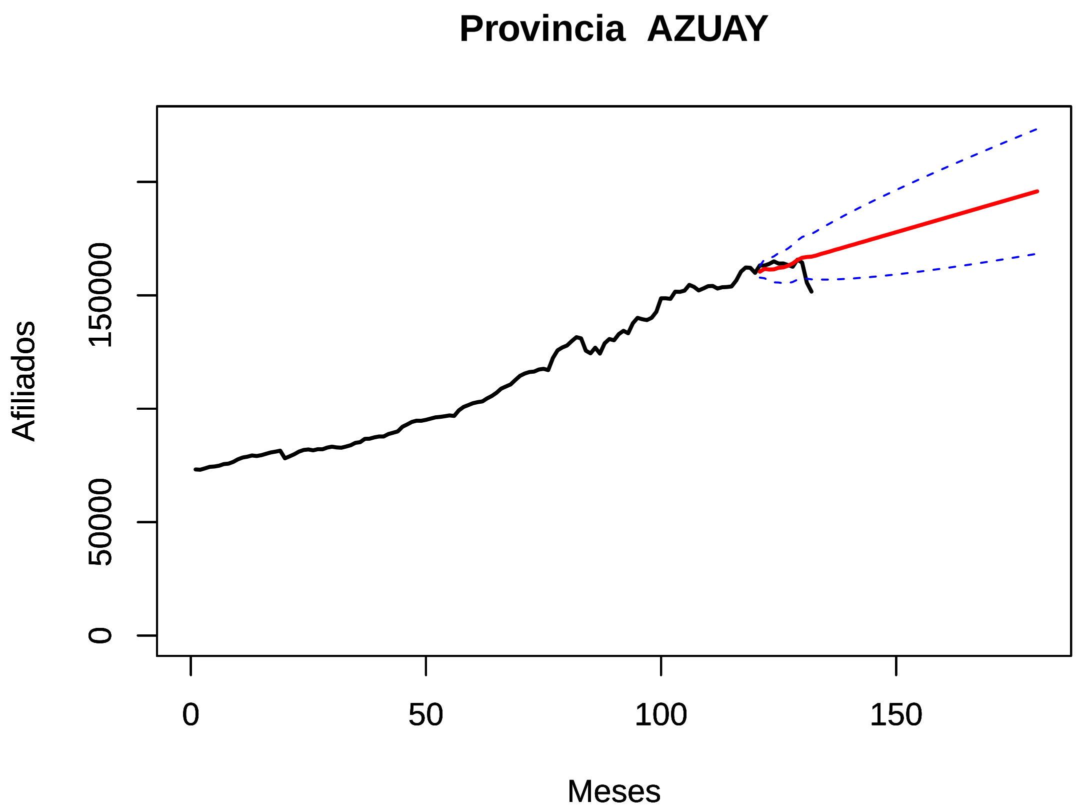
Una vez identificado el modelo a ajustarse sobre la serie correspondiente al número de afiliados de la provincia de Pichincha se procede al cálculo del mismo.

```
## Series: data[, 19]
## ARIMA(1,1,2)(0,0,2)[12] with drift
##
## Coefficients:
##           ar1     ma1      ma2    sma1    sma2      drift
##       -0.8265  0.6085  -0.3601  0.2209  0.1534  4313.3934
## s.e.   0.0708  0.1030   0.0948  0.1022  0.0890   786.5182
##
## sigma^2 estimated as 88992256:  log likelihood=-1258.76
## AIC=2531.51   AICc=2532.52   BIC=2550.97
```

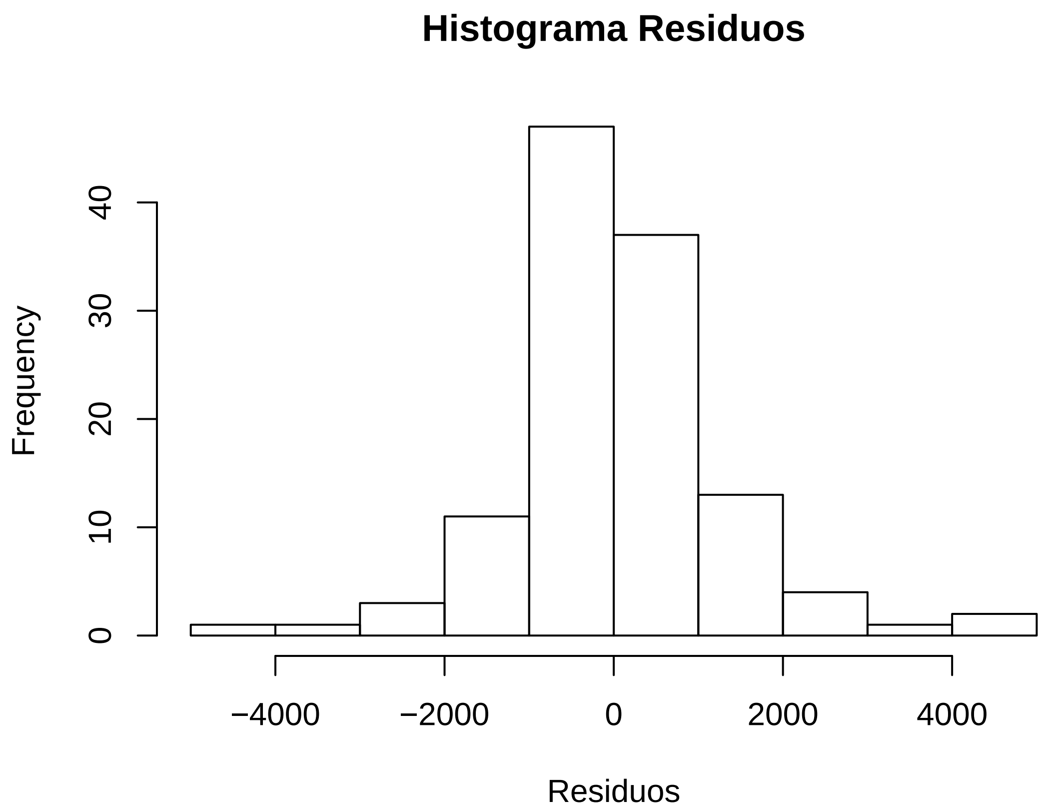


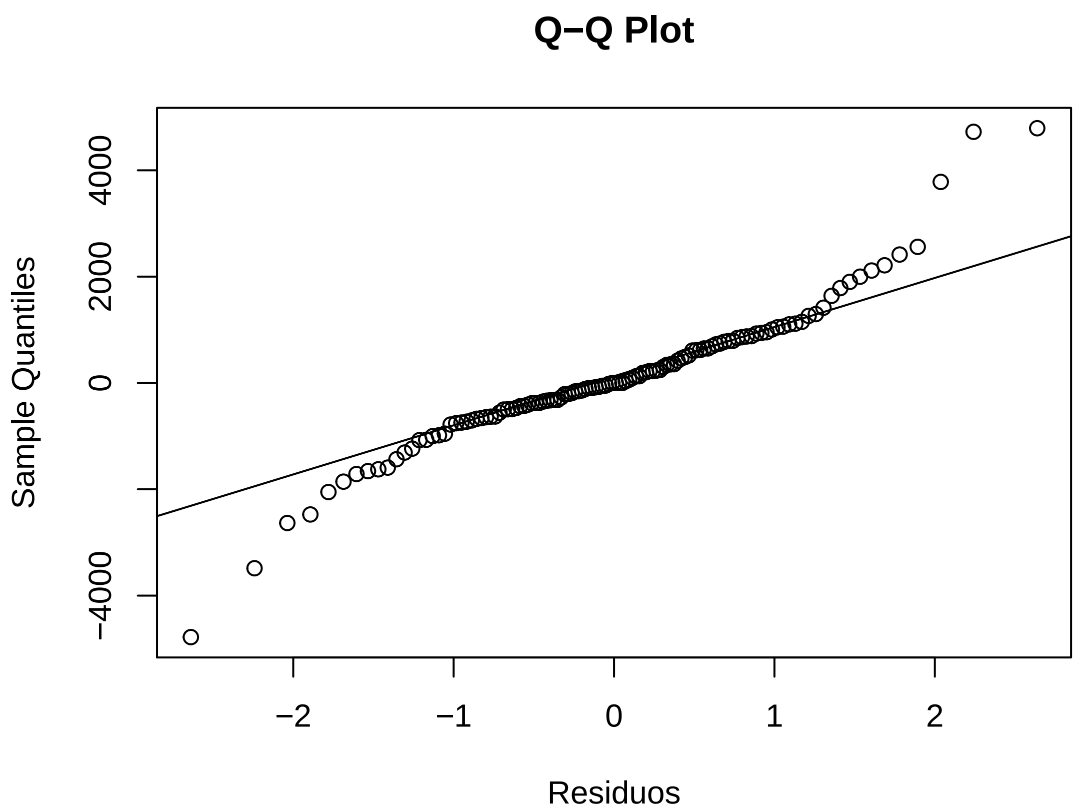
2.3. Ajuste de afiliados del resto de provincias

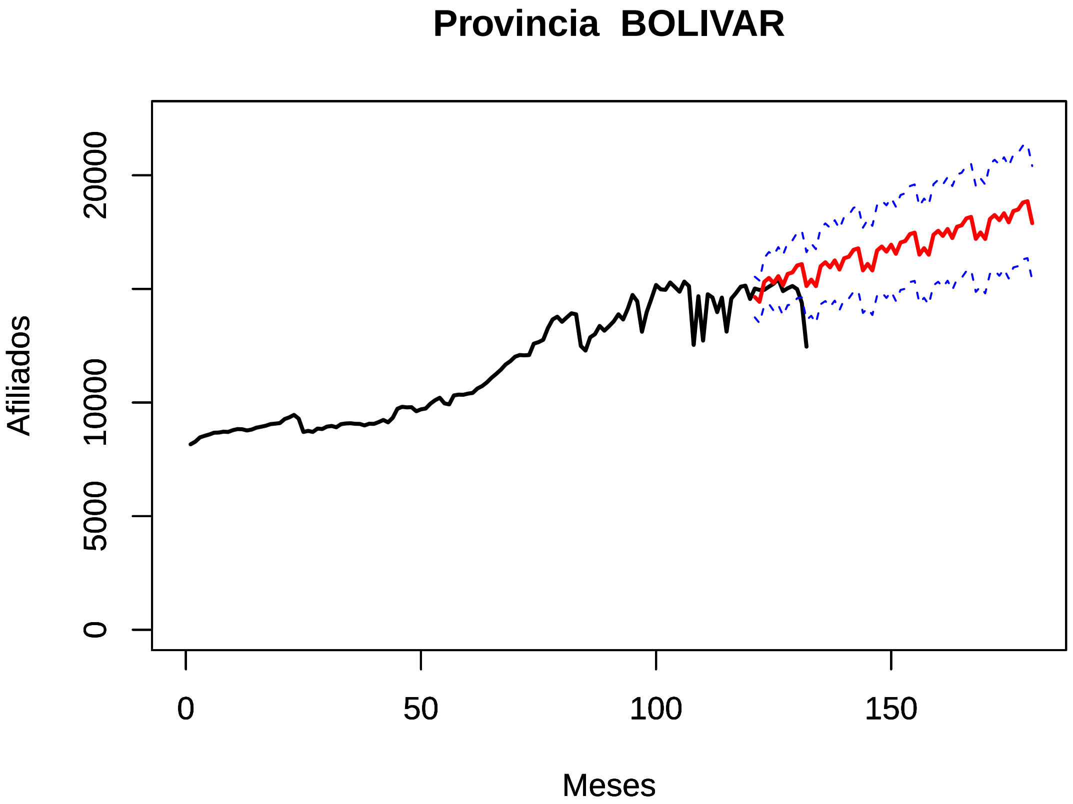
Una vez detallada la metodología a emplearse en la proyección del número de afiliados, se presentan los resultados obtenidos:



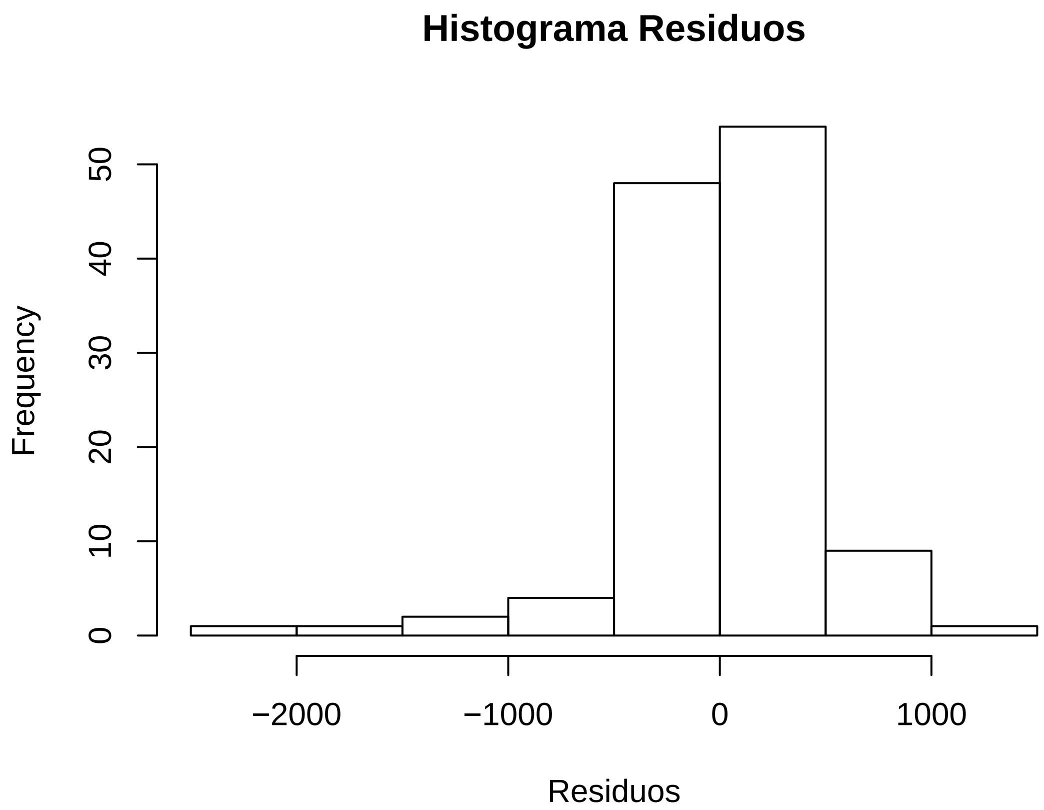

```
## Series: data[, i]
## ARIMA(2,1,2)(0,0,1)[12] with drift
##
## Coefficients:
##           ar1      ar2     ma1     ma2    sma1     drift
##       -1.3127  -0.7876  1.5596  0.8976  0.2793  600.9012
## s.e.   0.1199   0.1125  0.0966  0.1206  0.1167  163.7569
##
## sigma^2 estimated as 1750671:  log likelihood=-1024.2
## AIC=2047.78   AICc=2048.79   BIC=2067.23
```

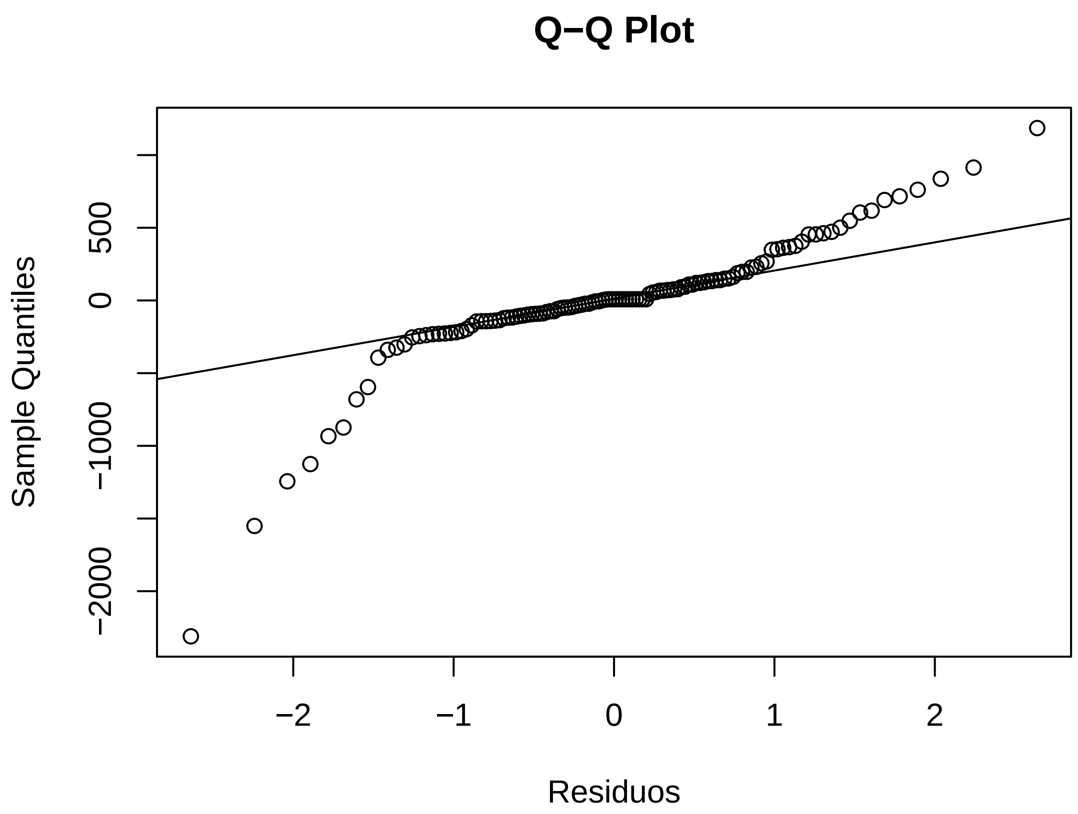

Histograma Residuos
Frequency
0
10
20
30
40
−4000
−2000
0
2000
4000
Residuos

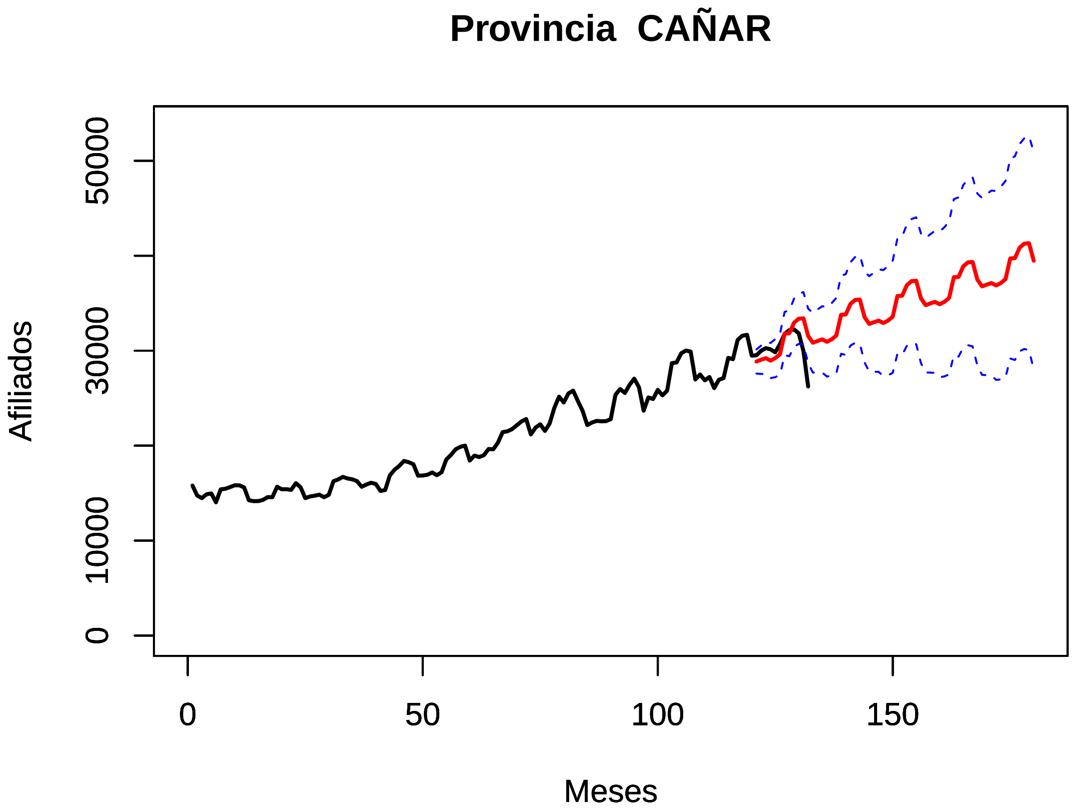

Q−Q Plot
Sample Quantiles
−4000
0
2000
4000
−2
−1
0
1
2
Residuos



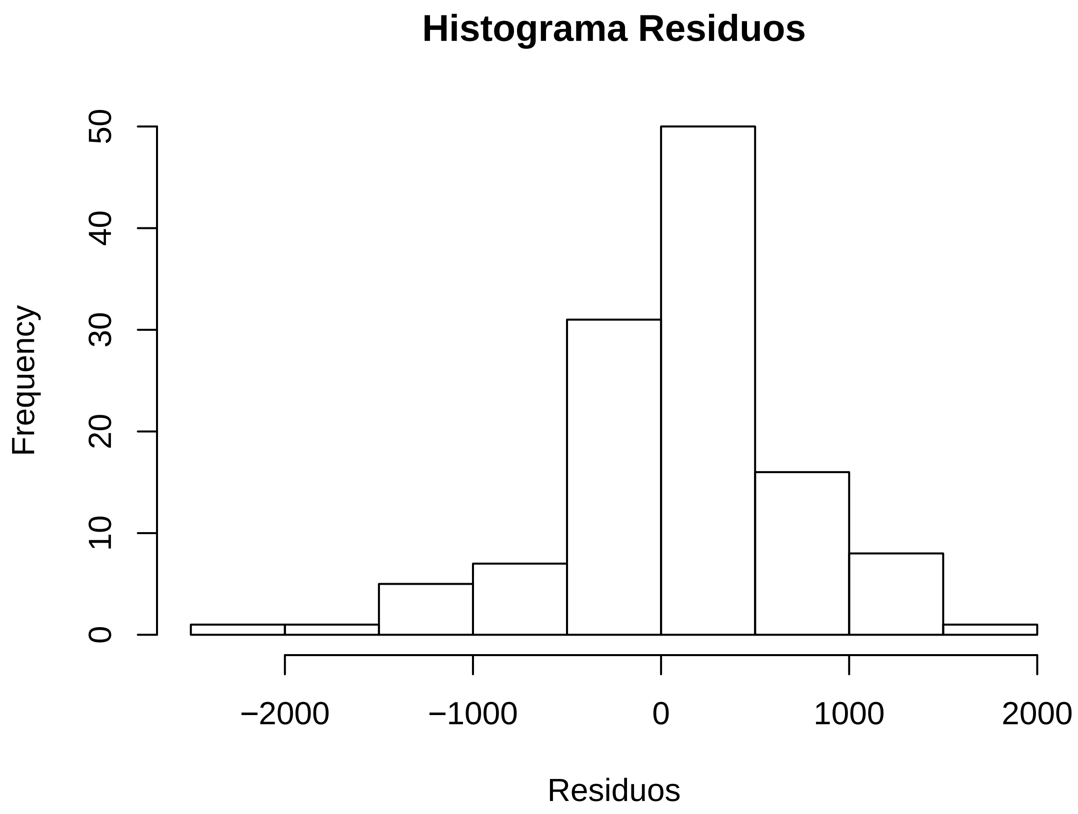

```
## Series: data[, i]
## ARIMA(1,0,2)(0,1,1)[12] with drift
##
## Coefficients:
##          ar1      ma1     ma2     sma1    drift
##       0.9370  -0.6243  0.2405  -0.6556  57.7468
## s.e.  0.0402   0.1011  0.0900   0.1059  13.4978
##
## sigma^2 estimated as 207452:  log likelihood=-818.15
## AIC=1648.3   AICc=1649.13   BIC=1664.39
```


Histograma Residuos
Frequency
0
10
20
30
40
50
−2000
−1000
0
1000
Residuos

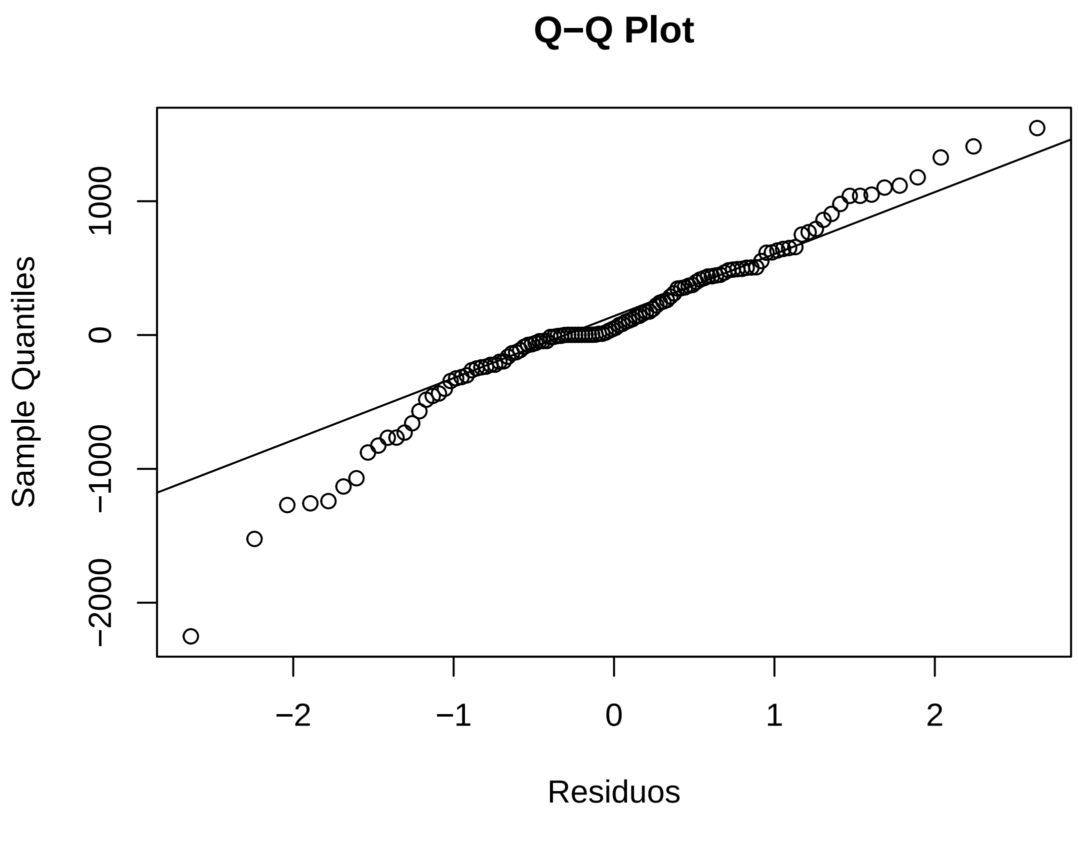

Q−Q Plot
Sample Quantiles
−2000
−1000
0
500
−2
−1
0
1
2
Residuos



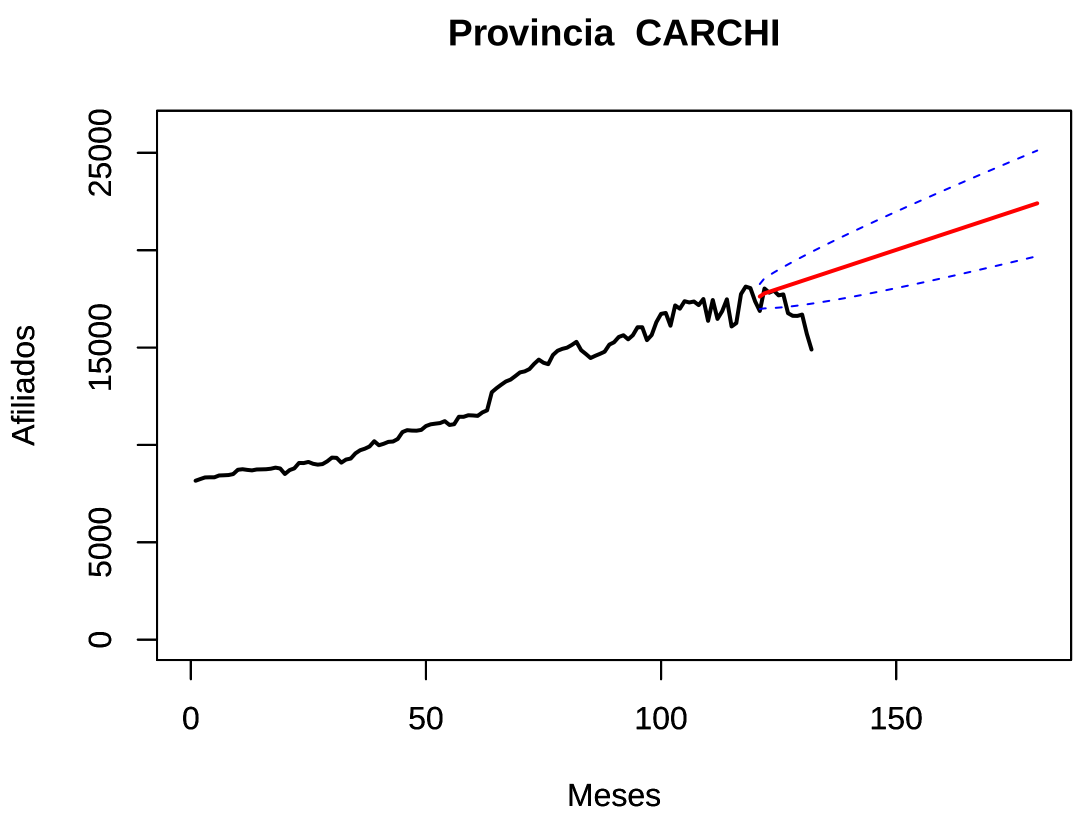

```
## Series: data[, i]
## ARIMA(1,1,1)(0,1,1)[12]
##
## Coefficients:
##          ar1      ma1     sma1
##       0.0120  -0.3981  -0.6018
## s.e.  0.2495   0.2297   0.0915
##
## sigma^2 estimated as 420574:  log likelihood=-847.4
## AIC=1702.8   AICc=1703.19   BIC=1713.49
```



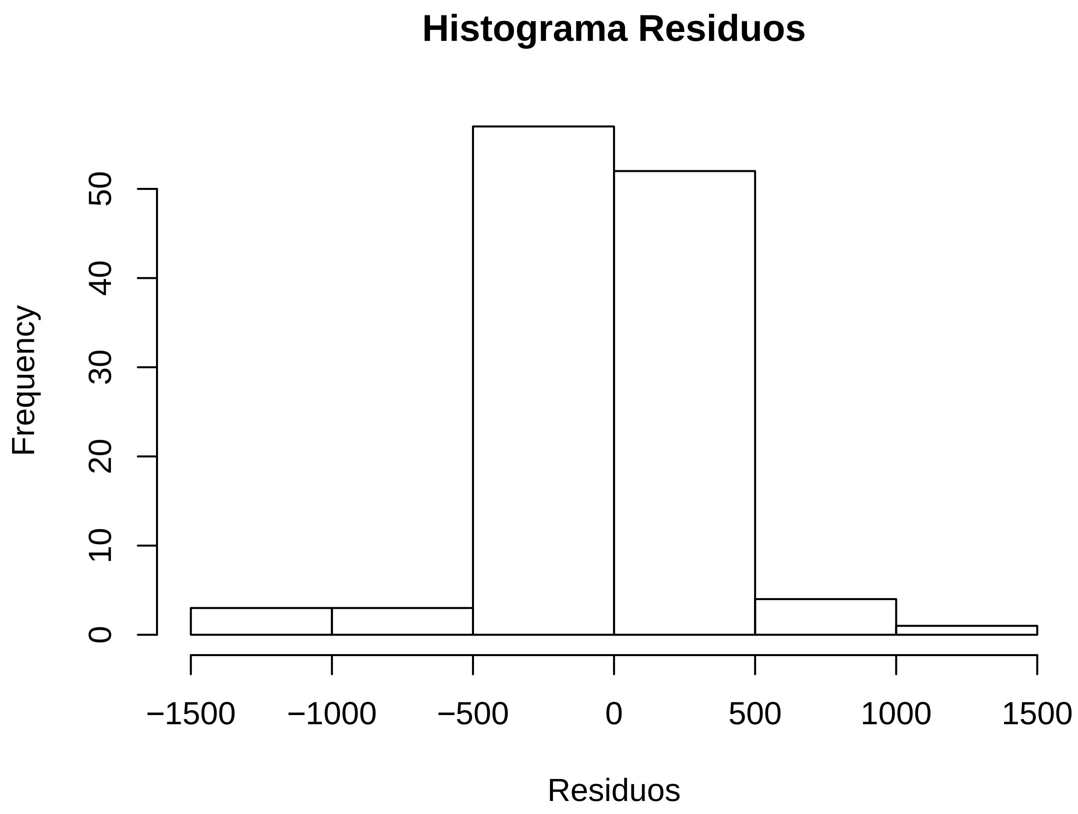
Histograma Residuos
Frequency
0
10
20
30
40
50
−2000
−1000
0
1000
2000
Residuos

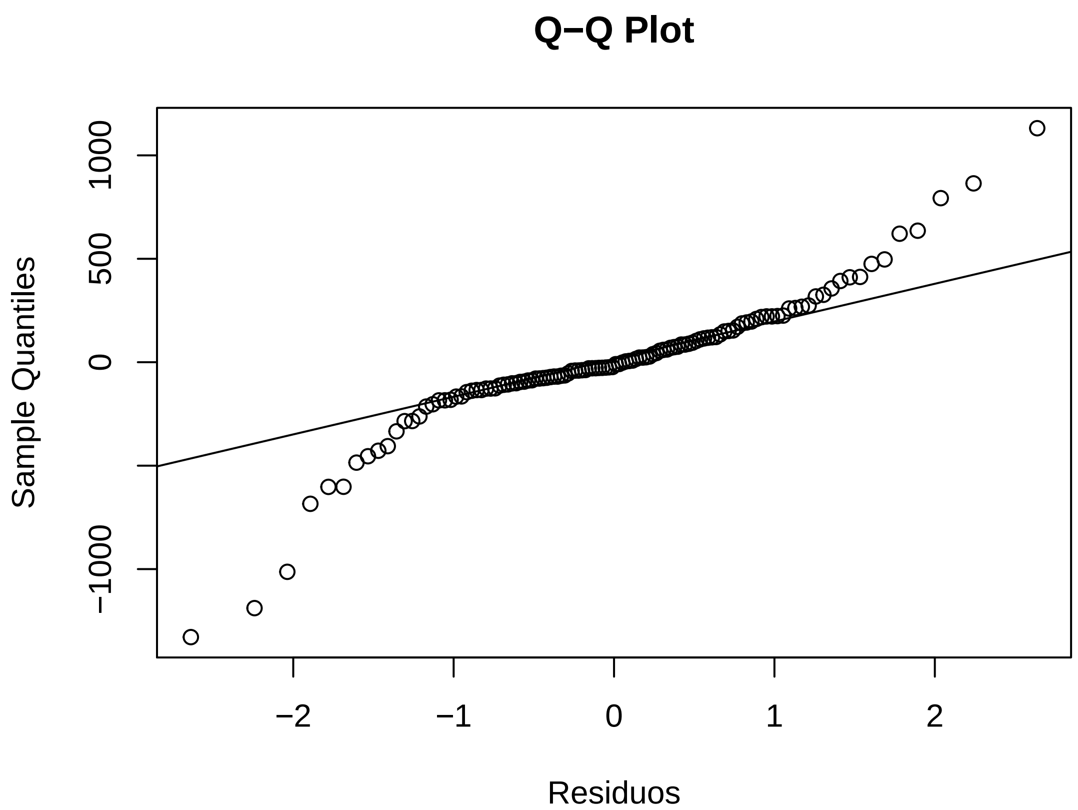


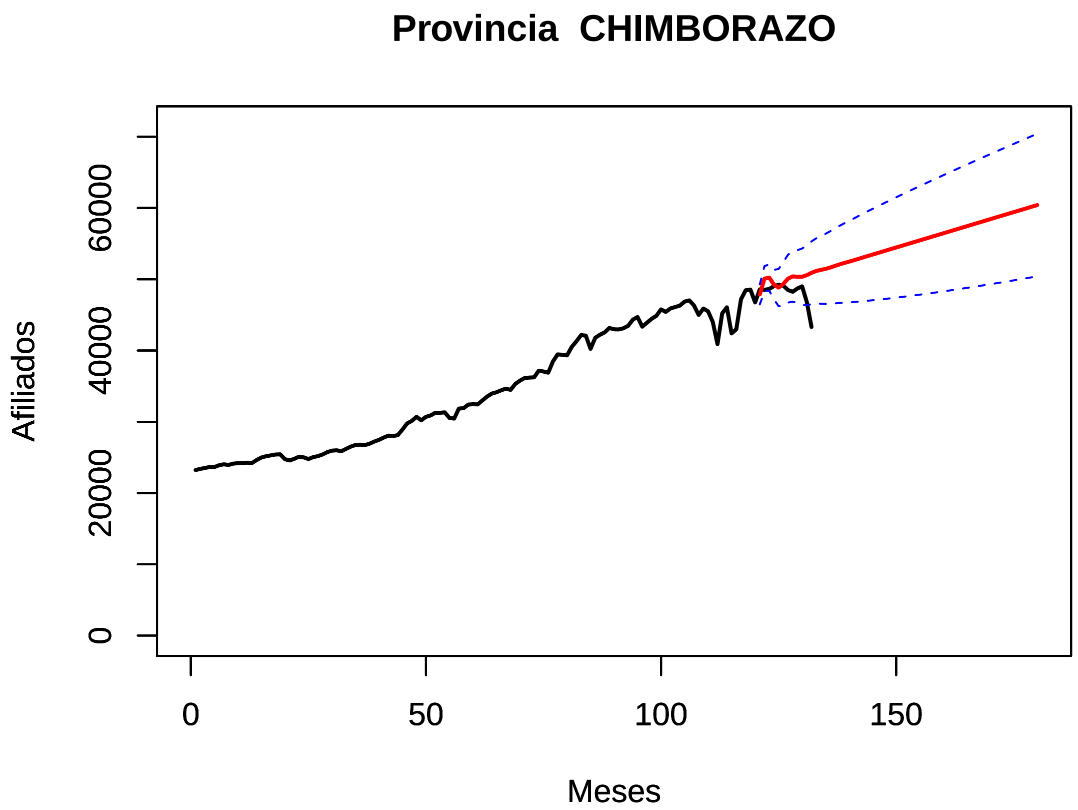
Q−Q Plot
Sample Quantiles
−2000
−1000
0
1000
−2
−1
0
1
2
Residuos



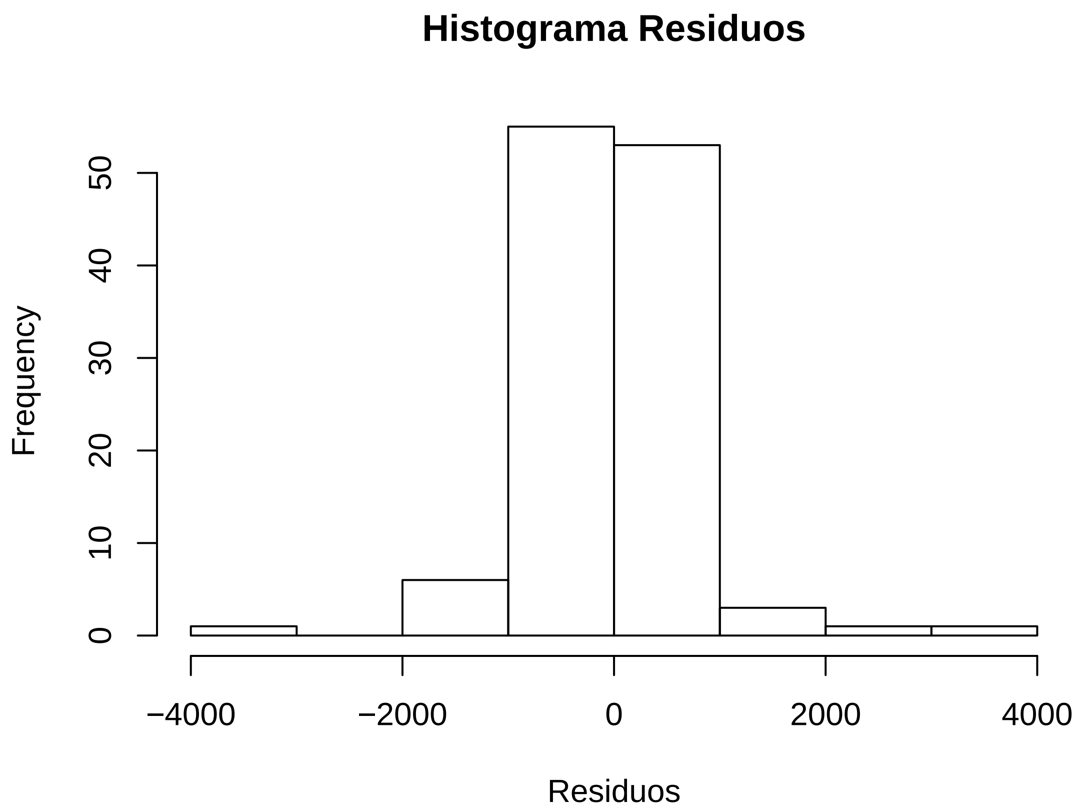

```
## Series: data[, i]
## ARIMA(0,1,2) with drift
##
## Coefficients:
##           ma1      ma2    drift
##       -0.3281  -0.1399  79.6781
## s.e.   0.0916   0.0897  16.1169
##
## sigma^2 estimated as 107074:  log likelihood=-858.04
## AIC=1724.08   AICc=1724.43   BIC=1735.19
```

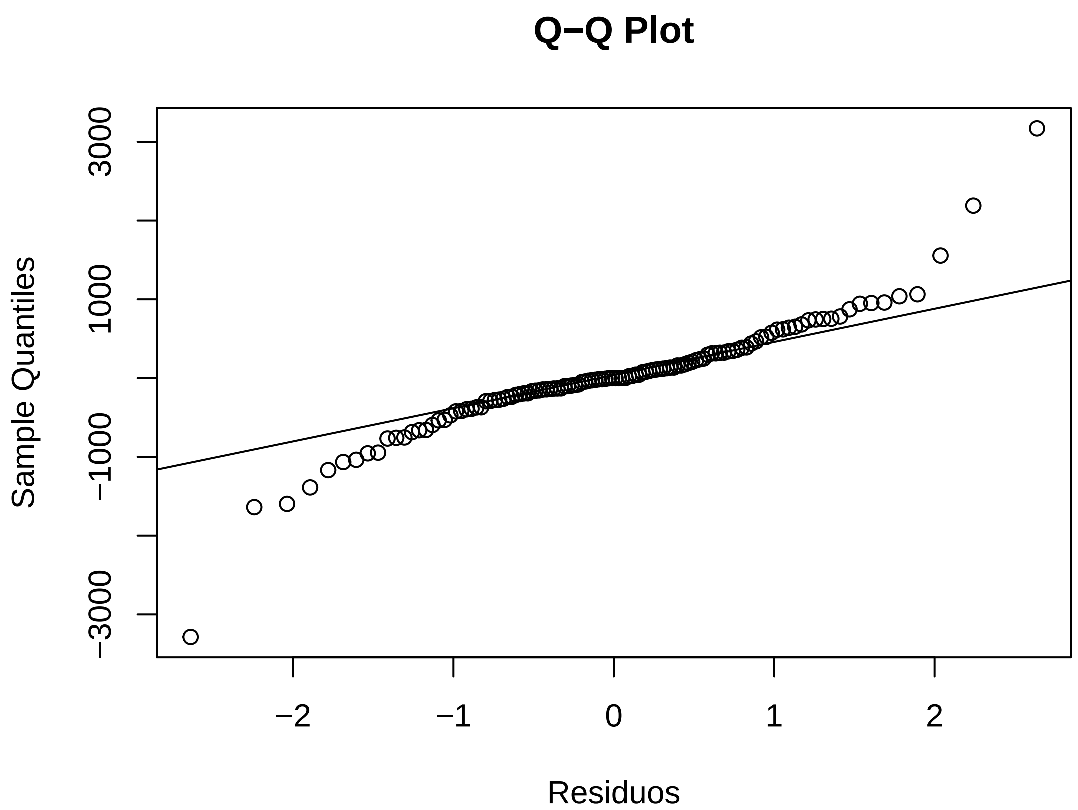

Histograma Residuos
Frequency
0
10
20
30
40
50
−1500
−1000
−500
0
500
1000
1500
Residuos

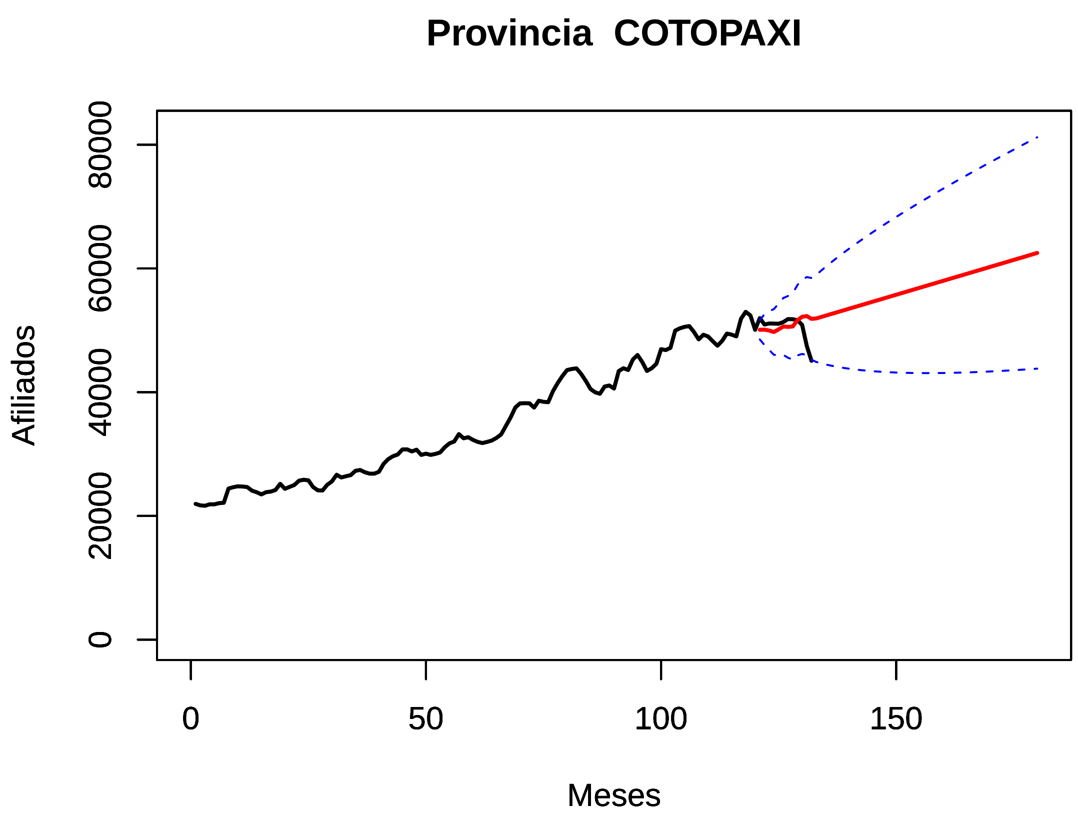

Q−Q Plot
Sample Quantiles
−1000
0
500
1000
−2
−1
0
1
2
Residuos



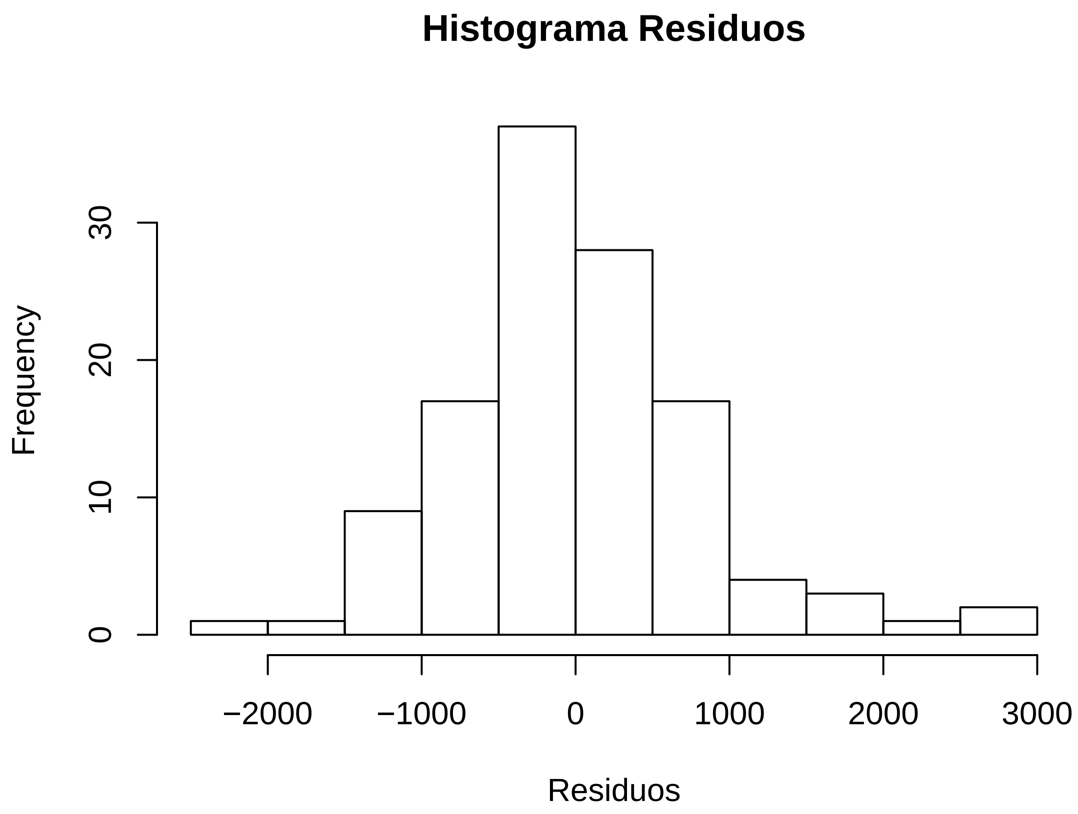

```
## Series: data[, i]
## ARIMA(2,1,3) with drift
##
## Coefficients:
##          ar1      ar2      ma1     ma2     ma3     drift
##       0.5599  -0.5733  -0.7919  0.2698  0.4577  198.3843
## s.e.  0.1072   0.1027   0.1032  0.1330  0.0923   59.7068
##
## sigma^2 estimated as 514123:  log likelihood=-951.29
## AIC=1902.95   AICc=1903.95   BIC=1922.4
```


Histograma Residuos
Frequency
0
10
20
30
40
50
−4000
−2000
0
2000
4000
Residuos

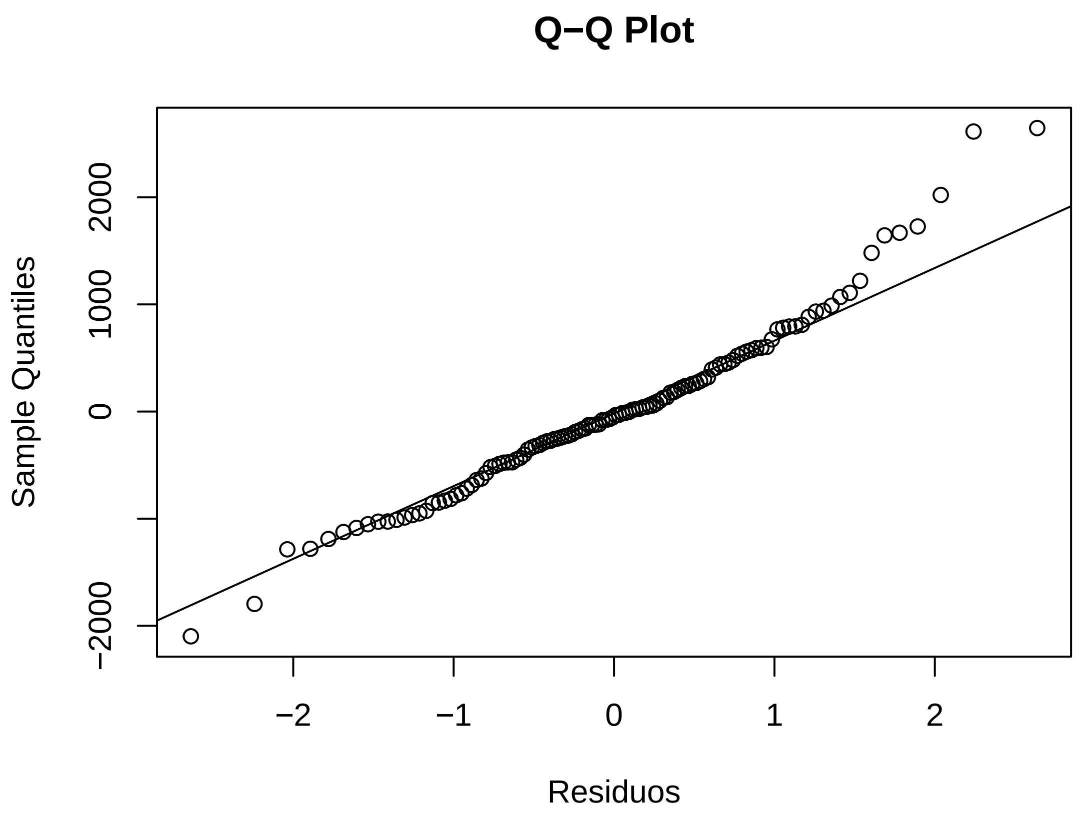

Q−Q Plot
Sample Quantiles
−3000
−1000
1000
3000
−2
−1
0
1
2
Residuos



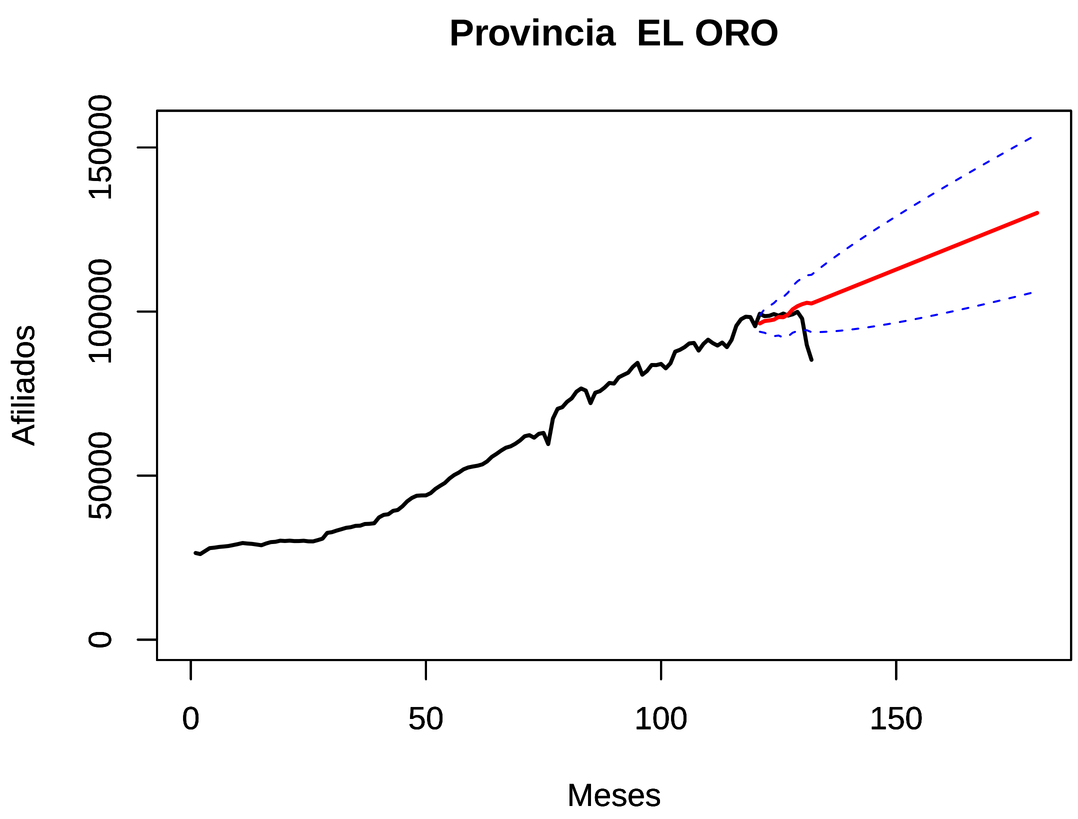

```
## Series: data[, i]
## ARIMA(0,1,1)(0,0,1)[12] with drift
##
## Coefficients:
##          ma1    sma1     drift
##       0.2365  0.3074  225.2014
## s.e.  0.0981  0.1050  115.3807
##
## sigma^2 estimated as 636452:  log likelihood=-964.62
## AIC=1937.23   AICc=1937.58   BIC=1948.35
```



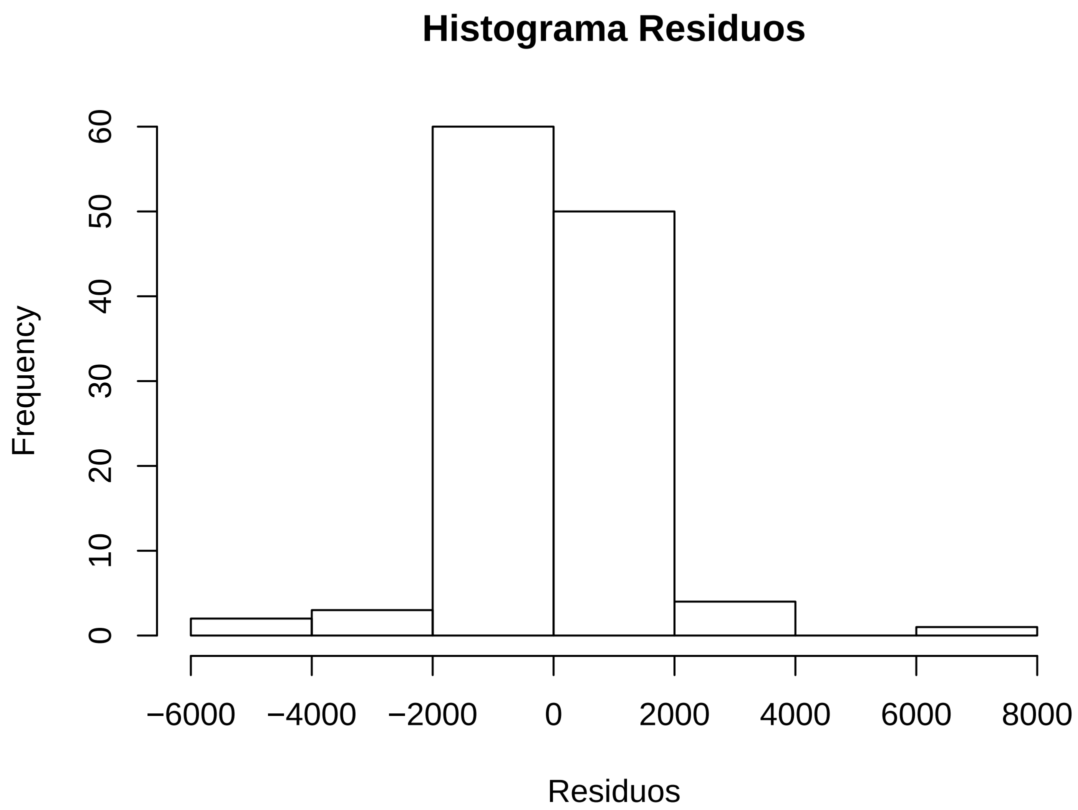
Histograma Residuos
Frequency
0
10
20
30
−2000
−1000
0
1000
2000
3000
Residuos

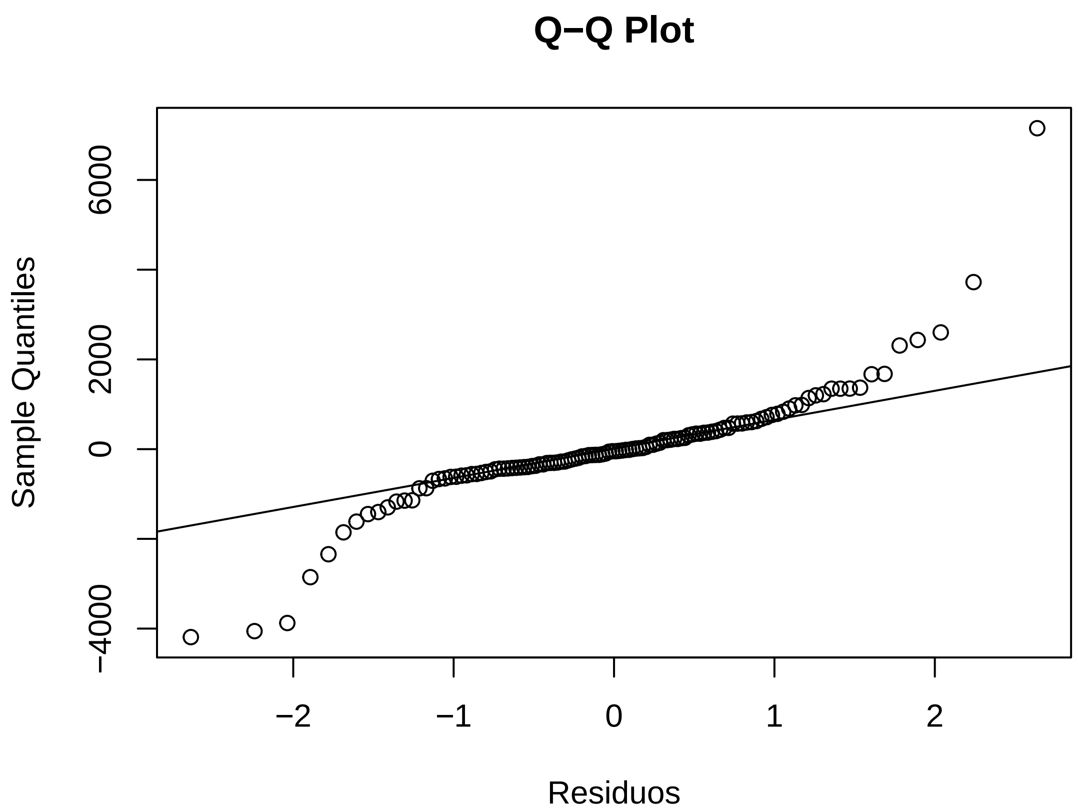


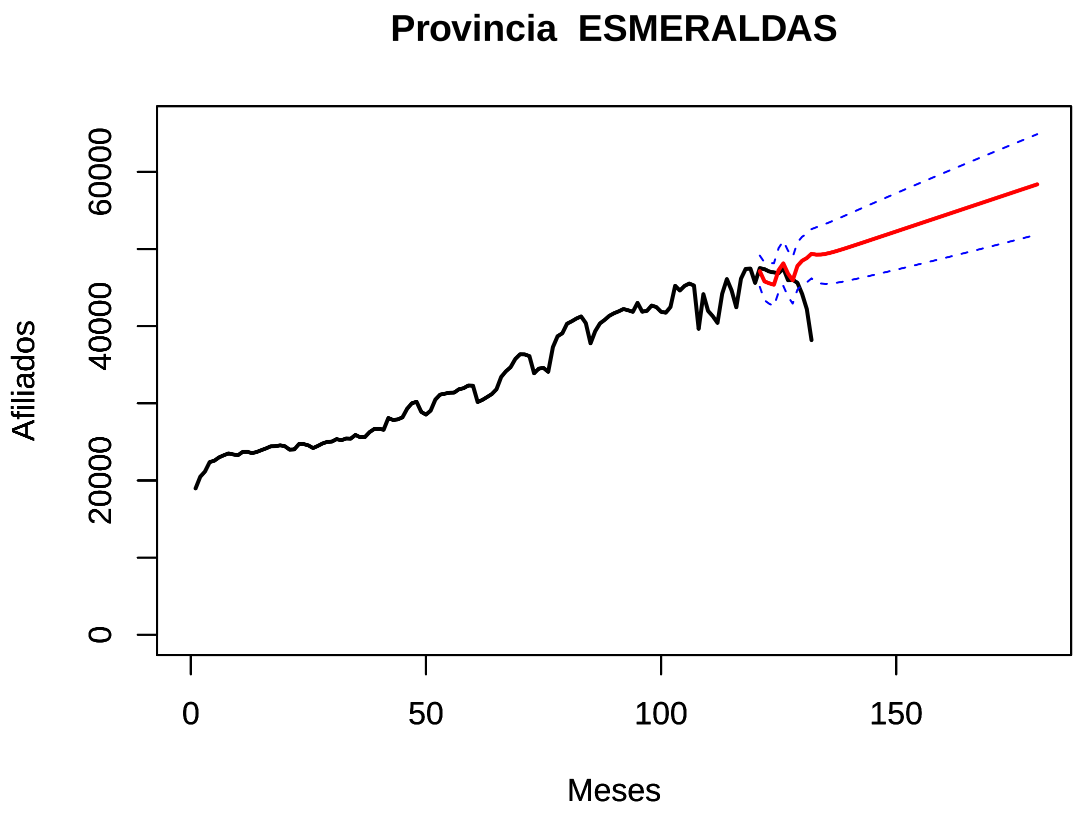
Q−Q Plot
Sample Quantiles
−2000
0
1000
2000
−2
−1
0
1
2
Residuos



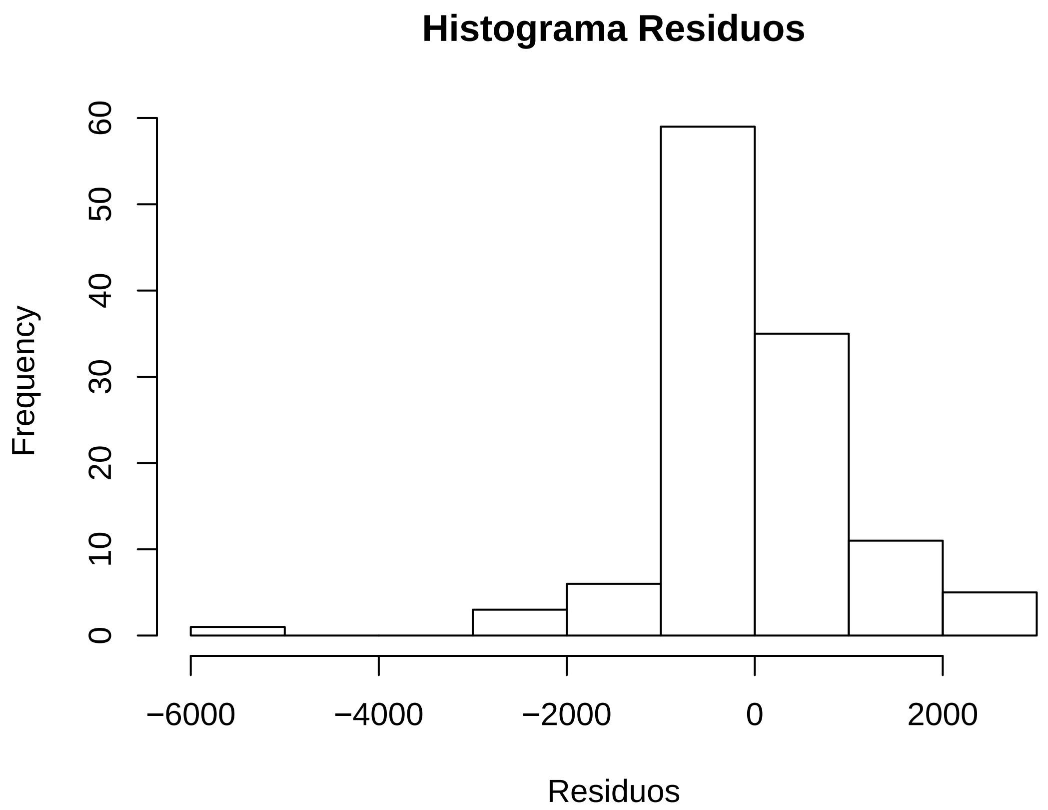

```
## Series: data[, i]
## ARIMA(0,1,0)(0,0,1)[12] with drift
##
## Coefficients:
##          sma1     drift
##        0.2749  575.2182
## s.e.   0.0876  147.2591
##
## sigma^2 estimated as 1656548:  log likelihood=-1021.38
## AIC=2048.76   AICc=2048.97   BIC=2057.1
```


Histograma Residuos
Frequency
0
10
20
30
40
50
60
−6000
−4000
−2000
0
2000
4000
6000
8000
Residuos

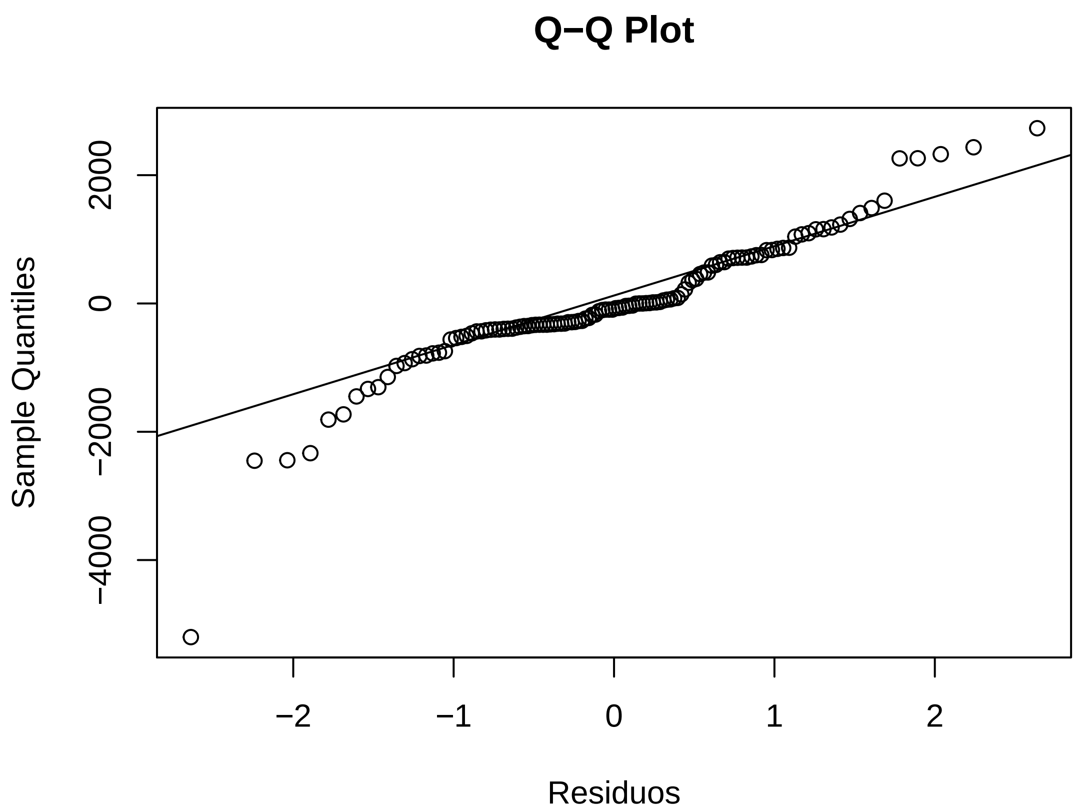

Q−Q Plot
Sample Quantiles
−4000
0
2000
6000
−2
−1
0
1
2
Residuos



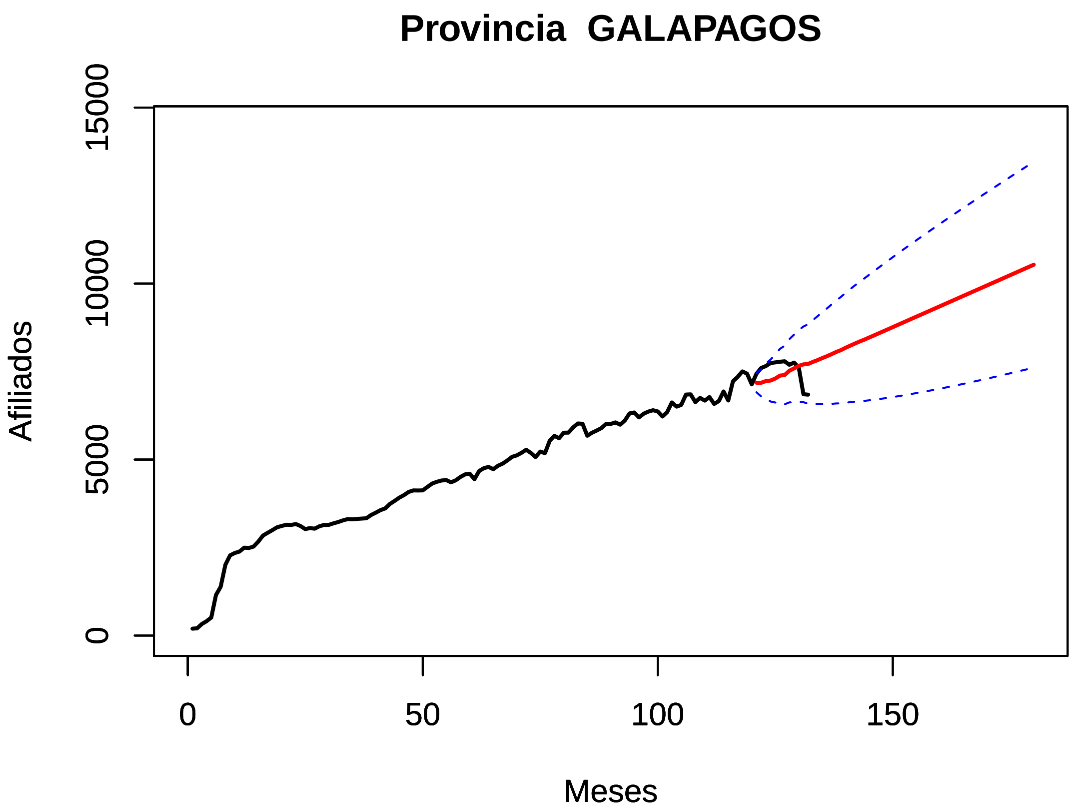

```
## Series: data[, i]
## ARIMA(1,1,1)(0,0,1)[12] with drift
##
## Coefficients:
##          ar1      ma1    sma1     drift
##       0.5797  -0.8943  0.5114  203.5063
## s.e.  0.0870   0.0558  0.1058   37.0427
##
## sigma^2 estimated as 1063292:  log likelihood=-994.53
## AIC=1984.9   AICc=1985.43   BIC=1998.79
```



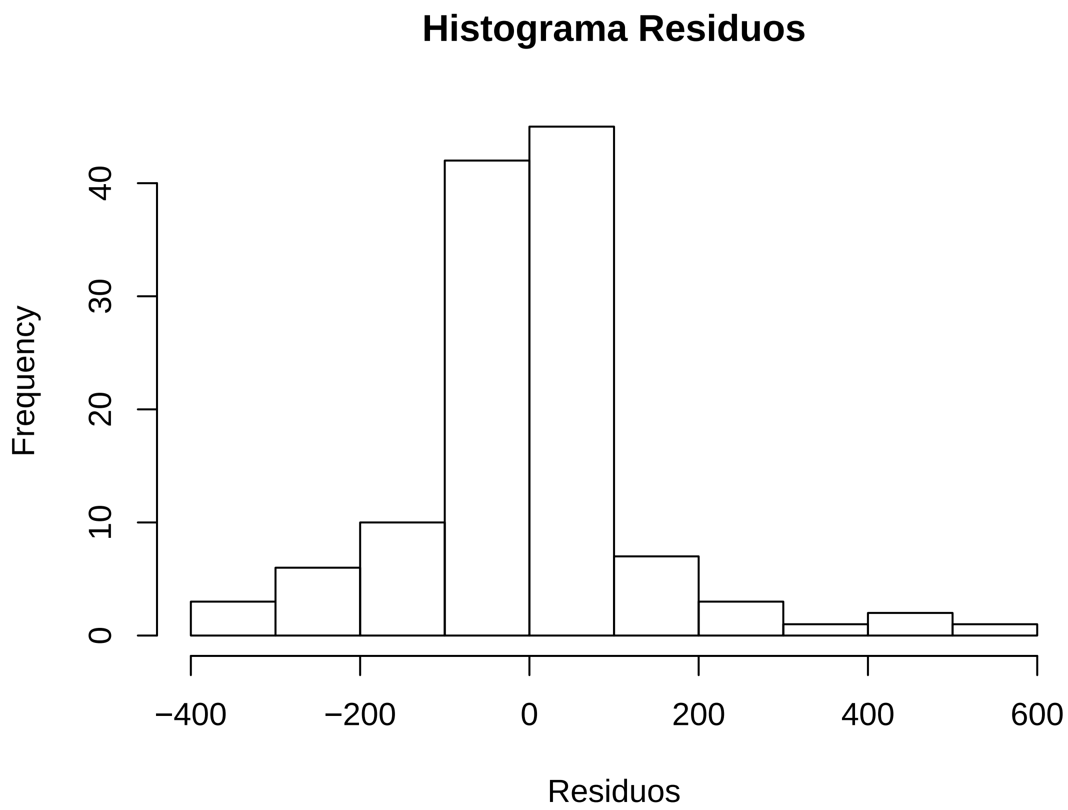
Histograma Residuos
Frequency
0
10
20
30
40
50
60
−6000
−4000
−2000
0
2000
Residuos

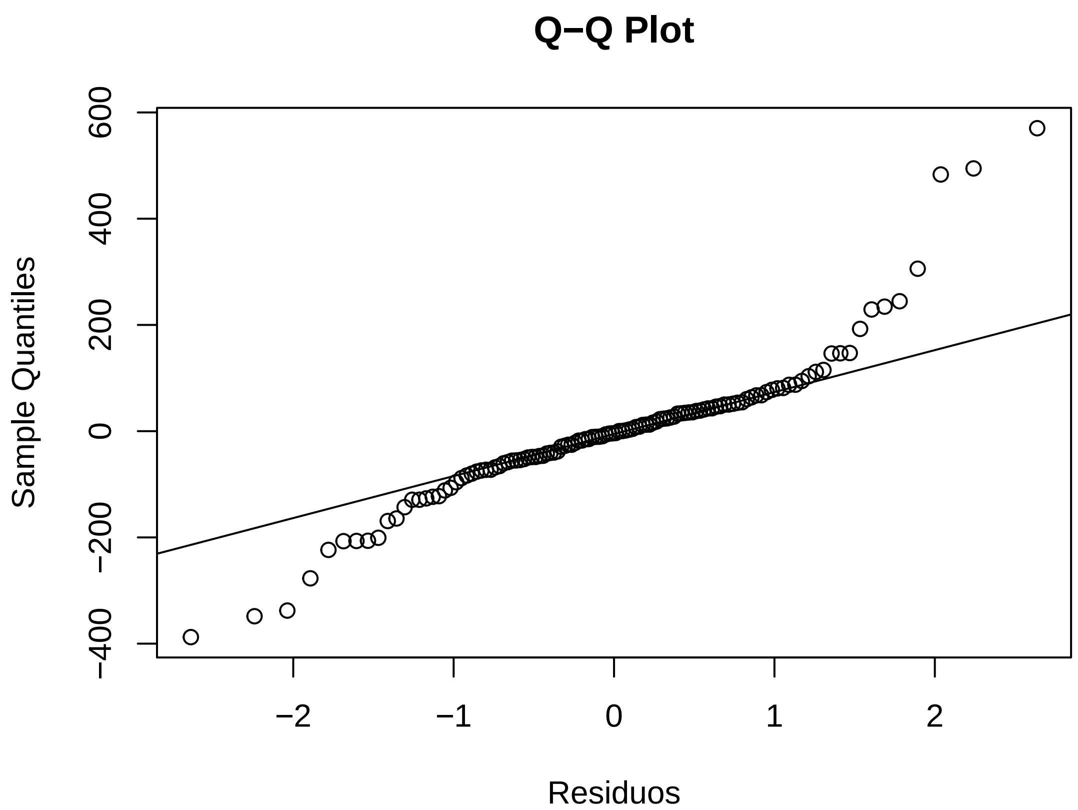


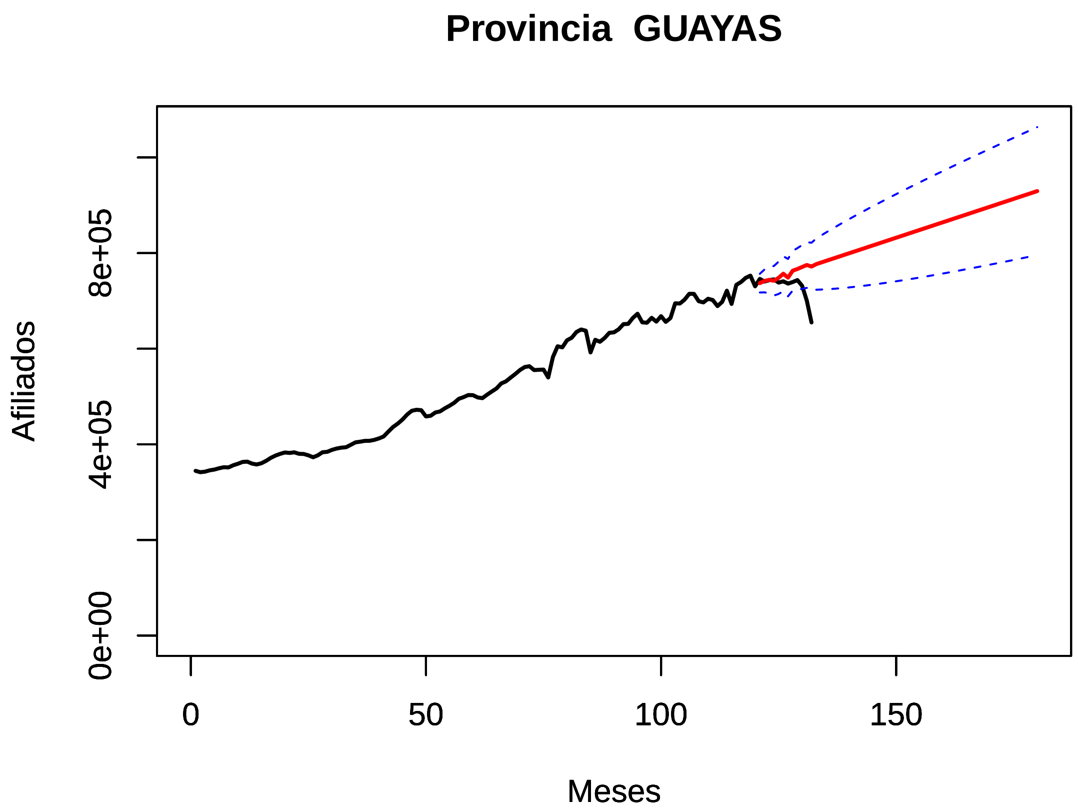
Q−Q Plot
Sample Quantiles
−4000
−2000
0
2000
−2
−1
0
1
2
Residuos



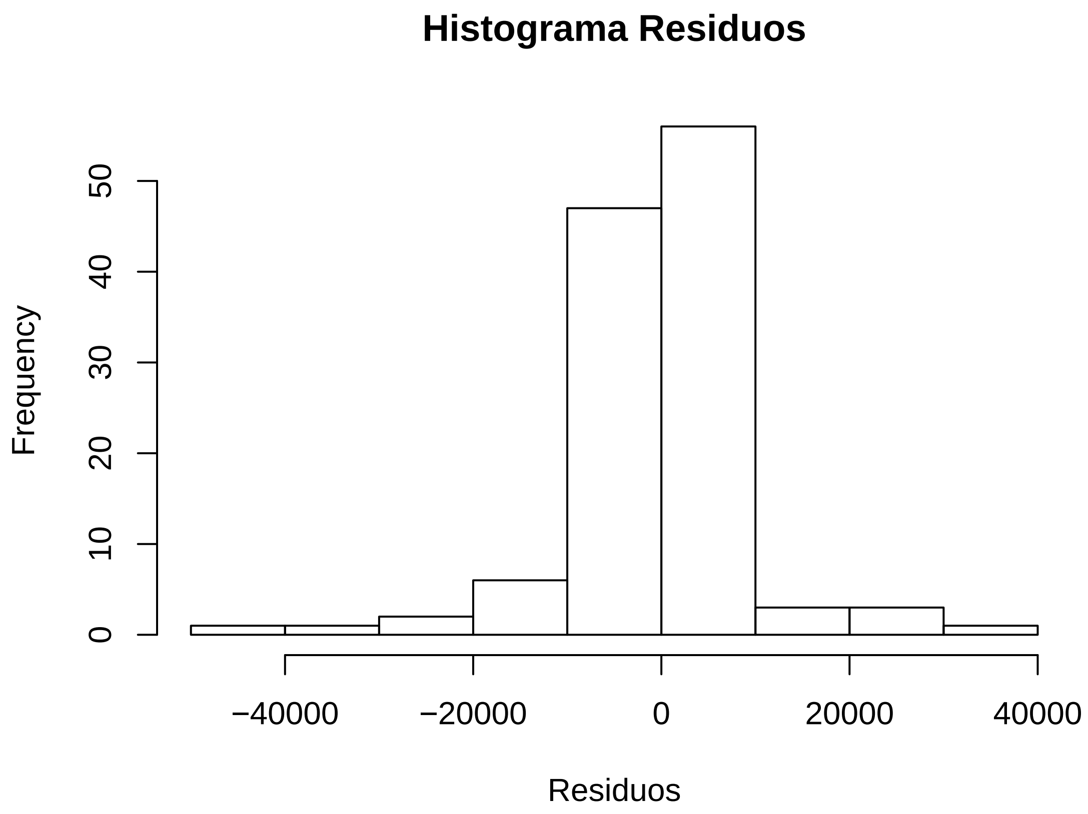

```
## Series: data[, i]
## ARIMA(2,1,1)(1,0,0)[12] with drift
##
## Coefficients:
##          ar1     ar2      ma1    sar1    drift
##       0.3235  0.1156  -0.2933  0.1298  59.1207
## s.e.  0.4523  0.0998   0.4543  0.1274  18.0832
##
## sigma^2 estimated as 19056:  log likelihood=-755.35
## AIC=1522.71   AICc=1523.46   BIC=1539.38
```


Histograma Residuos
Frequency
0
10
20
30
40
−400
−200
0
200
400
600
Residuos

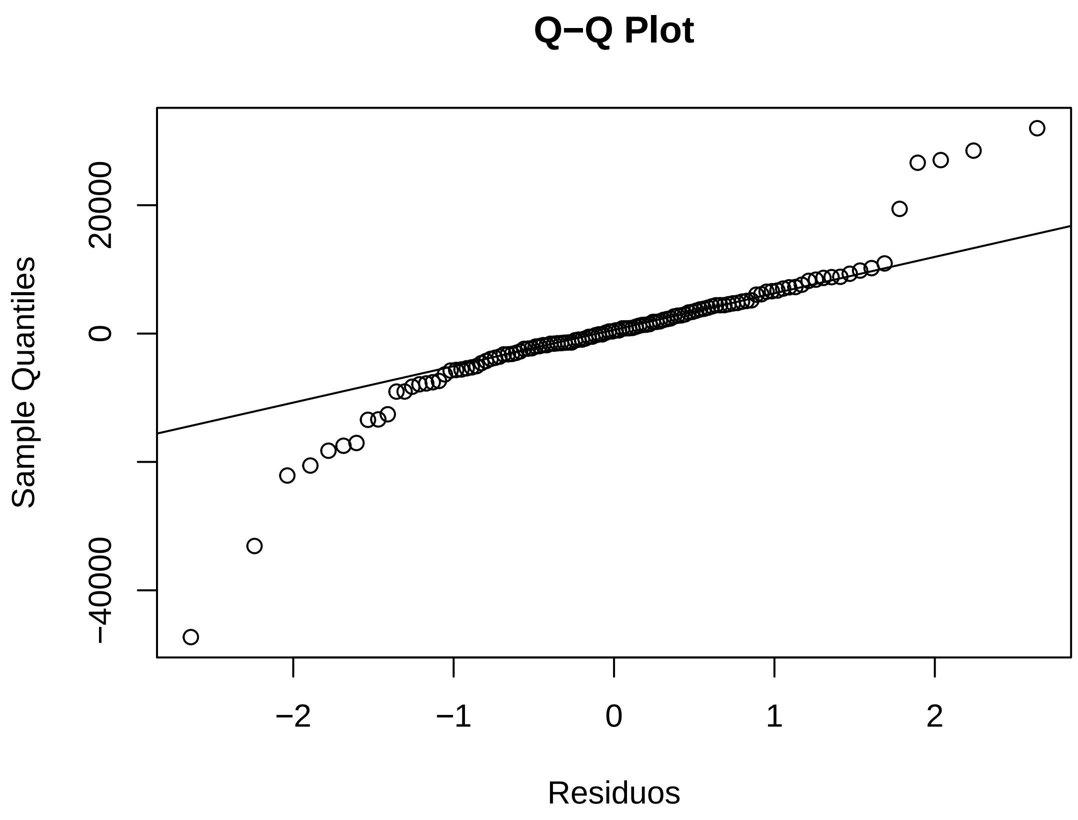

Q−Q Plot
Sample Quantiles
−400
−200
0
200
400
600
−2
−1
0
1
2
Residuos



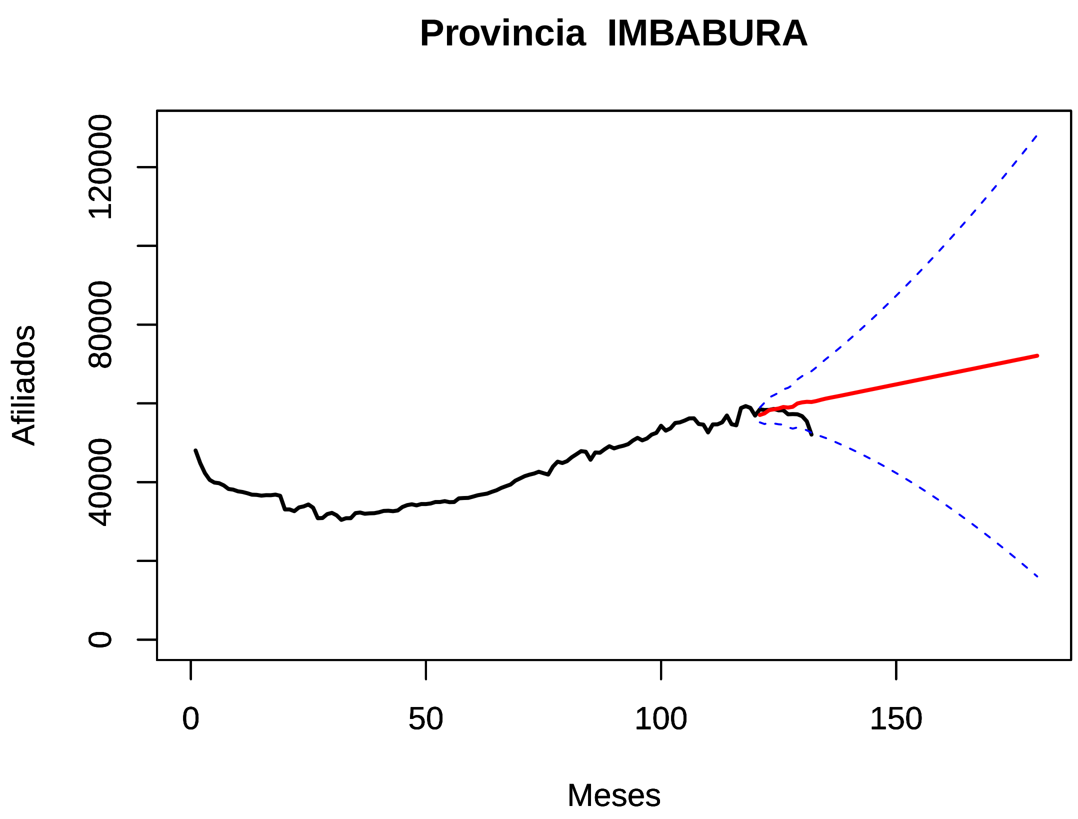

```
## Series: data[, i]
## ARIMA(0,1,1)(0,0,1)[12] with drift
##
## Coefficients:
##            ma1    sma1      drift
##        -0.2913  0.2949  3249.9255
## s.e.    0.0962  0.1032   821.4847
##
## sigma^2 estimated as 98937829:  log likelihood=-1264.84
## AIC=2537.68   AICc=2538.03   BIC=2548.79
```



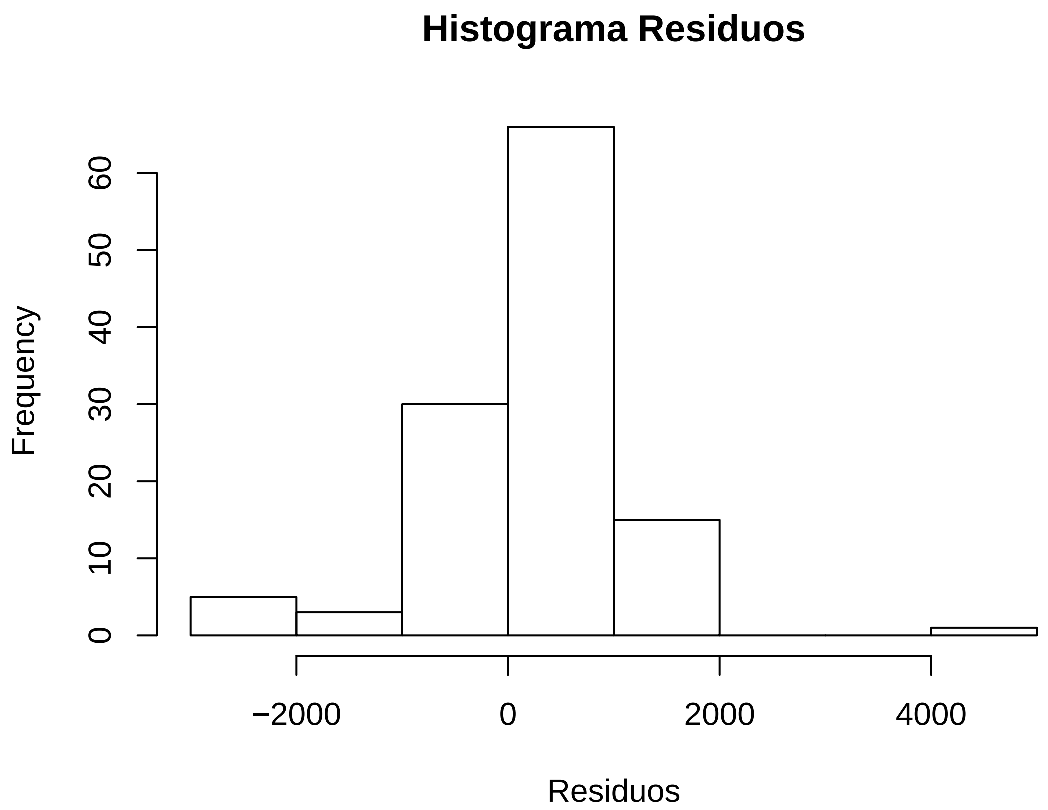
Histograma Residuos
Frequency
0
10
20
30
40
50
−40000
−20000
0
20000
40000
Residuos

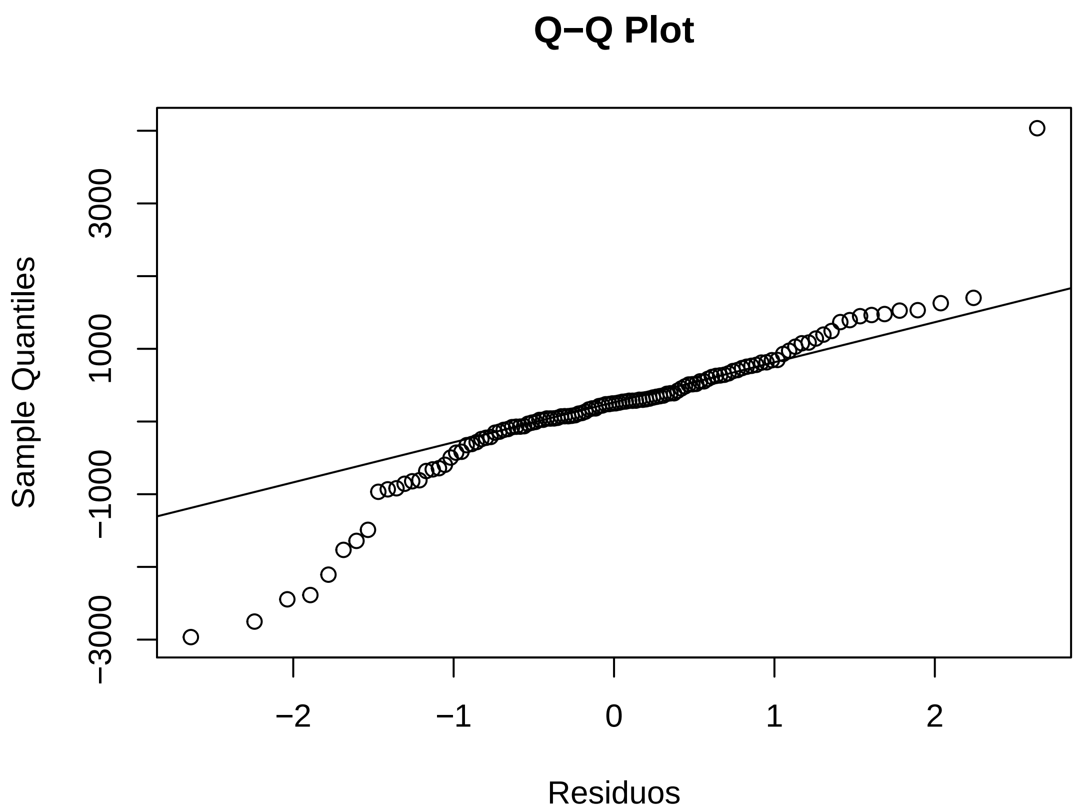


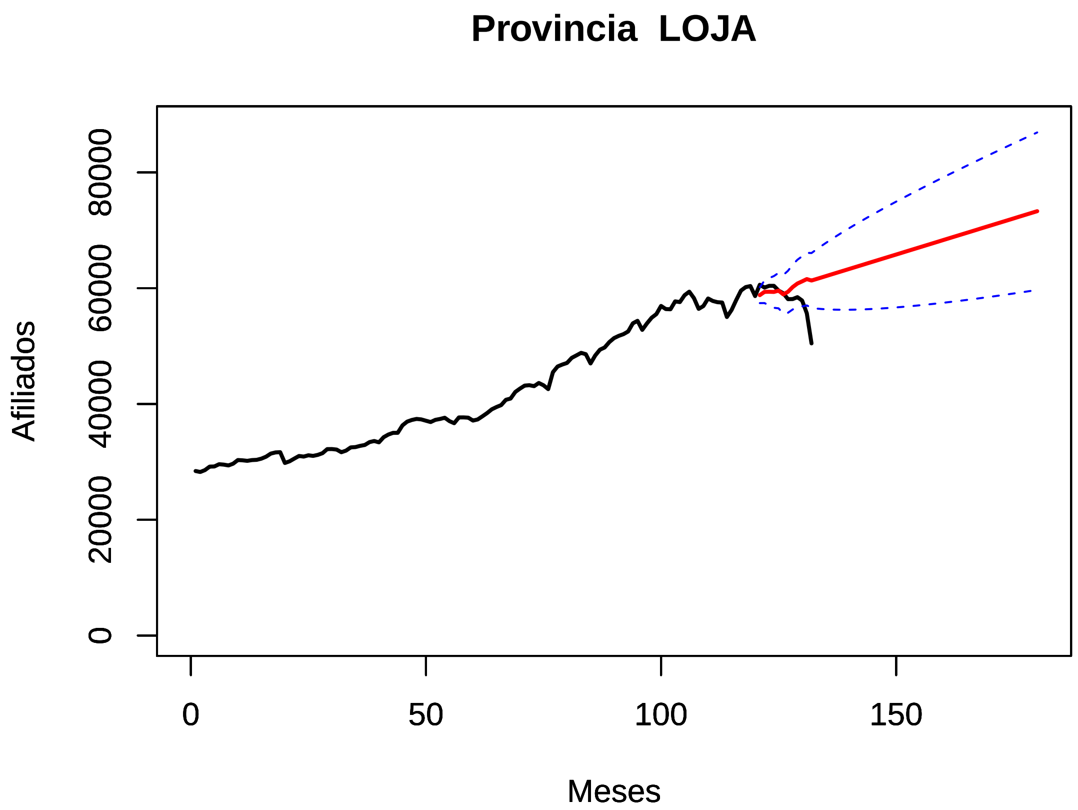
Q−Q Plot
Sample Quantiles
−40000
0
20000
−2
−1
0
1
2
Residuos



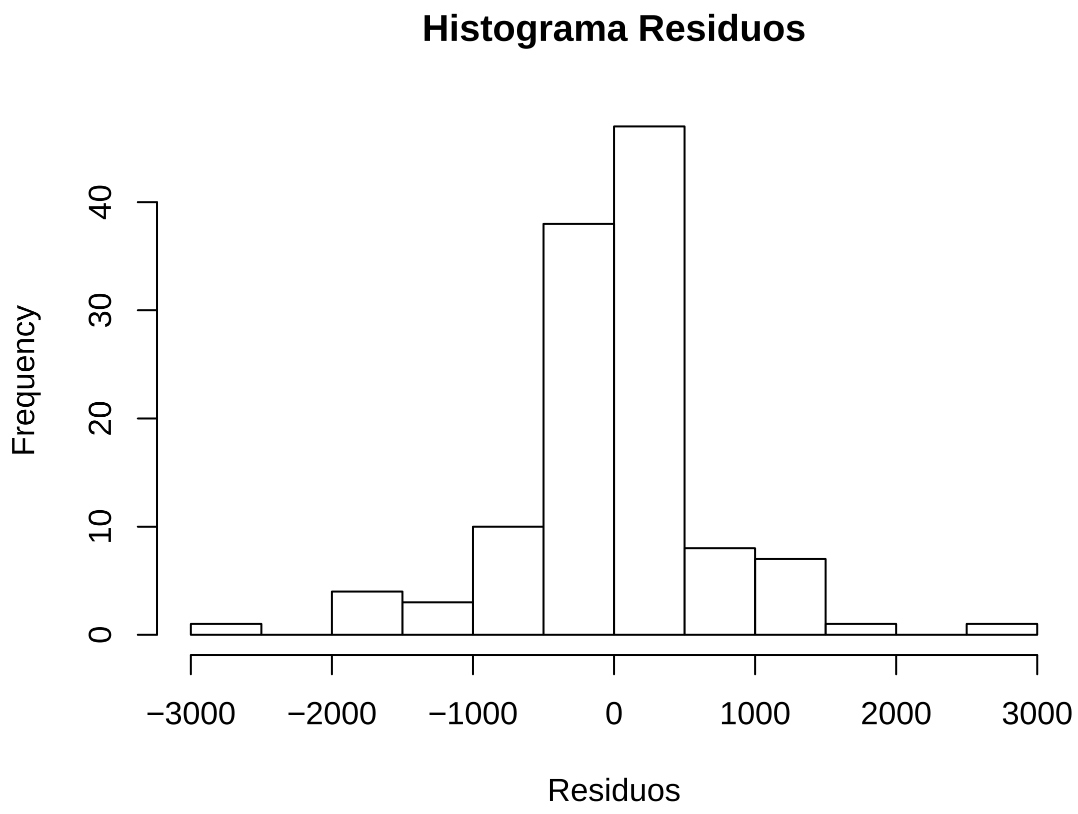

```
## Series: data[, i]
## ARIMA(3,2,2)(0,0,1)[12]
##
## Coefficients:
##          ar1      ar2      ar3      ma1     ma2    sma1
##       0.3792  -0.2286  -0.0757  -1.3271  0.4057  0.1456
## s.e.  0.3454   0.1043   0.1443   0.3338  0.2950  0.1346
##
## sigma^2 estimated as 906215:  log likelihood=-978.02
## AIC=1970.05   AICc=1971.06   BIC=1989.44
```


Histograma Residuos
Frequency
0
10
20
30
40
50
60
−2000
0
2000
4000
Residuos

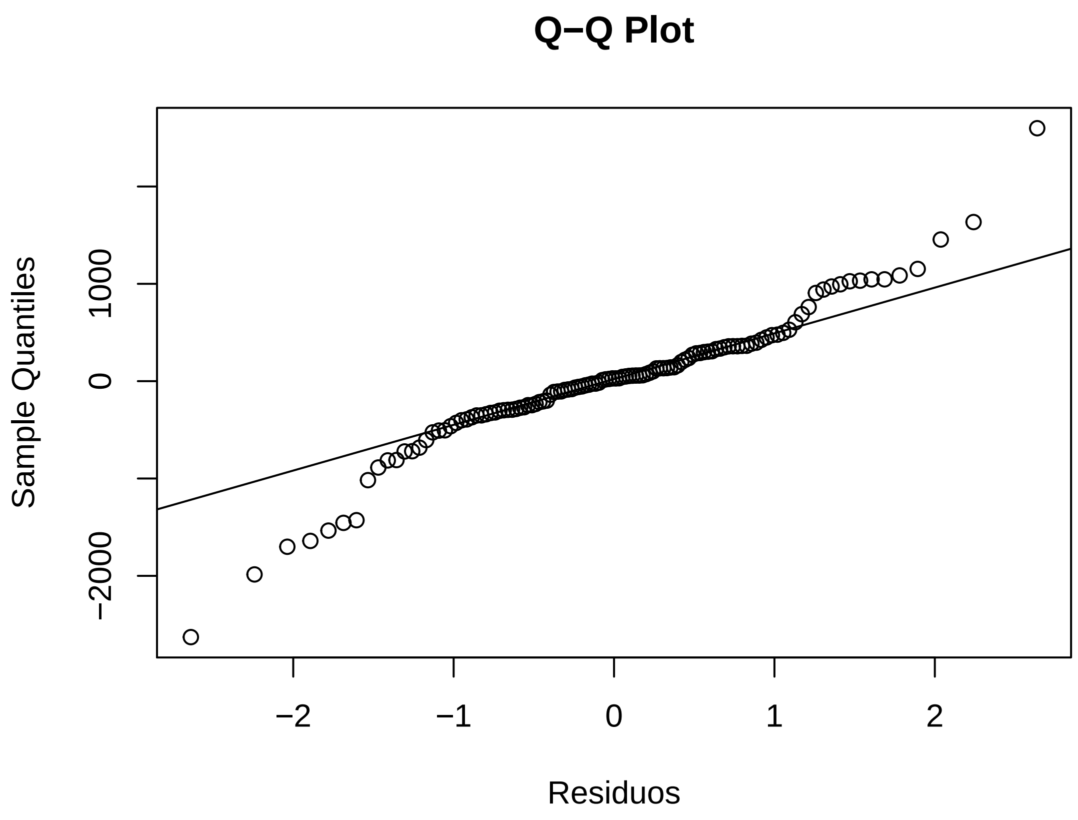

Q−Q Plot
Sample Quantiles
−3000
−1000
1000
3000
−2
−1
0
1
2
Residuos



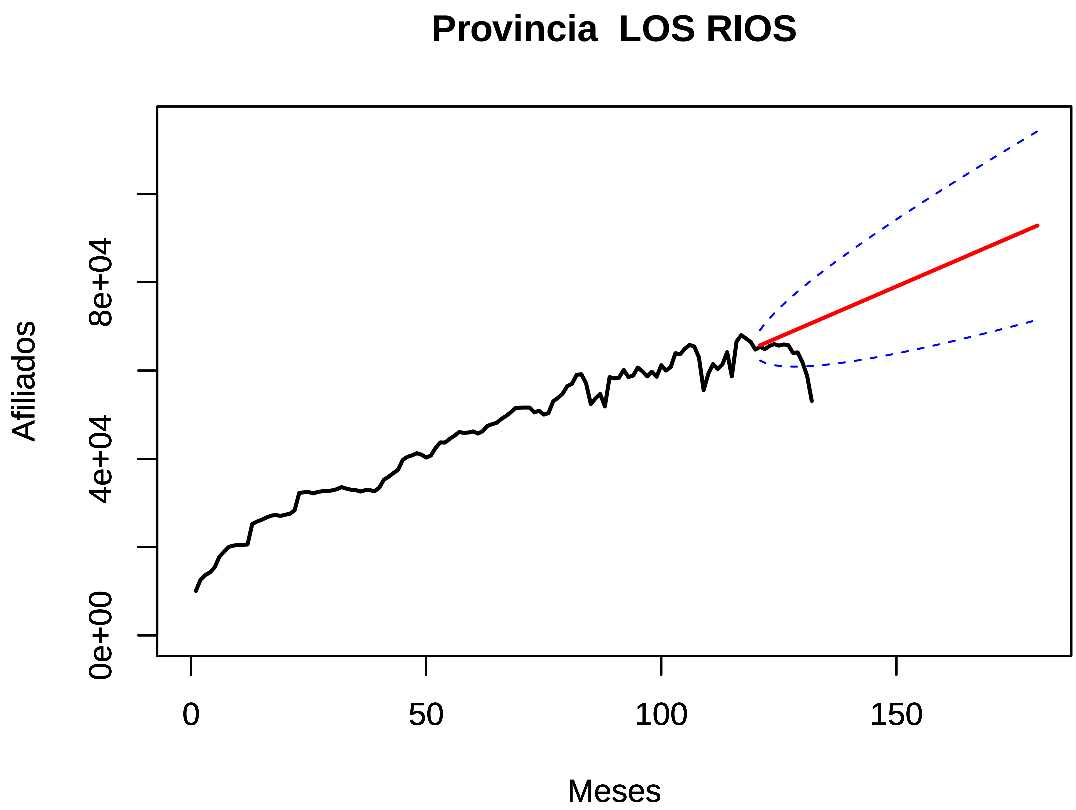

```
## Series: data[, i]
## ARIMA(0,1,0)(0,0,1)[12] with drift
##
## Coefficients:
##          sma1     drift
##        0.3438  249.5408
## s.e.   0.0923   83.9328
##
## sigma^2 estimated as 488130:  log likelihood=-948.96
## AIC=1903.92   AICc=1904.13   BIC=1912.26
```



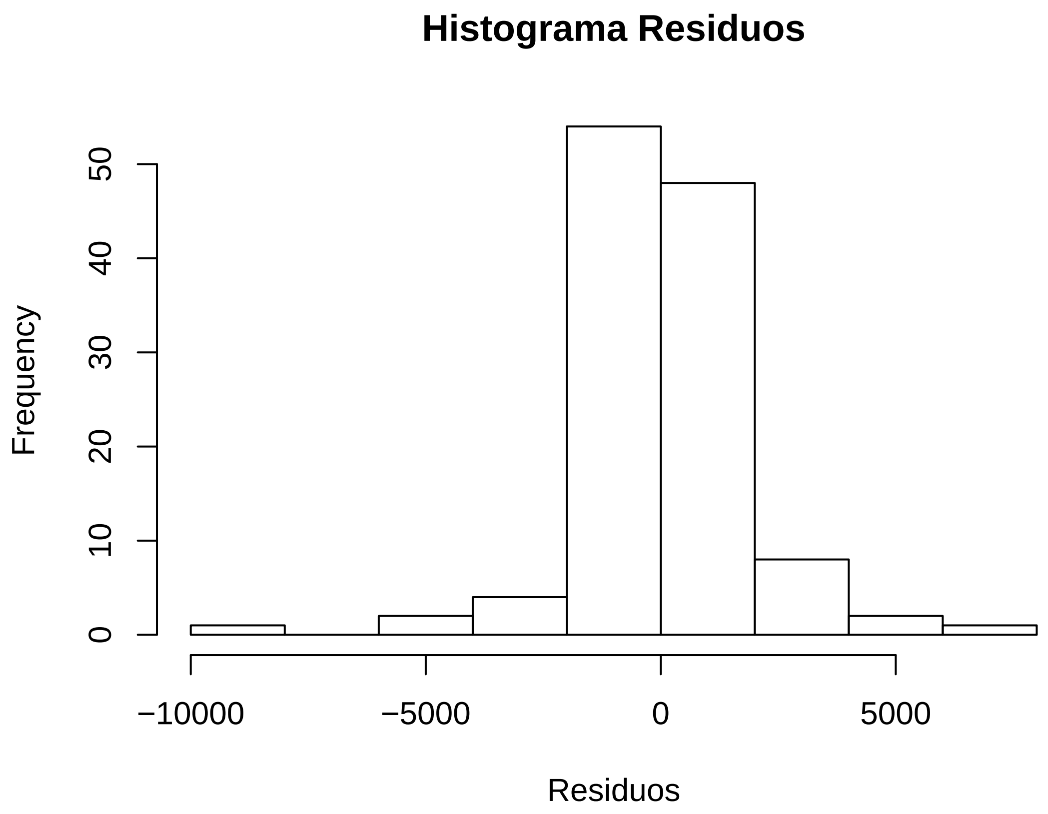
Histograma Residuos
Frequency
0
10
20
30
40
−3000
−2000
−1000
0
1000
2000
3000
Residuos

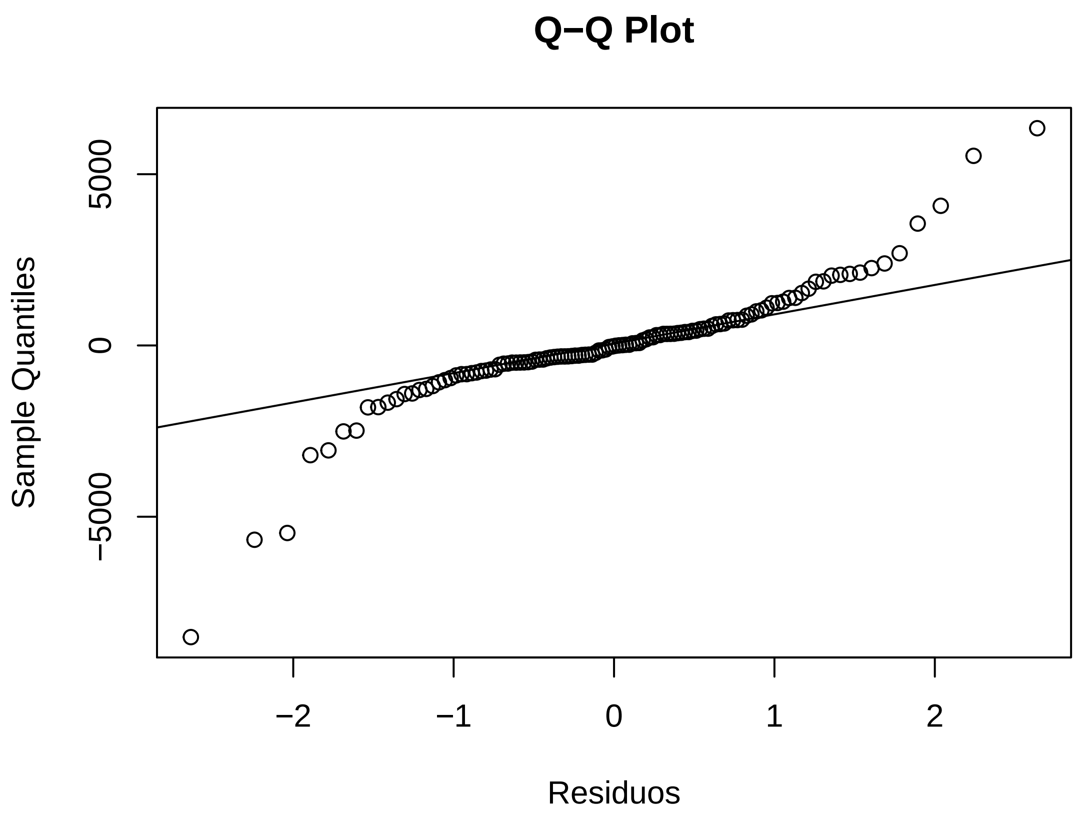


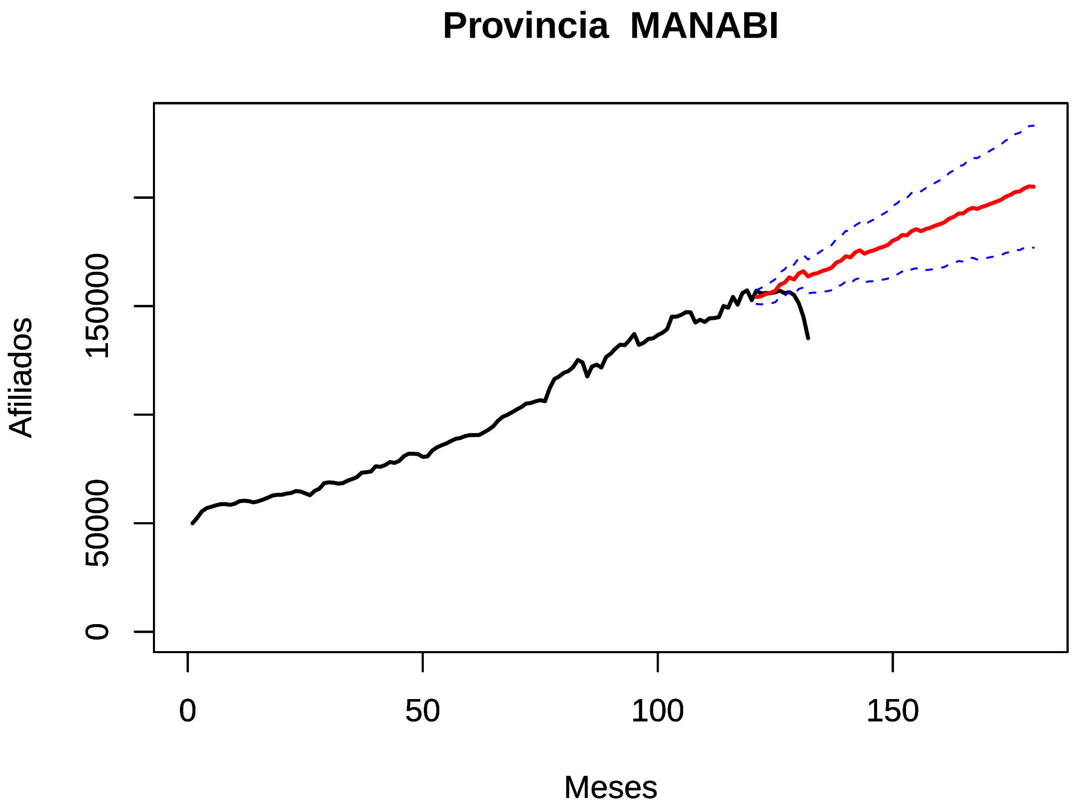
Q−Q Plot
Sample Quantiles
−2000
0
1000
−2
−1
0
1
2
Residuos



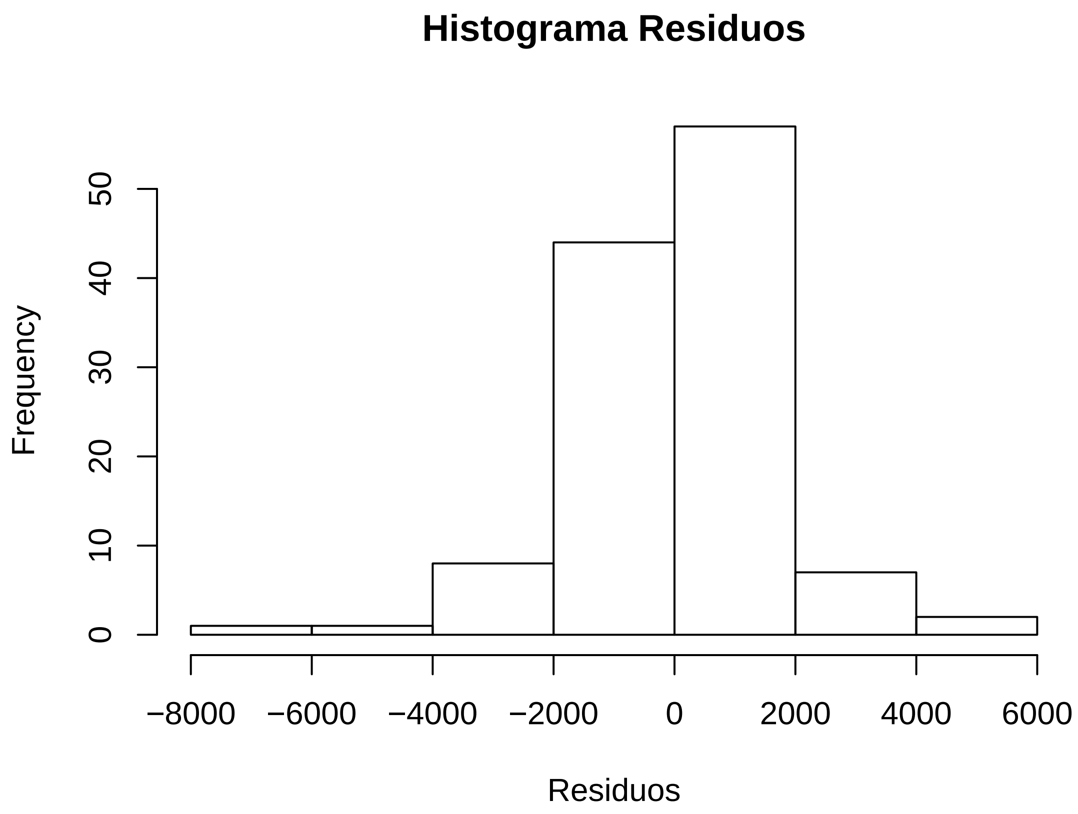

```
## Series: data[, i]
## ARIMA(0,1,1) with drift
##
## Coefficients:
##           ma1     drift
##       -0.2029  459.9023
## s.e.   0.0977  128.6499
##
## sigma^2 estimated as 3086349:  log likelihood=-1057.95
## AIC=2121.91   AICc=2122.12   BIC=2130.24
```

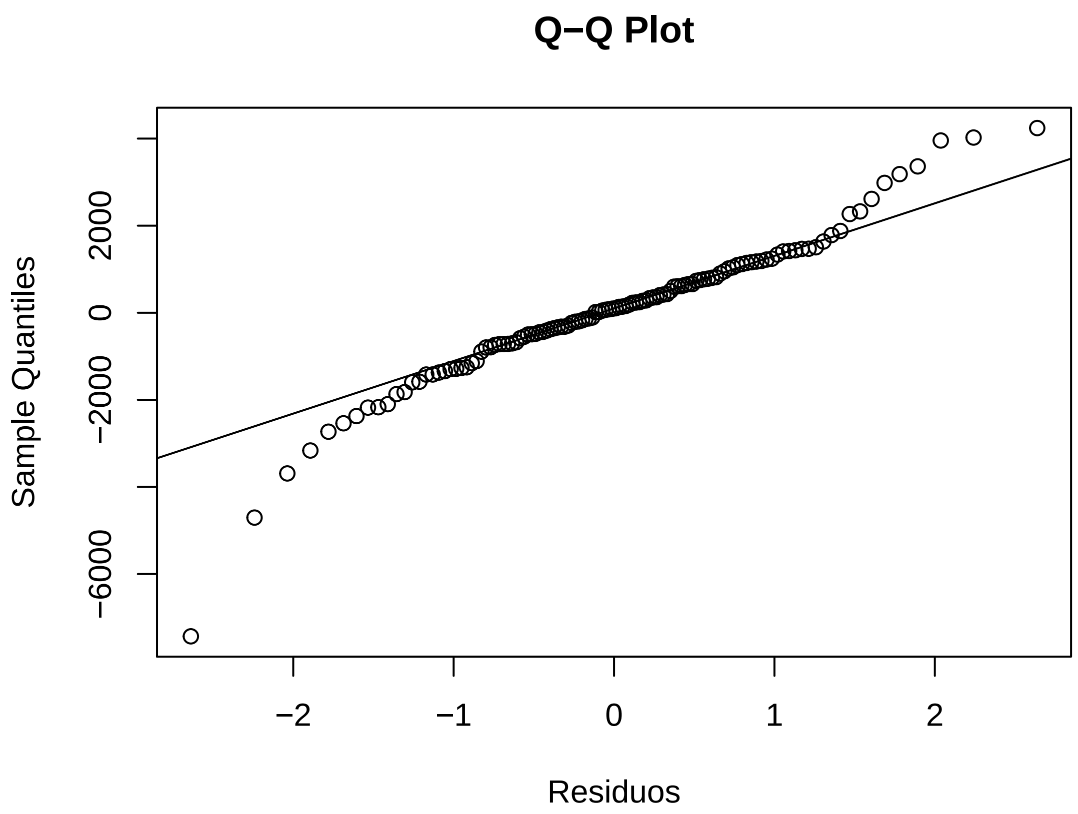

Histograma Residuos
Frequency
0
10
20
30
40
50
−10000
−5000
0
5000
Residuos

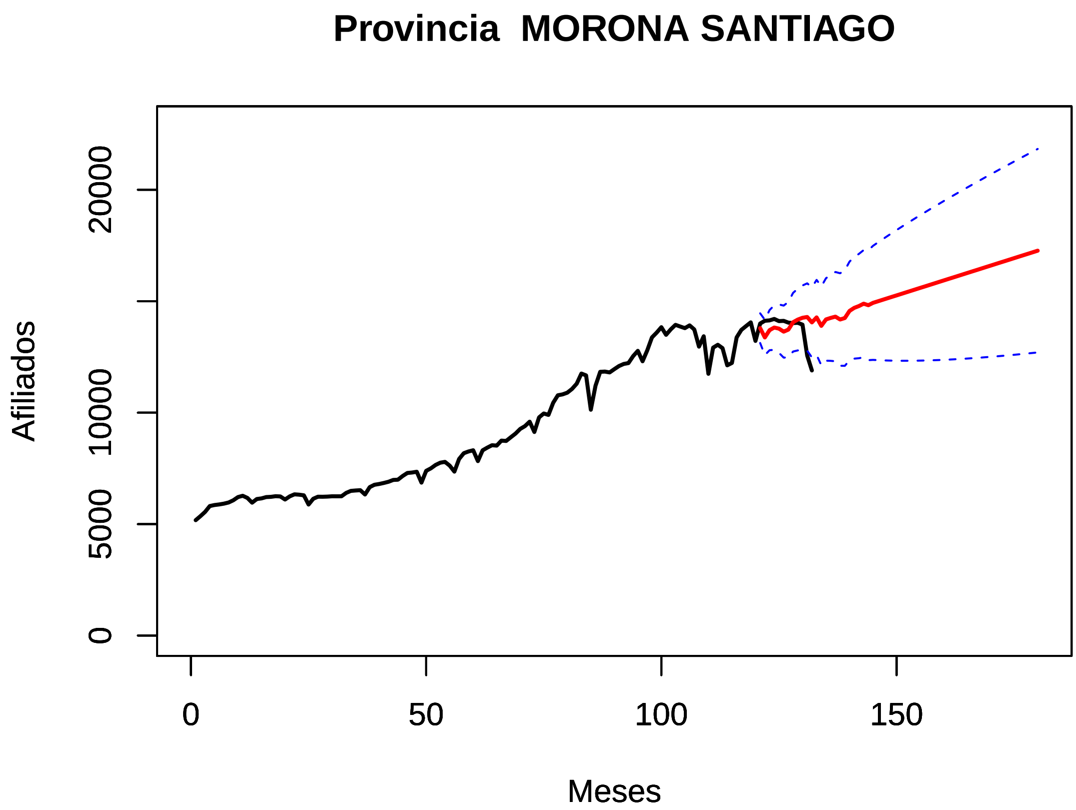

Q−Q Plot
Sample Quantiles
−5000
0
5000
−2
−1
0
1
2
Residuos



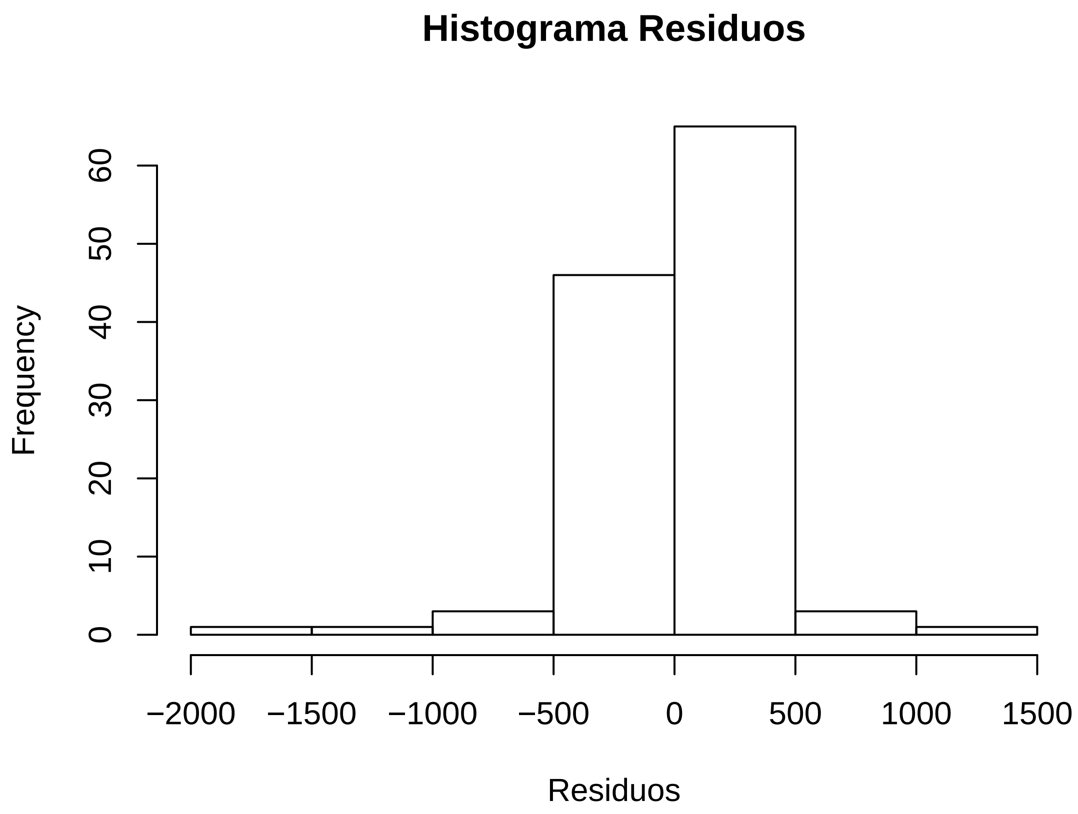

```
## Series: data[, i]
## ARIMA(2,1,0)(1,0,1)[12] with drift
##
## Coefficients:
##           ar1      ar2    sar1     sma1     drift
##       -0.3627  -0.1184  0.7409  -0.3345  836.3650
## s.e.   0.0976   0.0942  0.1345   0.1844  213.9424
##
## sigma^2 estimated as 2644289:  log likelihood=-1050.87
## AIC=2113.75   AICc=2114.5   BIC=2130.42
```

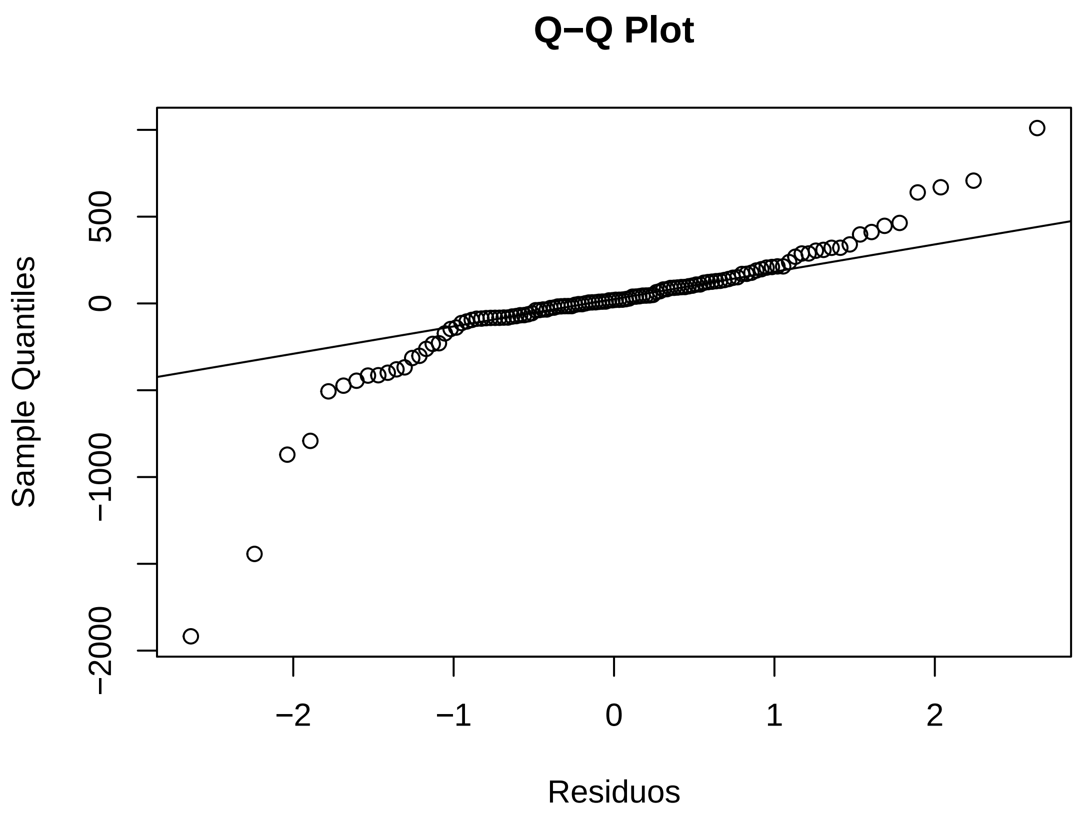

Histograma Residuos
Frequency
0
10
20
30
40
50
−8000
−6000
−4000
−2000
0
2000
4000
6000
Residuos

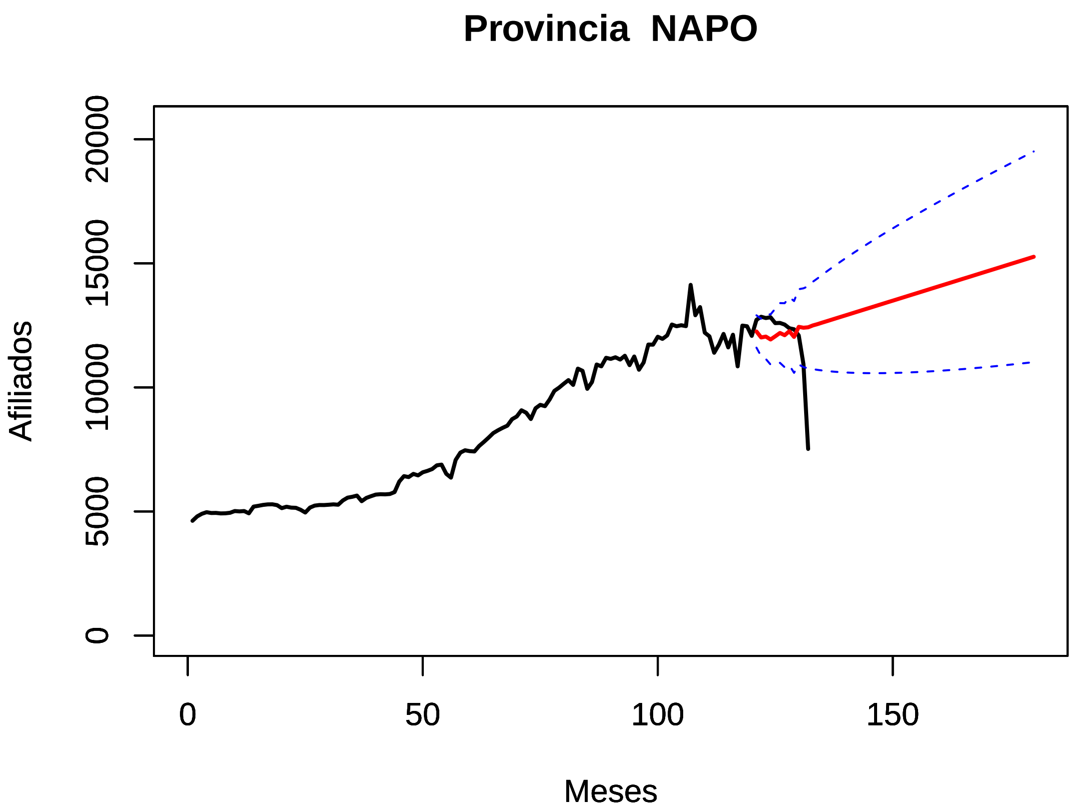

Q−Q Plot
Sample Quantiles
−6000
−2000
0
2000
−2
−1
0
1
2
Residuos



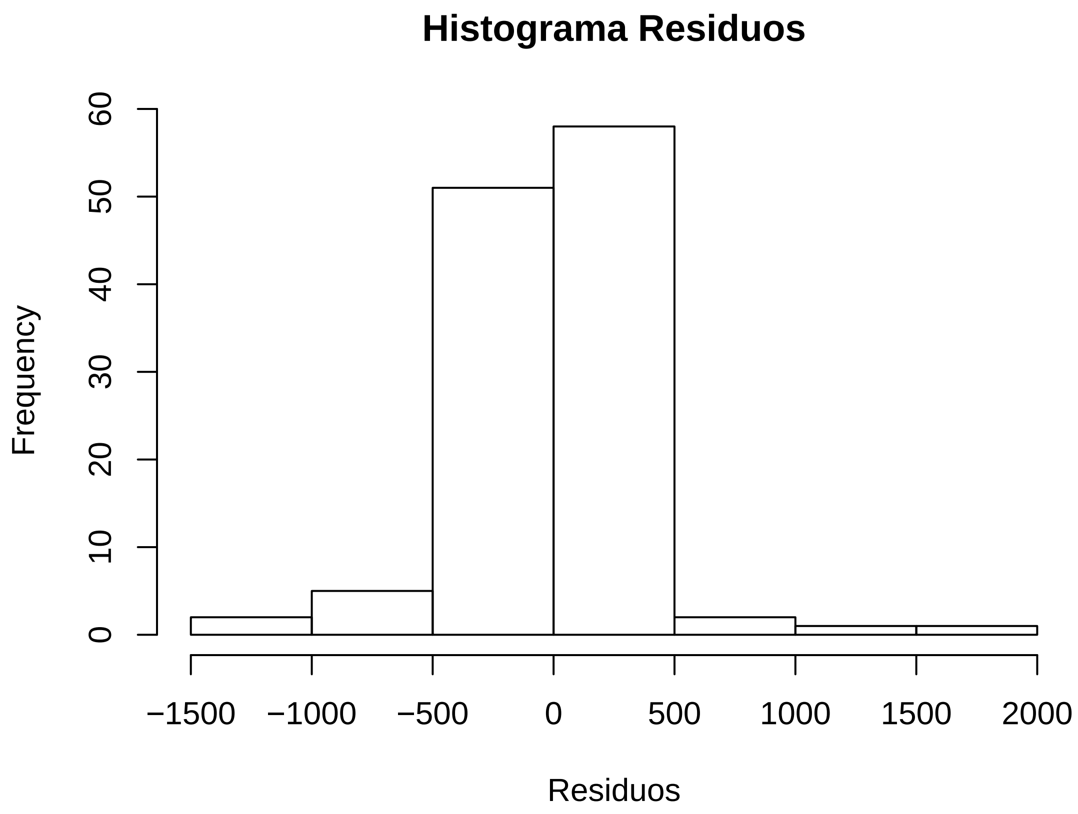

```
## Series: data[, i]
## ARIMA(0,1,1)(0,0,2)[12] with drift
##
## Coefficients:
##           ma1    sma1    sma2    drift
##       -0.3558  0.2654  0.2249  66.8423
## s.e.   0.0950  0.1185  0.1262  28.6697
##
## sigma^2 estimated as 115319:  log likelihood=-863.33
## AIC=1736.66   AICc=1737.2   BIC=1750.56
```

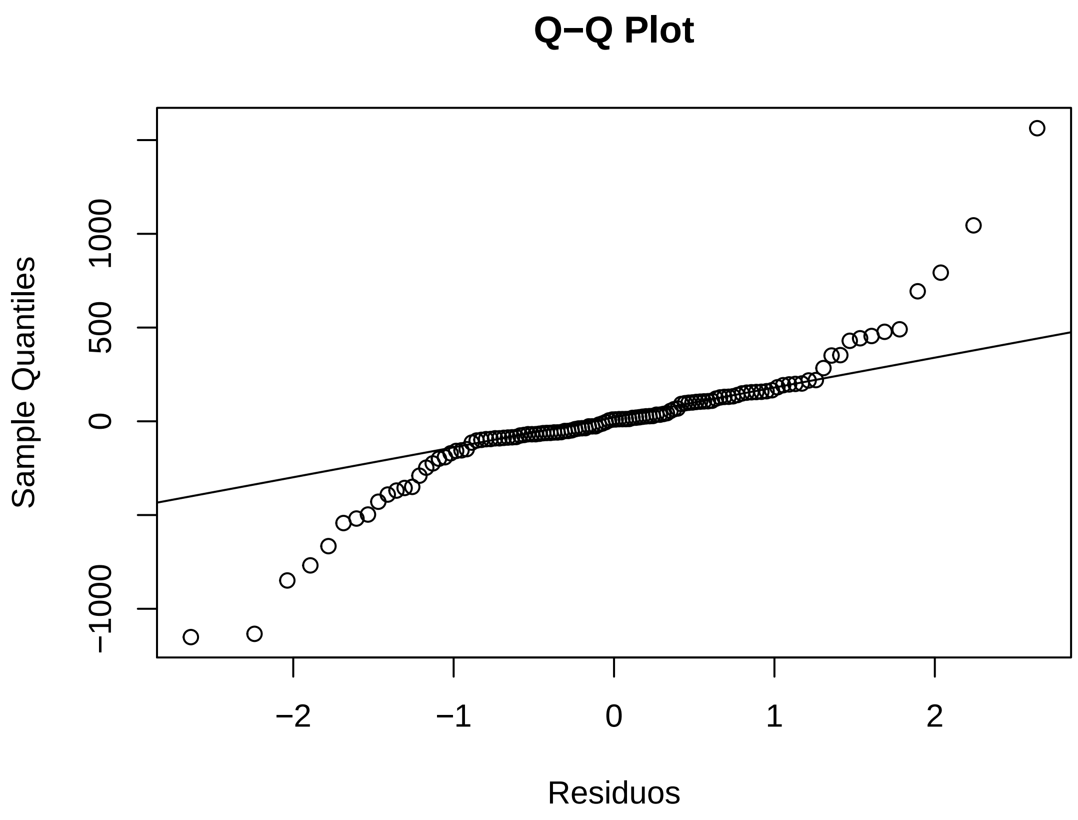

Histograma Residuos
Frequency
0
10
20
30
40
50
60
−2000
−1500
−1000
−500
0
500
1000
1500
Residuos

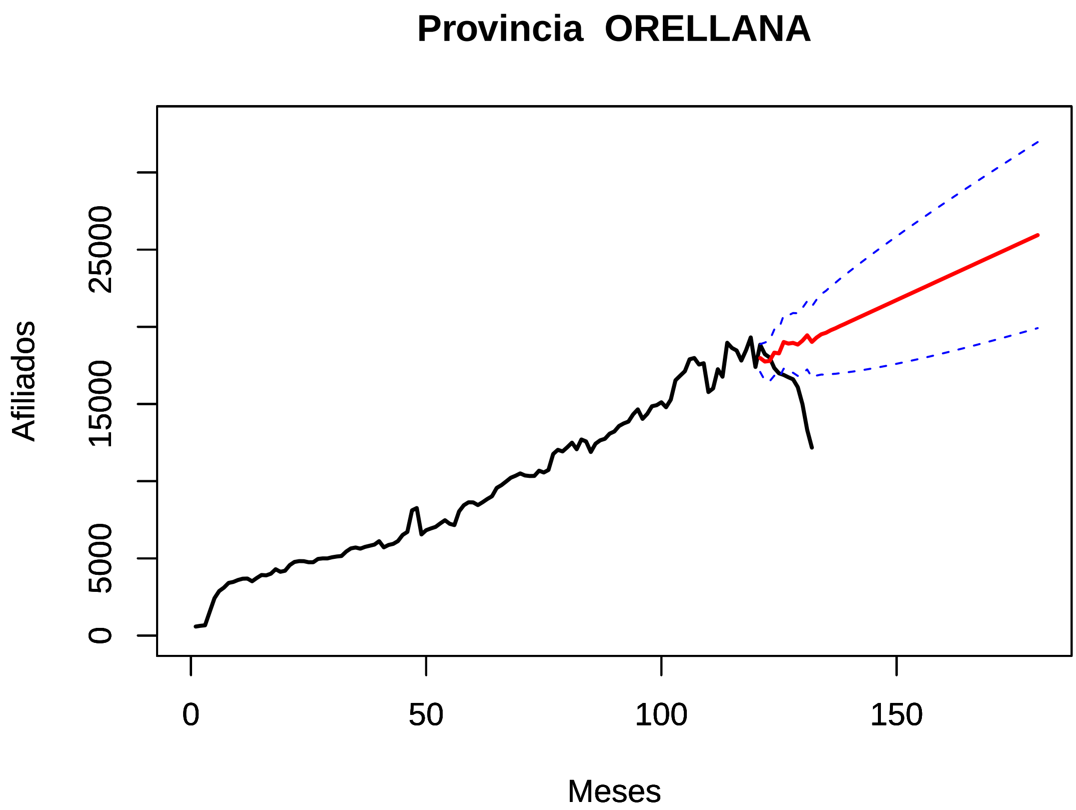

Q−Q Plot
Sample Quantiles
−2000
−1000
0
500
−2
−1
0
1
2
Residuos



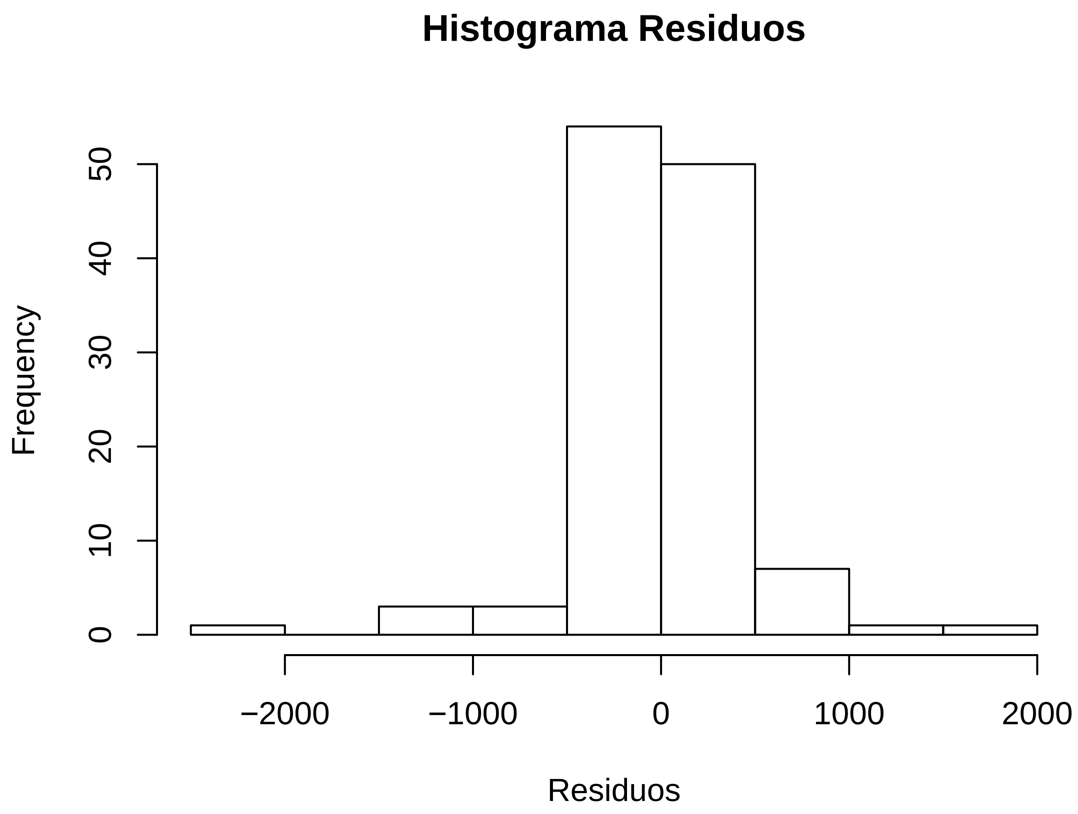

```
## Series: data[, i]
## ARIMA(1,1,0)(0,0,1)[12] with drift
##
## Coefficients:
##           ar1    sma1    drift
##       -0.4089  0.2191  58.9439
## s.e.   0.0833  0.1059  26.0632
##
## sigma^2 estimated as 110432:  log likelihood=-860.16
## AIC=1728.33   AICc=1728.68   BIC=1739.44
```

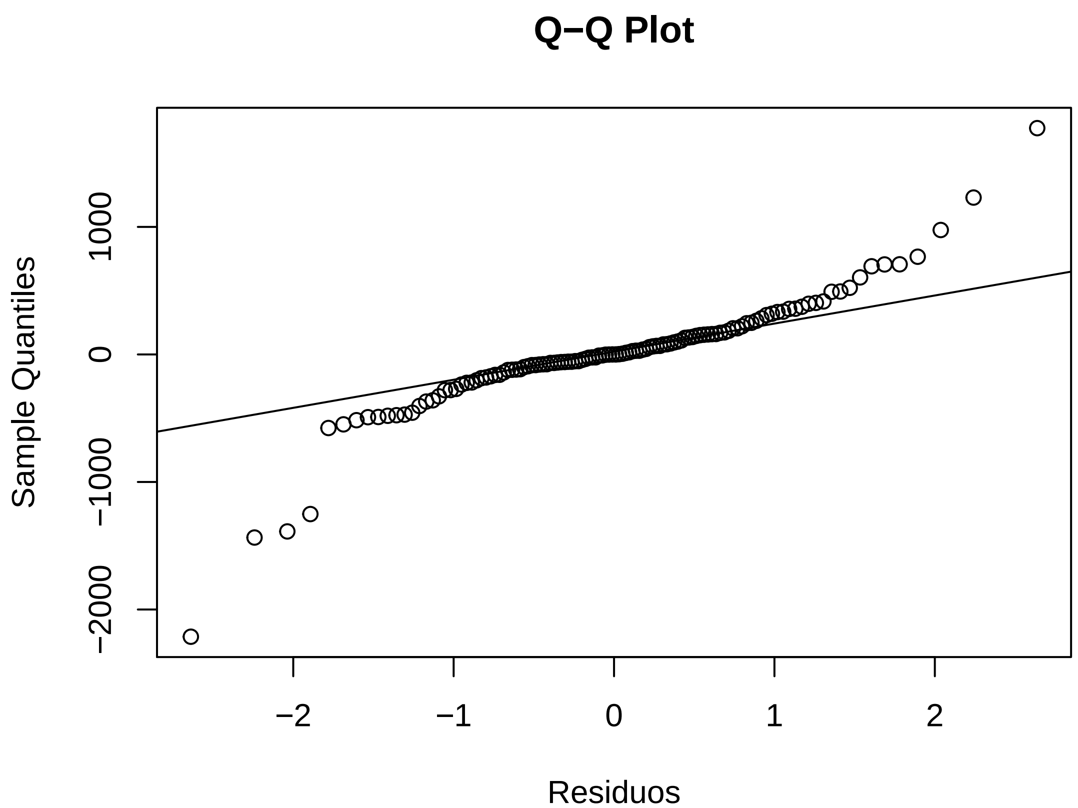

Histograma Residuos
Frequency
0
10
20
30
40
50
60
−1500
−1000
−500
0
500
1000
1500
2000
Residuos

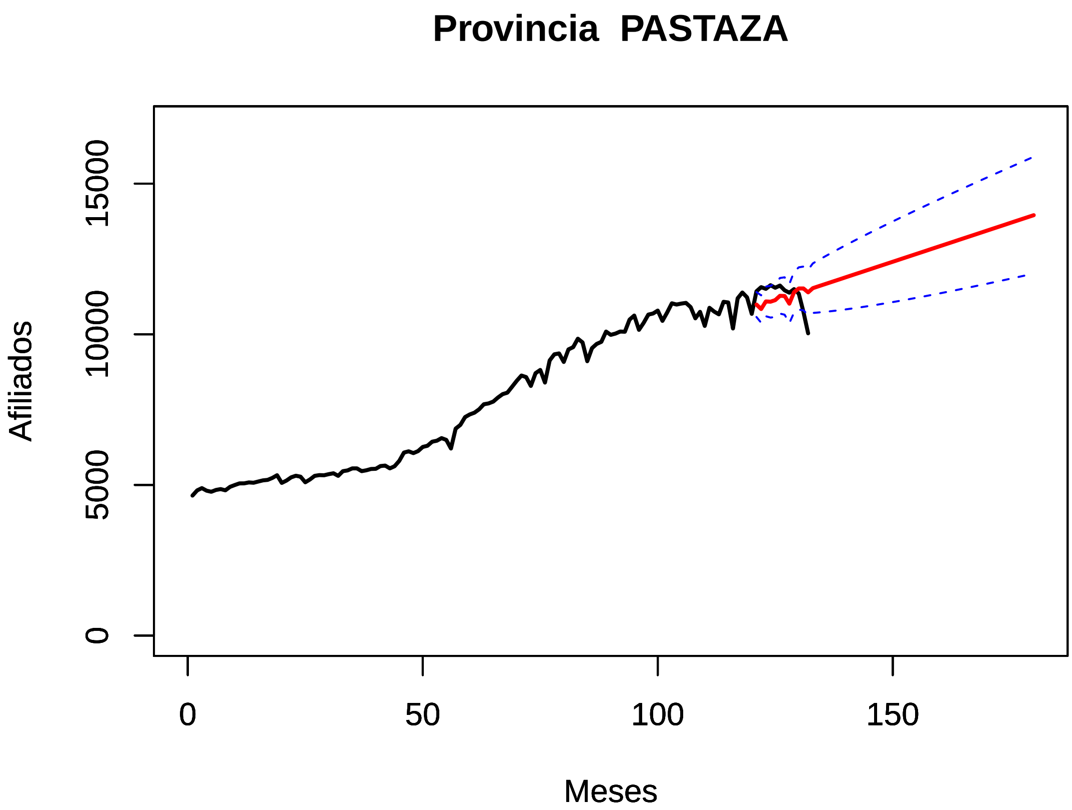

Q−Q Plot
Sample Quantiles
−1000
0
500
1000
−2
−1
0
1
2
Residuos



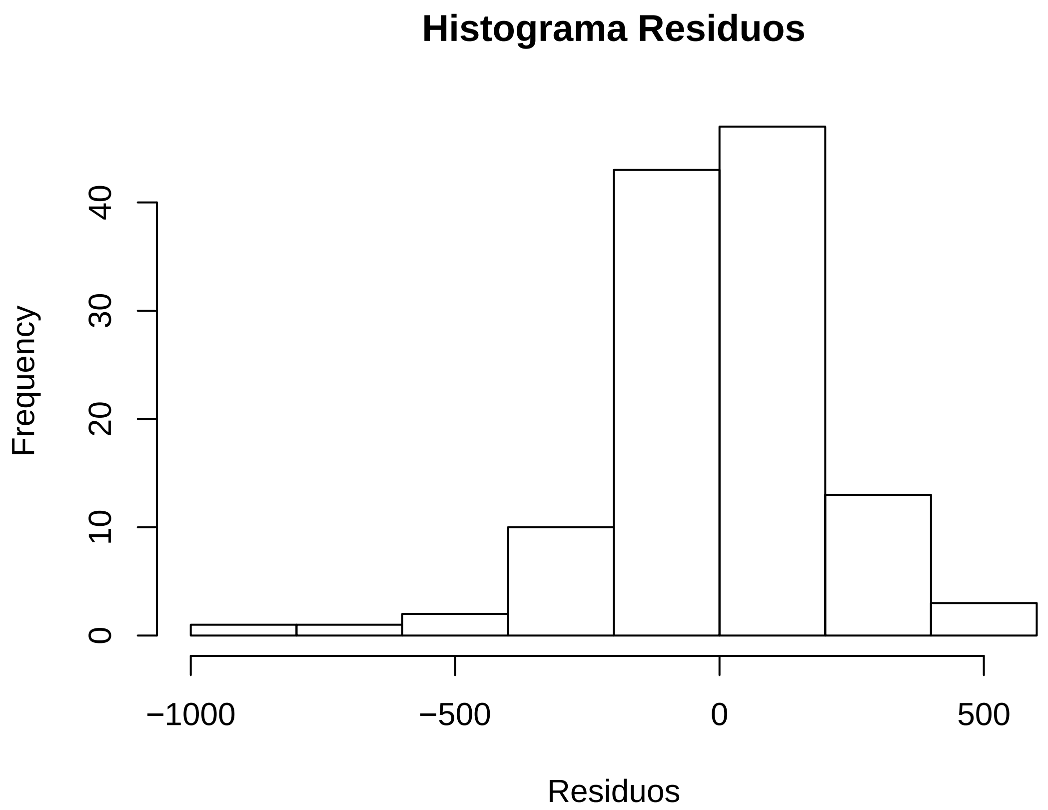

```
## Series: data[, i]
## ARIMA(1,1,2)(0,0,1)[12] with drift
##
## Coefficients:
##           ar1     ma1      ma2    sma1     drift
##       -0.5778  0.5061  -0.4189  0.2924  140.2573
## s.e.   0.1183  0.1450   0.1269  0.1349   36.9276
##
## sigma^2 estimated as 210776:  log likelihood=-898.24
## AIC=1795.96   AICc=1796.71   BIC=1812.63
```

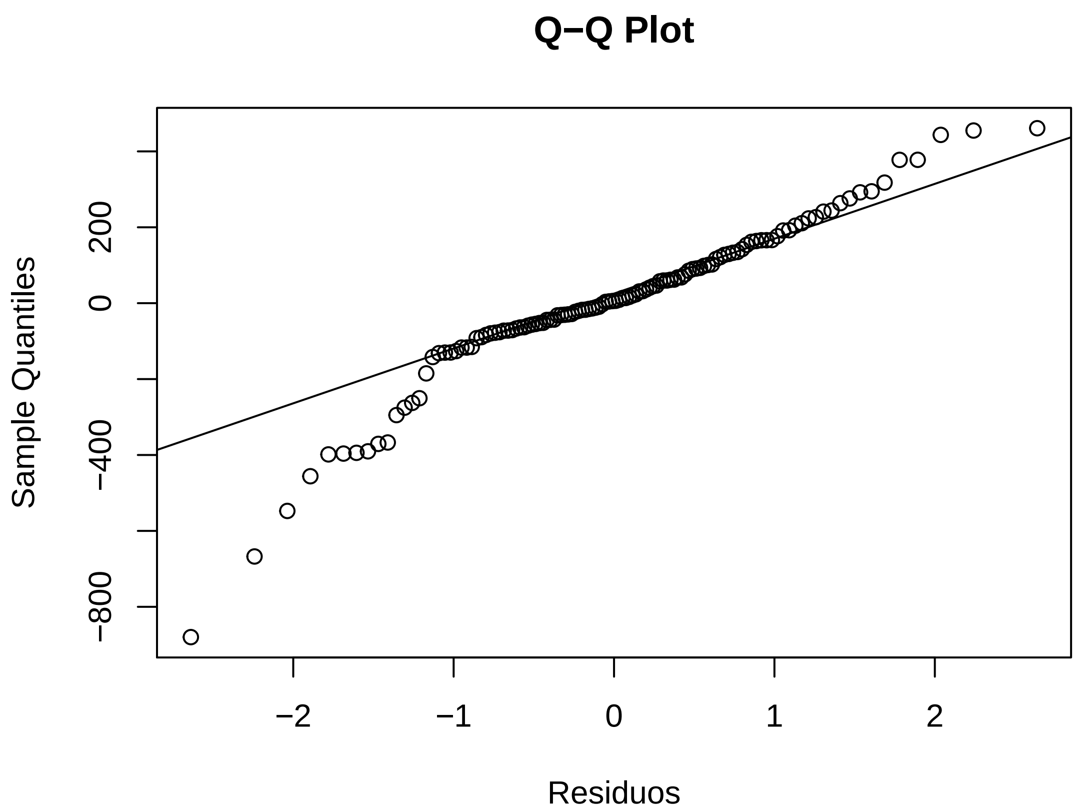

Histograma Residuos
Frequency
0
10
20
30
40
50
−2000
−1000
0
1000
2000
Residuos

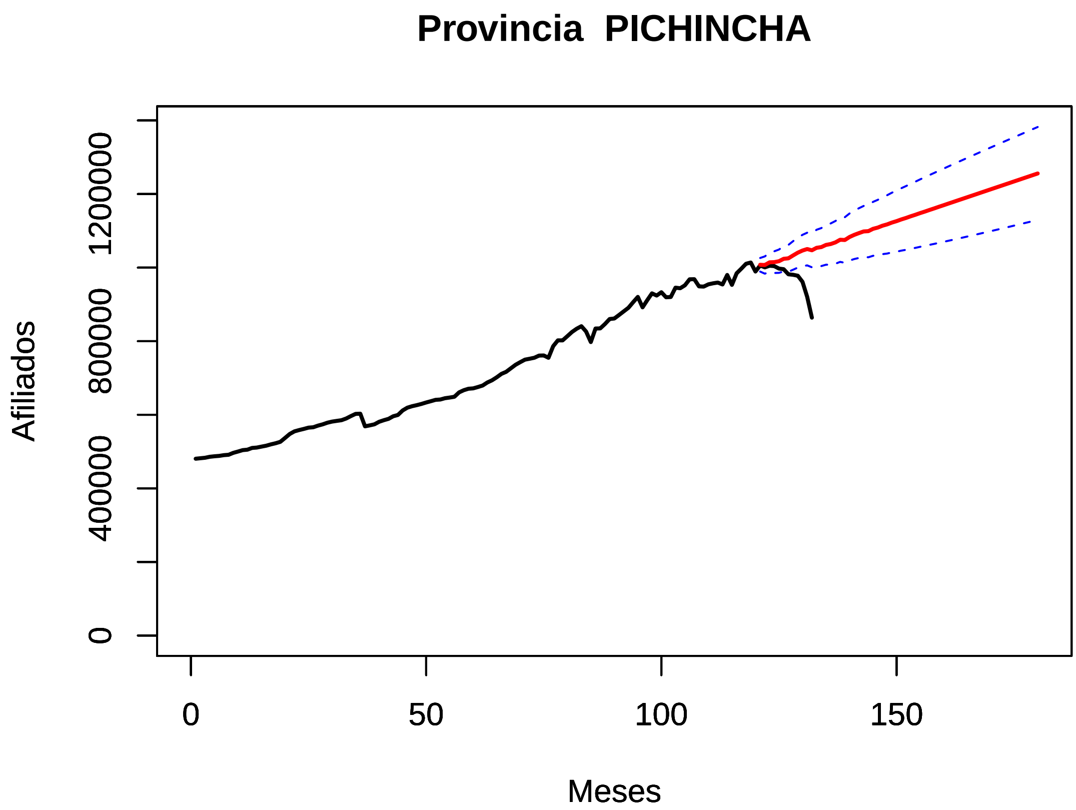

Q−Q Plot
Sample Quantiles
−2000
−1000
0
1000
−2
−1
0
1
2
Residuos



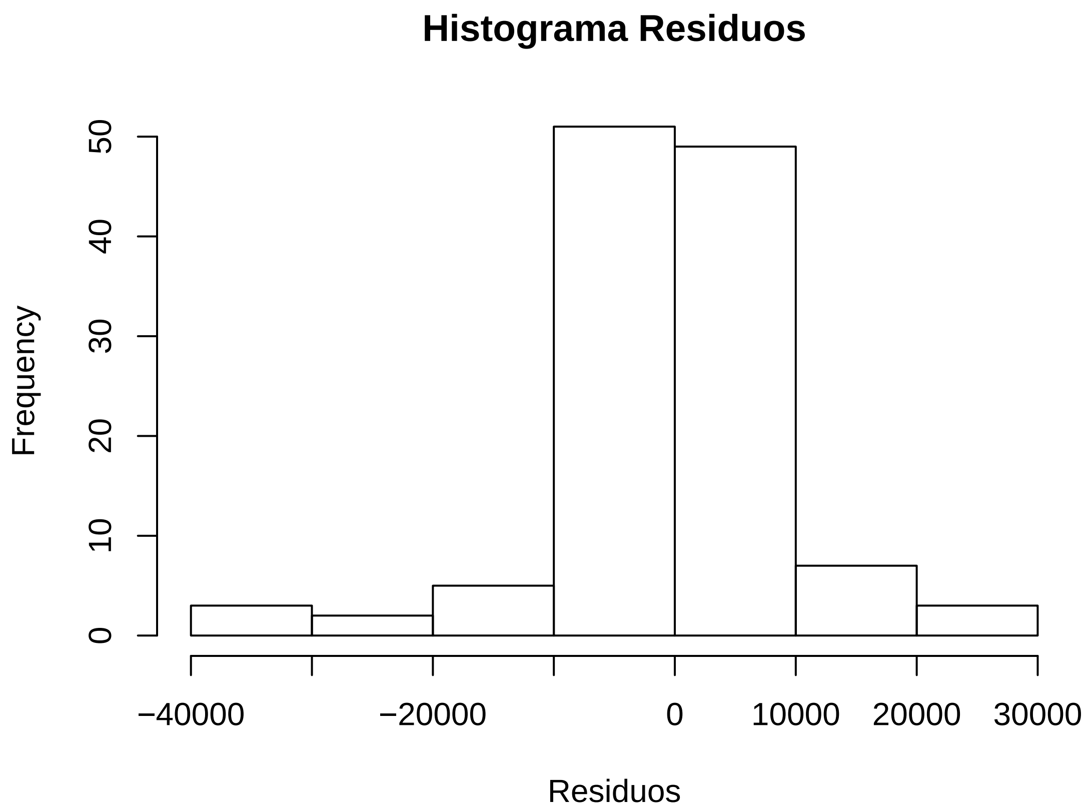

```
## Series: data[, i]
## ARIMA(0,1,1)(0,0,1)[12] with drift
##
## Coefficients:
##           ma1    sma1    drift
##       -0.5429  0.3556  51.5353
## s.e.   0.0761  0.1051  11.7900
##
## sigma^2 estimated as 44455:  log likelihood=-806.62
## AIC=1621.25   AICc=1621.6   BIC=1632.36
```

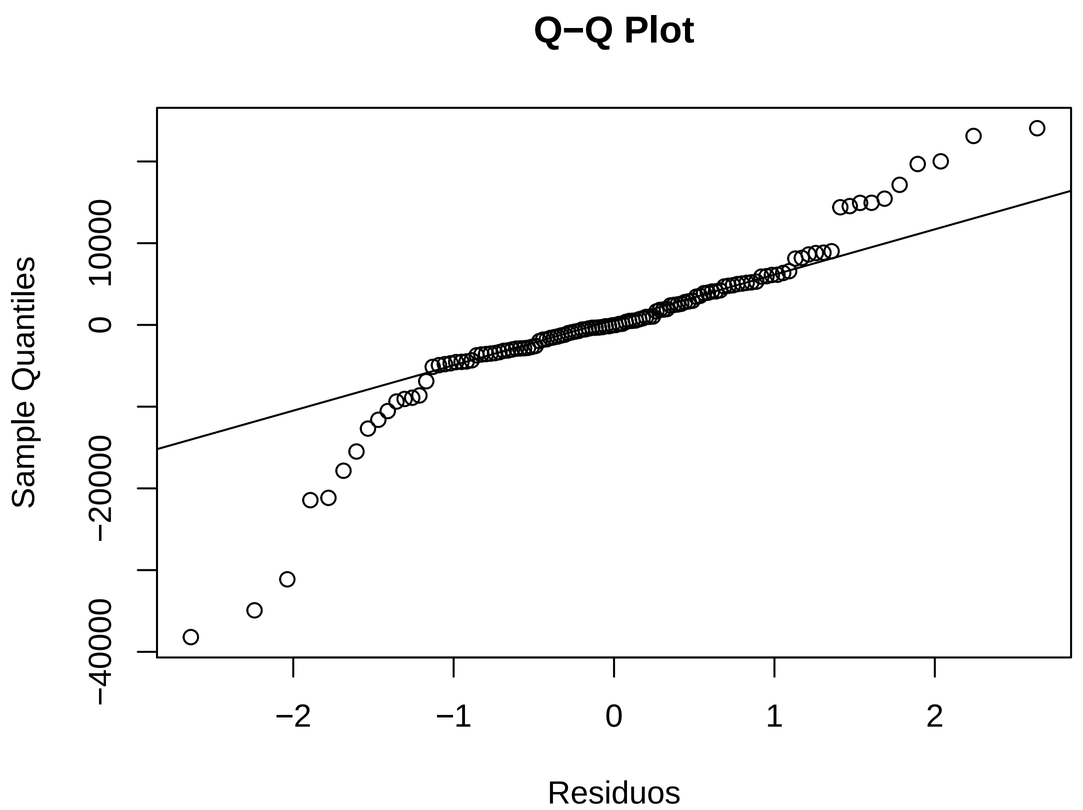

Histograma Residuos
Frequency
0
10
20
30
40
−1000
−500
0
500
Residuos

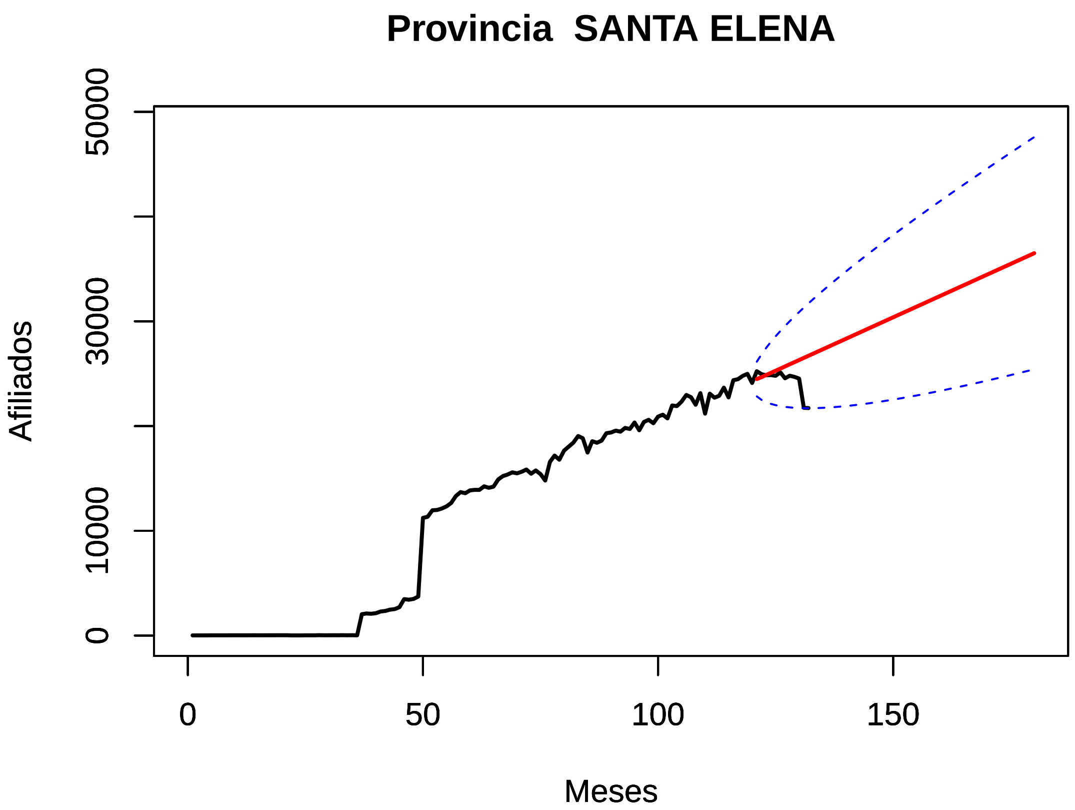

Q−Q Plot
Sample Quantiles
−800
−400
0
200
−2
−1
0
1
2
Residuos



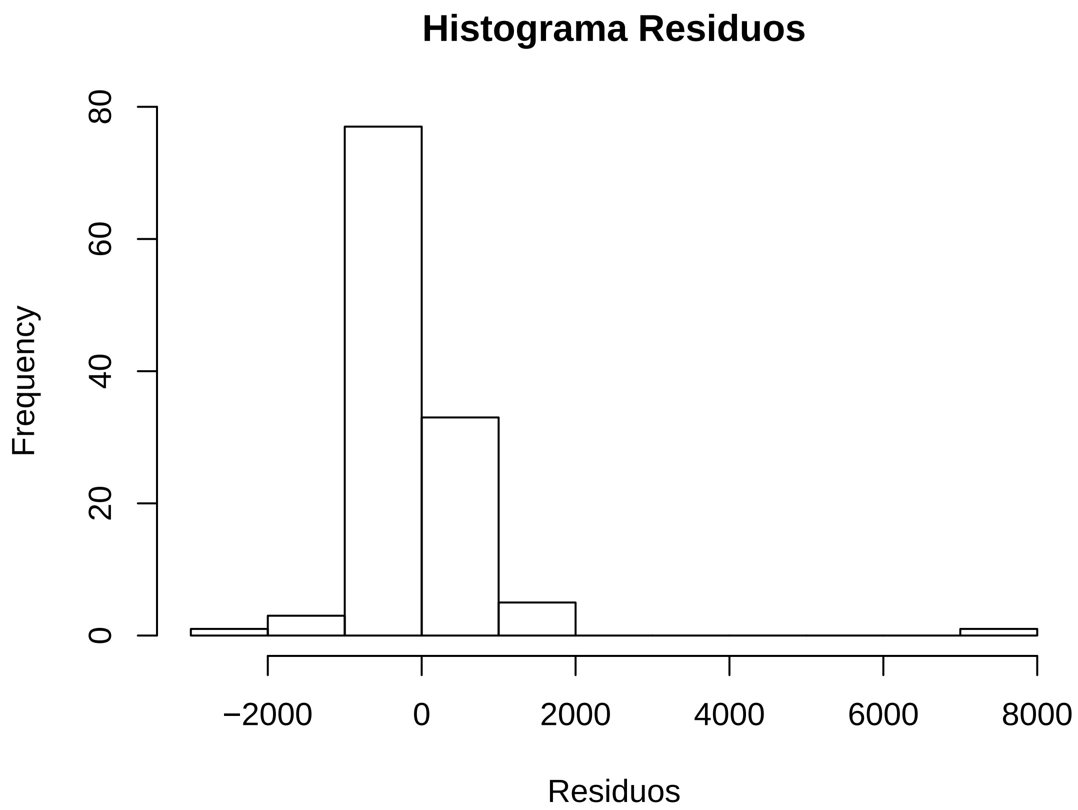

```
## Series: data[, i]
## ARIMA(1,1,2)(0,0,2)[12] with drift
##
## Coefficients:
##           ar1     ma1      ma2    sma1    sma2      drift
##       -0.8265  0.6085  -0.3601  0.2209  0.1534  4313.3934
## s.e.   0.0708  0.1030   0.0948  0.1022  0.0890   786.5182
##
## sigma^2 estimated as 88992256:  log likelihood=-1258.76
## AIC=2531.51   AICc=2532.52   BIC=2550.97
```

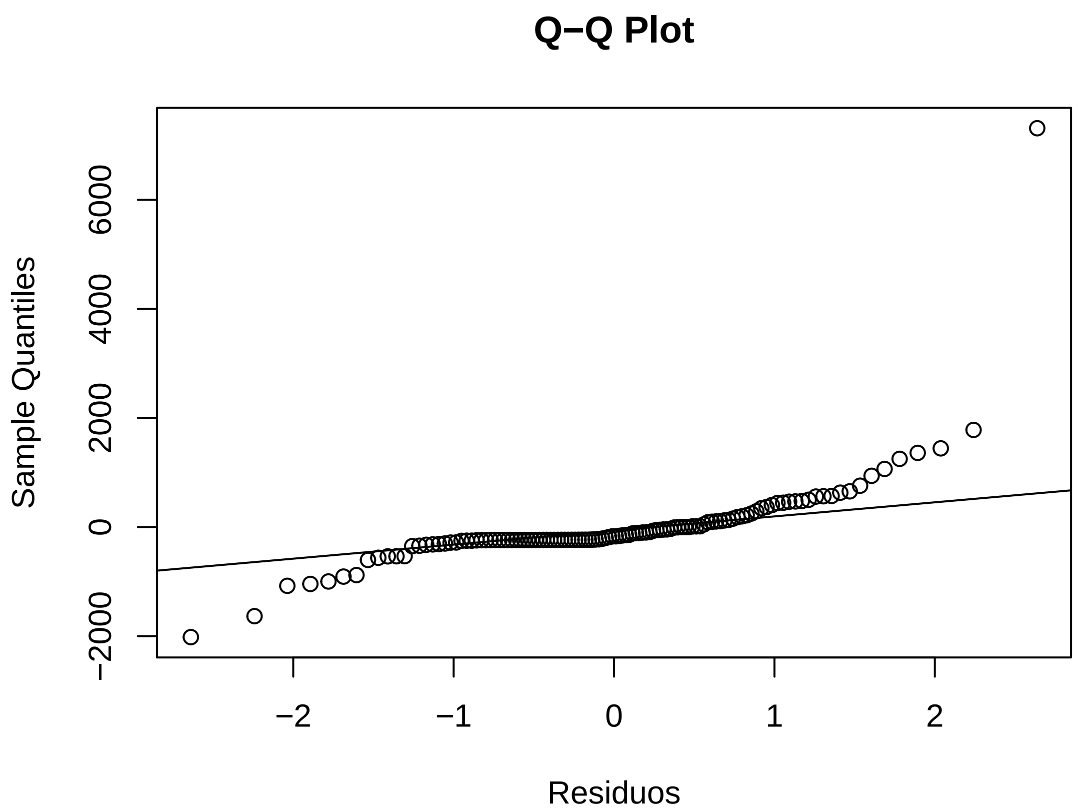

Histograma Residuos
Frequency
0
10
20
30
40
50
−40000
−20000
0
10000
20000
30000
Residuos

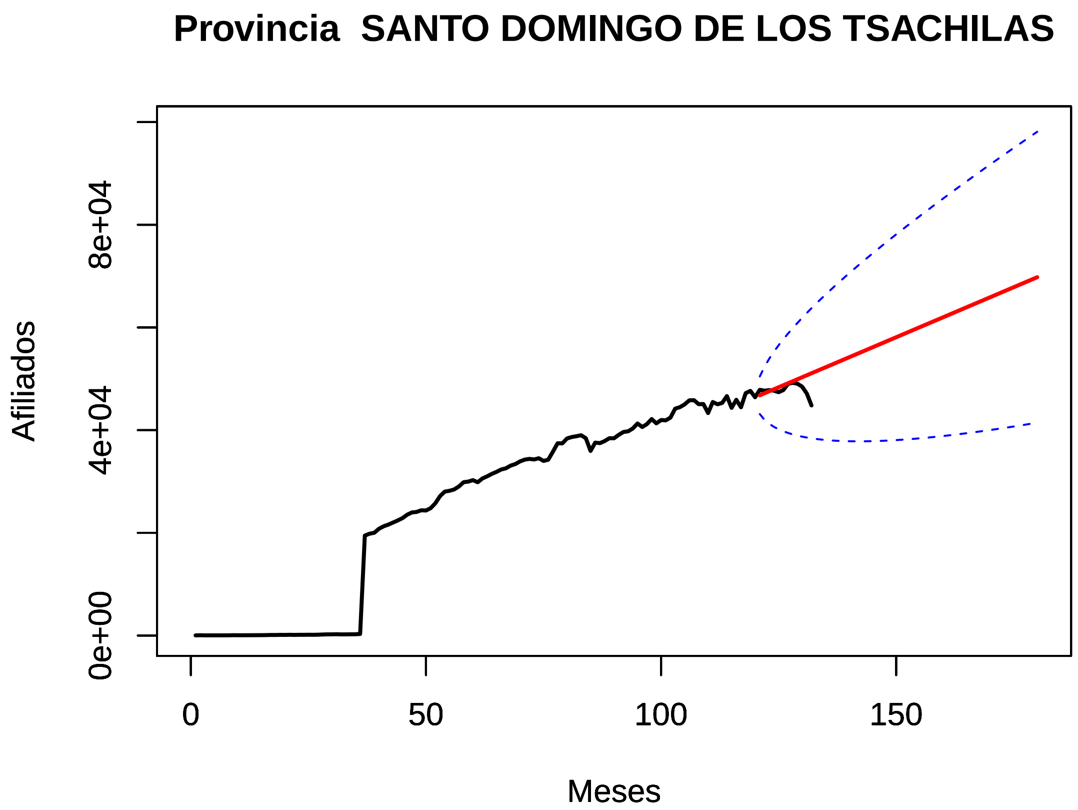

Q−Q Plot
Sample Quantiles
−40000
−20000
0
10000
−2
−1
0
1
2
Residuos



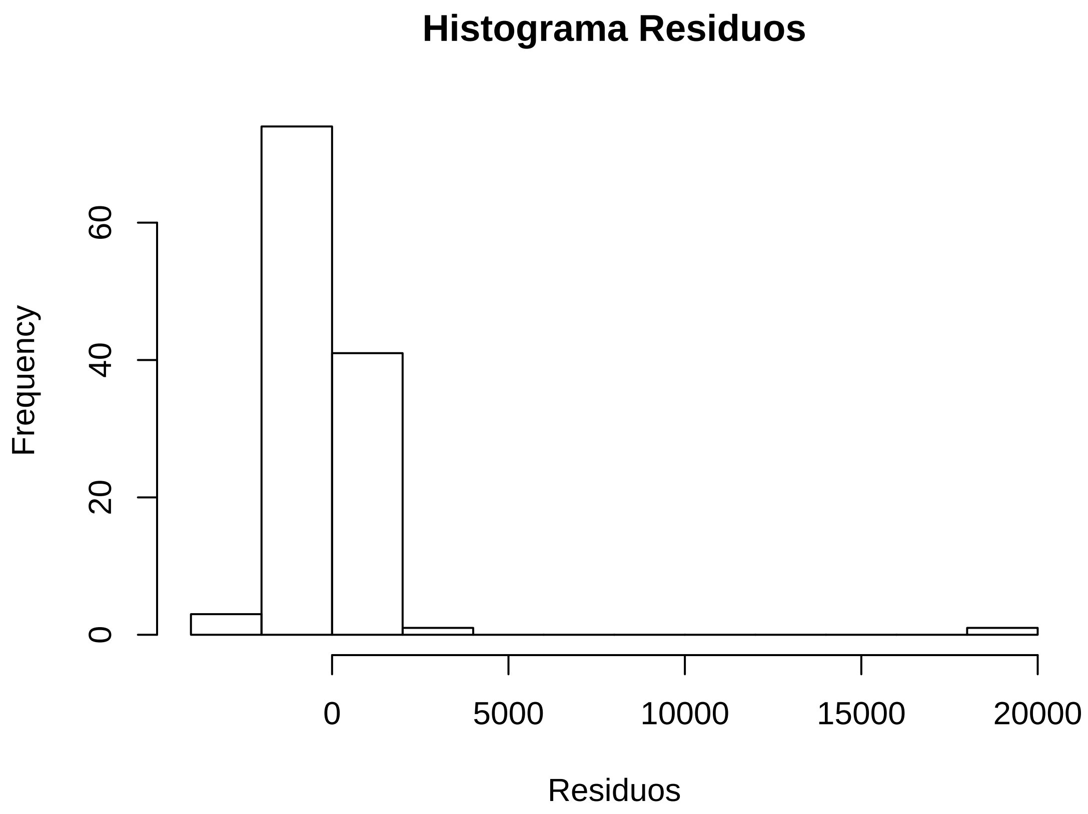

```
## Series: data[, i]
## ARIMA(1,1,0) with drift
##
## Coefficients:
##           ar1     drift
##       -0.1600  203.9975
## s.e.   0.0907   66.8343
##
## sigma^2 estimated as 713469:  log likelihood=-970.8
## AIC=1947.6   AICc=1947.81   BIC=1955.94
```

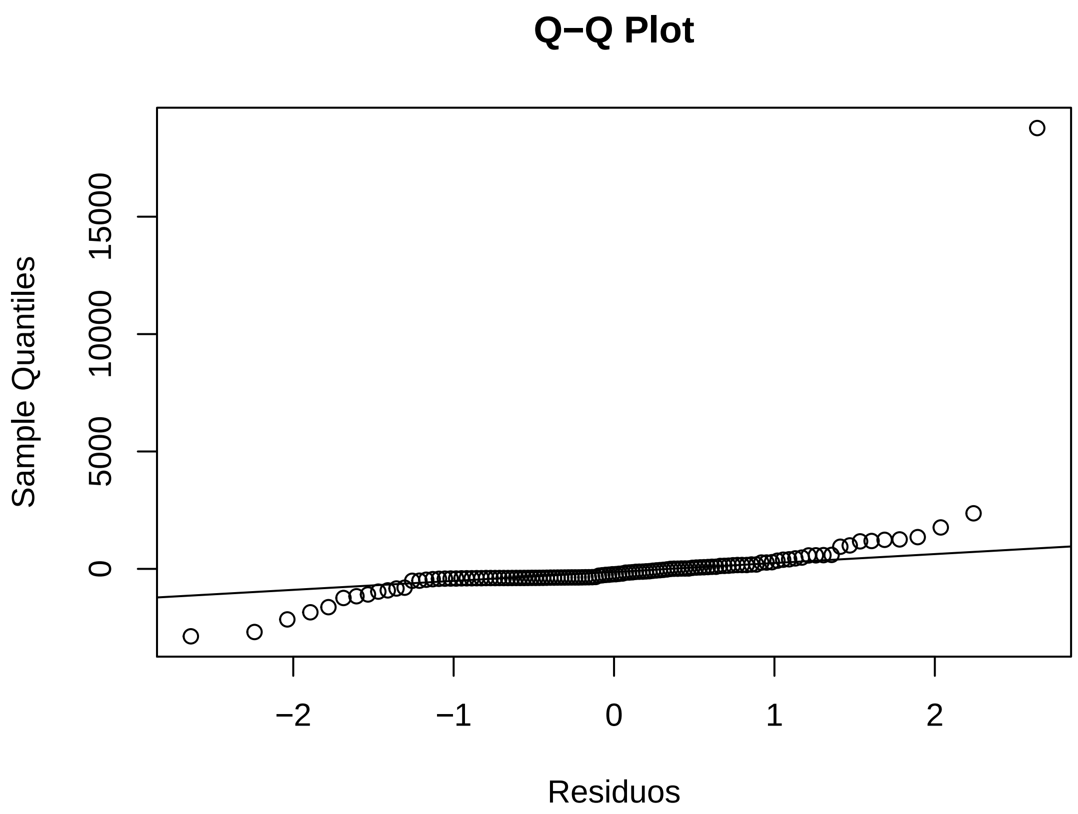

Histograma Residuos
Frequency
0
20
40
60
80
−2000
0
2000
4000
6000
8000
Residuos

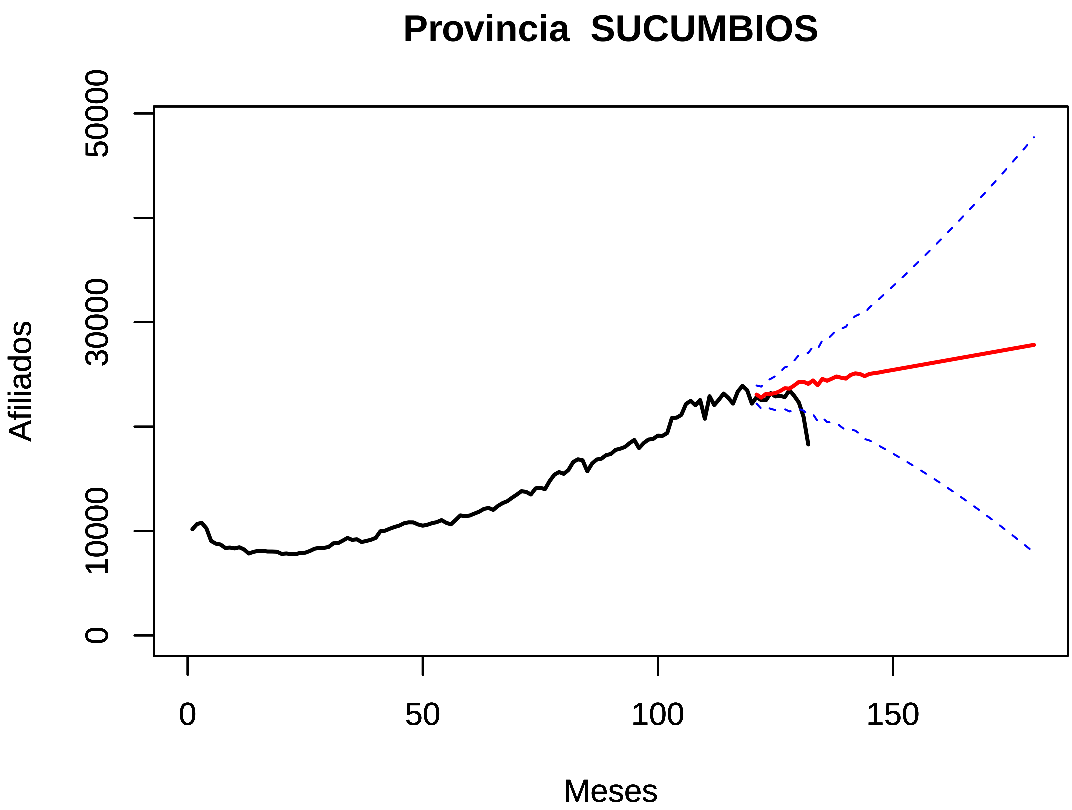

Q−Q Plot
Sample Quantiles
−2000
0
2000
4000
6000
−2
−1
0
1
2
Residuos



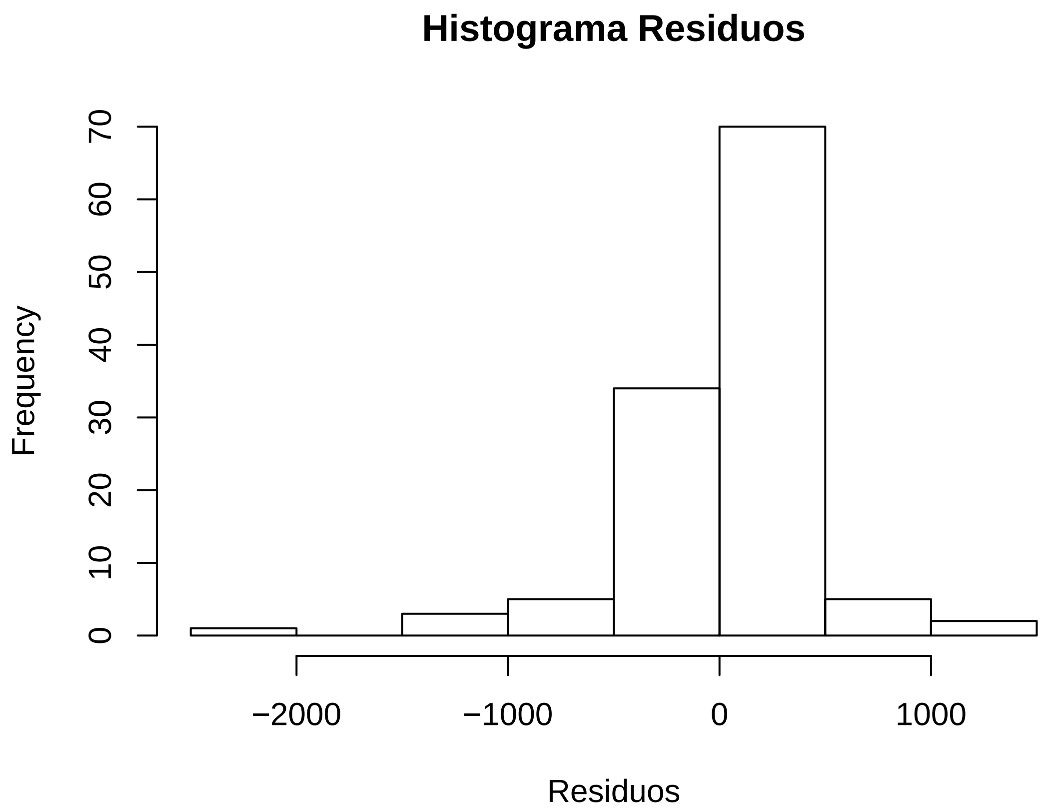

```
## Series: data[, i]
## ARIMA(0,1,0) with drift
##
## Coefficients:
##          drift
##       389.6050
## s.e.  171.0239
##
## sigma^2 estimated as 3480649:  log likelihood=-1065.09
## AIC=2134.17   AICc=2134.28   BIC=2139.73
```

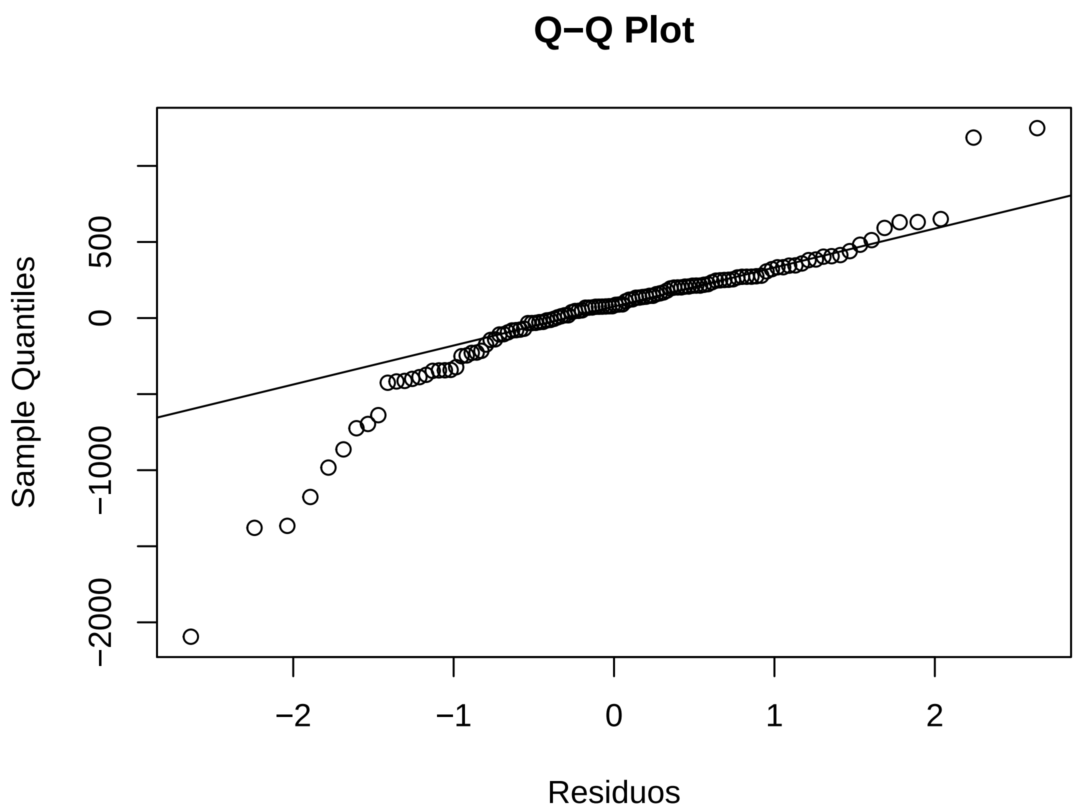

Histograma Residuos
Frequency
0
20
40
60
0
5000
10000
15000
20000
Residuos

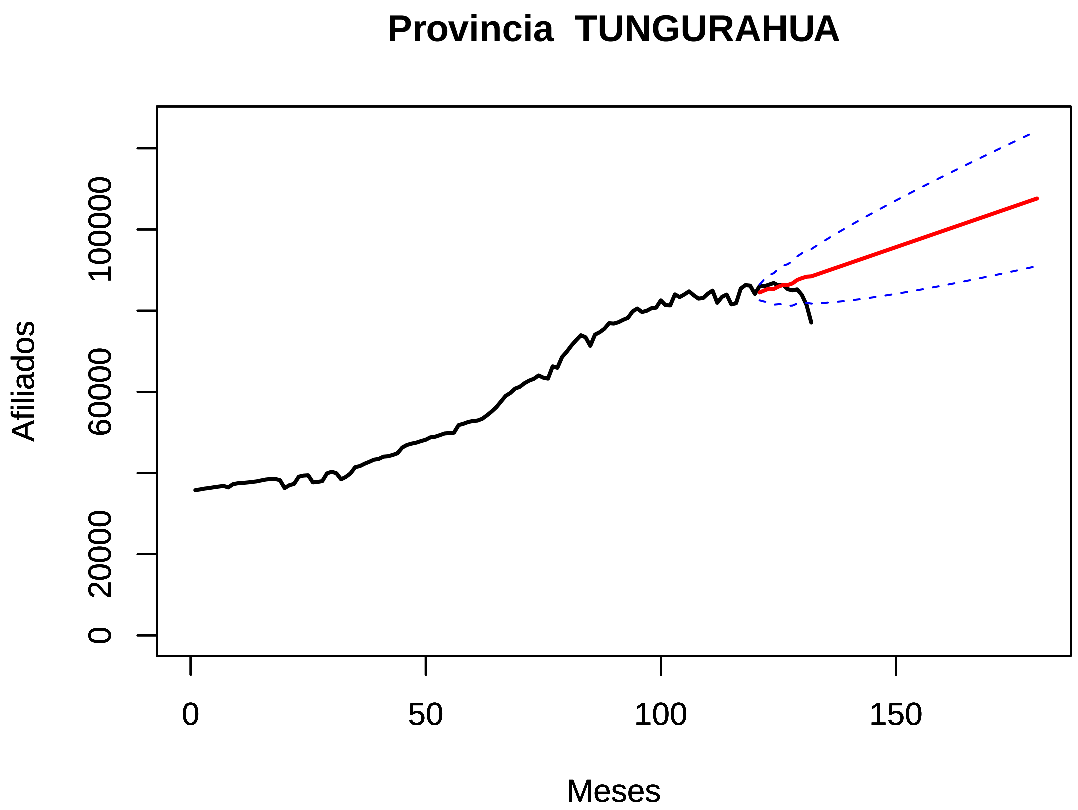

Q−Q Plot
Sample Quantiles
0
5000
10000
15000
−2
−1
0
1
2
Residuos



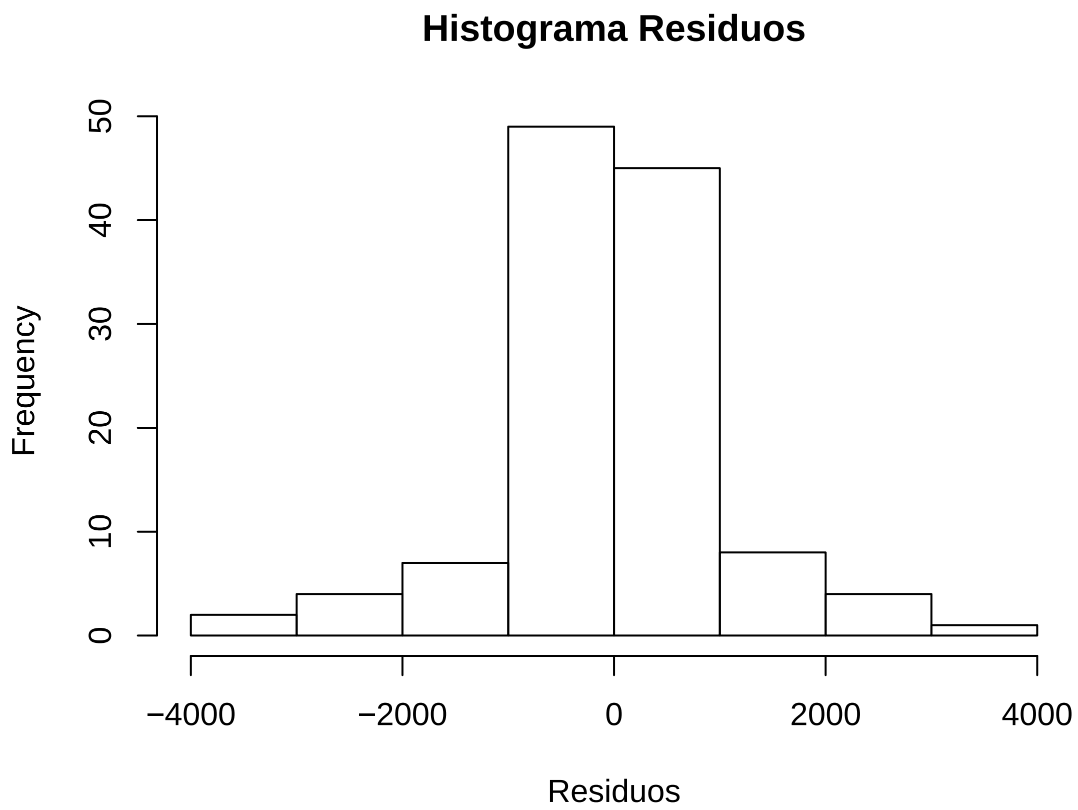

```
## Series: data[, i]
## ARIMA(2,2,1)(0,0,2)[12]
##
## Coefficients:
##           ar1      ar2      ma1    sma1    sma2
##       -0.3800  -0.1534  -0.9225  0.1800  0.2587
## s.e.   0.1086   0.1041   0.0469  0.1186  0.1428
##
## sigma^2 estimated as 195494:  log likelihood=-888.53
## AIC=1789.06   AICc=1789.82   BIC=1805.68
```

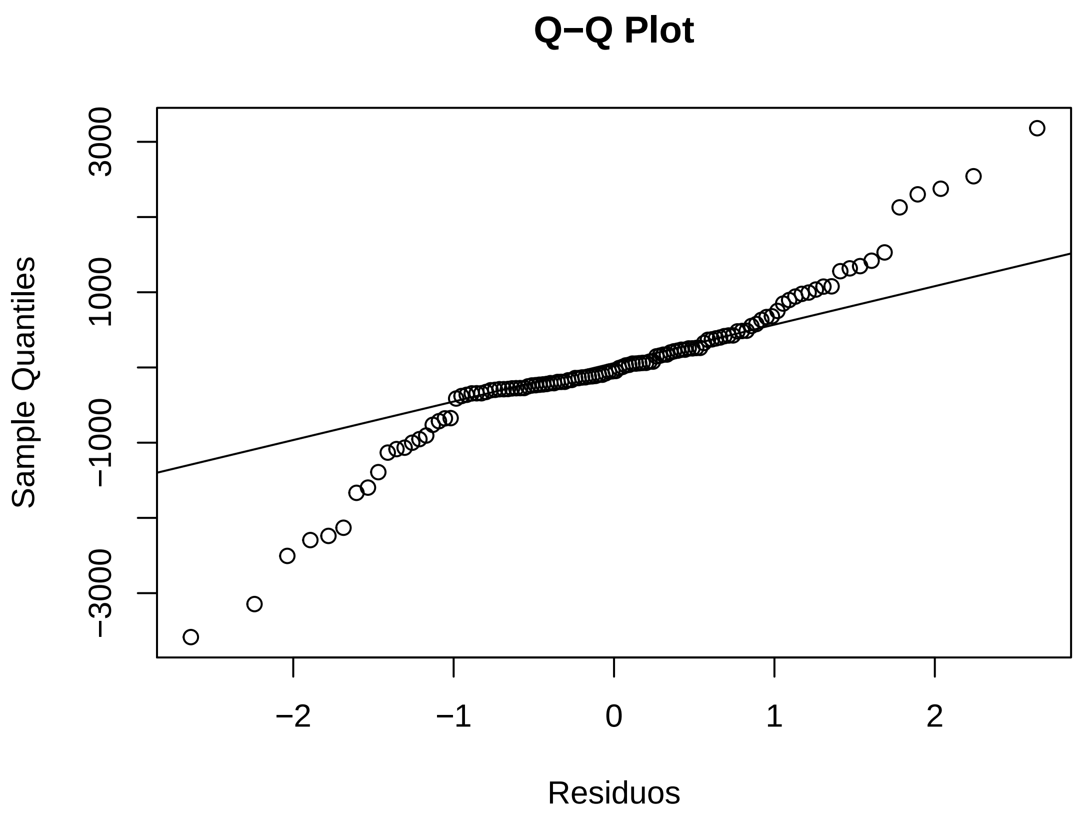

Histograma Residuos
Frequency
0
10
20
30
40
50
60
70
−2000
−1000
0
1000
Residuos

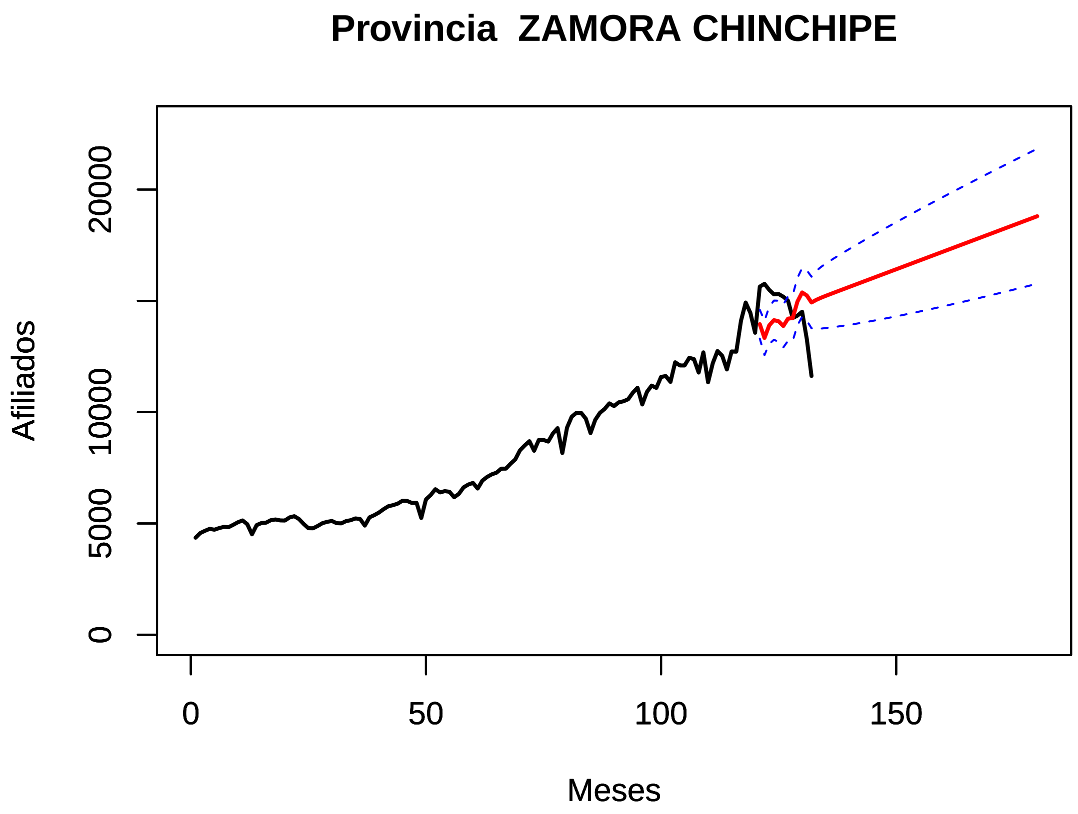

Q−Q Plot
Sample Quantiles
−2000
−1000
0
500
−2
−1
0
1
2
Residuos



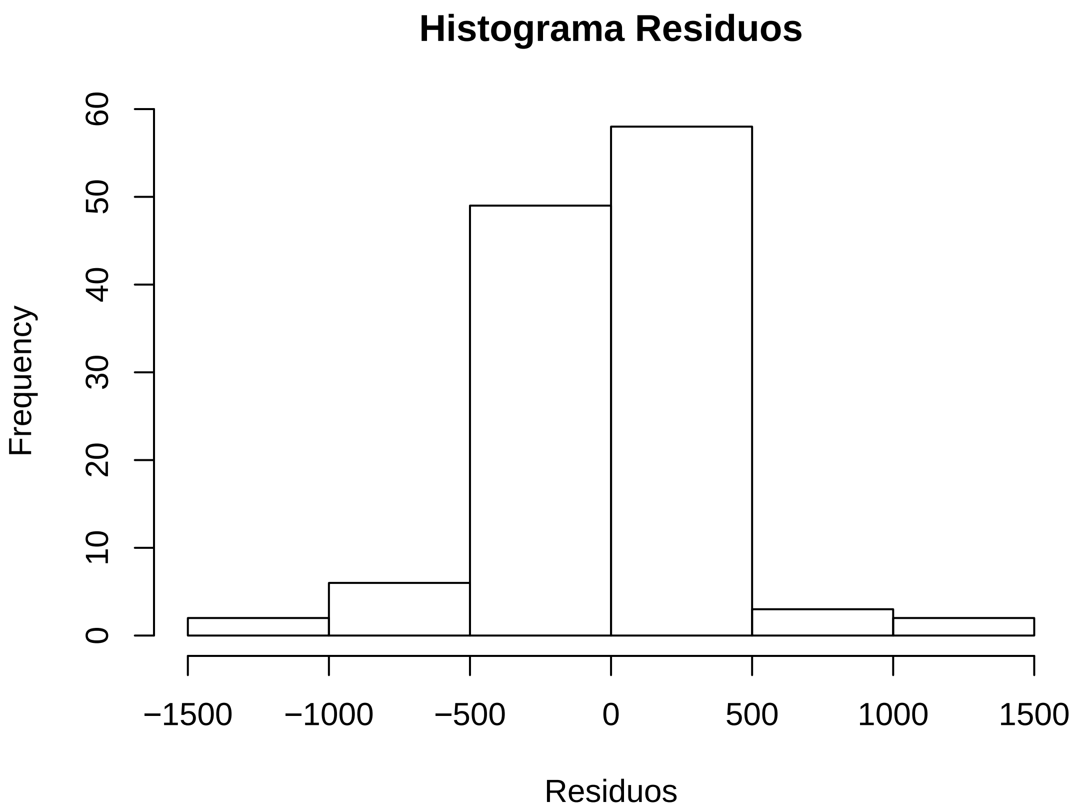

```
## Series: data[, i]
## ARIMA(0,1,0)(0,0,1)[12] with drift
##
## Coefficients:
##          sma1     drift
##        0.1360  399.5611
## s.e.   0.1007  101.7942
##
## sigma^2 estimated as 976849:  log likelihood=-989.59
## AIC=1985.19   AICc=1985.4   BIC=1993.53
```

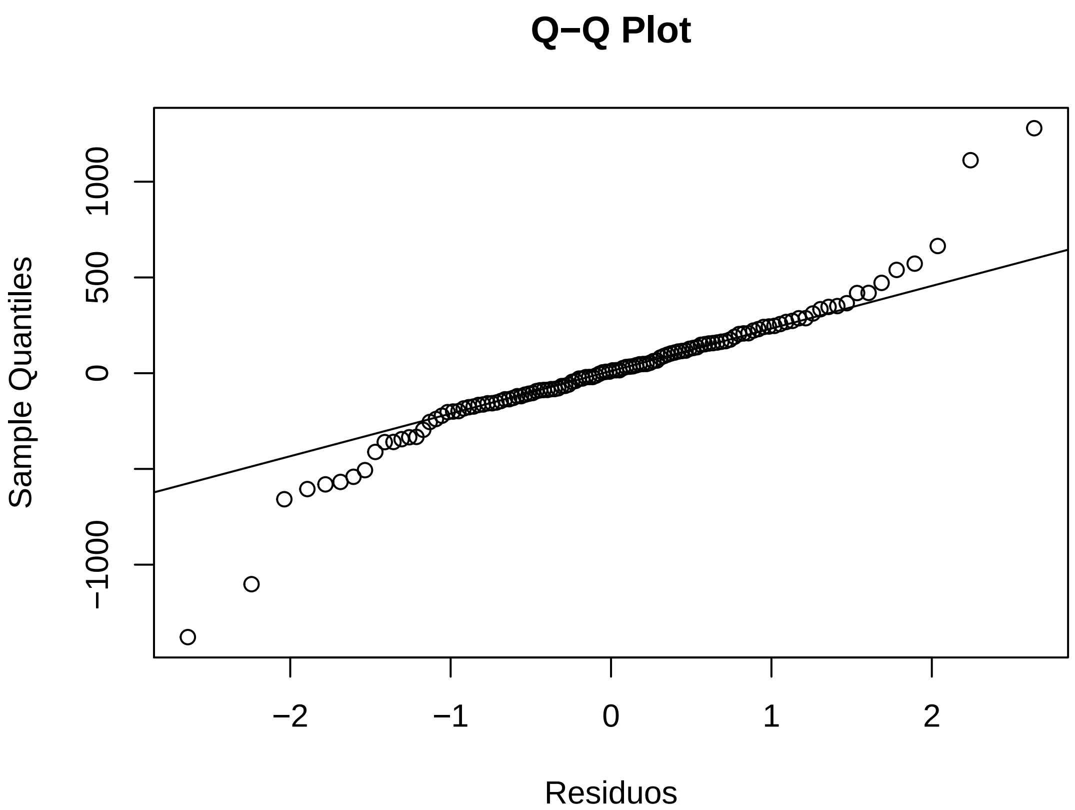

Histograma Residuos
Frequency
0
10
20
30
40
50
−4000
−2000
0
2000
4000
Residuos


Q−Q Plot
Sample Quantiles
−3000
−1000
1000
3000
−2
−1
0
1
2
Residuos




```
## Series: data[, i]
## ARIMA(1,1,1)(0,0,1)[12] with drift
##
## Coefficients:
##          ar1      ma1    sma1    drift
##       0.3897  -0.7511  0.5025  79.4536
## s.e.  0.1594   0.1111  0.1033  18.3544
##
## sigma^2 estimated as 109130:  log likelihood=-860.95
## AIC=1731.9   AICc=1732.44   BIC=1745.8
```





kable(matrix(rnorm(40), 5))