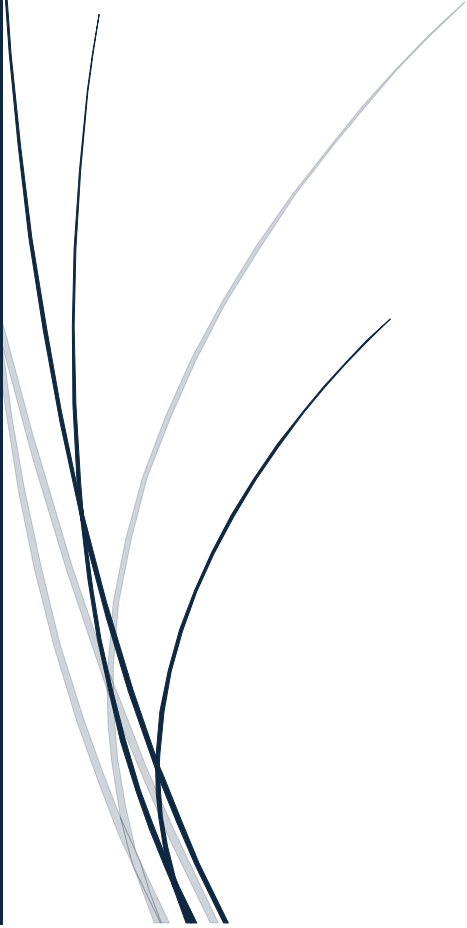


2-11-2024

Trabajo I: Series Temporales

TÉCNICAS ESTADÍSTICAS EN CIENCIA DE DATOS



XinYuan Zheng
GRUPO A

ÍNDICE

1. Análisis descriptivo	2
1.1. Introducción: Presentación de la serie a analizar. Representar la serie, comentar el gráfico. (0.5).....	2
1.2. Analizar la tabla de los estadísticos por unidad temporal. ¿Cuál es el número mínimo que ha habido? ¿Cuándo? ¿Cuál el máximo? (0.5)	3
1.3. Calcular los coeficientes de estacionalidad, mostrar su tabla y su representación junto con la componente irregular. ¿Qué coeficiente es el mayor? ¿Qué significa? ¿Cuál es el menor? ¿Cuál es su significado? (1).....	4
1.4. Representar la serie desestacionalizada, la estimación de la tendencia y el error. ¿Qué dirías sobre la tendencia? ¿Y sobre su comportamiento estacional? (0.5)	7
2. Encontrar el método de suavizado más adecuado a la serie teniendo en cuenta las características de la serie y comparando las medidas de bondad de ajuste. Para el método elegido:	8
2.1. Sobre las tablas de estimadores de los coeficientes del modelo ¿Cuánto vale la estimación de los parámetros de suavizado? ¿Qué podemos decir sobre los coeficientes del modelo de suavizado? Escribir las ecuaciones del modelo. (1)	9
2.2. Comentar los gráficos de las componentes del modelo en relación al apartado anterior. (0.5)	10
2.3. Calcular las medidas de bondad de ajuste y de comportamiento sobre los datos reservados para ello. Comentar su resultado. (0.5).....	11
2.4. Mostrar la gráfica donde aparecen los valores observados, los suavizados y las predicciones con los intervalos de confianza. (0.5)	11
2.5. Mostrar una tabla y gráfico en donde aparecen los valores observados, las predicciones y sus intervalos de confianza, sólo para el último periodo que habíamos reservado. ¿Cómo se comportan los intervalos de confianza? (1).....	12
3. Para ajustar un modelo ARIMA:	14
3.1. Representar la serie y los correlogramas. Hacer las diferenciaciones que sean necesarias. Decidir qué modelo puede ser ajustado. (0.5).....	14
3.2. Ajustar el modelo adecuado comprobando que sus residuales están incorrelados. (Sintaxis, tablas de los parámetros estimados y gráficos)	17
3.3. Escribir la expresión algebraica del modelo ajustado con los parámetros estimados. (0,5).....	22
3.4. Calcular las predicciones y los intervalos de confianza para las unidades de tiempo que se considere oportuno, dependiendo de la serie, siguientes al último valor observado. Representarlas gráficamente. (1).....	23
4. Comparar las predicciones obtenidas con cada uno de los métodos con los valores observados que habíamos reservado antes. Conclusiones. (1).....	25
5. ANEXO	27

1. Análisis descriptivo

1.1. Introducción: Presentación de la serie a analizar. Representar la serie, comentar el gráfico. (0.5)

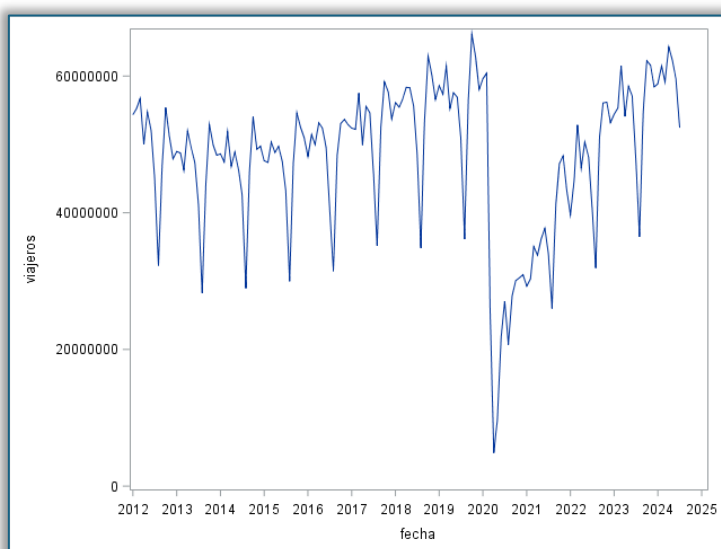
La serie a analizar en este estudio se trata de la cantidad de viajeros transportados(miles de personas) en metro en Madrid. La serie tiene recogido los datos historicos desde enero de 2012 hasta julio de 2024.

El paso previo para el análisis de la serie, se transforma la columna de meses del año en valor tipo fecha para SAS; y se cambia la escala del número de viajeros de miles de personas a unidades. Para no alterar los datos originales(“metro_madrid”), las 2 nuevas columnas creadas se han guardado en una nueva serie temporal denominado “metro”.

```
data series.metro(keep= fecha viajeros);  
set series.metro_madrid;  
fecha=intnx('month', '01JAN2012'd, _n_ - 1);  
format fecha date9.;  
viajeros=viajeros_transportados_miles_*1000;  
run;
```

Todo el estudio se trabajará con esta serie “metro”. En primer lugar, se mostrará la gráfica de la serie para conocer un poco sobre su comportamiento general.

```
proc sgplot data=series.metro;  
series x=fecha y=viajeros;  
xaxis interval=year;  
format fecha year4.;  
run;
```



Observando la gráfica de la serie metro, se puede reconocer un patrón cíclico durante el tiempo, por lo tanto se puede afirmar de que esta serie presenta comportamiento estacional en el tiempo.

En el año 2020, se puede apreciar una caída muy importante del número de viajero; este período de tiempo coincidía con el confinamiento debido a la pandemia de SARS-CoV-2. Aun así, se puede observar una clara tendencia creciente especialmente en el periodo postpandemia que recupera el número de viajeros transportado a nivel de antes de pandemia.

1.2. Analizar la tabla de los estadísticos por unidad temporal. ¿Cuál es el número mínimo que ha habido? ¿Cuándo? ¿Cuál el máximo? (0.5)

A partir del siguiente código, se puede mostrar la serie completa en función de unidad de tiempo¹.

```
proc sql;
select fecha, viajeros
from series.metro;
quit;
```

- Número mínimo de viajeros

```
proc sql;
select fecha, viajeros
from series.metro
where viajeros= (select min(viajeros) from series.metro);
quit;
```

fecha	viajeros
01APR2020	4865000

El número de viajeros transportado más bajo registrado en la serie fue de 4.865.000 personas aproximadamente en abril de 2020.

- Número máximo de viajeros

```
proc sql;
select fecha, viajeros
from series.metro
where viajeros= (select max(viajeros) from series.metro);
quit;
```

fecha	viajeros
01OCT2019	66151000

El récord de número de viajeros transportado fue en octubre de 2019 con 66.151.000 personas aproximadamente.

¹ La tabla completa se mostrará en el anexo.

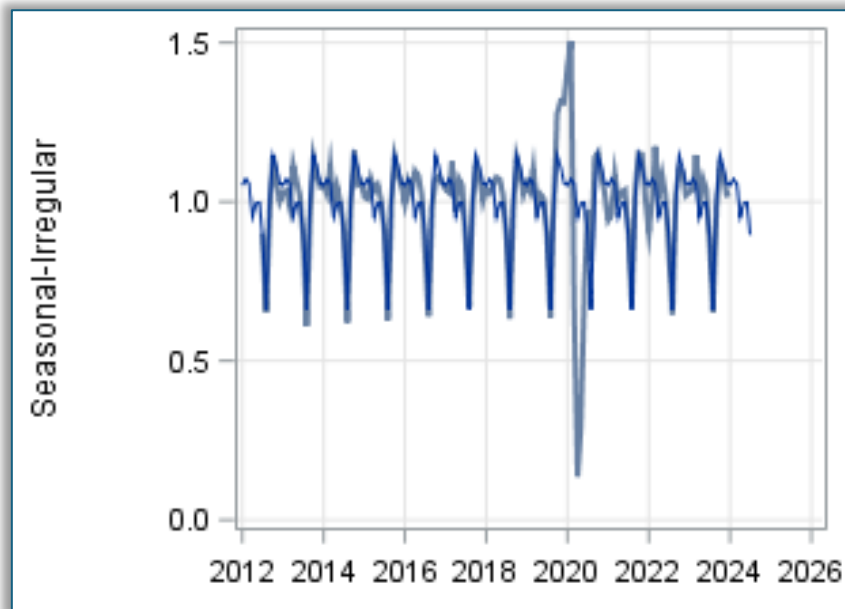
1.3. Calcular los coeficientes de estacionalidad, mostrar su tabla y su representación junto con la componente irregular. ¿Qué coeficiente es el mayor? ¿Qué significa? ¿Cuál es el menor? ¿Cuál es su significado? (1)

```
proc timeseries data=series.metro plots=(decomp series)
print=(seasons decomp);
id fecha interval=month;
var viajeros;
run;
```

Seasonal Decomposition for Variable viajeros						
Obs	Time	Season	Trend-Cycle	Seasonal	Irregular	Seasonally Adjusted
1	JAN2012	1	.	1.051584	.	51669672
2	FEB2012	2	.	1.071206	.	51605390
3	MAR2012	3	.	1.056866	.	53667157
4	APR2012	4	.	0.953613	.	52426946
5	MAY2012	5	.	0.994426	.	55081034
6	JUN2012	6	.	0.997237	.	52160130
7	JUL2012	7	49905708	0.896105	1.005640	50187195
8	AUG2012	8	49409208	0.660080	0.989850	48907698
9	SEP2012	9	48700708	1.006847	0.954783	46498612
10	OCT2012	10	48347500	1.146126	0.999543	48325396
11	NOV2012	11	48212708	1.108332	0.955654	46074654
12	DEC2012	12	47800292	1.057578	0.947667	45298769
13	JAN2013	1	47453667	1.051584	0.981293	46565943
14	FEB2013	2	47134667	1.071206	0.965144	45491720
15	MAR2013	3	46859917	1.056866	0.934161	43774702
16	APR2013	4	46648500	0.953613	1.168337	54501164
17	MAY2013	5	46495667	0.994426	1.071902	49838813
18	JUN2013	6	46469083	0.997237	1.021370	47462149
19	JUL2013	7	46476500	0.896105	0.992585	46131867
20	AUG2013	8	46408042	0.660080	0.922566	42814498
21	SEP2013	9	46587708	1.006847	0.943255	43944103
22	OCT2013	10	46605875	1.146126	0.990055	46142390
23	NOV2013	11	46362458	1.108332	0.971101	45022623
24	DEC2013	12	46290250	1.057578	0.989366	45798022
25	JAN2014	1	46300833	1.051584	0.998537	46233112
26	FEB2014	2	46383792	1.071206	0.954744	44284669
27	MAR2014	3	46498458	1.056866	1.055478	49078118
28	APR2014	4	46632542	0.953613	1.052499	49080724
29	MAY2014	5	46654250	0.994426	1.053753	49162041
30	JUN2014	6	46683083	0.997237	0.993919	46399211
31	JUL2014	7	46695417	0.896105	1.019334	47598213
32	AUG2014	8	46649875	0.660080	0.939770	43840131
33	SEP2014	9	46583792	1.006847	0.987575	46004992
34	OCT2014	10	46604375	1.146126	1.011486	47139662
35	NOV2014	11	46722542	1.108332	0.951605	44461419
36	DEC2014	12	46808125	1.057578	1.004964	47040483
37	JAN2015	1	46879542	1.051584	0.965561	45265048
38	FEB2015	2	46943333	1.071206	0.941873	44214654
39	MAR2015	3	47023417	1.056866	1.013291	47648419
40	APR2015	4	47084500	0.953613	1.087051	51183255
41	MAY2015	5	47241667	0.994426	1.058339	49997699
42	JUN2015	6	47426250	0.997237	1.004203	47625600
43	JUL2015	7	47505250	0.896105	1.013491	48146140
44	AUG2015	8	47702375	0.660080	0.952064	45415698
45	SEP2015	9	47855792	1.006847	0.979631	46880994
46	OCT2015	10	48020833	1.146126	0.992261	47649205
47	NOV2015	11	48313042	1.108332	0.979814	47337814
48	DEC2015	12	48508792	1.057578	0.993941	48214864
49	JAN2016	1	48488583	1.051584	0.946346	45886967
50	FEB2016	2	48444708	1.071206	0.991246	48020645
51	MAR2016	3	48553792	1.056866	0.973656	47274673
52	APR2016	4	48536292	0.953613	1.148649	55751147
53	MAY2016	5	48520958	0.994426	1.085502	52669593
54	JUN2016	6	48650250	0.997237	1.021007	49672256

55	JUL2016	7	48901167	0.896105	0.926916	45327274
56	AUG2016	8	49105125	0.660080	0.970311	47647245
57	SEP2016	9	49452208	1.006847	0.971002	48018207
58	OCT2016	10	49631583	1.146126	0.932546	46283736
59	NOV2016	11	49628583	1.108332	0.975603	48417816
60	DEC2016	12	49973250	1.057578	1.000877	50017095
61	JAN2017	1	50409917	1.051584	0.987922	49801062
62	FEB2017	2	50788125	1.071206	0.959828	48747863
63	MAR2017	3	51112708	1.056866	1.064694	54419381
64	APR2017	4	51537667	0.953613	1.015709	52347249
65	MAY2017	5	51961083	0.994426	1.075062	55861384
66	JUN2017	6	52165292	0.997237	1.050208	54784382
67	JUL2017	7	52356292	0.896105	0.980460	51333266
68	AUG2017	8	52646208	0.660080	1.011375	53245050
69	SEP2017	9	52740708	1.006847	0.987574	52085358
70	OCT2017	10	53052833	1.146126	0.972794	51609502
71	NOV2017	11	53518917	1.108332	0.973098	52079174
72	DEC2017	12	53675833	1.057578	0.946687	50814199
73	JAN2018	1	53833167	1.051584	0.991184	53358553
74	FEB2018	2	53934875	1.071206	0.959509	51751020
75	MAR2018	3	53970500	1.056866	0.991680	53521443
76	APR2018	4	54175750	0.953613	1.129616	61197803
77	MAY2018	5	54431125	0.994426	1.077008	58622777
78	JUN2018	6	54648750	0.997237	1.021198	55807208
79	JUL2018	7	54871417	0.896105	0.991569	54408798
80	AUG2018	8	55055083	0.660080	0.958455	52767835
81	SEP2018	9	55339083	1.006847	0.962202	53247401
82	OCT2018	10	55414500	1.146126	0.990474	54886628
83	NOV2018	11	55255042	1.108332	0.981272	54220230
84	DEC2018	12	55276250	1.057578	0.968029	53509034
85	JAN2019	1	55419500	1.051584	1.005607	55730213
86	FEB2019	2	55565000	1.071206	0.963516	53537792
87	MAR2019	3	55729792	1.056866	1.043601	58159684
88	APR2019	4	55975667	0.953613	1.035368	57955396
89	MAY2019	5	56222958	0.994426	1.029538	57883657
90	JUN2019	6	56395708	0.997237	1.011684	57054655
91	JUL2019	7	56498000	0.896105	1.006357	56857172
92	AUG2019	8	56667167	0.660080	0.965890	54734263
93	SEP2019	9	55334208	1.006847	1.009978	55886332
94	OCT2019	10	51773667	1.146126	1.114795	57717032
95	NOV2019	11	47677042	1.108332	1.188165	56648204
96	DEC2019	12	44217208	1.057578	1.241255	54884819
97	JAN2020	1	41758750	1.051584	1.357348	56681159
98	FEB2020	2	40121000	1.071206	1.405608	56394385
99	MAR2020	3	38294750	1.056866	0.652691	24994651
100	APR2020	4	35607583	0.953613	0.143274	5101652
101	MAY2020	5	32758583	0.994426	0.296046	9698059
102	JUN2020	6	30283417	0.997237	0.721132	21838344
103	JUL2020	7	27889708	0.896105	1.082982	30204047
104	AUG2020	8	25371042	0.660080	1.236466	31370433
105	SEP2020	9	24475917	1.006847	1.130561	27671526
106	OCT2020	10	26041000	1.146126	1.007396	26233587
107	NOV2020	11	28349375	1.108332	0.970481	27512522
108	DEC2020	12	30115917	1.057578	0.971211	29248895
109	JAN2021	1	31070333	1.051584	0.895784	27832299
110	FEB2021	2	31581500	1.071206	0.895618	28284944
111	MAR2021	3	32358417	1.056866	1.024723	33158410
112	APR2021	4	33628833	0.953613	1.053950	35443111
113	MAY2021	5	35086000	0.994426	1.034983	36313419
114	JUN2021	6	36343958	0.997237	1.040406	37812484
115	JUL2021	7	37291917	0.896105	1.018689	37988849
116	AUG2021	8	38329792	0.660080	1.027599	39387642
117	SEP2021	9	39676167	1.006847	1.031720	40934710
118	OCT2021	10	40948292	1.146126	1.005776	41184821
119	NOV2021	11	42071542	1.108332	1.036453	43605177
120	DEC2021	12	43101667	1.057578	0.949644	40931244
121	JAN2022	1	43818958	1.051584	0.860777	37718338
122	FEB2022	2	44345167	1.071206	0.943355	41833226
123	MAR2022	3	45005708	1.056866	1.110985	50000657
124	APR2022	4	45789875	0.953613	1.065594	48793395
125	MAY2022	5	46486333	0.994426	1.088926	50620169
126	JUN2022	6	47226000	0.997237	1.023537	48337568

127	JUL2022	7	48253833	0.896105	0.942658	45486853
128	AUG2022	8	49307042	0.660080	0.980410	48341100
129	SEP2022	9	50105875	1.006847	1.014055	50810090
130	OCT2022	10	50781667	1.146126	0.963538	48930041
131	NOV2022	11	51436750	1.108332	0.985231	50677067
132	DEC2022	12	52145833	1.057578	0.964708	50305490
133	JAN2023	1	52810958	1.051584	0.979884	51748601
134	FEB2023	2	53297500	1.071206	0.969215	51656734
135	MAR2023	3	53641083	1.056866	1.084696	58184285
136	APR2023	4	54050208	0.953613	1.049610	56731629
137	MAY2023	5	54531833	0.994426	1.078597	58817864
138	JUN2023	6	54974708	0.997237	1.040990	57228134
139	JUL2023	7	55378208	0.896105	0.964399	53406683
140	AUG2023	8	55819750	0.660080	0.990324	55279651
141	SEP2023	9	55976958	1.006847	0.972742	54451159
142	OCT2023	10	56302042	1.146126	0.964415	54298560
143	NOV2023	11	56885708	1.108332	0.976602	55554668
144	DEC2023	12	57145333	1.057578	0.966730	55244130
145	JAN2024	1	57435958	1.051584	0.974838	55990772
146	FEB2024	2	.	1.071206	.	57384858
147	MAR2024	3	.	1.056866	.	55948430
148	APR2024	4	.	0.953613	.	67391093
149	MAY2024	5	.	0.994426	.	62682406
150	JUN2024	6	.	0.997237	.	59622751
151	JUL2024	7	.	0.896105	.	58525503



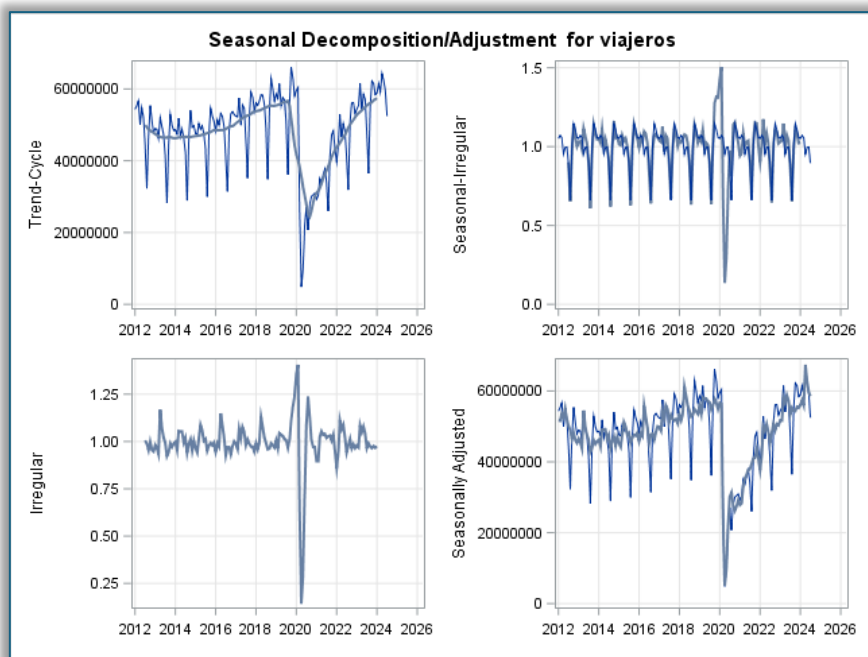
El coeficiente de estacionalidad más alta se asigna al mes de octubre con un índice de 1,15. Esto se podría interpretar como que en el mes de octubre se ejerce una influencia estacional positiva del 15% por encima del promedio de la serie temporal, como consecuencia de un aumento de número de viajeros que utilizan metro tras volver de las vacaciones para continuar sus trabajos, estudios, etc.

Por otra parte, el coeficiente de estacionalidad más baja es el mes de agosto, con un índice de 0,66. Por lo tanto, muestra una caída del 34% de la afluencia en los vagones en comparación con el promedio de la serie. Este suceso, se puede explicar desde el ámbito sociodemográfico de Madrid, que al ser una gran ciudad en la cual concentra grandes números de empresas , universidades, etc.; el mes de agosto coinciden con las vacaciones para los trabajadores y estudiantes, dando el resultado a la salida de la población de Madrid hacia otras comunidades autónomas o países para disfrutar de sus vacaciones.

1.4. Representar la serie desestacionalizada, la estimación de la tendencia y el error. ¿Qué dirías sobre la tendencia? ¿Y sobre su comportamiento estacional? (0.5)

```
proc timeseries data series.metro plots=(decomp series)
print=(seasons decomp);
id fecha interval=month;
var viajeros;
run;
```

La tabla con los datos de la serie desestacionalizado está presente en el apartado anterior. En este caso, solo analizaré la parte de la gráfica.

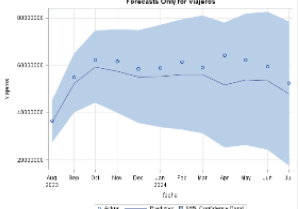


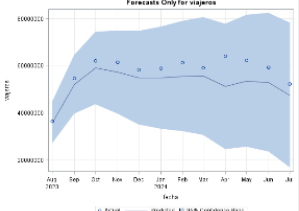
En general, la serie muestra una ligera tendencia creciente, exceptuando el periodo de la pandemia que sufrió una caída drástica de pasajeros transportados. Además, la serie también presenta un comportamiento estacional muy marcadas a lo largo de cada año registrado.

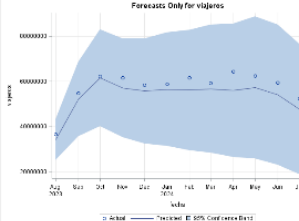
2. Encontrar el método de suavizado más adecuado a la serie teniendo en cuenta las características de la serie y comparando las medidas de bondad de ajuste. Para el método elegido:

En los dos modelos predictivos que vamos a construir, para comparar las predicciones con los datos reales vamos a utilizar la opción *BACK=n* para reservar los datos del último periodo completo observado y división en *TRAIN* y *TEST* en el modelo *ARIMA*

Debido a que la serie a estudiar presenta un comportamiento estacional, se mostrara el método de suavizado “seasonal”, “Holt-Winter Aditivo” y “Holt-Winter Multiplicativo” para realizar la comparación y obtener el método de suavizado más adecuado para la predicción.

Seasonal	MSE	R-Square	Gráfica
<pre>proc esm data=series.metro lead=12 back=12 print= all plots=all; id fecha interval=month; forecast viajeros /model=seasonal; run;</pre>	3.17487E13	0.37246457	

Holt-Winter Aditivo	MSE	R-Square	Gráfica
<pre>proc esm data=series.metro lead=12 back=12 print= all plots=all; id fecha interval=month; forecast viajeros /model=addwinters; run;</pre>	3.48347E13	0.31146761	

Holt-Winter Multiplicativo	MSE	R-Square	Gráfica
<pre>proc esm data=series.metro lead=12 back=12 print= all plots=all; id fecha interval=month; forecast viajeros /model=winters; run;</pre>	1.87301E13	0.62978765	

Tras realizar una comparación entre los 3 modelos, se puede observar que el método con menor error cuadrático medio y el coeficiente de R^2 más cercano al 1 es el método de suavizado Holt-Winter Multiplicativo.

Por lo tanto, el primer modelo predictivo seleccionado es el Holt-Winter Multiplicativo.

2.1. Sobre las tablas de estimadores de los coeficientes del modelo ¿Cuánto vale la estimación de los parámetros de suavizado? ¿Qué podemos decir sobre los coeficientes del modelo de suavizado? Escribir las ecuaciones del modelo. (1)

```
proc esm data=series.metro lead=12 back=12
print= all
plots=all;
id fecha interval=month;
forecast viajeros /model=winters;
run;
```

Winters Method (Multiplicative) Parameter Estimates				
Parameter	Estimate	Standard Error	t Value	Approx Pr > t
Level Weight	0.99900	0.06527	15.30	<.0001
Trend Weight	0.0010000	0.03161	0.03	0.9748
Seasonal Weight	0.99900	65.76701	0.02	0.9879

En este caso, el modelo de suavizado devuelve 3 valores estimados según los parámetros de nivel, tendencia y estacional. El peso de nivel muestra un coeficiente de 0,999 aprox., el peso del nivel de tendencia un 0,001 y el peso de la estacionalidad un 0,999 aprox.

Estos valores devueltos, nos indica que el modelo está muy influenciado por las observaciones más recientes de la serie gracias al alto peso del nivel junto con un p-value < 0,05 siendo este un valor significativo para el modelo. También se ajusta rápidamente a los patrones estacionales; aunque el p-value muestra que no es significativa, se puede deber a que la caída provocada en 2020, el modelo lo considere como un valor atípico, de allí vendrá el alto valor de error estándar. Sin embargo, este modelo no está capturando muy bien la tendencia de la serie valor del parámetro estimado bajo y p-value muy alto, probablemente considera que el componente de la tendencia no es lo suficientemente importante y por la culpa del año 2020, los datos hasta la última observación seguía sin superar los valores prepandemia, y por eso no muestra tendencia en la serie para este modelo.

- Ecuaciones del modelo:

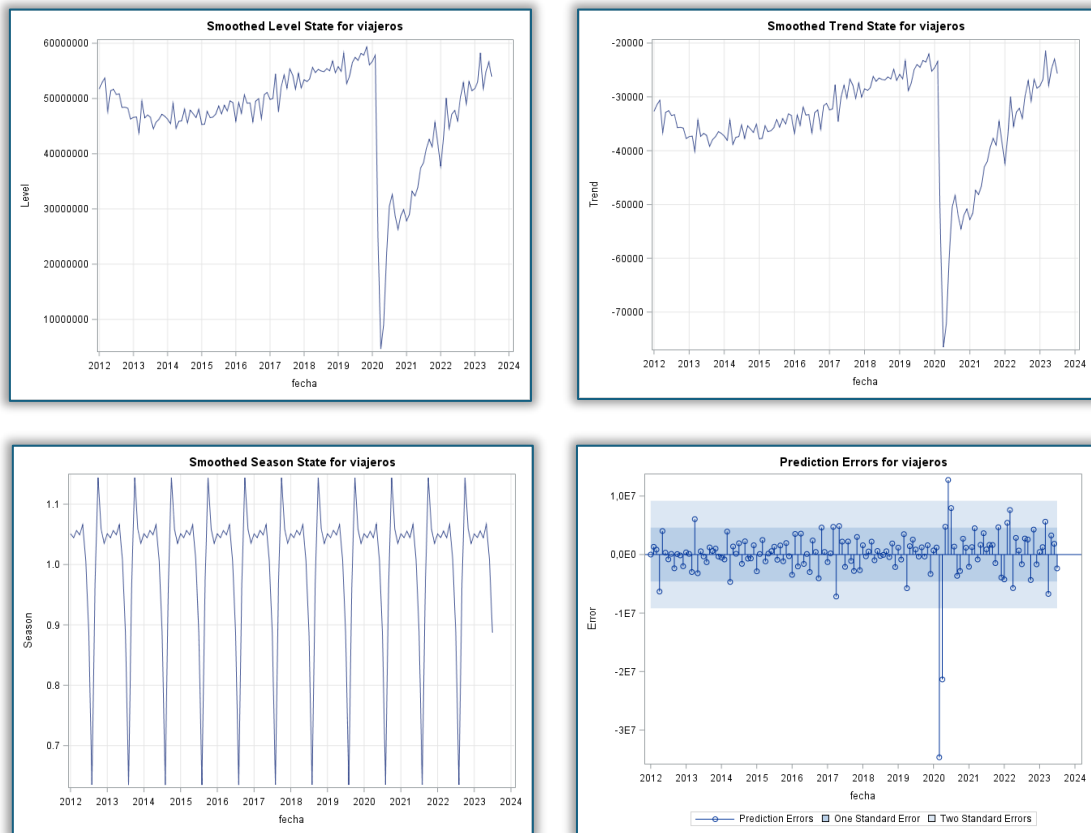
$$L_t = 0,999 * \frac{x_t}{S_{t-s}} + (1 - 0,999)(L_{t-1} + b_{t-1})$$

$$b_t = 0,001(L_t - L_{t-1}) + (1 - 0,001) * b_{t-1}$$

$$S_t = 0,999 * \frac{x_t}{L_t} + (1 - 0,999) * S_{t-s}$$

$$\hat{x}_{t+k} = (L_t + kb_t) * S_{t+k-s}$$

2.2. Comentar los gráficos de las componentes del modelo en relación al apartado anterior. (0.5)



En estas cuatro gráficas, se puede comprobar el estudio descrito en el apartado anterior. Por un lado, el modelo se ajusta de manera rápida a los cambios de datos recientes. Además, se puede observar un claro patrón de estacionalidad para el modelo, por lo que explica su alto valor estimado del parámetro de la estacionalidad.

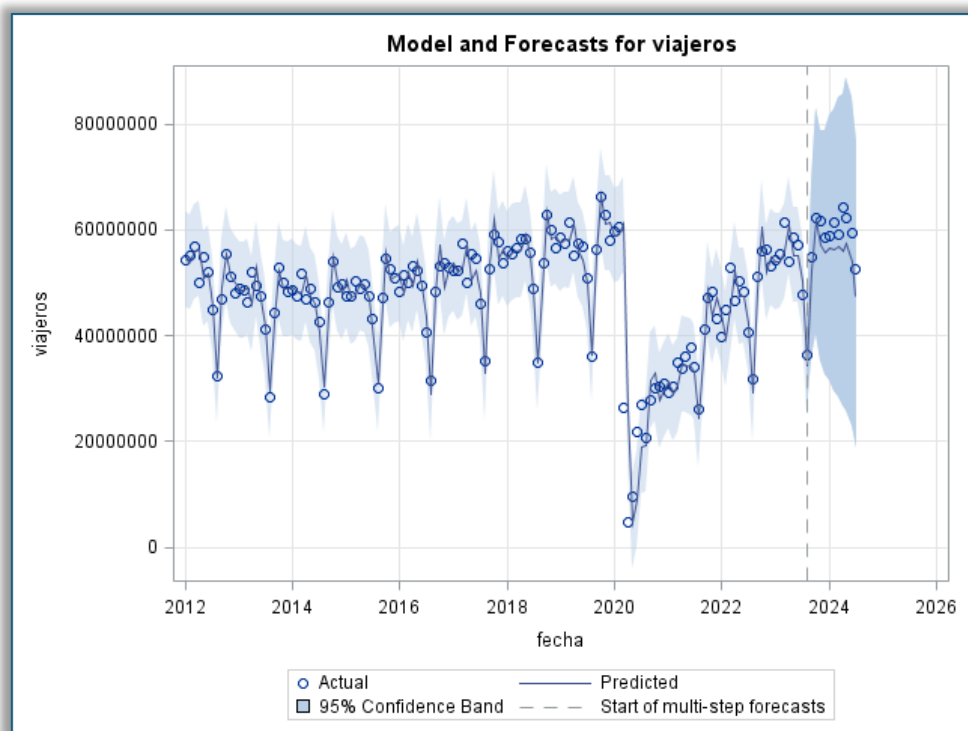
En la gráfica de la tendencia, debido a que el descenso del número de viajeros ha sucedido en una instancia por corto de tiempo con una caída brusca para el rango de tiempo que se utiliza en el estudio. Por ello, el modelo habrá considerado esta caída no como una tendencia de la serie, sino unas observaciones atípicas de la serie, y esto se puede contrastar con la gráfica de error en la cual muestra los puntos atípicos fuera del rango de en período del año de 2020.

2.3. Calcular las medidas de bondad de ajuste y de comportamiento sobre los datos reservados para ello. Comentar su resultado. (0.5)

Performance Statistics for Variable viajeros	
Statistic	Value
Number of Observations	12
Number of Observations Used	12
Mean Square Error	1.87301E13
Root Mean Square Error	4327823.41
R-Square	0.62978765

Para realizar la predicción con el modelo, se ha reservado los últimos 12 meses para su contrastación. Para la predicción se ha generado un error cuadrático medio de $1,87 \cdot 10^{13}$. Además, el valor de R^2 es de 0,63; por lo que se puede interpretar como que el modelo usado tiene la capacidad de explicar el 63% de la varianza del variable viajeros.

2.4. Mostrar la gráfica donde aparecen los valores observados, los suavizados y las predicciones con los intervalos de confianza. (0.5)



En esta gráfica se muestra los datos reales junto con la línea de predicción realizado por el modelo y las bandas de confianza, tanto para el periodo "train" como para el periodo reservado para el contraste "test".

2.5. Mostrar una tabla y gráfico en donde aparecen los valores observados, las predicciones y sus intervalos de confianza, sólo para el último periodo que habíamos reservado. ¿Cómo se comportan los intervalos de confianza? (1)

```
proc esm data=series.metro lead=12 back=12
```

```
print= all
```

```
plots=all
```

```
outfor=series.pred_hwm;
```

```
id fecha interval=month;
```

```
forecast viajeros /model=winters;
```

```
run;
```

```
proc sql;
```

```
create table series.pred_hwm_12 as
```

```
select *
```

```
from series.pred_hwm
```

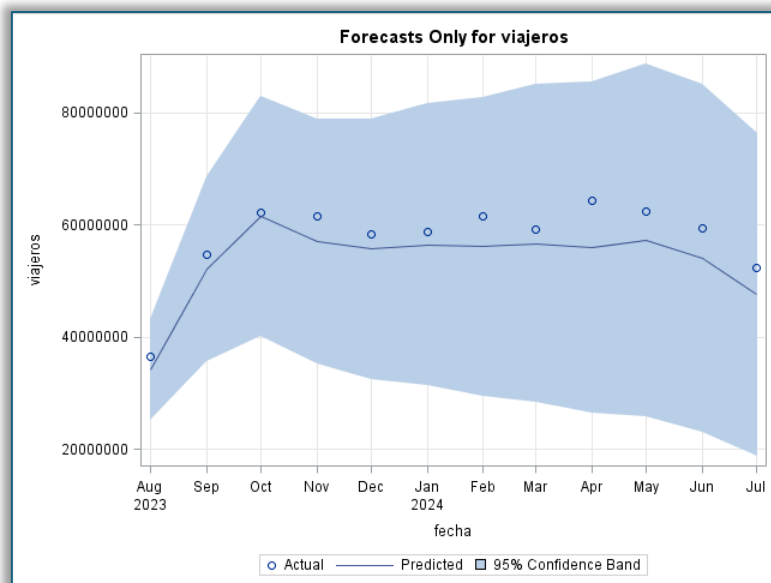
```
where fecha >= '01AUG2023'd;
```

```
quit;
```

```
proc print data=series.pred_hwm_12 label;
```

```
run;
```

Obs	Variable Name	fecha	Actual Values	Predicted Values	Lower Confidence Limits	Upper Confidence Limits	Prediction Errors	Prediction Standard Errors
1	viajeros	AUG2023	36489000	34249497.96	25250348.69	43248647.22	2239502.04	4591487.06
2	viajeros	SEP2023	54824000	52155437.74	35755368.88	68555506.60	2668562.26	8367535.83
3	viajeros	OCT2023	62233000	61606755.13	40225900.11	82987610.15	626244.87	10908799.95
4	viajeros	NOV2023	61573000	57024034.58	35260251.63	78787817.52	4548965.42	11104174.93
5	viajeros	DEC2023	58425000	55706803.64	32594113.77	78819493.51	2718196.36	11792405.40
6	viajeros	JAN2024	58879000	56529984.55	31379598.69	81680370.40	2349015.45	12832065.31
7	viajeros	FEB2024	61471000	56171919.40	29578281.93	82765556.88	5299080.60	13568431.71
8	viajeros	MAR2024	59130000	56729591.45	28369742.13	85089440.77	2400408.55	14469576.76
9	viajeros	APR2024	64265000	56098829.03	26609150.31	85588507.75	8166170.97	15046030.92
10	viajeros	MAY2024	62333000	57278844.76	25812987.31	88744702.21	5054155.24	16054303.91
11	viajeros	JUN2024	59458000	54078454.18	22998043.41	85158864.94	5379545.82	15857643.82
12	viajeros	JUL2024	52445000	47584982.02	18757902.01	76412062.03	4860017.98	14707964.15



La predicción realizada por este modelo es capaz de seguir los movimientos de los valores reales, es decir que la predicción es acertada; Por lo tanto, todos los valores de los últimos 12 meses están dentro de las bandas de confianzas.

De forma genérica, la predicción ha sido capaz de seguir la mayoría de los cambios presente de los datos reales; esto también se puede reflejar en la última columna del error estándar de predicción, en la cual va incrementando el error hasta el mes de mayo 2024 pero a partir de ese mes empieza a descender, por lo que podemos entender que el modelo de predicción sigue manteniendo cierta precisión en el pronóstico sin aumentar el error de forma incontrolada en el tiempo.

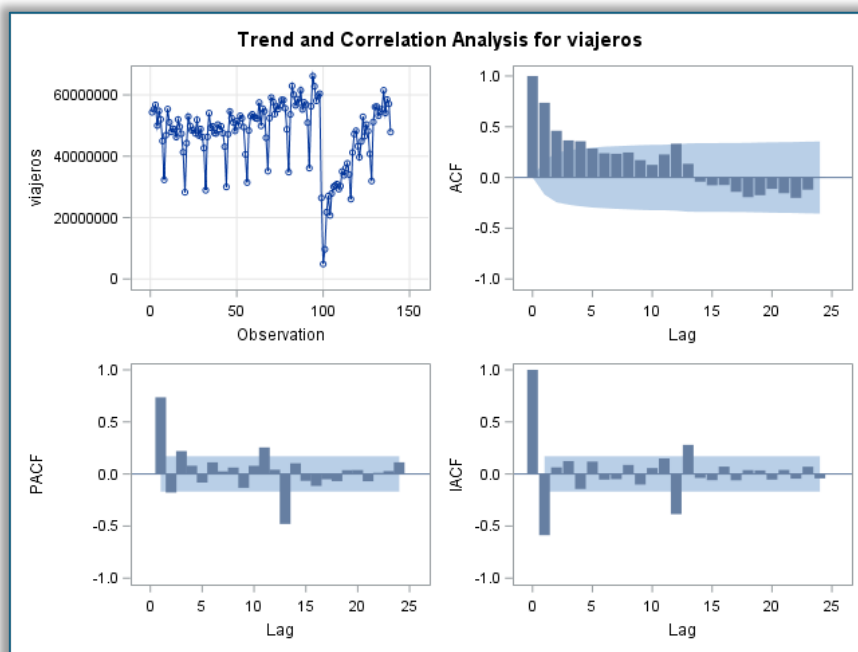
3. Para ajustar un modelo ARIMA:

```
data series.metro_train series.metro_test;  
set series.metro;  
if fecha<'01AUG2023'd then output series.metro_train;  
else output series.metro_test ;  
run;
```

Se divide la serie de “metro” en 2 series, una para el entrenamiento del modelo ARIMA denominado “train” y los últimos 12 meses de la serie, se reserva para contrastar con los valores resultantes de la predicción denominado “test”.

3.1. Representar la serie y los correlogramas. Hacer las diferenciaciones que sean necesarias. Decidir qué modelo puede ser ajustado. (0.5)

```
proc arima data=series.metro_train;  
identify var=viajeros nlag=24;  
run;
```



La gráfica de la serie muestra una ligera tendencia positiva que se observa incluso después del año 2020, donde hay una tendencia de recuperación a los niveles previos a la pandemia. Esta característica implica que la serie no presenta una media constante e igual a cero, lo cual es indicativo de una serie no estacionaria.

Además, en la ACF, se aprecia una caída lenta hasta llegar a cero, lo cual es un signo de la presencia de tendencia en la serie. Este comportamiento indica que los valores presentes aún están correlacionados con los valores pasados, mostrando persistencia en la tendencia.

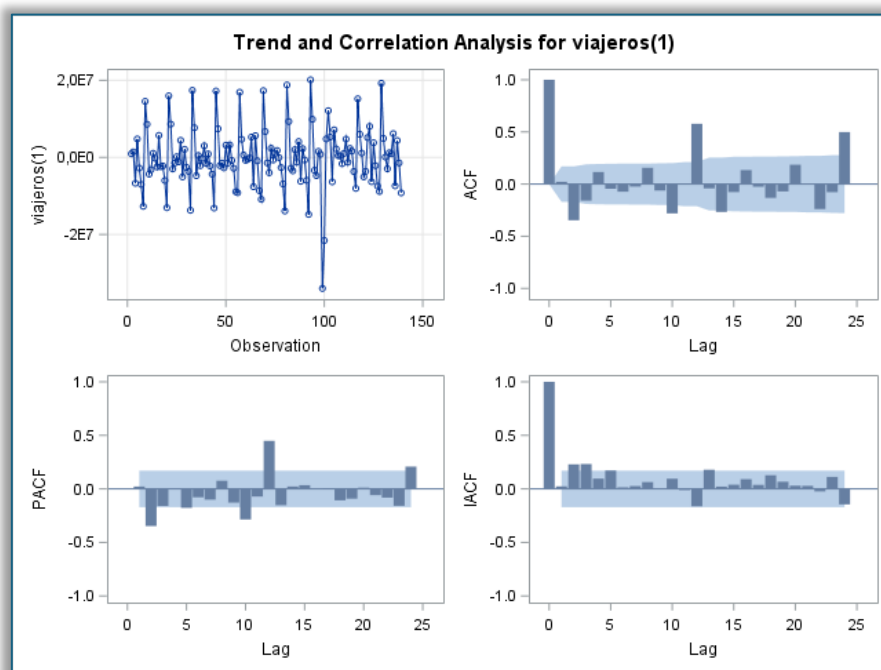
En cuanto a la PACF, se observa un corte claro después del primer lag, lo cual respalda la existencia de una tendencia en la serie. Asimismo, se nota un pico significativo en el lag 12, que sobresale de las bandas de confianza, lo que sugiere la presencia de estacionalidad con un periodo anual.

Por lo tanto, la presencia tanto de tendencia como de estacionalidad indica que es necesario aplicar una diferenciación en la parte estándar ($d=1$) para eliminar la tendencia, así como una diferenciación en la parte estacional ($D=1$) para eliminar el componente estacional. Esto nos permite ajustar adecuadamente el modelo y convertir la serie en estacionaria, condición necesaria para el uso del modelo ARIMA.

Efectuaremos en primer lugar una diferenciación en la parte estándar:

ARIMA (0,1,0)(0,0,0)₁₂

```
proc arima data=series.metro_train;  
identify var=viajeros(1) nlag=24;  
  
run;
```

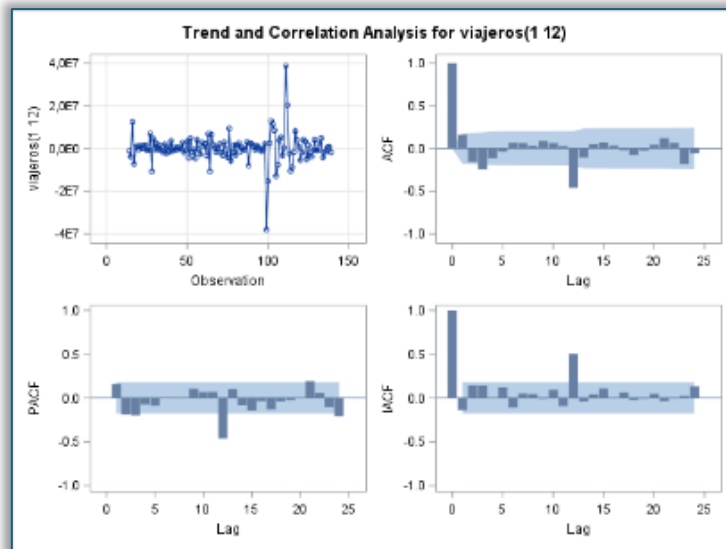


Una realizado la diferenciación en la parte estándar, los valores se mueven en torno a una media constante. Tras el proceso, se puede apreciar con más claridad que los lags más pronunciados están en el orden 12, por lo que se le vamos a añadir la diferenciación en la parte estacional para quitar el componente estacional en el modelo.

ARIMA (0,1,0)(0,1,0)₁₂

```
proc arima data=series.metro_train;
identify var=viajeros(1 12) nlag=24;

run;
```



Una vez quitado el componente tendencia y estacional, la serie ya se comporta como estacionario, con una media constante y sin patrones estacionales; cumpliendo así la condición para aplicar el modelo ARIMA. Además, se observa un lags que sobresale en el orden 12 en las gráficas de ACF y PACF, por lo que se tendrá que ajustar el modelo en la parte estacional.

Como estamos trabajando con datos de números de viajeros transportado por metro en Madrid, es menos probable que esté relacionado con el termino autorregresivo estacional (P=1) en la cual el número de viajeros de hace 12 meses esté directamente relacionado, ya que las condiciones se podrán haber cambiado como la pandemia, clima, etc.; que influiría directamente con el número de pasajeros que transporta. Lo más probable es que esté relacionado con la parte de media móvil estacional (Q=1), en la cual el error de hace 12 meses tenga alguna influencia en los números de viajeros actuales, en este caso una de las causas principales podría ser la caída de números en 2020.

Por lo tanto, lo más probable es que el modelo se ajuste mejor con un Q=1 antes que P=1. Pero se probará ambas para determinar cuál de ellos es mejor para el modelo.

Autocorrelation Check for White Noise									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	16.88	6	0.0097	0.160	-0.158	-0.245	-0.113	-0.036	0.068
12	49.25	12	<.0001	0.063	0.034	0.085	0.062	0.027	-0.461
18	53.00	18	<.0001	-0.106	0.048	0.072	0.036	-0.022	-0.074
24	61.91	24	<.0001	-0.025	0.043	0.119	0.066	-0.181	-0.060

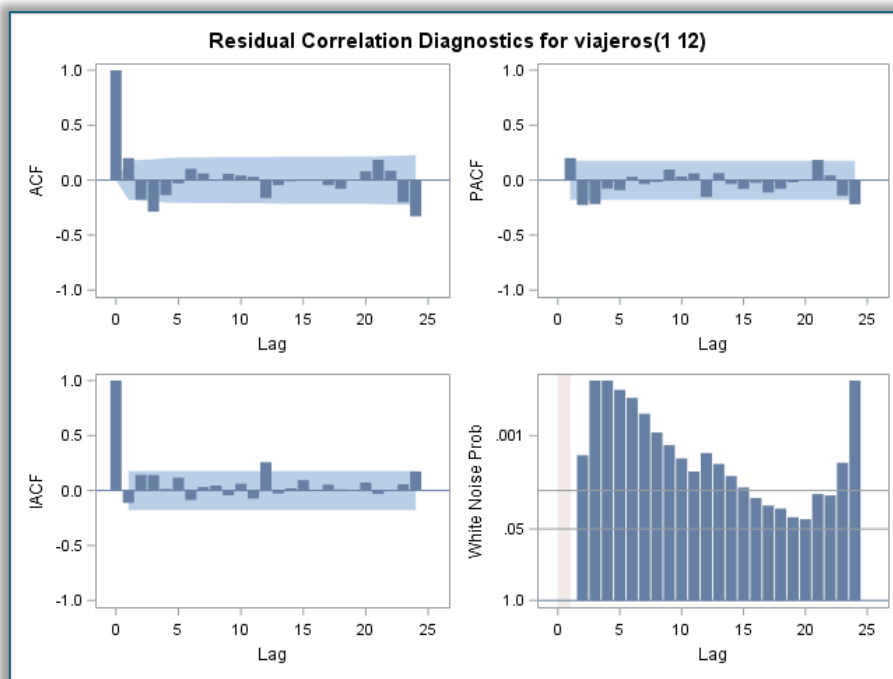
Con la tabla de autocorrelación se puede observar que el p-value es significativo siendo para todos los lags menor que 0,05. De esta forma rechaza la Ho de que los residuos sean ruidos blancos; por lo que en este caso presenta autocorrelación significativos. Esto significa que debemos seguir ajustando el modelo.

3.2. Ajustar el modelo adecuado comprobando que sus residuales están incorrelados. (Sintaxis, tablas de los parámetros estimados y gráficos)

Probaremos primero con P=1:

ARIMA (0,1,0)(1,1,0)₁₂

```
proc arima data=series.metro_train;
identify var=viajeros(1 12) nlag=24;
estimate p=(12);
run;
```



Fijando en las gráficas de ACF y PACF en los lags estacional de 12, se puede observar que ha mejorado el lag=12, pero como consecuencia de ella, ha salido un nuevo lag=24.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	24.12	5	0.0002	0.201	-0.177	-0.287	-0.136	-0.029	0.104
12	29.24	11	0.0021	0.061	0.008	0.058	0.043	0.030	-0.163
18	30.77	17	0.0213	-0.045	-0.009	-0.003	-0.002	-0.046	-0.078
24	61.57	23	<.0001	0.001	0.080	0.185	0.086	-0.200	-0.329

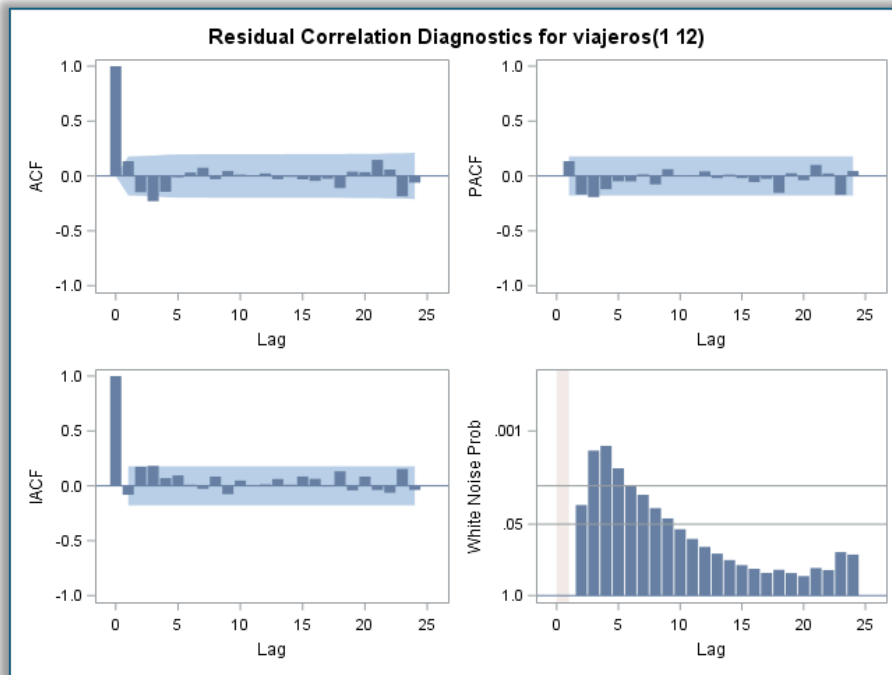
En cuanto a la autocorrelación, aunque ha mejorado los resultados, esta mejora sigue sin cambiar el resultado final, aun presentan autocorrelación en los lags.

Ahora vamos a probar con Q=1.

ARIMA (0,1,0)(0,1,1)₁₂

```
proc arima data=series.metro_train;
identify var=viajeros(1 12) nlag=24;
estimate q=(12);

run;
```



Se puede ver que ha mejorado los lags pronunciado en la parte estacional, ya los lags múltiplo de 12 están todos dentro de las bandas de confianzas.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	15.06	5	0.0101	0.135	-0.148	-0.230	-0.144	-0.012	0.032
12	16.29	11	0.1306	0.075	-0.029	0.044	0.012	-0.002	0.022
18	18.80	17	0.3404	-0.029	-0.008	-0.030	-0.045	-0.025	-0.111
24	29.00	23	0.1802	0.039	0.034	0.146	0.058	-0.186	-0.061

Revisando la tabla de autocorrelación, vemos que a partir del lag=6 el p-value ya son mayores que 0,05. Esto muestra que ya no presenta autocorrelación para estos lags y que los residuos ya empiezan a comportarse como ruidos blancos.

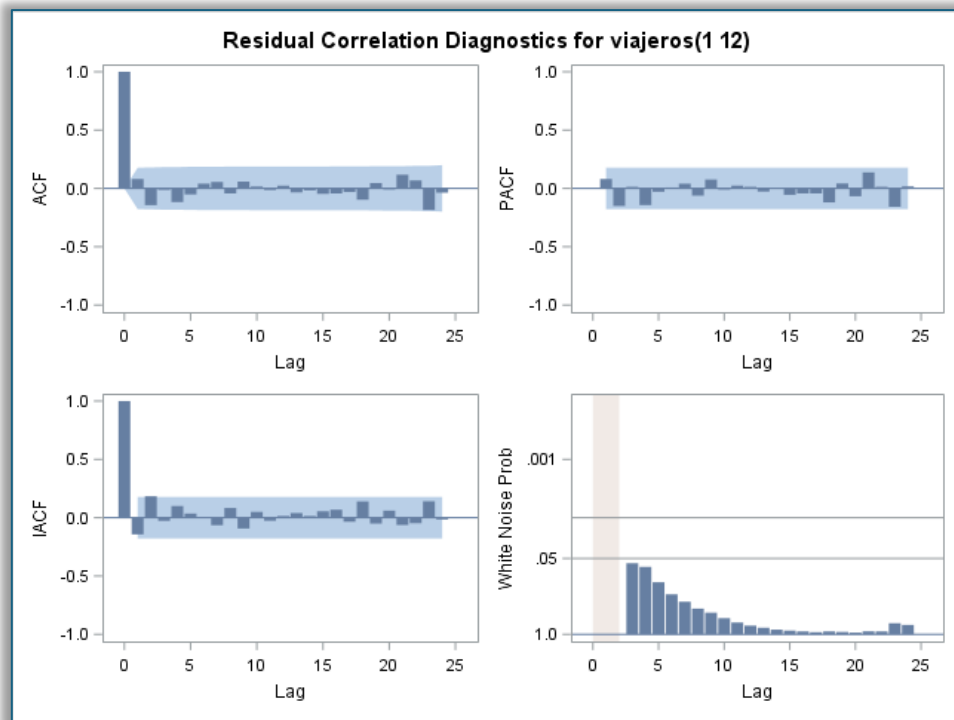
Por lo tanto, es mejor para el modelo utilizar la componente media móviles estacional (MA) en el modelo con (Q=1) antes que el componente autorregresivos(AR) con (P=1).

Ahora solo tendremos que centrarnos en el pico significativos en lag=3. Como no podemos determinar si este podría estar influencia más por el termino autorregresivo (AR) o media móviles(MA). Aunque suele funcionar mejor para lags pequeños el AR antes que MA, vamos a probar con ambos: p=3 y q=3 para comparar y sacar el mejor modelo.

Probaremos primero con q=3:

ARIMA (0,1,3)(0,1,1)₁₂

```
proc arima data=series.metro_train;
identify var=viajeros(1 12) nlag=24;
estimate q=(3)(12);
run;
```



Se puede observar que se ha mejorado el resultado, ahora ya todos los lags están dentro de las bandas de confianzas.

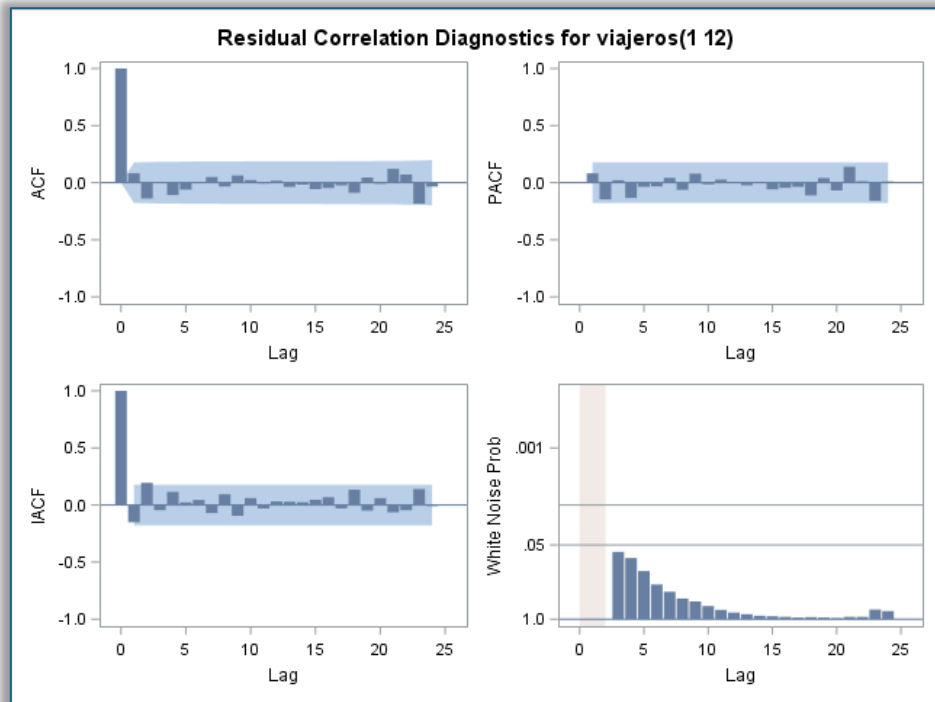
Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	5.88	4	0.2085	0.081	-0.143	-0.012	-0.116	-0.051	0.041
12	7.15	10	0.7117	0.055	-0.043	0.058	0.016	-0.013	0.023
18	9.49	16	0.8920	-0.034	-0.016	-0.047	-0.043	-0.031	-0.097
24	18.25	22	0.6911	0.046	-0.010	0.117	0.068	-0.186	-0.037

Fijándonos en la tabla de autocorrelación, se puede ver que p-value es mayor que 0,05 en todos los lags, por lo tanto los residuos se comportan como ruidos blancos y ya no existen autocorrelación entre ellos.

Ahora vamos a probar con p=3:

ARIMA (3,1,0)(0,1,1)₁₂

```
proc arima data=series.metro_train;
identify var=viajeros(1 12) nlag=24;
estimate p=(3) q=(12);
run;
```



Al igual que $q=3$, utilizando $p=3$, también se ha mejorado los resultados; todos los lags se encuentran dentro de las bandas de confianzas. Por lo tanto ya no queda dependencias en la serie que no haya sido capturado por el modelo.

Autocorrelation Check of Residuals									
To Lag	Chi-Square	DF	Pr > ChiSq	Autocorrelations					
6	5.39	4	0.2494	0.082	-0.138	-0.004	-0.108	-0.060	-0.003
12	6.55	10	0.7674	0.050	-0.033	0.063	0.024	-0.007	0.018
18	8.82	16	0.9207	-0.037	-0.017	-0.056	-0.047	-0.024	-0.089
24	17.70	22	0.7233	0.045	-0.010	0.120	0.071	-0.185	-0.034

Fijándonos en la tabla de autocorrelación, vemos que este modelo ofrece unos resultados mejores que el anterior. Los valores de p-values obtenido son mayores que 0,05, esto quiere decir que los residuos se muestran como un ruido blanco y ya no existe autocorrelación entre ellos. Además, los resultados obtenidos del p-value con $p=3$ es ligeramente mayor que $q=3$. Por lo tanto, a efectos similares de comportamiento en las gráficas, y una ligera mejora de los resultados de autocorrelación, el modelo más óptimo para realizar la predicción es la siguiente:

ARIMA (3,1,0)(0,1,1)₁₂

Ahora se va a mostrar algunos datos devuelto por el modelo seleccionado:

- Tabla de correlación de los parámetros estimados:

Correlations of Parameter Estimates			
Parameter	MU	MA1,1	AR1,1
MU	1.000	0.043	-0.005
MA1,1	0.043	1.000	0.088
AR1,1	-0.005	0.088	1.000

- Tabla de parámetros estimados

Conditional Least Squares Estimation					
Parameter	Estimate	Standard Error	t Value	Approx Pr > t 	Lag
MU	69413.0	65059.8	1.07	0.2881	0
MA1,1	0.95117	0.04399	21.62	<.0001	12
AR1,1	-0.23278	0.08823	-2.64	0.0094	3

Constant Estimate	85571.16
Variance Estimate	2.408E13
Std Error Estimate	4907358
AIC	4242.91
SBC	4251.419
Number of Residuals	126

- Otras informaciones del modelo ajustado:

Model for variable viajeros	
Estimated Mean	69413
Period(s) of Differencing	1,12

- Factores autorregresivos

Autoregressive Factors	
Factor 1:	1 + 0.23278 B**(3)

- Factores medias móviles:

Moving Average Factors	
Factor 1:	1 - 0.95117 B**(12)

3.3. Escribir la expresión algebraica del modelo ajustado con los parámetros estimados. (0,5)

- Expresión algebraico genérico para ARIMA (3,1,0)(0,1,1)₁₂

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3)(1 - B)(1 - B^{12})X_t = \hat{\delta} + (1 + \theta_1 B^{12})Z_t$$

- Sustituir las expresiones algebraicas de los factores autorregresivos y medias móviles

$$(1 - \phi_1 B - \phi_2 B^2 - \phi_3 B^3) = 1 + 0,23B^3$$

$$(1 + \theta_1 B^{12}) = 1 - 0,95B^{12}$$

- Añadimos en la ecuación principal

$$(1 + 0,23B^3)(1 - B)(1 - B^{12})X_t = 85571,16 + (1 - 0,95B^{12})Z_t$$

$$(1 - B)(1 - B^{12}) = 1 - B - B^{12} + B^{13}$$

$$(1 + 0,23B^3)(1 - B - B^{12} + B^{13})X_t = 85571,16 + (1 - 0,95B^{12})Z_t$$

- Despejar los paréntesis

$$1 * (1 - B - B^{12} + B^{13}) = 1 - B - B^{12} + B^{13}$$

$$0,23B^3 * (1 - B - B^{12} + B^{13}) = 0,23B^3 - 0,23B^4 - 0,23B^{15} + 0,23B^{16}$$

- Expresión parte X_t

$$1 - B - B^{12} + B^{13} + 0,23B^3 - 0,23B^4 - 0,23B^{15} + 0,23B^{16}$$

- Formamos la ecuación

$$(1 - B - B^{12} + B^{13} + 0,23B^3 - 0,23B^4 - 0,23B^{15} + 0,23B^{16}) * X_t = 85571,16 + (1 - 0,95B^{12})Z_t$$

$$X_t - BX_t - B^{12}X_t + B^{13}X_t + 0,23B^3X_t - 0,23B^4X_t - 0,23B^{15}X_t + 0,23B^{16}X_t = 85571,16 + Z_t - 0,95B^{12}Z_t$$

- Ecuación completa

$$X_t - X_{t-1} - X_{t-12} + X_{t-13} + 0,23X_{t-3} - 0,23X_{t-4} - 0,23X_{t-15} + 0,23X_{t-16} = 85571,16 + Z_t - 0,95Z_{t-12}$$

$$X_t = X_{t-1} - 0,23X_{t-3} + 0,23X_{t-4} + X_{t-12} - X_{t-13} + 0,23X_{t-15} - 0,23X_{t-16} + Z_t - 0,95Z_{t-12} + 85571,16$$

3.4. Calcular las predicciones y los intervalos de confianza para las unidades de tiempo que se considere oportuno, dependiendo de la serie, siguientes al último valor observado. Representarlas gráficamente. (1)

```
proc arima data=series.metro_train;
identify var=viajeros(1 12) nlag=24;
estimate p=(3) q=(12);
forecast lead=12 id=fecha interval=month out=series.pred_arima printall;
quit;

data union;
merge series.metro_test series.pred_arima(drop=viajeros); where fecha>='01AUG2023'd; by
fecha;
run;

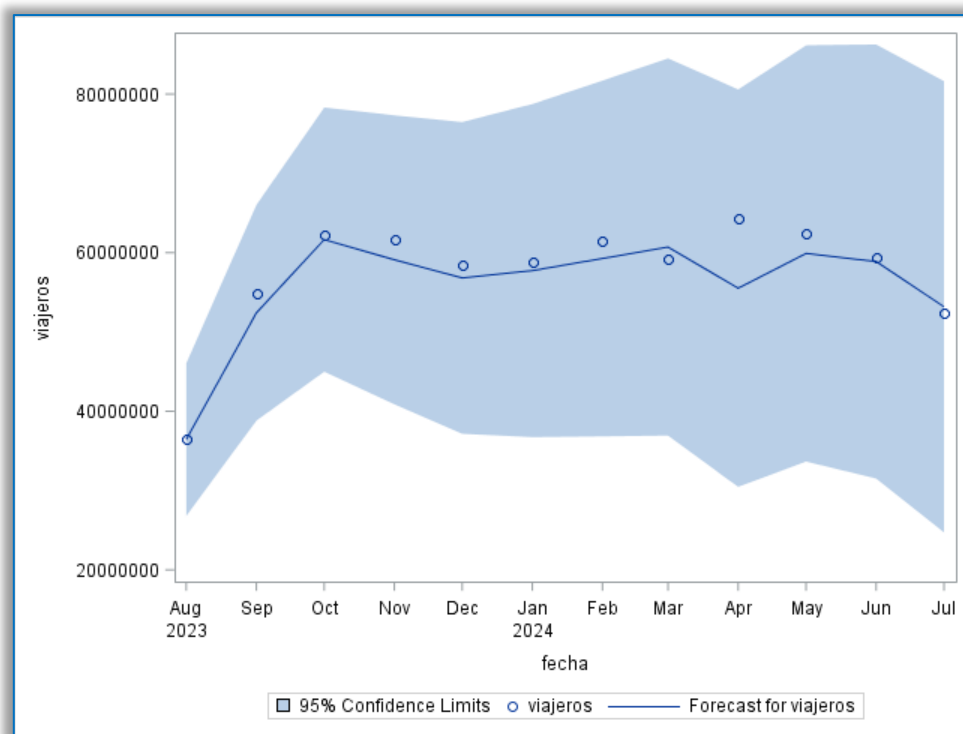
proc sgplot data=union; band Upper=u95 Lower=l95 x=fecha
/ LegendLabel="95% Confidence Limits"; scatter x=fecha y=viajeros; series x=fecha
y=forecast;
run;

proc print data=union label;
var fecha viajeros Forecast Std L95 U95;
run;

proc sql;
create table series.pred_arima_12 as
select fecha, viajeros, forecast, std, l95, u95
from union;
quit;
```


Obs	fecha	viajeros	Forecast for viajeros	Std Error of Forecast	Lower 95% Confidence Limit	Upper 95% Confidence Limit
1	01AUG2023	36489000	36429860.34	4907357.96	26811615.47	46048105.21
2	01SEP2023	54824000	52410507.76	6940052.19	38808255.42	66012760.11
3	01OCT2023	62233000	61652045.26	8499793.33	44992756.47	78311334.06
4	01NOV2023	61573000	59106592.85	9296331.24	40886118.43	77327067.27
5	01DEC2023	58425000	56815883.54	10029808.69	37157819.73	76473947.35
6	01JAN2024	58879000	57741042.58	10713185.82	36743584.21	78738500.94
7	01FEB2024	61471000	59266866.01	11446428.72	36832277.96	81701454.06
8	01MAR2024	59130000	60708866.79	12135448.51	36923824.77	84493908.82
9	01APR2024	64265000	55521400.69	12787395.78	30458565.51	80584235.87
10	01MAY2024	62333000	59902153.33	13389199.42	33659804.67	86144501.98
11	01JUN2024	59458000	58886779.81	13965093.33	31515699.84	86257859.77
12	01JUL2024	52445000	53171168.28	14518161.11	24716095.38	81626241.19

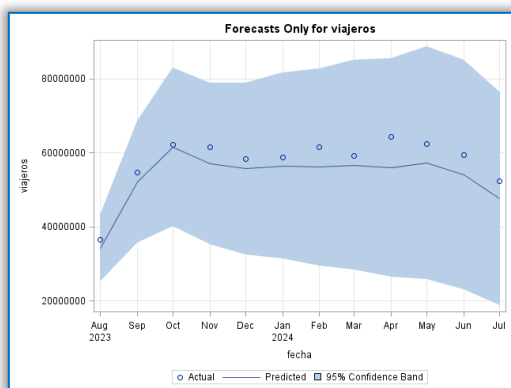
En esta tabla se muestra los valores de predicción de los últimos 12 meses con los valores reales de viajeros. Asimismo, se muestra el error estándar de predicción y las bandas de confianza al 95%.



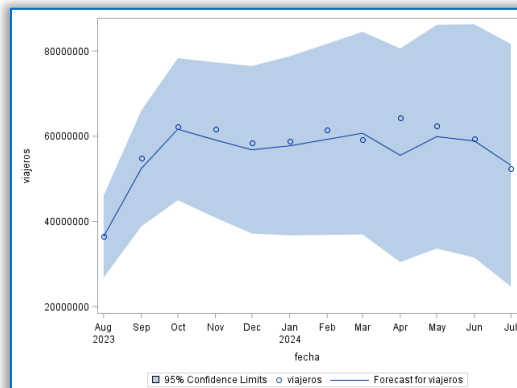
Esta gráfica muestra la predicción de viajeros en los últimos 12 meses juntos con los valores reales, en la cual se puede observar que exceptuando 1 punto en el mes de abril, en los demás puntos la predicción es capaz de seguir los movimientos de los datos reales. También muestra las bandas de confianzas al 95%.

4. Comparar las predicciones obtenidas con cada uno de los métodos con los valores observados que habíamos reservado antes. Conclusiones. (1)

Holt-Winter Multiplicativo



ARIMA (3,1,0)(0,1,1)₁₂



Realizando una comparación de las gráficas de predicción de ambos modelos, se puede observar que la predicción realizada mediante el método de ARIMA $(3,1,0)(0,1,1)_{12}$, ofrece una precisión más alta con los datos reales de los meses de predicción en comparación con el método de Holt-Winter Multiplicativo.

El método Holt-Winter Multiplicativo presenta una predicción algo más suavizado que también predice de forma correcta los datos de la serie. El método ARIMA, en el mes de abril parece ser que hubo algún cambio no detectado por el modelo que hizo que la predicción fuera inversa al movimiento de los datos reales de ese mes.

Aun así, el modelo de ARIMA $(3,1,0)(0,1,1)_{12}$ es el más preciso, con lo cual no quedaríamos con este modelo para la predicción futura.

Holt Winter Multiplicativo

Obs	Variable Name	fecha	Actual Values	Predicted Values	Lower Confidence Limits	Upper Confidence Limits	Prediction Errors	Prediction Standard Errors
1	viajeros	AUG2023	36489000	34249497.96	25250348.69	43248647.22	2239502.04	4591487.06
2	viajeros	SEP2023	54824000	52155437.74	35755368.88	68555506.60	2668562.26	8367535.83
3	viajeros	OCT2023	62233000	61606755.13	40225900.11	82987610.15	626244.87	10908799.95
4	viajeros	NOV2023	61573000	57024034.58	35260251.63	78787817.52	4548965.42	11104174.93
5	viajeros	DEC2023	58425000	55706803.64	32594113.77	78819493.51	2718196.36	11792405.40
6	viajeros	JAN2024	58879000	56529984.55	31379598.69	81680370.40	2349015.45	12832065.31
7	viajeros	FEB2024	61471000	56171919.40	29578281.93	82765556.88	5299080.60	13568431.71
8	viajeros	MAR2024	59130000	56729591.45	28369742.13	85089440.77	2400408.55	14469576.76
9	viajeros	APR2024	64265000	56098829.03	26609150.31	85588507.75	8166170.97	15046030.92
10	viajeros	MAY2024	62333000	57278844.76	25812987.31	88744702.21	5054155.24	16054303.91
11	viajeros	JUN2024	59458000	54078454.18	22998043.41	85158864.94	5379545.82	15857643.82
12	viajeros	JUL2024	52445000	47584982.02	18757902.01	76412062.03	4860017.98	14707964.15

ARIMA (3,1,0)(0,1,1)₁₂

Obs	fecha	viajeros	Forecast for viajeros	Std Error of Forecast	Lower 95% Confidence Limit	Upper 95% Confidence Limit
1	01AUG2023	36489000	36429860.34	4907357.96	26811615.47	46048105.21
2	01SEP2023	54824000	52410507.76	6940052.19	38808255.42	66012760.11
3	01OCT2023	62233000	61652045.26	8499793.33	44992756.47	78311334.06
4	01NOV2023	61573000	59106592.85	9296331.24	40886118.43	77327067.27
5	01DEC2023	58425000	56815883.54	10029808.69	37157819.73	76473947.35
6	01JAN2024	58879000	57741042.58	10713185.82	36743584.21	78738500.94
7	01FEB2024	61471000	59266866.01	11446428.72	36832277.96	81701454.06
8	01MAR2024	59130000	60708866.79	12135448.51	36923824.77	84493908.82
9	01APR2024	64265000	55521400.69	12787395.78	30458565.51	80584235.87
10	01MAY2024	62333000	59902153.33	13389199.42	33659804.67	86144501.98
11	01JUN2024	59458000	58886779.81	13965093.33	31515699.84	86257859.77
12	01JUL2024	52445000	53171168.28	14518161.11	24716095.38	81626241.19

Fijándonos en la tabla, podemos respaldar el análisis anterior con la columna de error estándar de la predicción de ambos modelos. Podemos observar que el modelo ARIMA (3,1,0)(0,1,1)₁₂ ofrece un error más pequeño comparado con el modelo Holt Winter Multiplicativo.

Por ejemplo en el mes de junio de 2024, el método de ARIMA (3,1,0)(0,1,1)₁₂ muestra un error de 13965093,33 viajeros, mientras que el modelo Holt Winter Multiplicativo muestra un error de 15857643,82 viajeros. La diferencia de error estándar que presenta entre ambos modelos para este mes, es de 1.892.550,49.

Dicho de otra forma, mediante el uso del modelo de ARIMA (3,1,0)(0,1,1)₁₂, estaríamos mejorando una precisión en la predicción de más del 10% de la cifra de viajeros, que traducido a cifra reales, sería unos 1,9 millones de viajeros transportados.

CONCLUSIÓN

De forma resumida, ambos modelos son capaces de predecir la serie de manera correcta y eficaz, pudiendo detectar en ambos casos la mayoría de las variaciones que sufren la serie a lo largo del tiempo.

Sin embargo, el modelo ARIMA (3,1,0)(0,1,1)₁₂, ofrece una precisión más alta a la hora de realizar la predicción, en este caso obteniendo un error estándar más bajos en comparación con el método Holt Winter Multiplicativo.

Por lo tanto, nos quedaríamos con ARIMA (3,1,0)(0,1,1)₁₂ como el modelo más óptimo para realizar la predicción de esta serie con los datos que disponemos actualmente.

5. ANEXO

- Serie completa en función de unidad de tiempo

fecha	viajeros
01JAN2012	54335000
01FEB2012	55280000
01MAR2012	56719000
01APR2012	49995000
01MAY2012	54774000
01JUN2012	52016000
01JUL2012	44973000
01AUG2012	32283000
01SEP2012	46817000
01OCT2012	55387000
01NOV2012	51066000
01DEC2012	47907000
01JAN2013	48968000
01FEB2013	48731000
01MAR2013	46264000
01APR2013	51973000
01MAY2013	49561000
01JUN2013	47331000
01JUL2013	41339000
01AUG2013	28261000
01SEP2013	44245000
01OCT2013	52885000
01NOV2013	49900000
01DEC2013	48435000
01JAN2014	48618000
01FEB2014	47438000
01MAR2014	51869000
01APR2014	46804000
01MAY2014	48888000
01JUN2014	46271000
01JUL2014	42653000
01AUG2014	28938000
01SEP2014	46320000
01OCT2014	54028000
01NOV2014	49278000
01DEC2014	49749000
01JAN2015	47600000
01FEB2015	47363000
01MAR2015	50358000
01APR2015	48809000
01MAY2015	49719000
01JUN2015	47494000
01JUL2015	43144000
01AUG2015	29978000
01SEP2015	47202000
01OCT2015	54612000
01NOV2015	52466000
01DEC2015	50991000
01JAN2016	48254000
01FEB2016	51440000

01MAR2016	49963000
01APR2016	53165000
01MAY2016	52376000
01JUN2016	49535000
01JUL2016	40618000
01AUG2016	31451000
01SEP2016	48347000
01OCT2016	53047000
01NOV2016	53663000
01DEC2016	52897000
01JAN2017	52370000
01FEB2017	52219000
01MAR2017	57514000
01APR2017	49919000
01MAY2017	55550000
01JUN2017	54633000
01JUL2017	46000000
01AUG2017	35146000
01SEP2017	52442000
01OCT2017	59151000
01NOV2017	57721000
01DEC2017	53740000
01JAN2018	56111000
01FEB2018	55436000
01MAR2018	56565000
01APR2018	58359000
01MAY2018	58296000
01JUN2018	55653000
01JUL2018	48756000
01AUG2018	34831000
01SEP2018	53612000
01OCT2018	62907000
01NOV2018	60094000
01DEC2018	56590000
01JAN2019	58605000
01FEB2019	57350000
01MAR2019	61467000
01APR2019	55267000
01MAY2019	57561000
01JUN2019	56897000
01JUL2019	50950000
01AUG2019	36129000
01SEP2019	56269000
01OCT2019	66151000
01NOV2019	62785000
01DEC2019	58045000
01JAN2020	59605000
01FEB2020	60410000
01MAR2020	26416000
01APR2020	4865000
01MAY2020	9644000
01JUN2020	21778000
01JUL2020	27066000
01AUG2020	20707000
01SEP2020	27861000
01OCT2020	30067000
01NOV2020	30493000
01DEC2020	30933000

01JAN2021	29268000
01FEB2021	30299000
01MAR2021	35044000
01APR2021	33799000
01MAY2021	36111000
01JUN2021	37708000
01JUL2021	34042000
01AUG2021	25999000
01SEP2021	41215000
01OCT2021	47203000
01NOV2021	48329000
01DEC2021	43288000
01JAN2022	39664000
01FEB2022	44812000
01MAR2022	52844000
01APR2022	46530000
01MAY2022	50338000
01JUN2022	48204000
01JUL2022	40761000
01AUG2022	31909000
01SEP2022	51158000
01OCT2022	56080000
01NOV2022	56167000
01DEC2022	53202000
01JAN2023	54418000
01FEB2023	55335000
01MAR2023	61493000
01APR2023	54100000
01MAY2023	58490000
01JUN2023	57070000
01JUL2023	47858000
01AUG2023	36489000
01SEP2023	54824000
01OCT2023	62233000
01NOV2023	61573000
01DEC2023	58425000
01JAN2024	58879000
01FEB2024	61471000
01MAR2024	59130000
01APR2024	64265000
01MAY2024	62333000
01JUN2024	59458000
01JUL2024	52445000