

Entrega Final - Series Cronológicas

Facultad de Ciencias Económicas y Administración - 2025 - UDeLaR

Leandro Berrueta, Lucca Frachelle, Cecilia Waksman

2025-06-28

El presente trabajo se desarrollará en base a la serie mensual de la cantidad de clientes con deuda vigente en el Banco Santander durante el período de Diciembre 2018 a Marzo 2025, proveniente de la central de riesgos que se informa al Banco Central del Uruguay (BCU). La serie se encuentra constituida, entonces, por 76 observaciones.

Dada la baja cantidad de observaciones disponibles se utilizará, a los efectos de identificar el modelo que logre representar el comportamiento de la serie, los datos de hasta Diciembre del 2024 inclusive (73 observaciones en total). Las tres observaciones restantes, referidas al año corriente, serán utilizadas con el fin de contrastar el desempeño de la predicción.

Gráfico de la Serie Temporal

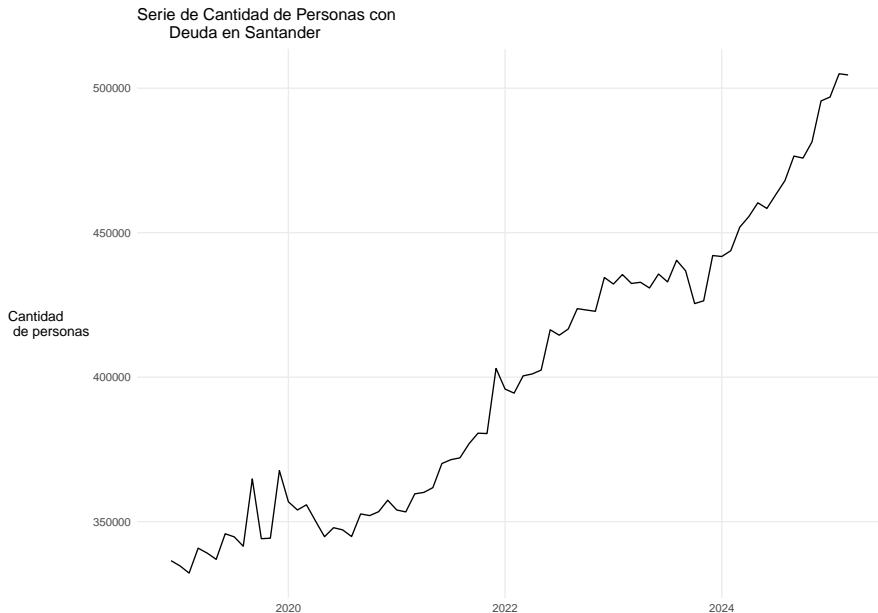
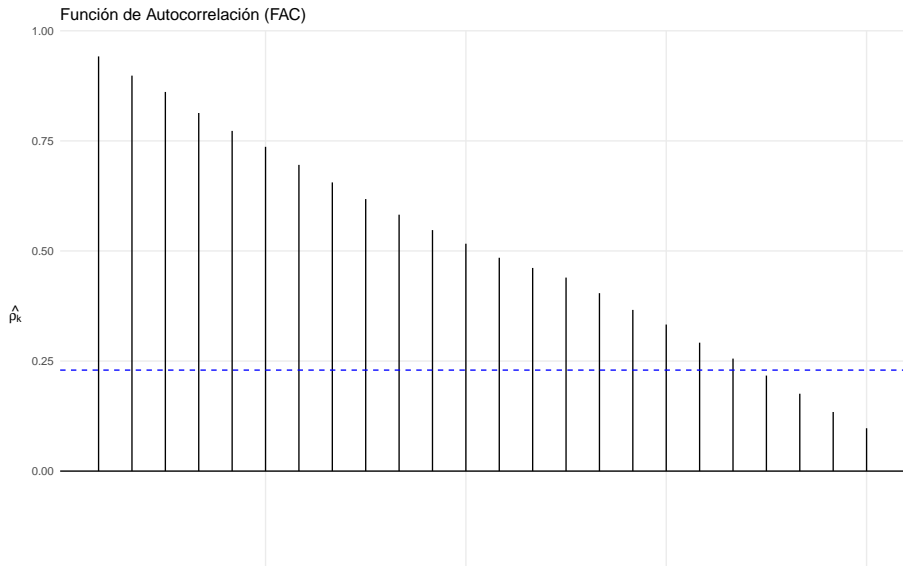


Table 1: Estadísticas Descriptivas de la Serie de Cantidad de Personas con Deuda

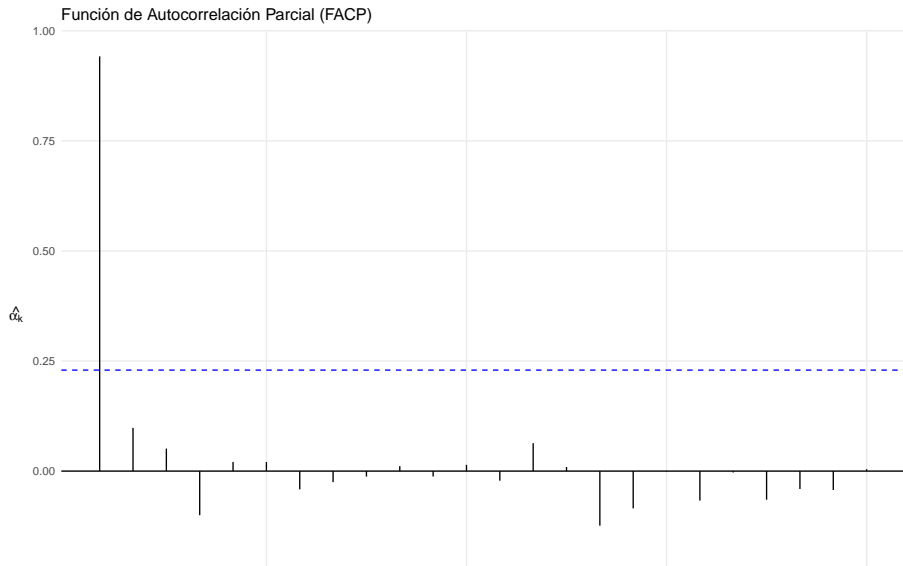
Estadística	Valor
Min.	332198
1st Qu.	353361
Median	394463
Mean	395286
3rd Qu.	432958
Max.	495587

Identificación del Modelo

Análisis en el Dominio del Tiempo: Función de Autocorrelación (FAC)

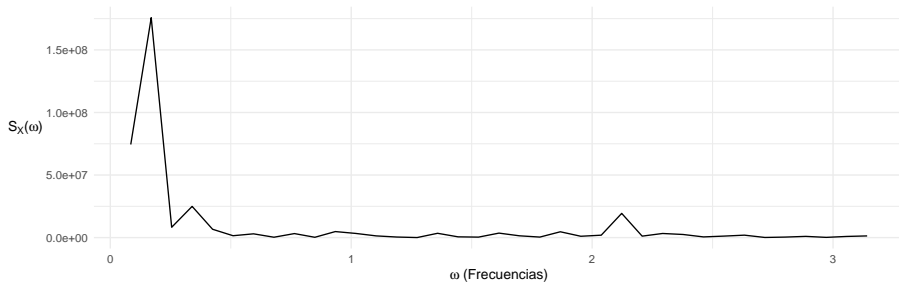


Análisis en el Dominio del Tiempo: Función de Autocorrelación Parcial (FACP)



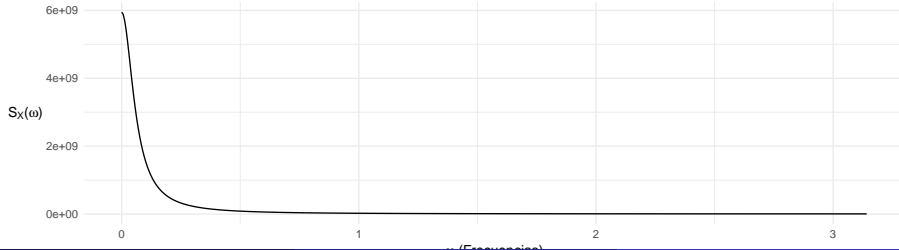
Análisis en el Dominio de Frecuencias de la Serie Original

Periodograma de la Serie



Periodograma Suavizado de la Serie

Método: AR(1)



- Contraste de Dickey-Fuller Aumentado (DF o DFA)
- Kwiatkowski-Phillips-Schmidt-Shin (KPSS).

Table 2: Resultados del Test de Dickey-Fuller Aumentado

	Estadístico	VC.1.	VC.5.	VC.10.
tau3 (con tendencia)	-0.8745181	-4.04	-3.45	-3.15
phi2	9.9484835	6.50	4.88	4.16
phi3	2.3435201	8.73	6.49	5.47

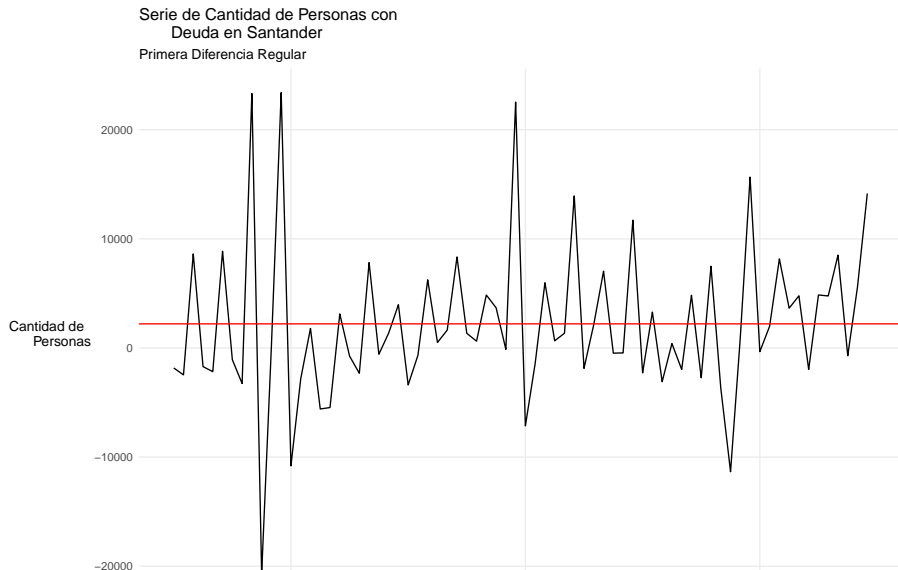
No se rechaza la Hipótesis Nula de que la serie presente una raíz unitaria a ninguno de los niveles de significación planteados. De esta manera se tiene un respaldo estadístico para aplicar la Primera Diferencia Regular.

Table 3: Resultados del Test KPSS (con tendencia)

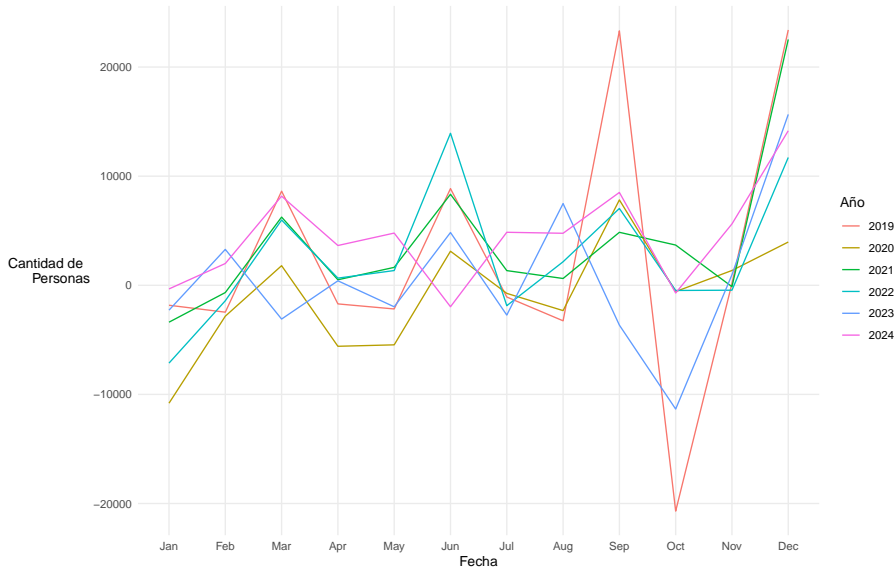
Item	Valor
Estadístico de Test	0.2567046
Valor Crítico 10%	0.1190000
Valor Crítico 5%	0.1460000
Valor Crítico 2.5%	0.1760000
Valor Crítico 1%	0.2160000

Como resultado se rechaza la Hipótesis Nula de que la serie sea Integrada de Orden 0, lo que nuevamente da un respaldo estadístico para la aplicación de la Primera Diferencia Regular en los datos.

Serie Diferenciada de acuerdo a la Primera Diferencia Regular

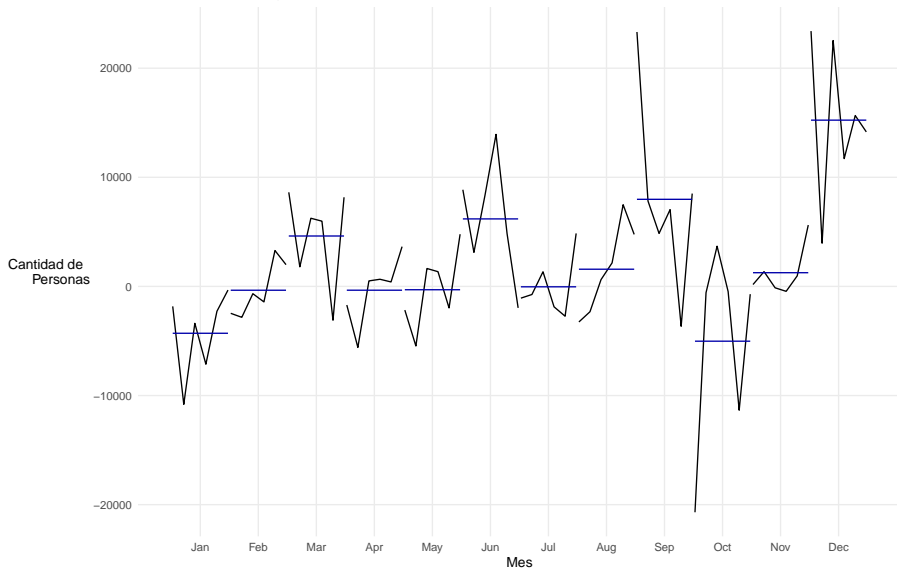


Serie de Cantidad de Personas con
Deuda en Santander
Primera Diferencia Regular



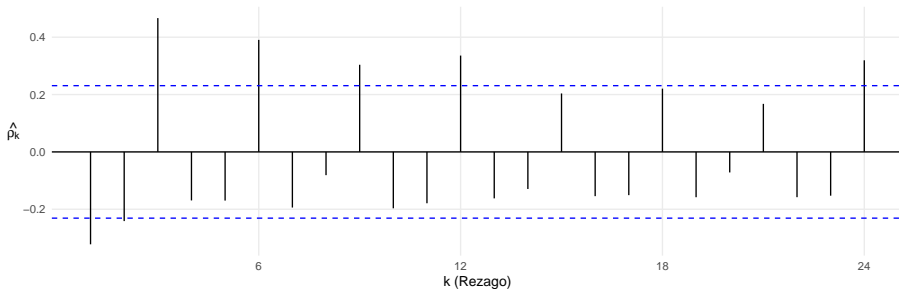
Separando los datos de acuerdo a los meses se desprende que los meses de

Serie de Cantidad de Personas con
Deuda en Santander
Primera Diferencia Regular

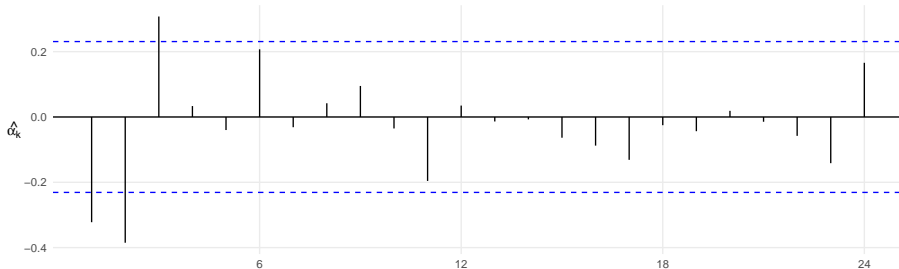


FAC y FACP de la Serie Diferenciada

FAC de la Serie Diferenciada

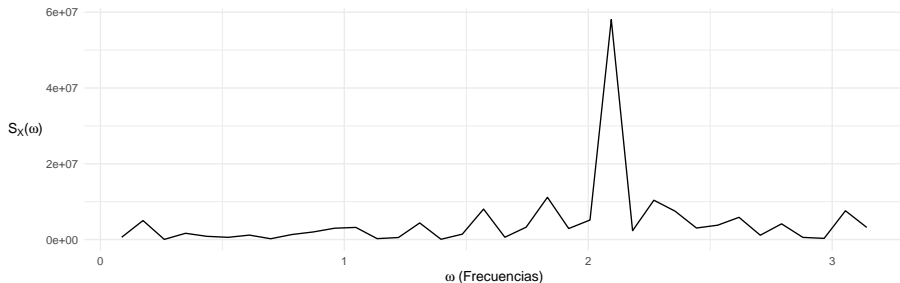


FACP de la Serie Diferenciada



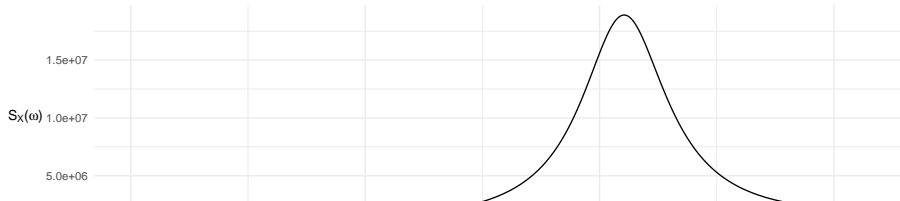
Dominio de Frecuencias: Análisis del Espectro de la Serie Diferenciada

Periodograma de la Serie Diferenciada



Periodograma Suavizado de la Serie Diferenciada

Método: AR(3)



Se procedió a llevar los contrastes de Dickey-Fuller Aumentado y KPSS a efectos de determinar si es necesaria la aplicación de una Segunda Diferencia Regular en la serie.

Table 4: Resultados del Test de Dickey-Fuller Aumentado

	Estadístico	VC.1.	VC.5.	VC.10.
tau2 (con constante)	-10.23177	-3.51	-2.89	-2.58
phi1	52.41887	6.70	4.71	3.86

Se rechaza la Hipótesis Nula de que la serie presente una raíz unitaria en todos los niveles de significación planteados.

Table 5: Resultados del Test KPSS (con constante)

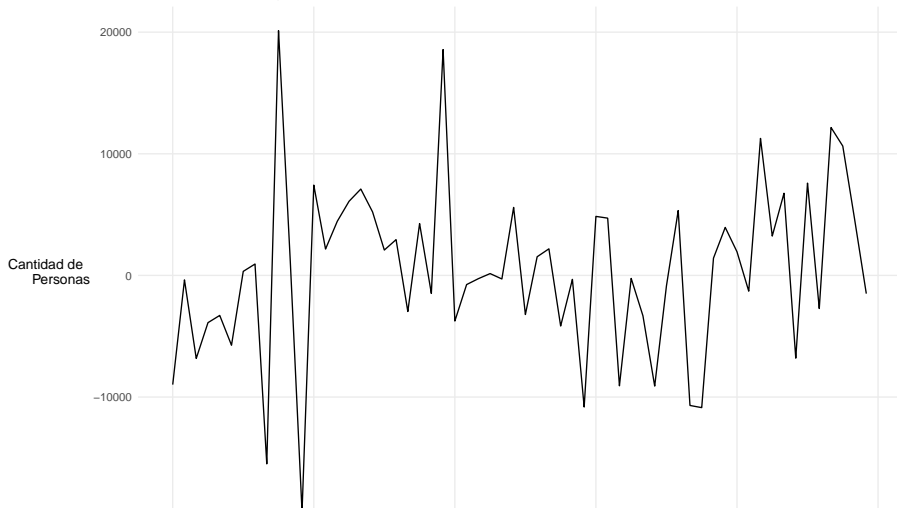
Item	Valor
Estadístico de Test	0.2987105
Valor Crítico 10%	0.3470000
Valor Crítico 5%	0.4630000
Valor Crítico 2.5%	0.5740000
Valor Crítico 1%	0.7390000

El no rechazo de la Hipótesis Nula de que la serie sea Integrada de Orden 0, lo que nuevamente da un respaldo estadístico para continuar trabajando con la serie diferenciada, sin aplicar una Segunda Diferencia Regular ($d = 2$).

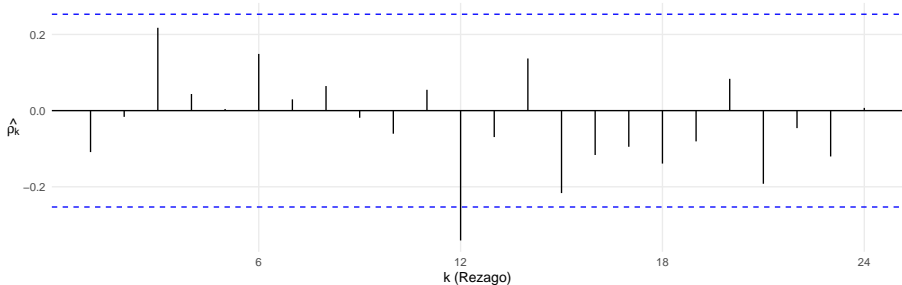
Serie Diferenciada de acuerdo a Primera Diferencia Regular y Primera Diferencia Estacional

Serie de Cantidad de Personas con
Deuda en Santander

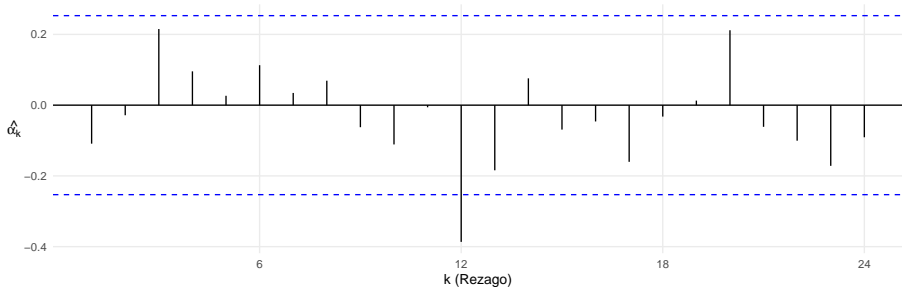
Primera Diferencia Regular y Primera Diferencia Estacional (Anual)



FAC de la Serie Diferenciada (Regular y Estacional)

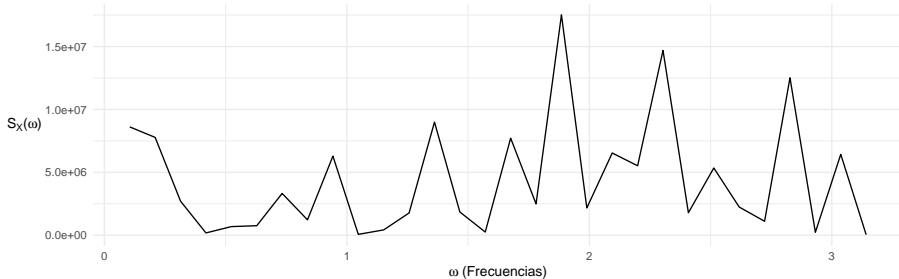


FACP de la Serie Diferenciada (Regular y Estacional)



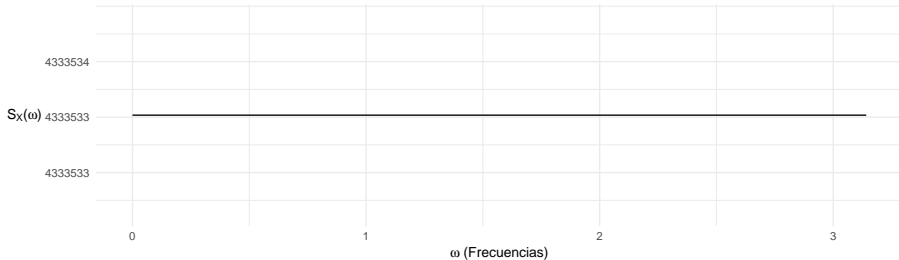
Observar la FAC y FACP en los múltiplos de 12. La significación tanto del

Periodograma de la Serie Diferenciada



Periodograma Suavizado de la Serie Diferenciada

Método: AR(0)



La Primera Diferencia Regular en conjunto con la Primera Diferencia

Modelo Final Propuesto: SARIMA(2,1,0)(1,1,0)

Characteristic	Beta ¹	SE	95% CI	p-value
ar1	0.00	0.000		
ar2	0.00	0.000		
ar3	0.52***	0.101	0.32, 0.72	<0.001

¹*p<0.05; **p<0.01; ***p<0.001

Abbreviations: CI = Confidence Interval, SE = Standard Error

Los residuos estandarizados indicaron la presencia de outliers. A su vez, se rechazó la Hipótesis Nula de Normalidad en los contrastes de Shapiro-Wilk y Jarque-Bera, lo que motivó la intervención de los siguientes puntos anómalos:

- Outlier Aditivo en Setiembre de 2019.
- Cambio Transitorio en Diciembre de 2019.
- Cambio Transitorio en Diciembre de 2021.
- Outlier Aditivo en Febrero de 2023.
- Outlier Aditivo en Agosto de 2023.
- Cambio Transitorio en Octubre de 2023.
- Cambio Transitorio en Junio de 2024.

Ajuste del Modelo Final

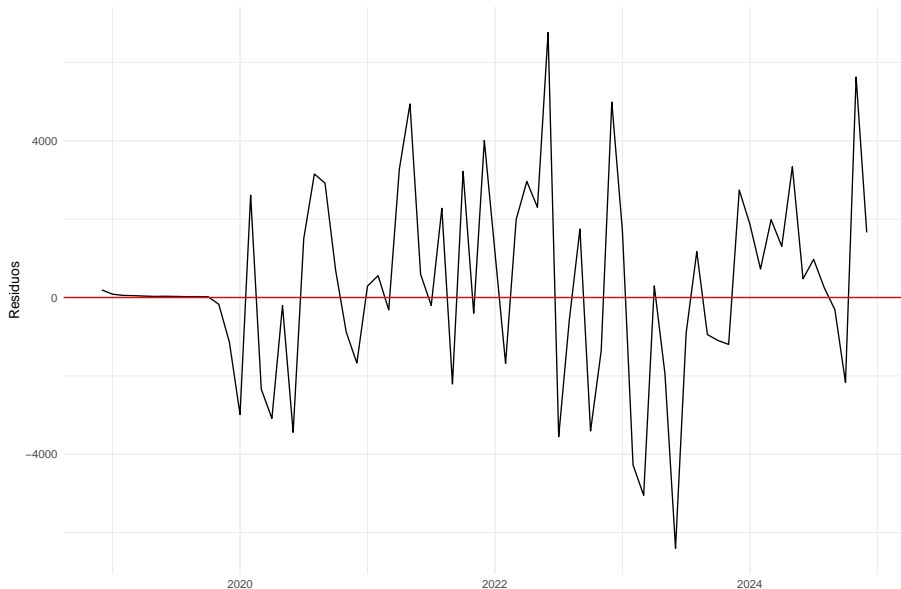
Characteristic	Beta ¹	SE	95% CI	p-value
ar1	0.00	0.000		
ar2	0.44***	0.124	0.19, 0.68	<0.001
sar1	-0.48**	0.150	-0.77, -0.18	0.002
AO10	19,189***	1,829	15,603, 22,774	<0.001
TC13	17,394***	2,542	12,412, 22,376	<0.001
TC37	11,534***	1,978	7,658, 15,411	<0.001
AO51	5,588***	1,557	2,536, 8,641	<0.001
AO57	7,679***	1,637	4,470, 10,888	<0.001
TC59	-11,055***	2,235	-15,436, -6,674	<0.001
TC67	-13,023***	2,419	-17,763, -8,282	<0.001

¹*p<0.05; **p<0.01; ***p<0.001

Abbreviations: CI = Confidence Interval, SE = Standard Error

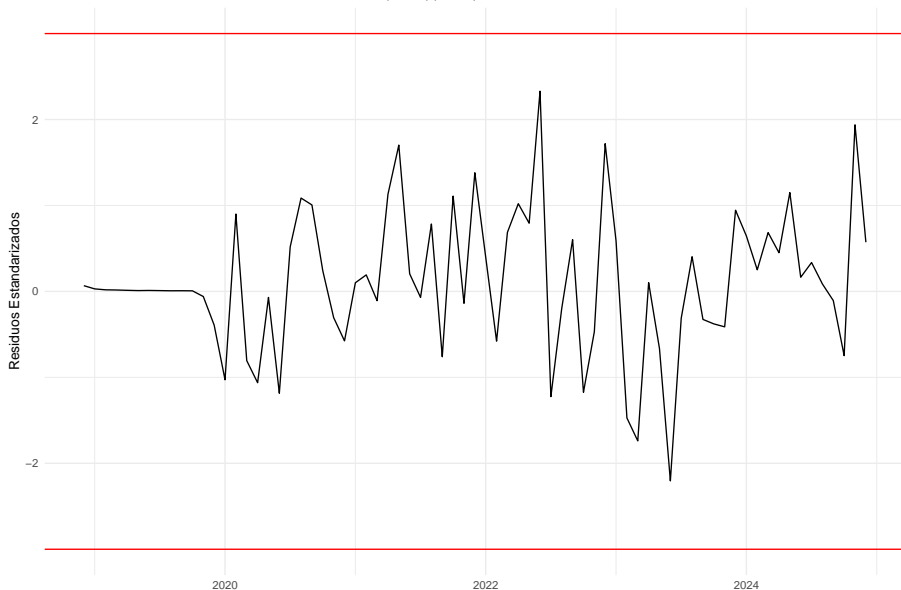
Diagnóstico de Residuos

Residuos del Modelo SARIMA(2,1,0)(1,1,0)



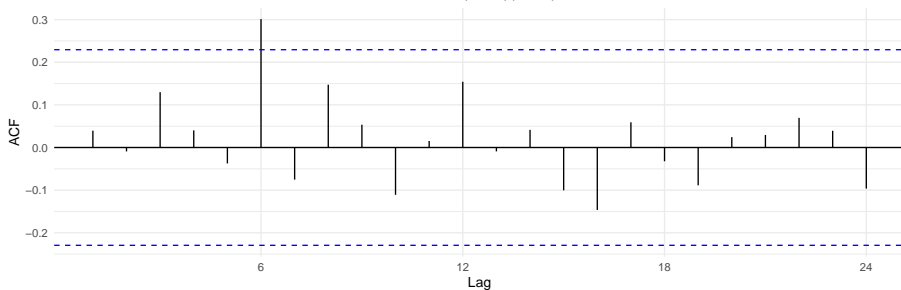
Análisis de Normalidad

Residuos Estandarizados del Modelo SARIMA(2,1,0)(1,1,0)



Homocedasticidad

FAC del Cuadrado de los Residuos del Modelo SARIMA(2,1,0)(1,1,0)



FACP del Cuadrado de los Residuos del Modelo SARIMA(2,1,0)(1,1,0)

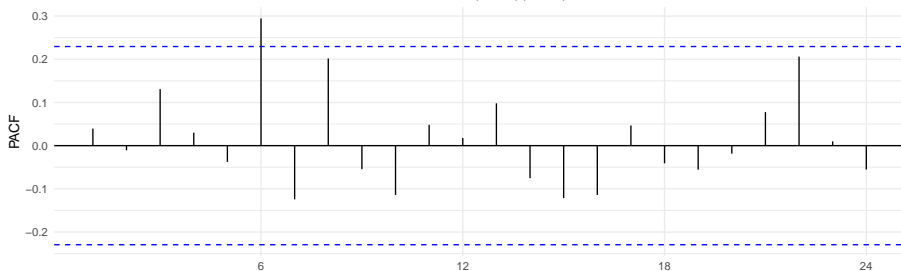


Table 10: Test de Ljung-Box para el Cuadrado de los Residuos del Modelo SARIMA(2,1,0)(1,1,0) (Rezagos 3, 6, 9, 12 y 24)

statistic	p.value	parameter	method
3.459295	0.0000000	0	Box-Ljung test
3.699865	0.2957504	3	Box-Ljung test
11.753917	0.0676892	6	Box-Ljung test
14.198213	0.1154471	9	Box-Ljung test
31.505574	0.0656457	21	Box-Ljung test

No obstante, los resultados del contraste de Ljung-Box hacen que se desestime esta idea.

Contraste de Media Nula de los Residuos

Table 11: Tests de Media Nula para Residuos del Modelo SARIMA(2,1,0)(1,1,0)

estimate	statistic	p.value	parameter
312.1445	1.100609	0.2747331	72

No se rechaza la Hipótesis Nula de que los residuos tengan una media distinta de 0.

El Modelo a utilizar es un SARIMA(2, 1, 0)(1, 1, 0) con $\phi_1 = 0$, que puede expresarse como:

$$(1 - L)(1 - L^{12})(1 - L^2\phi_2)(1 - L\Phi_1)Y_t = \epsilon_t$$

Donde Y_t denota a la serie de datos original, L al operador de rezagos y ϵ_t un ruido blanco. ϕ_2 y Φ_1 son los coeficientes respectivos de las partes AR del componente regular y estacional del Modelo.

Enero, Febrero y Marzo de 2025

Se realiza predicciones para los meses de Enero, Febrero y Marzo de 2025, a efectos de contrastarla con las tres observaciones que no se utilizaron a la hora de ajustar el Modelo.

Cantidad de Clientes con Deuda en Santander:
Predicción de 2025

Modelo SARIMA(2, 1, 0)(1, 1, 0)

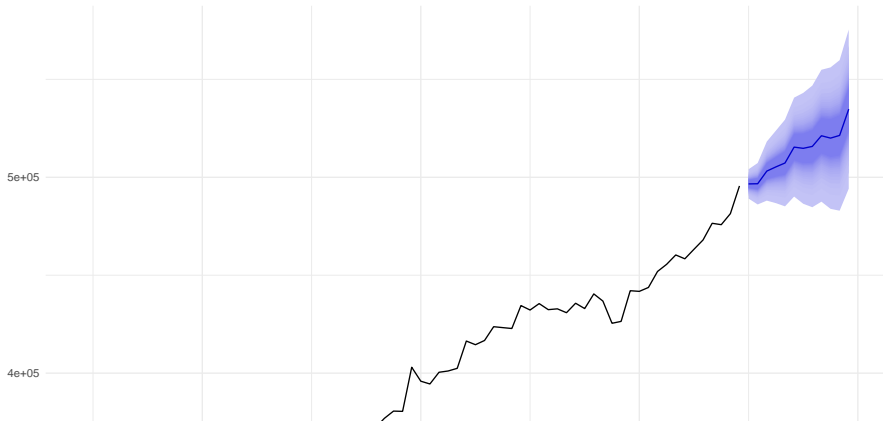


Table 12: Intervalos de Confianza al 50% de los Primeros 3 Meses de 2025 contra Realidad

Fecha	Límite Inferior 50%	Predicción	Límite Superior 50%	Realidad	Capturado
Enero - 2025	494652.4	496610.6	498568.8	496939	Si
Febrero - 2025	493933.3	496702.7	499472.1	504989	No
Marzo - 2025	499202.9	503148.9	507094.8	504575	Si

Table 13: Intervalos de Confianza al 75% de los Primeros 3 Meses de 2025 contra Realidad

Fecha	Límite Inferior 75%	Predicción	Límite Superior 75%	Realidad	Capturado
Enero -	493270.8	496610.6	499950.4	496939	Si

Métricas sobre los errores de predicción

Table 16: Métricas de Predicción - SARIMA(2, 1, 0)(1, 1, 0)

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	312.1445	2426.6771	1788.9850	0.0757	0.4437	0.0719	0.1244	NA
Test set	3346.9394	4858.1333	3346.9390	0.6632	0.6632	0.1344	- 0.6559	1.0427

2024 y 2025

También se realiza predicciones para el 2024 y los primeros tres meses de 2025, utilizando el mismo modelo propuesto pero entrenándolo con las observaciones disponibles hasta diciembre de 2023.

Cantidad de Clientes con Deuda en Santander:
Predicción de 2024 y Principios 2025

Modelo SARIMA(2, 1, 0)(1, 1, 0)

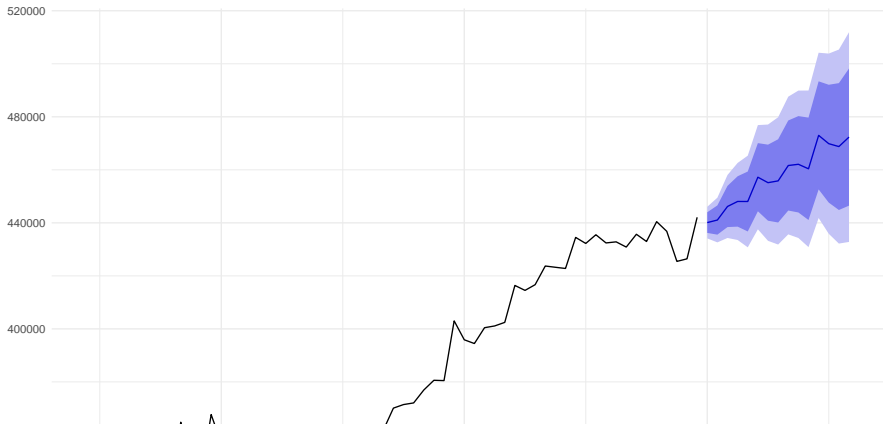


Table 17: Métricas de Predicción - SARIMA(2, 1, 0)(1, 1, 0)

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	107.9246	2461.7561	768.2810	0.0339	0.4512	0.0753	0.1116	NA
Test set	14604.1741	18096.4071	4604.1731	0.0075	3.0075	0.6222	0.7409	2.9412

A partir de Junio de 2024

Considerando una muestra de entrenamiento con límite en Junio de 2024, tal que se incorpore el cambio transitorio identificado para ese mes.

Cantidad de Clientes con Deuda en Santander:

Predicción a partir de Junio de 2024

Modelo SARIMA(2, 1, 0)(1, 1, 0)

5e+05

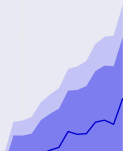
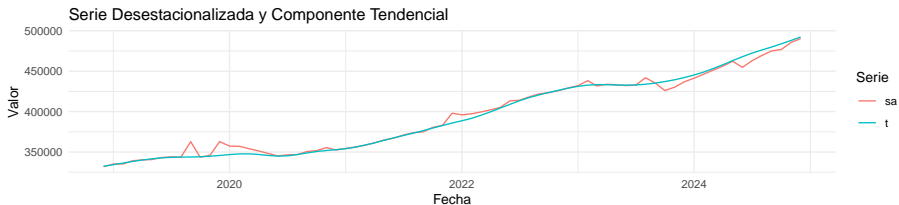


Table 18: Métricas de Predicción - SARIMA(2, 1, 0)(1, 1, 0)

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1	Theil's U
Training set	247.9815	2393.6611	743.7550	0.0638	0.4397	0.076	0.1366	NA
Test set	5983.7858	8183.2176	172.5201	1.2035	1.2432	0.269	0.6178	1.1937

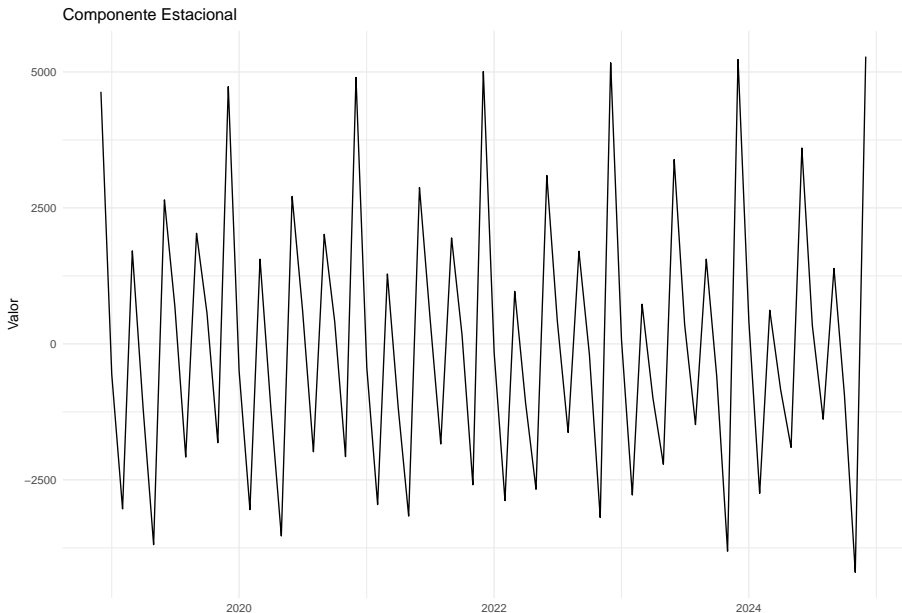
Mediante la función `x13` (asociada al método estadístico del mismo nombre) del paquete `RJDemetra` se llevó a cabo la descomposición de la serie (de acuerdo a términos aditivos), tomando en cuenta el Modelo Final planteado anteriormente.

Como resultado se obtuvo la descomposición de la serie, lo que permitió computar una serie desestacionalizada.

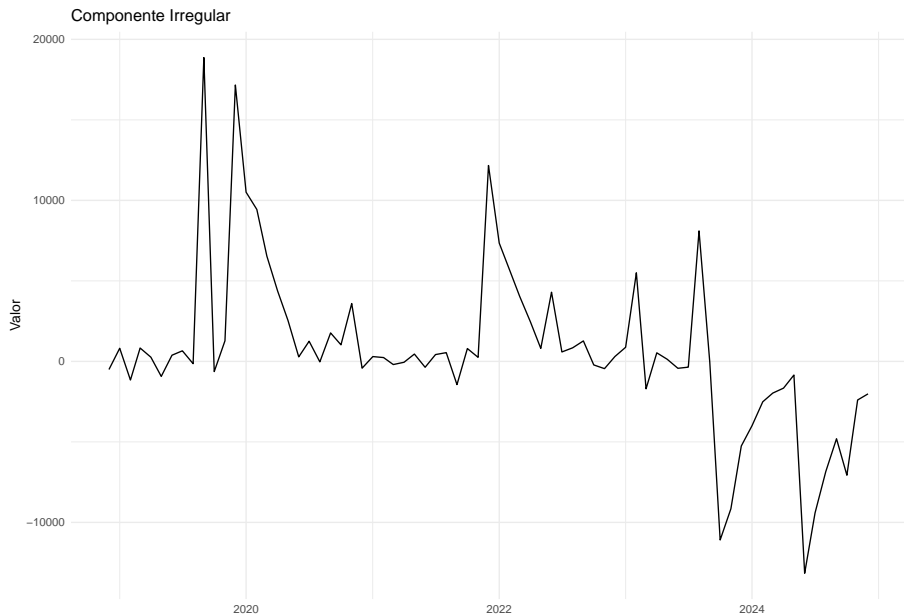


En los dos primeros gráficos se puede notar como la serie se ha vuelto más

Componente estacional

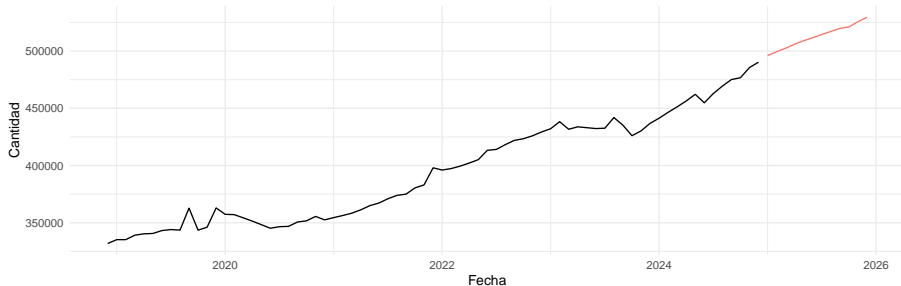


Componente irregular

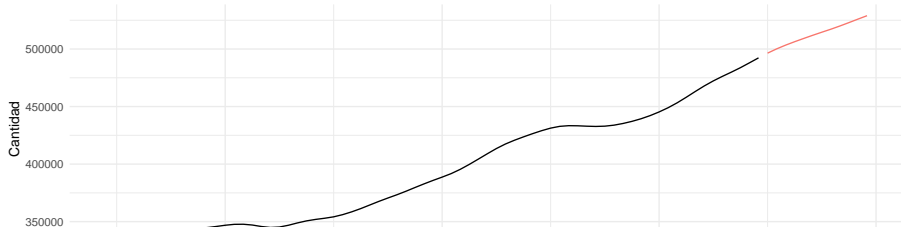


A las predicciones realizadas en el punto anterior, se le agrega las predicciones de los componentes tendencial así como de la serie desestacionalizada:

Serie Desestacionalizada con Predicción de 2025



Componente Tendencial con Predicción de 2025



Respecto al Modelo SARIMA(2, 1, 0)(1, 1, 0) se debe tener en cuenta que:

- Se logra el no rechazo de los Contrastes de Normalidad utilizando siete intervenciones por atípicos, lo que dada la baja cantidad de observaciones disponibles puede resultar un número no deseable.
- El supuesto de residuos no autocorrelacionados puede no estar cumpliéndose, tal y como se refleja en los p-valores del Contraste de Ljung-Box, que resultan cercanos al 5%.
- Dadas dos muestras de entrenamiento se observó la tendencia a subestimar por parte del modelo.

No obstante lo anterior, esta modelización presenta los siguientes puntos a destacar:

- El Modelo SARIMA resultante es de bajo orden, lo que va de la mano con la idea de que los modelos de este tipo sean los mejores a la hora de predecir, además de resultar más parsimoniosos.
- Dado que no se presentan problemas de homocedasticidad y a que, como se había mencionado al principio del presente trabajo, la transformación logarítmica no logra homogeneizar la Varianza de la serie, se descarta la aplicación de dicha transformación.
- Se logra predecir con éxito a un bajo nivel de confianza dos de tres valores que tomó la serie en 2025: los asociados a los meses de Enero y Marzo. En el caso de Febrero se subpredice el valor que toma efectivamente la serie.