UAB
Universitat Autònoma de Barcelona

Dipòsit digital de documents de la UAB

This is the **published version** of the bachelor thesis:

Grasa Falgà, Judit. *Anàlisi estadística de l'activitat hotelera a Catalunya (2015-2025) mitjançant sèries temporals*. Treball de Final de Grau (Universitat Autònoma de Barcelona), 2026 (Economia). 148 pàg.

This version is available at https://ddd.uab.cat/record/337006

under the terms of the  license.

Universitat Autònoma de Barcelona

Treball de Final de Grau

Facultat d'Economia i Empresa

TÍTOL: Anàlisi estadística de l'activitat hotelera a Catalunya (2015-2025) mitjançant sèries temporals

AUTORA: Judit Falgà Grasa

GRAU: Economía

TUTORA: Maria Dolors Márquez Cebrián

DATA: 26/05/2026

RESUM

Una de les principals amenaces que el sector turístic ha d'afrontar a escala global és la vulnerabilitat davant les crisis i les fluctuacions econòmiques. Amb l'objectiu d'atenuar-ne l'impacte, el disseny i la implementació de diverses estratègies de planificació resulten necessaris, tant en l'àmbit nacional com internacional. No obstant això, aquestes mesures solen centrar-se en els agregats macroeconòmics generals i poden arribar a descuidar els trets i les característiques del comportament hoteler. En l'última dècada ha crescut significativament el pes d'aquest sector a Catalunya, consolidant-la com una de les destinacions més visitades d'Europa. A més, la forta dependència de les temporades altes i l'impacte de la pandèmia de la COVID-19 una intensitat que varia segons la regió.

Partint d'aquest context, l'estudi analitza la trajectòria del sector turístic durant el període 2015-2025 per a Catalunya i les seves quatre províncies: Barcelona, Girona, Lleida i Tarragona. Amb aquesta finalitat, s'ha construït una base de dades mensual que integra les pernoctacions, el nombre de viatgers i la despesa mitja diària dels turistes estrangers. Metodològicament, s'han aplicat models de regressió lineal (RLS) per poder captar la tendència i estacionalitat; així com models ARIMA, amb l'objectiu d'analitzar la dinàmica temporal i elaborar prediccions a curt termini.

Els resultats mostren que la ruptura estructural del març de 2020 va provocar una contracció de l'activitat del sector propera al 70%. Malgrat la recuperació progressiva que va permetre superar els nivells prepandèmics el 2024, l'anàlisi provincial fa visible certes divergències: mentre Barcelona, Girona i Tarragona acceleren les seves trajectòries en el post-COVID, Lleida presenta una desacceleració que apunta a un possible canvi estructural en la seva demanda turística. Pel que fa a la despesa mitja diària, s'observa un increment sostingut al llarg del període, amb una contracció durant la pandèmia i una notable recuperació posterior que apunta a un canvi en el perfil del turista.

Paraules clau: sector hoteler, Catalunya, pernoctacions, viatgers, sèries temporals, ARIMA, COVID-19, estacionalitat.

Agraïments: vull agrair a la meva tutora, Maria Dolors Márquez Cebrián, la seva dedicació i orientació al llarg de tot el procés d'elaboració d'aquest treball.

TAULA DE CONTINGUTS

RESUM....2
1. INTRODUCCIÓ....4
1.1. El sector hoteler en el context econòmic de Catalunya....4
1.2. L'impacte de la COVID-19....5
1.3. Definició del problema d'investigació, justificació i objectius....6
1.4. Preguntes d'investigació....7
2. MARC TEÒRIC....8
2.1. Determinants de la demanda turística i models de sèries temporals....8
2.2. Estacionalitat de les destinacions mediterrànies....9
2.3. L'impacte de les crisis sobre el turisme....9
2.4. Indicadors útils per mesurar l'activitat hotelera....10
3. METODOLOGÍA....11
3.1. Construcció de la base de dades....11
3.2. Estadística descriptiva....12
3.3. Propietats de les sèries: estacionarietat....13
4. ANÀLISI DEL MODEL ECONOMÈTRIC....15
4.1. Model 1: Tendència temporal....15
4.2. Model 2: Incorporació de l'estacionalitat....15
4.3. Model 3: Ruptura estructural. El xoc de la COVID-19....15
4.4. Anàlisi dels períodes pre i post COVID....16
4.5. Model ARIMA....16
4.6. Diagnòstic, inferència i limitacions dels models....16
5. RESULTATS EMPÍRICS....18
5.1. Les pernoctacions hoteleres....18
5.1.1. L'ajust global i la capacitat explicativa dels models....18
5.1.2. Tendència temporal i anàlisi per subperíodes....19
5.1.3. Estructura de l'estacionalitat....20
5.1.4. Impacte estimat de la COVID-19....21

5.1.5. Diagnòstic i anàlisi d'ARIMA per les pernoctacions..... 22
5.2. Els viatgers estrangers.....24
5.2.1. L'ajust global i la capacitat explicativa dels models.....24
5.2.2. Tendència temporal i anàlisi per subperíodes..... 25
5.2.3. Estructura de l'estacionalitat.....26
5.2.4. Impacte estimat de la COVID-19.....27
5.2.5. Diagnòstic i anàlisi d'ARIMA pels viatgers.....27
5.3. Anàlisi de la despesa mitjana diària dels turistes.....29
5.3.1. L'ajust global i la capacitat explicativa dels models.....29
5.3.2. Tendència temporal i anàlisi per subperíodes..... 29
5.3.3. Impacte estimat de la COVID-19 sobre la despesa mitja diària..... 30
5.3.4. Diagnòstic i anàlisi d'ARIMA per la despesa.....30
6. CONCLUSIONS.....32
BIBLIOGRAFÍA..... 34
ANNEXOS.....38
Annex 1: Anàlisi de les pernoctacions hoteleres a Catalunya..... 38
Annex 2: Anàlisi dels viatgers a Catalunya..... 75
Annex 3: Anàlisi de la despesa mitja diària de Catalunya..... 111

1. INTRODUCCIÓ

1.1. El sector hoteler en el context econòmic de Catalunya

El turisme és un dels pilars estructurals de l'economia catalana, i l'allotjament hoteler esdevé el nucli d'aquest sistema productiu. Més enllà del seu pes directe en el PIB, actua com a indicador clau del turisme del territori i destaca pel seu efecte multiplicador: el creixement del flux de viatgers estimula la restauració, el comerç, el transport i la resta de serveis associats a la demanda turística. D'altra banda, l'evolució de variables com les pernoctacions, el nombre de viatgers o la despesa dels turistes no només donen una idea aproximada del flux turístic, sinó també d'altres transformacions econòmiques que afecten el conjunt del país.

En aquest context, la divergència entre territoris és determinant, ja que Catalunya opera en funció del destí: Barcelona sol concentrar el volum més gros de pernoctacions amb amplis perfils turístics; Girona i Tarragona presenten una elevada estacionalitat concentrada a l'estiu; i Lleida actua com a destí turístic per a aquells perfils que cerquen muntanya i esquí, amb una demanda distribuïda en diferents èpoques. Aquesta diversitat resulta especialment rellevant per a l'anàlisi de la demanda hotelera, donant-li un enfocament comparatiu.

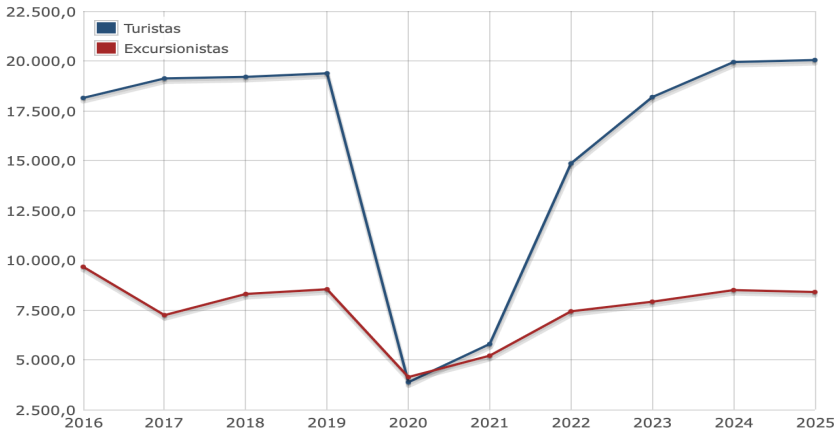
Metodològicament, el sector hoteler ofereix dades amb sèries mensuals de l'Enquesta d'Ocupació Hotelera (EOH) de l'INE. Aquestes dades, d'elevada qualitat estadística, faciliten anar més enllà de l'anàlisi descriptiva, permetent identificar els canvis estructurals i la resposta del sector davant de pertorbacions mitjançant models economètrics.

1.2. L'impacte de la COVID-19

Durant el període previ a la pandèmia, concretament entre els anys 2015 i 2019, l'activitat hotelera a Catalunya va definir una trajectòria expansiva explicada per: la recuperació econòmica posterior a la crisi financera internacional de 2008-2012, l'increment de la mobilitat global i els nous hàbits de consum.

Tal com s'observa al Gràfic 1, l'evolució del nombre de turistes a Catalunya va ser sostinguda fins l'any 2019, on se'n va registrar aproximadament 19,3 milions, prop del 20% del total de visitants a Espanya[1].

Gràfic 1: Visitants estrangers. Turistes i excursionistes (milers). Catalunya. 2016-2025



Font: *elaboració de l'INE a partir de dades de l'Idescat.*

Tanmateix, l'impacte sobtat de la COVID-19 va alterar aquest comportament: l'estat d'alarma, decretat el 14 de març de 2020, va suposar el tancament obligatori dels establiments hotelers, fins al punt que l'abril del mateix any es van registrar pràcticament 0 pernoctacions[2]. Això va provocar un descens del flux de viatgers de gairebé un 70% respecte a l'any anterior[3]. Catalunya se'n va veure més afectada per dos motius estructurals: la seva elevada dependència del turisme internacional, que representa al voltant del 70% de les pernoctacions totals; i la concentració del xoc en la temporada alta de la demanda.

La recuperació es va iniciar cap a setembre de 2021, amb l'eliminació de les restriccions de mobilitat internacional, tot i que el ritme va estar condicionat per cada territori. L'any 2024 es va confirmar la normalització l'activitat del sector i, finalment, el 2025 Catalunya va superar per primera vegada els 20 milions de turistes estrangers, assolint un nou rècord històric, i la despesa dels visitants en el principat va augmentar fins a 24.807 M€[4]. Aquesta evolució converteix el període 2015-2025 en un escenari d'especial interès ja que permet examinar tres etapes diferenciades: una expansió, un col·lapse i una recuperació gradual.

---

[1] INE (2020). Estadística de Movimientos Turísticos en Fronteras (FRONTUR). Sèrie anual 2019; Idescat (2020). Turisme estranger a Catalunya.

[2] INE (2024). Enquesta d'Ocupació Hotelera; Idescat (2025). Viatgers i pernoctacions en establiments hotelers a Catalunya.

[3] AQR-Lab, Universitat de Barcelona (2020). Impacto de la crisis turística originada por el COVID-19 sobre la economía catalana. Recuperat de https://aecr.org/en/impacto-de-la-crisis-turistica-originada-por-el-covid-19-sobre-la-economia-catalana/

[4] Ara. (2026). Rècord històric: Catalunya va superar els 20 milions de turistes estrangers el 2025. Recuperat de: https://es.ara.cat/economia/record-historico-cataluna-supero-20-millones-turistas-extranjeros-2025_1_5636901.html

1.3. Definició del problema d'investigació, justificació i objectius

Malgrat el gran volum d'informació disponible sobre el sector turístic català, encara existeixen certes limitacions a l'hora de comprendre el comportament dels principals indicadors de l'activitat hotelera, especialment amb un canvi estructural. En el cas estudiat, la irrupció de pandèmia de la COVID-19 va causar una ruptura sobtada en el comportament habitual de les sèries temporals, alterant tant les tendències com els patrons estacionals observats fins aquell moment.

En aquest sentit, resulta d'especial interés analitzar fins a quin punt els patrons identificats abans de la crisi sanitària es mantenen en el període posterior, així com estudiar la capacitat dels models economètrics per captar aquestes alteracions i generar previsions fiables a curt termini. És essencial, per tant, separar el xoc dels components estructurals de la sèrie per avaluar la trajectòria real del sector: en el cas contrari, qualsevol model estimat generarà estimadors esbiaixats i les prediccions no seràn gaire rellevants estadísticament.

Els objectius específics de la recerca pretenen analitzar descriptivament l'evolució de les variables d'activitat hotelera per a Catalunya i les seves quatre províncies; descompondre les sèries en tendència, estacionalitat i components irregulars; estimar l'impacte de la pandèmia i avaluar les diferències entre territoris; construir models ARIMA estacionals per generar prediccions a curt termini; i analitzar l'evolució de la despesa mitja diària dels turistes estrangers com a indicador del valor econòmic del turisme.

En conclusió, la motivació principal d'aquest treball és contribuir a una millor comprensió dels moviments del sector hoteler català durant el període 2015–2025, mitjançant l'estudi de sèries temporals mensuals dels principals indicadors del sector: nombre de viatgers, pernoctacions i despesa mitja diària dels turistes estrangers. D'aquesta manera, poder obtenir resultats útils per planificar i estimar la demanda turística per a la presa de decisions tant d'administracions públiques com d'agents econòmics del sector.

1.4. Preguntes d’investigació

El treball planteja tres preguntes d'investigació principals:

- Com ha evolucionat l'activitat hotelera a Catalunya i les seves províncies entre 2015 i 2025, especialment en el context de la crisi de la COVID-19?
- En quina mesura l'estacionalitat, la tendència i la ruptura estructural de la COVID-19 expliquen conjuntament la variabilitat de les sèries mensuals de pernoctacions, viatgers i despesa mitjana diària dels turistes estrangers?
- Fins a quin punt és possible predir el comportament a curt termini de l’activitat hotelera mitjançant models de sèries temporals?

2. MARC TEÒRIC

Nombrosos estudis han analitzat la demanda turística i l'activitat hotelera des d'una perspectiva economètrica, destacant la utilitat dels models de sèries temporals per explicar i predir el comportament del sector. En aquest context, la literatura acadèmica ha destacat quins son els factors que influeixen en la demanda turística, l'impacte de les crisis i la importància de factors com l'estacionalitat per analitzar l'estructura del sector.

2.1. Determinants de la demanda turística i models de sèries temporals

La demanda turística, des d'una perspectiva econòmica, està en funció de variables com el nivell de renda del país d'origen, els preus relatius entre la destinació i les seves alternatives, els costos de transport i, en menor mesura, el tipus de canvi. Partint d'aquest concepte, els treballs de Haiyan Song i Gang Li (2008) demostren que els models econòmics basats en sèries temporals i regressions són eines eficients per analitzar la demanda turística internacional, i que les variables esmentades expliquen una part molt significativa de la seva variabilitat, destacant-ne l'elevada elasticitat respecte a la renda. La perspectiva és coherent amb els resultats de Stephen F. Witt i Christine A. Witt (1995), que indiquen la importància dels factors macroeconòmics per determinar els fluxos turístics: la demanda turística del sector creix més ràpidament que la renda en moments d'expansió i es contrau amb major intensitat durant les recessions.

Pel que fa a la predicció estadística, la metodología de Box i Jenkins (1970), actualitzada per Box, Jenkins, Reinsel i Ljung (2015), ofereix un punt de referència pel treball, identificant els ordres del model a partir de les funcions d'autocorrelació (ACF) i autocorrelació parcial (PACF), l'estimació per màxima versemblança i el diagnòstic residual. Hamilton (1994) i Enders (2015) aporten el fonament teòric sobre estacionarietat per a qualsevol anàlisi de sèries temporals rigorosa. Rob J. Hyndman i George Athanasopoulos (2021) proposen l'ús de models ARIMA i d'altres tècniques per capturar la tendència i estacionalitat. De manera similar, Clive W. J. Granger (2001) estableix les bases de la causalitat i la predicció en aquest tipus de models.

En l'àmbit de Catalunya, Torres Prats (2023) aplica una metodologia similar a la d'aquest treball per analitzar l'impacte pandèmic sobre el turisme de Barcelona, obtenint resultats comparables.

2.2. Estacionalitat de les destinacions mediterrànies

L'estacionalitat és un element estructural del turisme, principalment en destinacions amb costa com Catalunya. Kulendran i Stephen F. Witt (2003) evidencien que la demanda turística presenta patrons estacionals estables al llarg dels anys i, generalment, força resistents als xocs econòmics. Aquesta estabilitat justifica l'ús de models que incorporin components estacionals, com variables dummy mensuals, com s'ha fet en els Models 1, 2 i 3 d'aquest treball.

D'acord amb Granger (2001), l'anàlisi de sèries temporals exigeix diferenciar el component estacional determinista del l'estocàstic. Les variables analitzades responen bé al primer enfocament gràcies al comportament repetitiu dels seus pics i valls. En canvi, en els models ARIMA s'imposa una diferenciació de període dotze (d12) per corregir la no-estacionarietat.

2.3. L'impacte de les crisis sobre el turisme

Pel que fa a l'impacte de les crisis sobre el turisme, la literatura mostra evidència sobre la vulnerabilitat del sector. Egon Smeral (2009) analitza la crisi financera global i conclou que la demanda turística es contrau significativament en períodes de recessió, amb una substitució cap a destinacions més properes i econòmiques. Aquests resultats són coherents amb els de Haiyan Song et al. (2011), que mostren que els xocs econòmics tenen un gran impacte sobre la demanda turística.

D'altra banda, la pandèmia de la COVID-19 ha generat una extensa literatura. Dogru et al. (2020) demostren que la pandèmia va provocar una caiguda sobtada de la demanda turística global, tenint en compte el grau de dependència del turisme internacional. Stefan Gössling et al. (2020) assenyalen que la COVID-19 ha estat un dels xocs més importants en la història del turisme modern, afectant tant l'oferta com la demanda simultàniament. A més, numerosos estudis més actuals han analitzat la recuperació del sector turístic després de la pandèmia i han determinat que aquesta és progressiva i desigual: el turisme domèstic es caracteritza per una recuperació més ràpida que en l'àmbit internacional. Aquest patró és coherent amb l'evidència empírica observada en diferents països europeus.

Finalment, Rob Law et al. (2019) destaquen que els models economètrics estructurals continuen sent preferibles quan l'objectiu central és la interpretació econòmica de les variables dependents. Aquesta conclusió justifica l'elecció dels models de RLS i ARIMA.

2.4. Indicadors útils per mesurar l'activitat hotelera

Per tal de mesurar l'activitat i la demanda del sector hoteler, diversos indicadors es poden tindre en compte. Song, Witt i Li (2009) identifiquen les pernoctacions com la mesura estàndard per a la comparació internacional, per la seva capacitat de capturar tant el volum de visitants com la intensitat de l'estada. El nombre de viatgers, en canvi, proporciona informació sobre els fluxos d'entrada però no distingeix entre estades breus i prolongades. L'estada mitjana, el quocient entre pernoctacions i viatgers, és un indicador clau del perfil de consum turístic: valors elevats s'associen al turisme vacacional de costa, mentre que valors reduïts caracteritzen el turisme cultural i de negocis. La despesa mitja diària dels turistes, provista per l'Institut d'Estadística de Catalunya (Idescat) en valors mensuals, mesura el valor econòmic (en euros) generat pel flux turístic internacional. En el present treball s'utilitza per comparar la intensitat de la despesa de forma més precisa que la despesa total, ja que elimina l'efecte del volum de visitants.

3. METODOLOGÍA

3.1. Construcció de la base de dades

Les sèries temporals són una eina clau en l'anàlisi econòmic, ja que permeten estudiar l'evolució d'una variable al llarg del temps a partir d'observacions ordenades. Tal com assenyalen Hamilton (1994) i Enders (2015), aquest tipus de dades és útil perquè permet identificar patrons i diferenciar-los de fluctuacions aleatòries. En el context del sector turístic, les sèries temporals resulten especialment rellevants per analitzar l'activitat hotelera, degut a la forta dependència de factors estacionals.

A l'hora de construir la base de dades, s'ha tingut en compte dues fonts: L'INE i l'Idescat. Les dades primàries provenen de la Encuesta de Ocupación Hotelera (EOH), elaborada mensualment per l'Institut Nacional de Estadística (INE) i, amb dades per a Catalunya, per l'Institut d'Estadística de Catalunya (Idescat). L'EOH recull informació sobre el nombre de viatgers i pernoctacions als establiments hotelers de totes les categories, per comunitats autònomes, províncies i procedència dels visitants. Les dades sobre la despesa s'han obtingut de l'Idescat, que proporciona estimacions mensuals de la despesa mitjana diària dels visitants estrangers per persona i dia.

Taula 1. Resum de variables utilitzades

Variable	Descripció	Tipus	Unitat	Font
Viatgers (Y_t)	Total de persones allotjades en establiments hotelers	Variable dependent	Nombre de persones	INE (EOH)
Pernoctacions (Y_t)	Total de nits en establiments hotelers	Variable dependent	Nombre de pernoctacions	INE (EOH)
Despesa turística (Y_t)	Despesa mitjana diària dels turistes estrangers	Variable dependent	Euros	Idescat
Índex temporal (t)	Variable contínua de 1 a 132	Tendència	Temps	Elaboració pròpia
Dummies mensuals (D_{mt})	11 indicadors mensuals (base: gener)	Estacionalitat	Mesos	Elaboració pròpia
Dummy COVID-19 ($COVID_t$)	Impacte de la pandèmia 1 per 2020-03 fins 2021-09, 0 per la resta d'observacions	Ruptura estructural	[0, 1]	Elaboració pròpia

Font: *elaboració pròpia a partir de les especificacions metodològiques del treball.*

A partir d'aquestes, s'ha construït una base de dades específica per a Catalunya i les seves quatre províncies: Barcelona, Girona, Lleida i Tarragona. Per a les variables de pernoctacions i viatgers, les dades cobreixen el període de gener de 2015 a desembre de 2025, amb un total de 132 observacions mensuals per a cadascun dels cinc territoris. Per a la variable de despesa mitja diària, les dades disponibles comencen el gener de 2016 i s'estenen fins al desembre de 2025, representant 120 observacions mensuals a nivell de Catalunya.

El procés de preparació de les dades ha implicat: transposar les columnes per transformar el format amplada a format longitudinal, assignant a cada observació un identificador de data en format YYYY-MM per tal de ser llegible per RStudio; generar una variable temporal t que pren valors d' 1 (gener 2015 o 2016) fins a 132 o 120; construir 11 variables dummy mensuals (amb gener com a referència: dm2..dm12) i una variable fictícia (dummy) COVID-19, que pren valor 1 per als 19 mesos compresos entre 2020-03 i 2021-09, i 0 per la resta.

3.2. Estadística descriptiva

La Taula 2 presenta els estadístics descriptius de les tres variables principals per al conjunt de Catalunya al llarg del període d'estudi i dels subperíodes. L'elevada desviació típica en totes les variables, explicada per: l'estacionalitat intensa, que genera una variabilitat recurrent entre els mesos de l'any; i el xoc pandèmic, que introdueix observacions pròximes a zero el 2020 i el primer semestre de 2021. El valor mínim de zero en totes les variables correspon al mes d'abril de 2020, quan la paralització total de l'activitat hotelera va deixar pràcticament buits els establiments. El màxim de pernoctacions (8.994.168) s'assoleix a l'etapa pre-COVID. No obstant, el màxim de viatgers (2.723.647) es dona al post-COVID; el mateix passa amb la despesa (271), situació que reflecteix la recuperació postpandèmica i la continuació del creixement a llarg termini.

Taula 2: Estadístics descriptius. Catalunya, 2015–2025

Variable	Temps	Mitjana	Desviació típica	Mín.	Màx.
Pernoctacions	2015-2025	4178438	2465848	0	8994168
	Pre-COVID	4587035	2298746	1737082	8994168
	Post-COVID	4700768	2283016	1461340	8900068

Taula 2: Continuació

Viatgers	2015-2025	1499741	684006,400	0	2723647
	Pre-COVID	1604354	549410,800	747378	2586101
	Post-COVID	1715084	608340,200	649122	2723647
Despesa mitjana diària (€)	2016-2025	180,458	40,267	0	271
	Pre-COVID	180,960	19,812	130	221
	Post-COVID	199,980	32,001	119	271

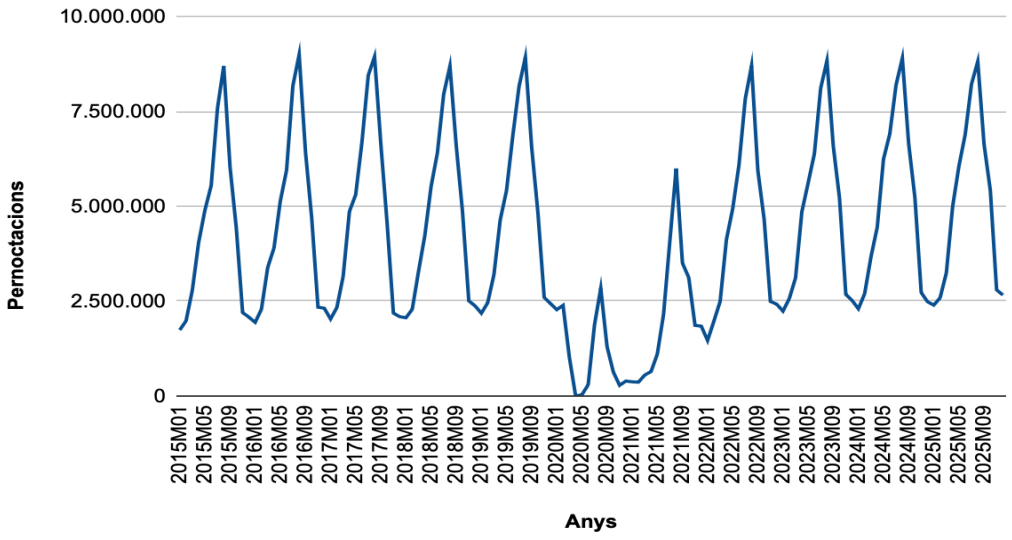
Font: *Elaboració pròpia a partir de dades de l'INE, l'Idescat i RStudio.*

3.3. Propietats de les sèries: estacionarietat

Des d'un punt de vista teòric, una sèrie temporal es pot descompondre en diversos components: la tendència, que reflecteix el moviment a llarg termini de la sèrie; l'estacionalitat, que recull els patrons recurrents de freqüència anual. Addicionalment, poden existir components cíclics així com un component aleatori que recull variacions imprevisibles (Hamilton, 1994), en aquest cas seria la pendèmia de 2020.

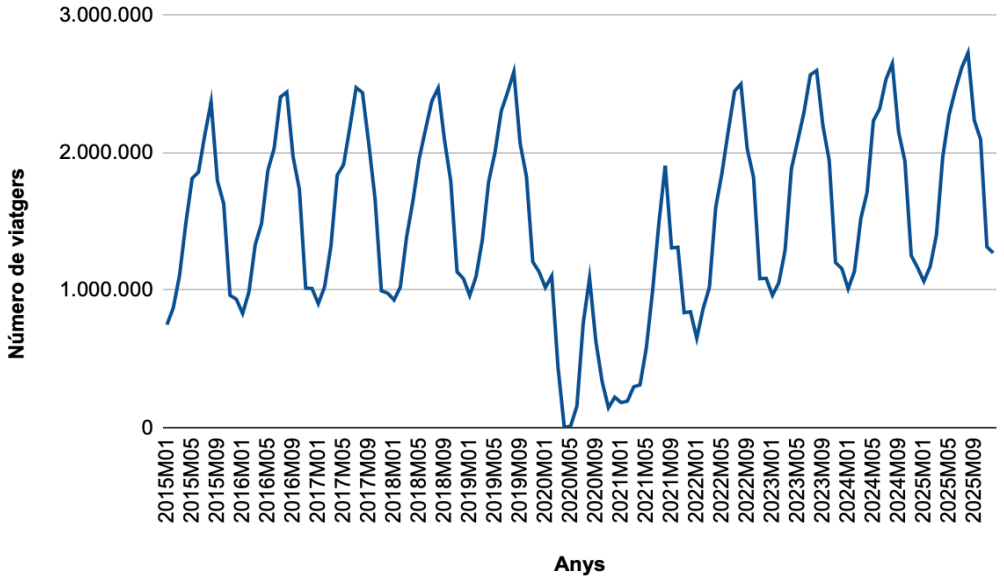
Des d'un punt de vista analític, l'observació i la interpretació de les dades s'ha realitzat amb el programari RStudio, utilitzant regressions lineals aplicades a dades de sèries temporals i els paquets forecast, tseries, lmtest, sandwich, strucchange i modelsummary.

Gràfic 2: Evolució de les pernoctacions mensuals. Catalunya, 2015–2025.



Font: *Elaboració pròpia a partir de dades de l'INE i l'Enquesta d'Ocupació Hotelera (EOH).*

Gràfic 3: Evolució del nombre de turistes (milers). Catalunya, 2015–2025.



Font: *Elaboració pròpia a partir de dades de l'INE i Idescat.*

Tal com s'observa al Gràfic 2, l'evolució de les pernoctacions mensuals a Catalunya va ser sostinguda fins al 2019, presentant un patró estacional molt marcat amb pics recurrents entre juny i setembre; el Gràfic 3 detalla l'evolució del nombre de viatgers per mes, i també presenta una ruptura brusca el 2020. Tot i el descens sobtat d'ambdues sèries, la recuperació posterior mostra una trajectòria ascendent, que es diferencia depenent del territori estudiat.

4. ANÀLISI DEL MODEL ECONOMÈTRIC

L'anàlisi empírica d'aquest treball es basa en l'estimació de models de regressió lineals aplicats a dades de sèries temporals. L'objectiu és poder descriure, explicar i predir l'evolució de l'activitat hotelera a Catalunya i les seves províncies. Aquesta metodología permet capturar els principals components de les sèries: tendència, estacionalitat i possibles ruptures estructurals; a més de les prediccions a curt termini dels models ARIMA estacionals.

4.1. Model 1: Tendència temporal

En primer lloc, es considera un model de regressió lineal simple amb tendència temporal:

$$Y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

On $\boldsymbol{Y}_t$ representa la variable dependent (pernoctacions, viatgers o despesa) en el període $\boldsymbol{t}$. L'índex $\boldsymbol{t}$, és una variable de temps que captura la tendència i pren valors de 1 a 132. El paràmetre $\boldsymbol{\beta}_1$ captura el creixement o decreixement estructural mensual de la variable presa. Per últim, $\boldsymbol{\varepsilon}_t$ és el terme d'error. Aquí és d'esperar que l'ajust (R^2) del Model 1 sigui molt baix perquè ni el component estacional ni la pandèmia queden captats.

4.2. Model 2: Incorporació de l'estacionalitat

Ara, sabent que el sector turístic es caracteritza per una forta estacionalitat, s'amplia el model anterior mitjançant la inclusió d'11 variables fictícies mensuals, amb gener com a mes base:

$$Y_t = \beta_0 + \beta_1 t + \Sigma\ (m=2..12)\ \gamma_m D_{mt} + \varepsilon_t$$

On $\boldsymbol{D}_{mt}$ són les variables dummy que prenen valor 1 si l'observació correspon al mes **m** i 0 en cas contrari; s'exclou el mes de referència per evitar multicol·linealitat perfecta. Els coeficients γ_m capturen les diferències estacionals respecte al mes base. Amb aquest model s'espera que la inclusió de l'estacionalitat millori en gran mesura l'ajust global (R^2), confirmant que l'estacionalitat és un factor clau que explica la variabilitat de les sèries.

4.3. Model 3: Ruptura estructural. El xoc de la COVID-19

Per analitzar l'impacte de la pandèmia, s'incorpora una altre variable dummy que captura el xoc associat a l'any 2020:

$$Y_t = \beta_0 + \beta_1 t + \Sigma \gamma_m D_{mt} + \delta COVID_t + \varepsilon_t$$

On $\boldsymbol{COVID}_t$ pren valor 1 durant la pandèmia (de març de 2020 al setembre de 2021) i 0 en cas contrari. El coeficient $\boldsymbol{\delta}$ mesura l'impacte mitjà de la COVID-19 sobre la variable dependent analitzada, $\boldsymbol{Y}_t$. Després d'incorporar aquest condicionant, es preveu que el xoc pandèmic sigui molt significatiu en la variabilitat del model, incrementant l'R^2.

4.4. Anàlisi dels períodes pre i post COVID

Per tal de determinar si la pandèmia ha modificat significativament la trajectòria del sector, s'ha d'estimar els models de tendència i estacionalitat per separat, en dos subperíodes: pre-COVID (de gener de 2015 a febrer de 2020; 62 observacions) i post-COVID (d'octubre de 2021 a desembre de 2025; 51 observacions). El període intermig (de març de 2020 a setembre de 2021; 19 observacions), queda exclòs de l'estimació per evitar que les observacions de la crisi pertorbin l'estimació de les tendències.

$$Y_t = \beta_0 + \beta_1 t + \Sigma\gamma_m D_{mt} + \varepsilon_t$$

Comparar el coeficient $\boldsymbol{\beta}_1$ entre subperíodes permet caracteritzar si el creixement mensual del sector s'ha accelerat o desaccelerat en el post-COVID, i si les diferències de creixement entre provinces s'han ampliat o reduït.

4.5. Model ARIMA

Tot i que els models de regressió lineals generalment capturen bé la tendència i estacionalitat, passen per alt un detall crucial: la relació i dependència que existeix entre un mes i el següent. Per corregir-ho i analitzar l'evolució de les dades, s'apliquen models ARIMA(p,d,q)(P,D,Q)12: en aquesta estructura, la part (p,d,q) recull el comportament general de la sèrie (l'efecte autoregressiu, la integració i la mitjana mòbil), mentre que la part (P,D,Q)12 replica aquests mateixos components per al patró estacional anual.

4.6. Diagnòstic, inferència i limitacions dels models

Els models de regressió lineal s'estimen mitjançant Mínims Quadrats Ordinaris (MQO). No obstant això, els problemes d'autocorrelació (test Durbin-Watson) i heteroscedasticitat (test Breusch-Pagan) en els residus són força habituals. Per garantir la validesa de la inferència, s'utilitzen errors estàndard robustos a l'heteroscedasticitat, de tipus HC1, i la correcció de Newey-West (amb 12 retards). D'altra banda, per verificar l'estabilitat temporal dels

paràmetres es realitza el test de canvi estructural CUSUM, que permet justificar la introducció de la variable dummy de la COVID-19.

Per últim, caldria destacar quines son les dues principals limitacions metodològiques. La primera: la modelització de la ruptura COVID com una dummy binària constant simplifica massa el xoc, una especificació amb dummies trimestrals probablement hauria capturat millor la dinàmica de la caiguda i la recuperació gradual. La segona: fixar l'inici de la fase de recuperació a partir de l'octubre de 2021 podria passar per alt canvis estructurals ja que el període de crisi podría ser considerat més extens.

5. RESULTATS EMPÍRICS

Els resultats estan organitzats segons la variable dependent analitzada: les pernoctacions, els viatgers i la despesa mitja diària dels turistes estrangers. Per a cada variable s'ha examinat la capacitat explicativa dels models (R^2 ajustat), la interpretació econòmica dels coeficients estimats i les especificitats de l'ARIMA.

5.1. Les pernoctacions hoteleres

5.1.1. L'ajust global i la capacitat explicativa dels models

L'anàlisi de les pernoctacions hoteleres és un factor clau per comprendre la intensitat de l'activitat turística; a més, és la variable que ha presentat els resultats més robustos de l'estudi. El Model 3, que incorpora la tendència, l'estacionalitat i l'impacte de la COVID-19, assoleix un R^2 de 0,940 i un R^2 ajustat de 0,933 amb una F-estadística de 142.371. Això es tradueix com que el conjunt de variables independents aconsegueixen explicar el 93,3% de la variabilitat de les pernoctacions a Catalunya.

Taula 3. Comparació de models. Pernoctacions a Catalunya.

Paràmetre	Model 1 (T)	Model 2 (T+S)	Model 3 (T+S+COVID)
R^2	0,002	0,654	0,940
R^2 ajustat	-0,006	0,619	0,933
F-estadístic	0,242	18,747***	142,371***
Tendència β_1 (pernoct./mes)	2.777	1.739	3.887***
COVID δ (pernoct./mes)	-	-	-3.780.837***
Coef. agost (γ_8)	-	6.110.434***	6.439.109***
Durbin-Watson	-	-	0,639
ARIMA	-	-	(3,0,0)(2,1,0)[12]

Font: *elaboració pròpia. *** p<0,001; ** p<0,01; * p<0,05; no significatiu.*

La Taula 3 recull els resultats de l'output d'RStudio per a les pernoctacions de Catalunya i fa visible que hi ha una clara progressió entre els models: el Model 1, amb la tendència com a única variable independent, assoleix un R^2 ajustat de -0,006, revelant que la tendència per si sola no té cap poder explicatiu que sigui rellevant; el Model 2, en afegir les variables

fictícies estacionals, incrementa fins a 0,619; i el Model 3, amb l'incorporació de la dummy COVID, arriba fins a 0,933, indicant que xoc de la pandèmia és de gran rellevància.

Taula 4. Comparació de models. Pernoctacions. Catalunya i províncies.

Indicador	Catalunya	Barcelona	Girona	Tarragona	Lleida
R² (M1. Tendència)	0,002	-0,004	0,001	0,000	0,018
R² (M2. T+Estacionalitat)	0,654	0,384	0,830	0,818	0,754
R² (M3. T+E+COVID)	0,940	0,925	0,954	0,934	0,897
R² ajustat (M3)	0,933	0,917	0,948	0,927	0,885
Error estànd. residual (M3)	636.061	309.683	165.049	195.685	23.796
F-estadístic (M3)	142,371***	112,358***	186,333***	128,200***	78,892***
p-valor F (M3)	< 0,001	< 0,001	< 0,001	< 0,001	< 0,001

Font: *elaboració pròpia. *** p<0,001; ** p<0,01; * p<0,05; no significatiu.*

La Taula 4 evidencia que Girona i Tarragona son les províncies amb els ajustos globals més elevats (R² ajustat M3: 0,948 i 0,927 respectivament), degut a la major regularitat dels seus patrons. Barcelona obté un ajust lleugerament inferior (0,917) per la major variabilitat temporal de la demanda. Lleida presenta el valor més baix (0,885), indicant que el turisme de muntanya i rural conté components que l'afecten no tinguts en compte.

5.1.2. Tendència temporal i anàlisi per subperíodes

Un cop s'ha dut a terme l'estimació del Model 3, ja tenint controlats l'estacionalitat i el xoc de la pandèmia, el coeficient de tendència temporal és $\boldsymbol{\beta_1 = 3.887}$ pernoctacions mensuals addicionals a Catalunya (Taula 3), valor estadísticament significatiu ($p<0,001$). En altres paraules, el sector creixia en 3.887 pernoctacions addicionals cada mes entre 2015 i 2025.

L'anàlisi per subperíodes (Taula 5), posa en evidència resultats reveladors: la tendència mensual puja de 9.088 pernoctacions al pre-COVID fins a 21.373 al post-COVID. Aquesta acceleració és consistent a les províncies catalanes: Barcelona concentra l'acceleració més intensa; Girona té un increment de gairebé un 400%, podent reflectir tant l'efecte rebot postpandèmic com nous perfils de visitants; Tarragona segueix un patró similar; i Lleida és

l'excepció, mostrant una desacceleració en el post-COVID, passant de 583 a 394 pernoctacions mensuals, apuntant a canvis estructurals en la de la seva demanda turística.

Taula 5. Anàlisi dels subperíodes. Pernoctacions. Catalunya i províncies.

INDICADOR	Catalunya		Barcelona		Girona		Tarragona		Lleida	
	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID
Observacions (n)	62	51	62	51	62	51	62	51	62	51
R^2	0.993	0.986	0.989	0.958	0.994	0.991	0.992	0.990	0.964	0.975
R^2 ajustat	0.992	0.982	0.986	0.945	0.992	0.988	0.990	0.986	0.956	0.967
Error estándar residual (milers)	207.429	307.274	95.510	197.497	63.563	79.650	74.952	85.074	13.139	11.379
Coef. de la tendència (β_{time})	9.088 ***	21.373 ***	6.418 ***	13.610 ***	0.844 *	4.114 ***	1.244 **	3.255 ***	0.583 ***	0.394 ***
Mes màx. impacte	Agost	Agost	Agost	Agost	Agost	Agost	Agost	Agost	Agost	Agost

Font: *elaboració pròpia amb RStudio. *** p<0,001; ** p<0,01; * p<0,05*

5.1.3. Estructura de l'estacionalitat

Tal com mostra la Taula 6, els coeficients de les dummies mensuals reproduïxen el patró estacional esperat en una destinació mediterrània. Per al conjunt de Catalunya, el coeficient de l'agost (γ_8 = 6.439.109, p < 0,001) implica que, en condicions normals i amb tendència constant, l'agost registra 6,4 milions de pernoctacions addicionals respecte al mes de referència (gener). És detectable un patró: els mesos d'estiu (juny a setembre) concentren la demanda, amb un pic absolut a l'agost; els mesos de tardor i hivern mostren coeficients molt menors. D'altra banda, la resta de territoris es comporten de manera força divergent: Girona presenta la concentració estacional estival més extrema; Tarragona segueix un patró similar; Barcelona, en canvi, té una distribució temporal molt més homogènia, mantenint nivells previsibles; i Lleida presenta una distribució singular, amb un pic a l'estiu moderat i certs increments durant l'hivern, coherents amb el turisme d'esquí i de muntanya.

Taula 6. Comparació de models. Coeficients per mesos. Catalunya i províncies.

Paràmetre	Catalunya	Barcelona	Girona	Tarragona	Lleida
β_0 Constant	2.013.089***	1.506.042***	235.124***	113.979*	157.944***
β_1 Tendència (time)	3.887***	2.687***	519 n.s.	361 n.s.	320***
γ_2 Febrer	261.525 n.s.	158.984 n.s.	59.756 n.s.	13.944 n.s.	28.841***
γ_3 Març	1.145.804***	661.019***	289.137***	182.836**	12.812 n.s.
γ_4 Abril	2.126.607***	970.630***	645.043***	578.144***	-67.210***
γ_5 Maig	2.988.986***	1.342.024***	836.526***	889.481***	-79.046***
γ_6 Juny	3.888.259***	1.518.412***	1.151.499***	1.261.617***	-43.269***
γ_7 Juliol	5.561.864***	2.020.244***	1.786.318***	1.722.319***	32.983***
γ_8 Agost ★	**6.439.109*** **	**2.237.529*** **	**2.106.382*** **	**1.968.288*** **	**126.910*** **
γ_9 Setembre	4.124.608***	1.566.189***	1.271.348***	1.315.636***	-28.565***
γ_{10} Octubre	2.385.881***	1.126.415***	687.553***	634.163***	-62.251***
γ_{11} Novembre	294.486 n.s.	276.015**	61.150 n.s.	66.155 n.s.	-108.835***
γ_{12} Desembre	195.599 n.s.	91.567 n.s.	48.838 n.s.	66.547 n.s.	-11.353 n.s.
δ COVID-19	**-3.780.837*** **	**-2.267.303*** **	**-731.322*** **	**-706.026*** **	**-76.186*** **
ADF (estacionarietat pernoctacions)	p = 0.010 (estacionaria)	p = 0.241 (no estacionaria)	p = 0.010 (estacionaria)	p = 0.010 (estacionaria)	p = 0.344 (no estacionaria)

Font: *elaboració pròpia amb RStudio. *** p<0,001 | ** p<0,01 | * p<0,05 | n.s = no significatiu.*

★ *= mes de màxim impacte estacional.*

5.1.4. Impacte estimat de la COVID-19

Taula 7. Impacte estimat de la COVID-19 sobre pernoctacions. Model 3.

Territori	δ (pernoct./mes)	Err. estànd. robust	R^2 ajustat M3	ARIMA
Catalunya	-3.780.837***	(159.321)	0,933	(3,0,0)(2,1,0)[12]
Barcelona	-2.267.303***	(77.570)	0,917	(2,0,0)(2,1,0)[12]
Girona	-731.322***	(41.342)	0,948	(3,0,0)(0,1,1)[12]
Tarragona	-706.026***	(49.015)	0,927	(3,0,0)(2,1,0)[12]
Lleida	-76.186***	(5.961)	0,885	(1,0,0)(2,1,0)[12]

Font*: elaboració pròpia. *** p<0,001.*

La Taula 7 fa visible que el coeficient de la dummy COVID és negatiu i de gran rellevància estadística en tots els territoris. Per a Catalunya, δ = -3.780.837 pernoctacions mensuals (error estàndard robust: 159.321; p < 0,001). Per tant, durant la còvid les pernoctacions mensuals van ser en promig 3,78 milions inferiors al que hauria predit el model en absència de pandèmia. Barcelona absorbeix la major part del xoc (δ = -2.267.303). Girona (δ = -731.322) i Tarragona (δ = -706.026) mostren impactes similars en magnitud. Finalment, Lleida presenta el coeficient menys negatiu (δ = -76.186), consistent amb la seva oferta orientada als seus perfils turístics.

Els tests de canvi estructural (Taula 8) confirmen la presència d'inestabilitat a les sèries analitzades (p < 0.05), indicant que la pandèmia no va ser únicament una fluctuació temporal, sinó una ruptura significativa en la trajectòria de les variables estudiades.

Taula 8. Canvi estructural (CUSUM)

	Catalunya	**Barcelona**	**Girona**	**Tarragona**	**Lleida**
S	1.241	1.298	1.087	1.241	1.238
p-valor	0.004	0.002	0.016	0.012	0.004

Font: *elaboració pròpia amb RStudio.*

5.1.5. Diagnòstic i anàlisi d'ARIMA per les pernoctacions

L'anàlisi de la qualitat dels models estimats revela dues incidències tècniques rellevants que condicionen la inferència: l'heteroscedasticitat, amb el test de Breusch-Pagan (BP), que retorna un valor de 60,353, indicant la presència d'heteroscedasticitat en els residus; i l'autocorrelació, amb el test de Durbin-Watson (DW = 0,639 per a Catalunya), confirmant una autocorrelació positiva significativa i evidenciant que el model de RLS no capta tota la dinàmica temporal de la sèrie. Per corregir-ho i garantir la robustesa de la inferència, s'han calculat els errors estàndard mitjançant l'estimador HC1 (Taula 9).

Taula 9. Errors robustos. Coeficients.

Variable	Coeficient estimat	Error estàndandar robust	Valor t	p-valor	Significació
Tendència temporal (time)	3887,23	751,45	5,17	<0,001	***
Febrer (dm2)	261525,23	262696,39	1,00	0,322	-
Març (dm3)	1145804,41	287375,18	3,99	<0,001	***
Abril (dm4)	2126607,19	239042,78	8,90	<0,001	***
Maig (dm5)	2988985,69	277525,44	10,77	<0,001	***
Juny (dm6)	3888258,55	315416,57	12,33	<0,001	***
Juliol (dm7)	5561863,51	309327,06	17,98	<0,001	***
Agost (dm8)	6439108,56	308585,15	20,87	<0,001	***
Setembre (dm9)	4124607,61	267712,12	15,41	<0,001	***
Octubre (dm10)	2385880,52	266903,60	8,94	<0,001	***
Novembre (dm11)	294485,47	260744,47	1,13	0,261	-
Desembre (dm12)	195599,25	271189,08	0,72	0,472	-
COVID	-3780837,49	327615,30	-11,54	<0,001	***

Font: *elaboració pròpia amb RStudio. *** p<0,001 | ** p<0,01 | * p<0,05.*

L'anàlisi de les funcions d'autocorrelació dels residus (ACF i PACF) de Catalunya mostra que la majoria dels retards es troben dins de les bandes de confiança del 95%. Aquest resultat valida la robustesa del model, demostrant que s'ha extret correctament tota la dinàmica temporal de la sèrie i que no queda informació rellevant sense capturar.

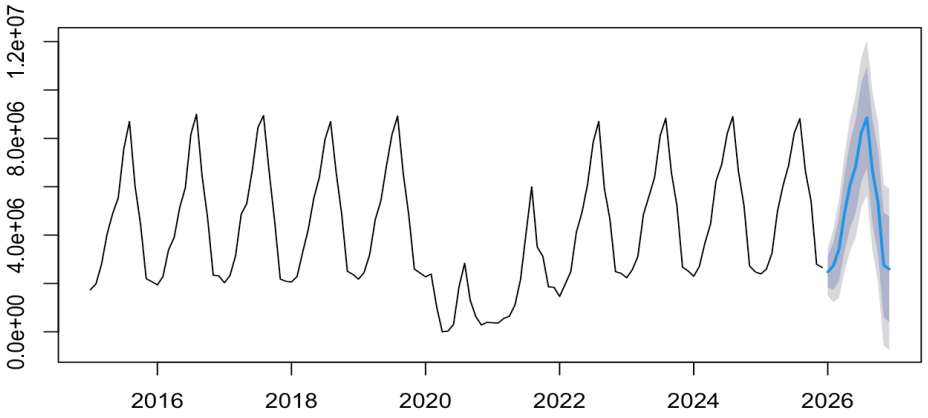
Els resultats del test d'Augmented Dickey-Fuller (ADF), visibles a la Taula 6, confirmen una que el comportament estadístic de les sèries va en funció del territori. Sèries estacionàries ($p<0,05$): Catalunya, Girona i Tarragona (0,010) presenten una mitjana constant, la qual cosa permet la modelització directa sense necessitat de diferenciació. Sèries no estacionàries ($p>0,05$): Barcelona (0,241) i Lleida (0,344) requereixen un procés de diferenciació ($d\geq1$) per eliminar la tendència i garantir l'estabilitat dels models ARIMA.

La selecció automàtica ARIMA amb la comanda "*auto.arima()*" per a les pernoctacions de Catalunya, produeix un model ARIMA(3,0,0)(2,1,0)[12]. Aquest resultat indica que el nivell actual de pernoctacions està influenciat pels valors observats durant els tres mesos anteriors,

demostrant una elevada dependència temporal a curt termini. A més, el model incorpora un component estacional anual que permet captar la recurrència dels patrons turístics al llarg del temps, considerant la influència dels mateixos períodes dels anys anteriors.

Pel que fa a les províncies, Barcelona presenta una estructura similar però lleugerament més simple, definida per un model ARIMA(2,0,0)(2,1,0)[12], suggerint una dependència temporal una mica menys intensa. Girona, en canvi, segueix un model ARIMA(3,0,0)(0,1,1)[12], incorporant un component de mitjana estacional que indica que els efectes estacionals tendeixen a corregir-se de manera més ràpida al llarg del temps. Finalment, Lleida presenta l'especificació més senzilla, ARIMA(1,0,0)(2,1,0)[12], suggerint una dinàmica temporal menys complexa i una evolució de la demanda més estable respecte a la resta de territoris analitzats. En conjunt, aquests resultats evidencien que el comportament de les pernoctacions presenta una marcat component estacional amb diferències territorials associades als perfils turístics de cada província.

Gràfic 4. Previsió de les pernoctacions de Catalunya. ARIMA.



Font: *elaboració pròpia amb RStudio.*

5.2. Els viatgers estrangers

5.2.1. L'ajust global i la capacitat explicativa dels models

Les sèries de viatgers demostren un patró d'ajust comparable al de les pernoctacions: el Model 3, presenta un R^2 ajustat global lleugerament inferior, de 0,922 per a Catalunya. La progressió entre models és força consistent: el Model 1 obté un R^2 ajustat de 0,004, el Model 2 puja fins a 0,485 i el Model 3 assoleix 0,922. Girona i Tarragona tornen a ser les províncies amb els ajustos més elevats (R^2: 0,940 i 0,931 respectivament); Barcelona els segueix (0,899); i Lleida presenta, de nou, el valor més baix (0,854).

Taula 10. R² ajustat per a la variable viatgers. Catalunya i províncies.

Territori	M1 (T)	M2 (T+S)	M3 (T+S+COVID)
Catalunya	0,004	0,485	0,922
Barcelona	0,004	0,176	0,899
Girona	0,005	0,776	0,940
Tarragona	-0,003	0,799	0,931
Lleida	0,016	0,593	0,854

Font: *elaboració pròpia.*

5.2.2. Tendència temporal i anàlisi per subperíodes

Un cop controlades l'estacionalitat i el xoc de la pandèmia (Model 3), el coeficient de tendència temporal és $\beta_1 = 2.368$ persones/mes a Catalunya (Taula 11), valor estadísticament significatiu ($p<0,001$). En altres paraules, el sector creixia en 2.368 persones addicionals cada mes entre 2015 i 2025. En analitzar els subperíodes, la taxa de creixement mensual es gairebé duplica, passant de 4.754 viatgers/mes al pre-COVID a 9.004 viatgers/mes al post-COVID. La tendència de Lleida al post-COVID (191 viatgers/mes) és pràcticament la meitat de la pre-COVID (354 viatgers/mes), reforçant la hipòtesi d'un canvi estructural en la composició de la demanda.

Taula 11. Anàlisi dels subperíodes. Viatgers. Catalunya i províncies.

INDICADOR	Catalunya		Barcelona		Girona		Tarragona		Lleida	
	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID	Pre-COVID	Post-COVID
Observacions (n)	62	51	62	51	62	51	62	51	62	51
R^2	0.989	0.982	0.982	0.945	0.988	0.990	0.988	0.988	0.948	0.962
R^2 ajustat	0.987	0.976	0.978	0.927	0.985	0.986	0.985	0.984	0.935	0.950
Error estándar residual (milers)	62.947	93.286	32.432	70.301	20.553	21.586	19.860	21.758	5.371	4.796
Coef. de la tendència (β_{time})	4.754 ***	9.004 ***	3.178 ***	6.988 ***	0.798 ***	1.174 ***	0.424 ***	0.651 ***	0.354 ***	0.191 ***
Mes màx. impacte	Agost	Agost	Juliol	Juliol	Agost	Agost	Agost	Agost	Agost	Agost

Font: *elaboració pròpia amb RStudio.*

5.2.3. Estructura de l'estacionalitat

Els coeficients de les dummies mensuals (Taula 12) reproduïxen el patró estacional esperat en una destinació mediterrània. Per al conjunt de Catalunya, el coeficient de l'agost (γ_8 = 1.595.656, $p < 0{,}001$) implica que, en condicions normals i amb tendència constant, l'agost registra 1,6 milions de viatgers addicionals respecte al mes de referència (gener). Torna a succeir el mateix que amb les pernoctacions, és detectable un patró: els mesos d'estiu (juny a setembre) concentren la demanda, amb un pic absolut a l'agost; els mesos de tardor i hivern mostren coeficients molt menors; i la resta de territoris es comporten de manera dispar: ara, Barcelona presenta la concentració estacional estival més extrema, seguida per Girona i Tarragona; Lleida, en canvi, presenta una distribució singular, amb un pic d'estiu baix (50.799 viatgers mensuals addicionals a l'agost) i disminucions marcades a l'hivern.

Taula 12. Comparació de models. Coeficients per mesos. Catalunya i províncies.

Paràmetre	Catalunya	Barcelona	Girona	Tarragona	Lleida
β_0 **Constant**	808.517***	615.716***	88.912***	39.456***	64.432***
β_1 **Tendència (time)**	2.368***	1.365***	569***	305***	129***
γ_2 **Febrer**	112.706 n.s.	68.290 n.s.	26.755 n.s.	5.870 n.s.	11.791***
γ_3 **Març**	400.294***	237.870***	98.105***	53.981***	10.338**
γ_4 **Abril**	692.879***	325.282***	195.886***	189.096***	-17.385***
γ_5 **Maig**	950.868***	453.793***	263.036***	252.027***	-17.988***
γ_6 **Juny**	1.161.101***	484.627***	341.522***	335.650***	-699 n.s.
γ_7 **Juliol**	1.459.013***	579.962***	437.415***	418.794***	22.842***
γ_8 **Agost ★**	**1.595.656*****	**574.373*****	**511.692*****	**458.813*****	**50.779*****
γ_9 **Setembre**	1.118.126***	466.002***	338.736***	308.256***	5.132 n.s.
γ_{10} **Octubre**	780.830***	394.084***	209.717***	186.114***	-9.085**
γ_{11} **Novembre**	147.443*	127.740***	20.298 n.s.	32.819*	-33.414***
γ_{12} **Desembre**	121.953 n.s.	54.620 n.s.	23.487 n.s.	42.732**	1.114 n.s.
δ **COVID-19**	**-1.236.493*****	**-828.054*****	**-203.264*****	**-168.668*****	**-36.507*****
Test ADF	p = 0.079 (no estacionària)	p = 0.369 (no estacionària)	p = 0.010 (estacionària)	p = 0.010 (estacionària)	p = 0.345 (no estacionària)

Font: *elaboració pròpia amb RStudio.*

*Significació: *** p<0,001 | ** p<0,01 | * p<0,05 | n.s = no significatiu. ★ = màxim impacte estacional.*

5.2.4. Impacte estimat de la COVID-19

L'impacte estimat sobre el nombre de viatgers segueix la mateixa estructura territorial que el de les pernoctacions. Per al conjunt de Catalunya, la reducció mensual estimada durant el període pandèmic és de 1.236.493 viatgers/mes ($p < 0,001$). Així mateix, Barcelona concentra de nou la major part d'aquest impacte (δ = -828.054 viatgers/mes), mentre que Lleida registra la reducció mínima (δ = -36.507 viatgers/mes). Tot i això, cal destacar que la caiguda en el nombre de nits pernoctades va ser proporcionalment major que la caiguda en el nombre d'entrades. Això suggereix que, a més de la reducció del volum de visitants, la pandèmia va comportar un escurçament de les estades.

Taula 13. Impacte estimat de la COVID-19. Viatgers. Model 3.

Territori	δ (pernoct./mes)	Error estàndard robust	R^2 ajustat M3
Catalunya	-1.236.493***	(47.718)	0,922
Barcelona	-828.054***	(28.372)	0,899
Girona	-203.264***	(11.275)	0,940
Tarragona	-168.668***	(11.135)	0,931
Lleida	-36.507***	(2.497)	0,854

***Font**: elaboració pròpia. *** p<0,001.*

Els tests de canvi estructural (Taula 14), evidencien la presència de canvis estructurals en totes les sèries analitzades. Els p-valors són inferiors a 0,05 en tots els casos, fet que permet rebutjar la hipòtesi d'estabilitat dels models.

Taula 14. Canvi estructural (CUSUM)

	Catalunya	Barcelona	Girona	Tarragona	Lleida
S	1.294	1.337	1.111	1.135	1.345
p-valor	0.002	0.002	0.013	0.011	0.01

***Font:** elaboració pròpia amb RStudio.*

5.2.5. Diagnòstic i anàlisi d'ARIMA pels viatgers

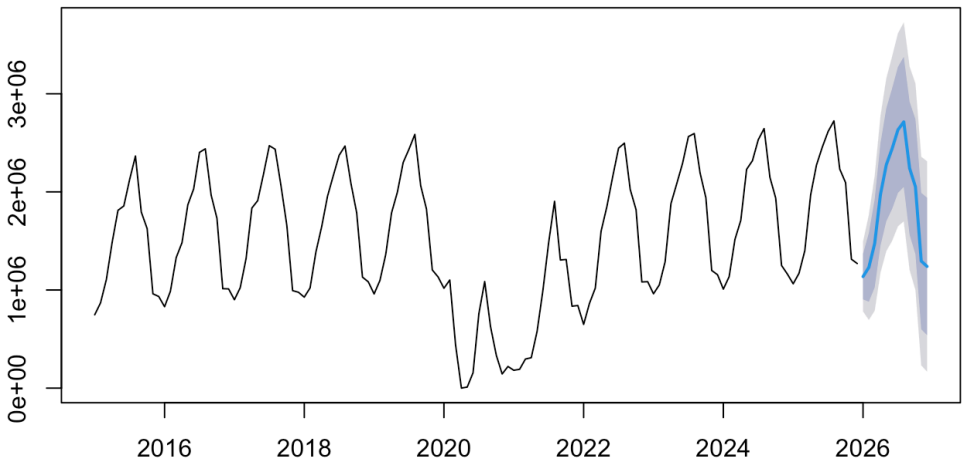
El test de Durbin-Watson (DW = 0,736 per a Catalunya) evidencia autocorrelació positiva significativa en els residus dels models de regressió. Aquesta autocorrelació invalida els errors estàndard convencionals, motiu pel qual s'han emprat errors robustos HC1 (Annex 2).

L'anàlisi de les funcions d'autocorrelació dels residus (ACF i PACF) de Catalunya revela una sèrie no estacionària amb una clara estacionalitat marcada. L'ACF presenta un decaïment molt lent amb un patró cíclic evident, mentre que la PACF confirma una elevada dependència del passat. Aquests resultats tècnics justifiquen la necessitat d'aplicar diferenciació (d) per assolir l'estacionarietat i l'ús de variables dummy estacionals dins del model ARIMA.

Els resultats del test d'Augmented Dickey-Fuller (ADF), a la Taula 12, confirmen que el comportament estadístic de les sèries torna a anar en funció del territori. Sèries estacionàries ($p<0,05$): Girona (0,010) i Tarragona (0,010). Sèries no estacionàries ($p>0,05$): Catalunya (0,079), Barcelona (0,369) i Lleida (0,345) requereixen un procés de diferenciació ($d\geq1$).

La selecció automàtica ARIMA amb la comanda "*auto.arima()*" pels viatgers presenta una estructura menys complexa que la identificada per a les pernoctacions. En el cas de Catalunya, la funció "*auto.arima()*" selecciona un model ARIMA(3,0,0)(2,1,0)[12], mentre que Girona, Tarragona i Lleida convergeixen cap a models ARIMA(1,0,0)(2,1,0)[12], incorporant únicament un terme autoregressiu no estacional.

Gràfic 5. Previsió dels viatgers a Catalunya. ARIMA.



Font: *elaboració pròpia amb RStudio.*

Aquesta diferència és rellevant: el flux de viatgers presenta una menor dependència temporal respecte a les pernoctacions. Mentre que l'arribada de visitants tendeix a respondre de manera més immediata als canvis econòmics o conjunturals, les pernoctacions mostren una persistència superior, associada a la durada de l'estada i als patrons de comportament dels

turistes. Això indica que la decisió de viatjar pot variar més ràpidament, mentre que l'estada mitjana tendeix a mantenir una major estabilitat al llarg del temps.

5.3. Anàlisi de la despesa mitjana diària dels turistes

Com s'ha mencionat a la secció 3, anomenada Metodología, la despesa mitjana diària dels turistes estrangers només ha tingut en compte dades per a Catalunya durant el període 2016-2025. Altrament, aquest terme dona una idea sobre el valor econòmic del turisme independentment del volum de visitants. Els estadístics descriptius (Taula 2) indiquen una despesa mitja de 180,46 euros per turista i dia, amb una desviació típica de 40,27 euros.

5.3.1. L'ajust global i la capacitat explicativa dels models

El Model 3 de la despesa mitja diària assoleix un R^2 ajustat de 0,550, valor inferior als obtinguts per les pernoctacions (0,933) i els viatgers (0,922). Aquesta diferència es deu a que la despesa depèn d'altres variables no tingudes en compte en el model, com ara: el volum, la composició dels mercats emissors, els preus relatius i els tipus de canvi.

Taula 15. Comparació de models. Despesa mitja diària. Catalunya. 2016–2025.

Paràmetre	Model 1 (T)	Model 2 (T+S)	Model 3 (T+S+COVID)
R^2	0,118	0,228	0,599
R^2 ajustat	0,111	0,142	0,550
F-estadístic	15,836***	2,636***	12,204***
Tendència β1 (€/mes)	0,398***	0,401***	0,398***
COVID δ (€/mes)	-	-	-67,548***
Coef. juny (γ_6) ★	-	43,093**	49,863***
Durbin-Watson	-	-	1,034
ARIMA	-	-	ARIMA(0,1,0)

Font: *elaboració pròpia. *** p<0,001; ** p<0,01. ★ = mes de màxim impacte estacional.*

5.3.2. Tendència temporal i anàlisi per subperíodes

El coeficient de tendència del Model 3 (Taula 15) és β1 = 0,398 euros/mes (p<0,001), indicant un increment mensual moderat però significatiu al llarg del període. L'anàlisi per subperíodes mostra una acceleració notable: la tendència mensual passa de 0,849 euros/mes al pre-COVID a 1,028 euros/mes al post-COVID, indicant novament un canvi en el perfil del turista estranger cap a segments de major poder adquisitiu. A diferència de les pernoctacions i els viatgers, amb el màxim impacte estacional a l'agost (γ_8), en la despesa

mitjana diària es produeix al juny (γ_6 = 49,863 euros addicionals respecte a gener, p<0,001). Això vol dir que els turistes de juny, tot i ser en nombre inferior als d'agost, presenten un perfil de despesa que pot ser més alt.

5.3.3. Impacte estimat de la COVID-19 sobre la despesa mitja diària

El coeficient de la dummy COVID per a la despesa mitja diària és δ = -67,548 euros/mes (error estàndard: 6,814; p<0,001). Això vol dir que, durant el període del xoc, la despesa mitja diària dels turistes estrangers va ser en promig 67,55 euros inferior al que hauria predit el model en absència de la crisi sanitària.

Els tests de canvi estructural (CUSUM), evidencien la presència de canvis estructurals en la sèrie analitzada: el p-valor és inferior a 0,05 (S = 0.97458, p-value = 0.04085), fet que permet rebutjar la hipòtesi d'estabilitat del model.

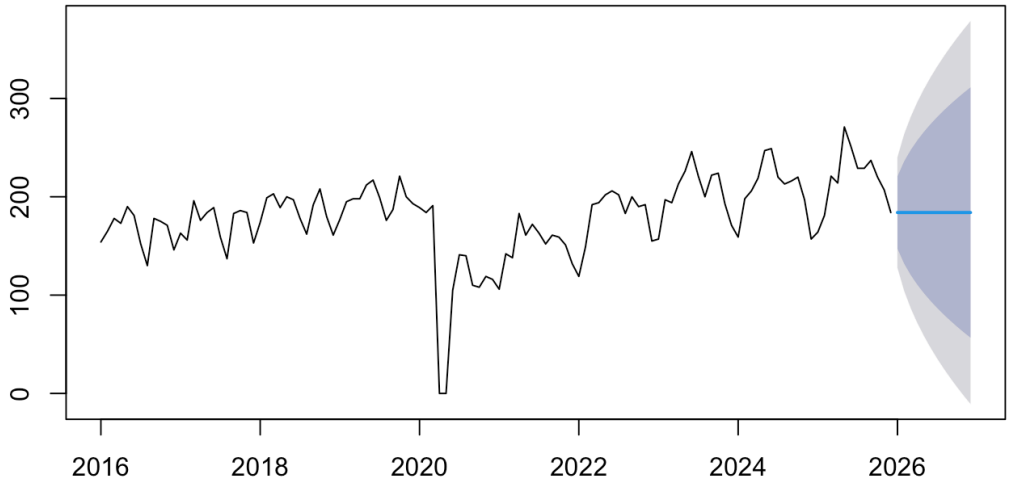
5.3.4. Diagnòstic i anàlisi d'ARIMA per la despesa

El test de Breusch-Pagan presenta un valor de BP = 26,361 (p-valor = 0,0152), fet que indica la presència d'heteroscedasticitat en els residus. El test de Durbin-Watson té un valor de DW = 1,034 (p-valor < 0,001), confirmant l'existència d'autocorrelació positiva significativa. Això suggereix que el model de regressió lineal estimat no capta completament la dinàmica temporal de la sèrie. Per corregir aquestes limitacions i assegurar una inferència estadística més robusta, s'han recalculat els errors estàndard amb l'estimador robust HC1 (Annex 3).

L'anàlisi de les funcions ACF i PACF mostra una clara dependència temporal de la sèrie. L'ACF presenta una disminució gradual de les correlacions, indicant persistència temporal, mentre que la PACF mostra un primer retard fortament significatiu. Aquest patró és coherent amb una estructura autoregressiva. Altrament, el resultat del test d'Augmented Dickey-Fuller (ADF), amb p>0,05 (valor-p = 0,081), demostra que la sèrie es no estacionària i requereix un procés de diferenciació ($d \geq 1$).

La selecció automàtica ARIMA per la despesa, presenta una estructuraque correspon a una especificació molt simple: ARIMA(0,1,0). Aquest resultat indica l'absència de components autoregressius fent que la sèrie sigui d'ordre 1 per aconseguir l'estacionarietat. Això suggereix una menor dependència temporal comparada amb les altres variables explicatives.

Gràfic 6. Previsió de la despesa mitja diària. ARIMA.



Font: *elaboració pròpia amb RStudio.*

6. CONCLUSIONS

El present treball ha analitzat l'evolució de l'activitat hotelera a Catalunya i les seves quatre províncies entre els anys 2015 i 2025, mitjançant models de regressió lineal amb components que reflecteixen la tendència, l'estacionalitat i la ruptura estructural, provocada pel xoc de la pandèmia de la COVID-19. També s'han generat prediccions del comportament de les variables explicatives (les pernoctacions, els viatgers i la despesa mitja diària) a curt termini, fent ús de models ARIMA estacionals. L'ús combinat d'ambdós models ha permés esbrinar els patrons temporals i les diferències territorials significatives de les variables.

Pel que fa a la primera pregunta de recerca, els resultats mostren una trajectòria clarament dividida en tres etapes: fins el 2019, el sector va experimentar un creixement sostingut; a partir del 14 de març de 2020, la irrupció de la pandèmia provoca una contracció sobtada de la demanda; i, finalment, a partir de la tardor de 2021, concretament al setembre, s'observa una recuperació progressiva, arribant a superar, en certs casos, els nivells prepandèmics durant principalment els anys 2024 i 2025. Realitzar els tests CUSUM ha sigut clau: es confirma l'existència d'una ruptura estructural associada a la COVID-19 en totes les sèries analitzades, es confirma la hipòsesi inicial.

D'una banda, les variables pernoctacions i viatgers presenten un ajust estadístic molt elevat, amb valors de R^2 ajustat superiors al 0,90, indicant que els components del model expliquen força bé aquestes dues variables independents. La incorporació de l'estacionalitat i de la variable fictícia vinculada a la pandèmia han incrementat significativament la capacitat explicativa dels models, fet que evidencia la importància tant dels patrons estacionals com dels xocs externs en la dinàmica del turisme hoteler.

De l'altra, l'anàlisi provincial revela certes divergències entre els territoris estudiats. Girona i Tarragona presenten una estacionalitat molt marcada, amb una concentració d'activitat durant els mesos d'estiu. Barcelona es manté més estable al llarg de l'any, gràcies a l'ampli abanic de perfils turístics que visiten la demarcació constantment. Finalment, Lleida exhibeix uns moviments més irregulars, vinculats al turisme rural i de muntanya, tenint una recuperació postpandèmica més lenta i moderada.

En relació amb la despesa mitjana diària dels turistes estrangers, els resultats demostren que és un paràmetre més complex que les pernoctacions o els viatgers. Tot i que al principi s'hi observa una tendència creixent i estadísticament significativa, el model presenta un poder explicatiu més limitat: l'R^2 ajustat és de 0,550. Aquest fet suggereix que la despesa turística està condicionada per altres factors que no s'han incorporat en el model: la variable sol estar vinculada a la composició dels mercats emissors, els tipus de canvi, els preus relatius o els canvis en el perfil dels visitants. En conseqüència, el fet d'haver omès qualsevol d'aquests factors ha provocat que el model dissenyat sigui menys robust i, el model ARIMA seleccionat, convergeix cap a un ARIMA(0,1,0), indicant una capacitat predictiva reduïda. Tot i això, els resultats permeten concloure que els models de sèries temporals constitueixen una eina útil per a la predicció a curt termini de variables vinculades al volum de l'activitat turística, encara que presentin limitacions concretes quan s'apliquen a variables econòmiques més complexes.

Des d'una perspectiva política, els resultats obtinguts permeten extreure diverses implicacions pel sector turístic. En primer lloc, la forta estacionalitat observada en determinats territoris reforça la necessitat de promoure estratègies de desestacionalització. En segon lloc, el cas de Lleida posa de manifest la necessitat d'adaptar les polítiques turístiques als diversos territoris del país i analitzar, amb més profunditat, aquells que presentin resultats més dispersos. Finalment, l'increment de la despesa mitjana després de la pandèmia podria ser el resultat d'un canvi en el perfil turístic del viatger.

Aquest estudi també presenta limitacions: la principal és la simplificació de l'impacte de la COVID-19 mitjançant una variable fictícia binària [0, 1], responsable de no permetre captar la complexitat del procés de tornar a la normalitat del sector. Futurs treballs podrien incorporar models amb múltiples punts de canvi, variables macroeconòmiques addicionals, en el cas de la despesa, o metodologies més flexibles o complexes que permetin analitzar amb més precisió les diferències temporals i territorials.

BIBLIOGRAFÍA

Articles en revistes acadèmiques i capítols

Arbulú, I., Lozano, J., i Rey-Maquieira, J. (2021). Tourism and economic recovery: The role of domestic demand in Spain. Tourism Economics, 27(6), 1291–1310. [Document en línia] Disponible en: https://doi.org/10.1177/1354816620941946

Benítez-Aurioles, B. (2021). El turismo interior en España ante la crisis de la COVID-19. Boletín Económico de ICE, (3139), 43–59. [Document en línia] Disponible en: https://doi.org/10.32796/bice.2021.3139.7285

Dogru, T., Bulut, U., i Sirakaya-Turk, E. (2020). Modelling tourism demand: The impact of COVID-19. Tourism Management Perspectives, 35, 100676. [Document en línia] Disponible en: https://doi.org/10.1016/j.tmp.2020.100676

Gössling, S., Scott, D., i Hall, C. M. (2020). Pandemics, tourism and global change: A rapid assessment of COVID-19. Journal of Sustainable Tourism, 29(1), 1–20. [Document en línia] Disponible en: https://doi.org/10.1080/09669582.2020.1758708

Kulendran, N., i Witt, S. F. (2003). Leading indicator tourism forecasts. Tourism Management, 24(5), 503–510. [Document en línia] Disponible en: https://doi.org/10.1016/S0261-5177(03)00010-4

Law, R., Li, G., Fong, D. K. C., i Han, X. (2019). Tourism demand forecasting: A deep learning approach. Annals of Tourism Research, 75, 410–423. [Document en línia] Disponible en: https://doi.org/10.1016/j.annals.2019.01.014

Lim, C. (1997). Review of international tourism demand models. Annals of Tourism Research, 24(4), 835–849. [Document en línia] Disponible en: https://doi.org/10.1016/S0160-7383(97)00049-2

Smeral, E. (2009). The impact of the financial and economic crisis on European tourism. Journal of Travel Research, 48(1), 3–13. [Document en línia] Disponible en: https://doi.org/10.1177/0047287509336332

Song, H., i Li, G. (2008). Tourism demand modelling and forecasting: A review of recent research. Tourism Management, 29(2), 203–220. [Document en línia] Disponible en: https://doi.org/10.1016/j.tourman.2007.07.016

Witt, S. F., i Witt, C. A. (1995). Forecasting tourism demand: A review of empirical research. International Journal of Forecasting, 11(3), 447–475. [Document en línia] Disponible en: https://doi.org/10.1016/0169-2070(95)00591-7

Llibres i manuals metodològics

Box, G. E. P., i Jenkins, G. M. (1970). Time series analysis: Forecasting and control. Holden-Day.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., i Ljung, G. M. (2015). Time series analysis: Forecasting and control (5a ed.). Wiley.

Enders, W. (2015). Applied econometric time series (4a ed.). Wiley.

Granger, C. W. J. (2001). Essays in econometrics: Collected papers of Clive W. J. Granger. Cambridge University Press.

Hamilton, J. D. (1994). Time series analysis. Princeton University Press.

Hyndman, R. J., i Athanasopoulos, G. (2021). Forecasting: Principles and practice (3a ed.). OTexts. https://otexts.com/fpp3/

Song, H., Witt, S. F., i Li, G. (2009). The advanced econometrics of tourism demand. Routledge.

Wooldridge, J. M. (2013). Introductory econometrics: A modern approach (5a ed.). South-Western Cengage Learning.

TFGs i estudis aplicats

AQR-Lab, Universitat de Barcelona. (2020). Impacto de la crisis turística originada por el COVID-19 sobre la economía catalana. Asociación Española de Ciencia Regional. [Document en línia] Disponible en: https://aecr.org/en/impacto-de-la-crisis-turistica-originada-por-el-covid-19-sobre-la-economia-catalana/

Torres Prats, M. (2023). Anàlisi de l'impacte de la Covid-19 en l'activitat turística de Barcelona [Treball de Final de Grau, Universitat Autònoma de Barcelona]. Dipòsit Digital de Documents de la UAB. https://ddd.uab.cat/pub/tfg/2023/tfg_2536109/Torres_Prats_TFG.pdf

Grendy Santiago Huacón Cabrera (2020). Impacto del tráfico aéreo sobre los niveles de contaminación atmosférica. Un análisis para la ciudad de Barcelona. [Treball de Final de Grau, Universitat Autònoma de Barcelona]. Facilitat per la tutora, extret del Dipòsit Digital de Documents de la UAB.

Fonts estadístiques

Ara. (febrer de 2026). Rècord històric: Catalunya va superar els 20 milions de turistes estrangers el 2025. [Document en línia] Disponible en: https://es.ara.cat/economia/record-historico-cataluna-supero-20-millones-turistas-extranjeros-2025

El Periódico. (26 de desembre de 2024). El turisme puja la seva aportació a l'economia fins al 12,3% del PIB. [Document en línia] Disponible en: https://www.elperiodico.cat/ca/economia/20241226/turisme-puja-seva-aportacio-l-112940625

Generalitat de Catalunya, Departament d'Empresa i Treball. (2024). Estructura productiva: turisme. [Document en línia] Disponible en: https://economia.gencat.cat/ca/ambits-actuacio/economia-catalana/trets/estructura-productiva/turisme/

Idescat. (2025). Indicadors bàsics de l'activitat turística a Catalunya. [Document en línia] Disponible en: https://www.idescat.cat/pub/?id=turhot

INE. (2025). Encuesta de Ocupación Hotelera. Base de dades mensual 2015–2025. [Document en línia] Disponible en: https://www.ine.es/jaxiT3/Tabla.htm?t=2074

INE. (2025). Coyuntura Turística Hotelera (EOH/IPH/IRSH). Nota de premsa anual 2024. [Document en línia] Disponible en: https://www.ine.es/dyngs/Prensa/CTH1224.htm

Nota sobre l'ús d'intel·ligència artificial

Amb l'autorització de la tutora del treball, s'ha fet ús d'eines d'intel·ligència artificial (IA) com a suport auxiliar en dues tasques concretes: la revisió i contraste de fonts bibliogràfiques, i la polítura linguística i formal del text. En cap cas la IA ha generat contingut analític, estadístic o interpretatiu. Tots els models economètrics, les anàlisis empíriques, les interpretacions dels resultats i les conclusions són de l'autora del treball, Judit Falgà Grasa.

ANNEXOS

A continuació es presenten tres documents amb el codi principal emprat en l'anàlisi econometrica i els seus resultats. El codi és completament reproduïble a partir de les dades de l'EOH descarregades de l'INE o l'Idescat, i inclou els passos d'importació, transformació, estimació de models, diagnòstic i predicció.

L'annex 1 és la sortida d'RStudio de les pernoctacions a Catalunya. L'annex 2 és la sortida d'RStudio pel nombre de viatgers de Catalunya. L'annex 3 és la sortida d'RStudio per a la despesa mitja diària de Catalunya. S'ha replicat el matex procediment estadístic per a totes les tres variables i per a cada una de les províncies amb les seves corresponents bases de dades però, per volum, només s'inclou l'agregat dels territoris, Catalunya.

Annex 1: Anàlisi de les pernoctacions hoteleres a Catalunya

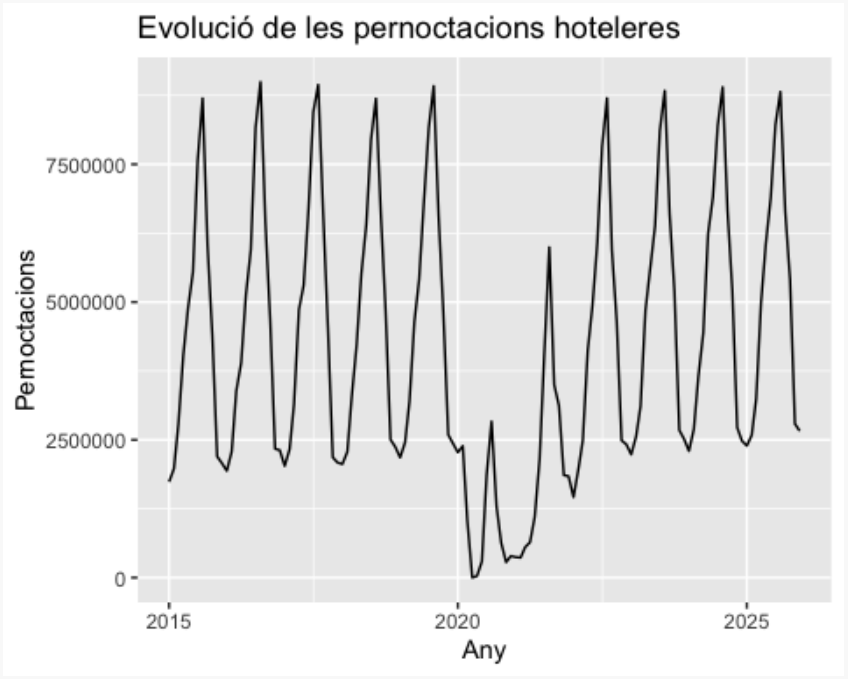
Introducció

Aquest document analitza l'evolució de les pernoctacions hoteleres a Catalunya mitjançant tècniques de sèries temporals i regressió economètrica. L'objectiu és identificar la tendència estructural, l'estacionalitat mensual i l'impacte de la COVID-19 sobre l'activitat hotelera.

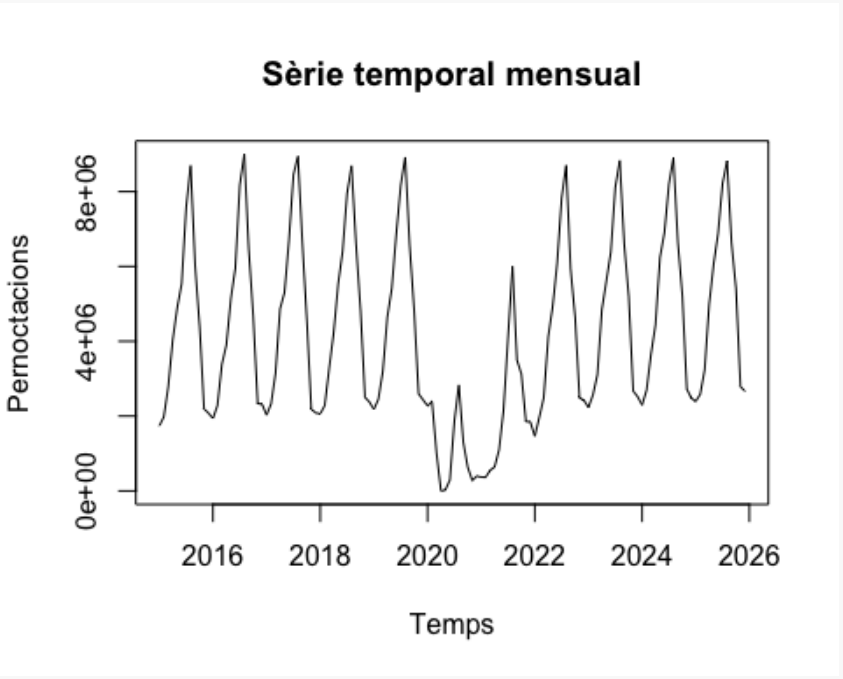
```
########################
# 1. PAQUETS
########################



library(readxl)
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────────────────── tidyverse 2.0.0
──
## ✔ dplyr     1.2.1     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.4     ✔ tibble    3.2.1
## ✔ lubridate 1.9.5     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────────────────────────
```

```
tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(lubridate)
library(forecast)

## Registered S3 method overwritten by 'quantmod':
##   method           from
##   as.zoo.data.frame zoo

library(tseries)
library(lmtest)

## Loading required package: zoo
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(sandwich)
library(modelsummary)

#########################
# 2. IMPORTAR DADES
#########################


datos <- read_excel("/Users/judit/Downloads/PernoctacionsC.xlsx")


# Ordenar cronològicament
datos <- datos %>%
  arrange(Any, mes)

mean(datos$Pern)

## [1] 4178438

sd(datos$Pern)
```

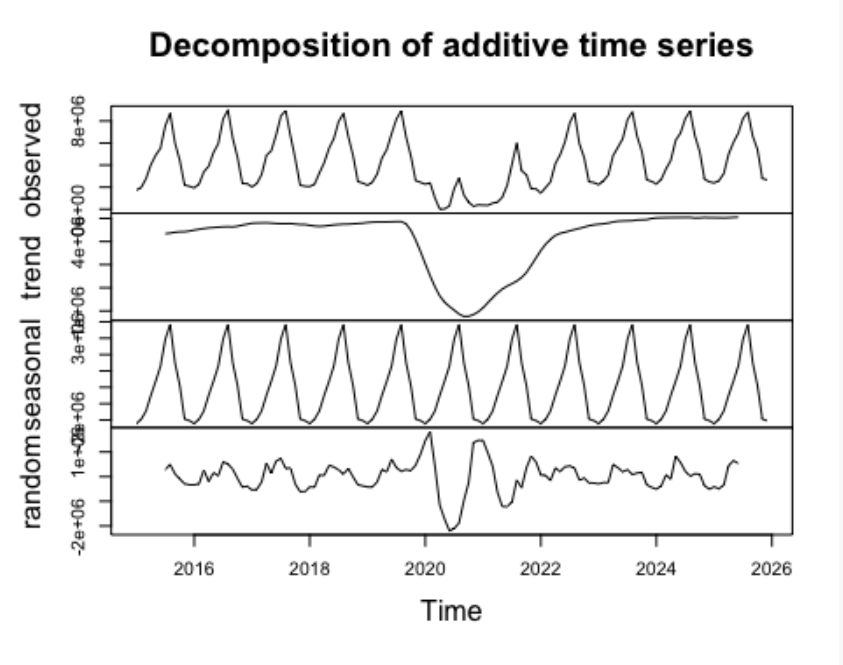
```
## [1] 2465848

min(datos$Pern)

## [1] 0

max(datos$Pern)

## [1] 8994168

# Data mensual
datos$fecha <- as.Date(paste(datos$Any, datos$mes, "01", sep="-"))

# Tendència temporal
datos$time <- 1:nrow(datos)

# Convertir mes en factor (gener com a base)
datos$mes_factor <- factor(datos$mes)

#########################
# 4. VARIABLE COVID
#########################

# Dummy COVID: març 2020 - setembre 2021
datos$Covid <- ifelse(datos$fecha >= as.Date("2020-03-01") &
                      datos$fecha <= as.Date("2021-09-30"), 1, 0)

#########################
# 5. ANÀLISI DESCRIPTIVA
#########################

ggplot(datos, aes(x = fecha, y = Pern)) +
 geom_line(color = "black") +
 labs(title = "Evolució de les pernoctacions hoteleres",
      x = "Any",
      y = "Pernoctacions")
```



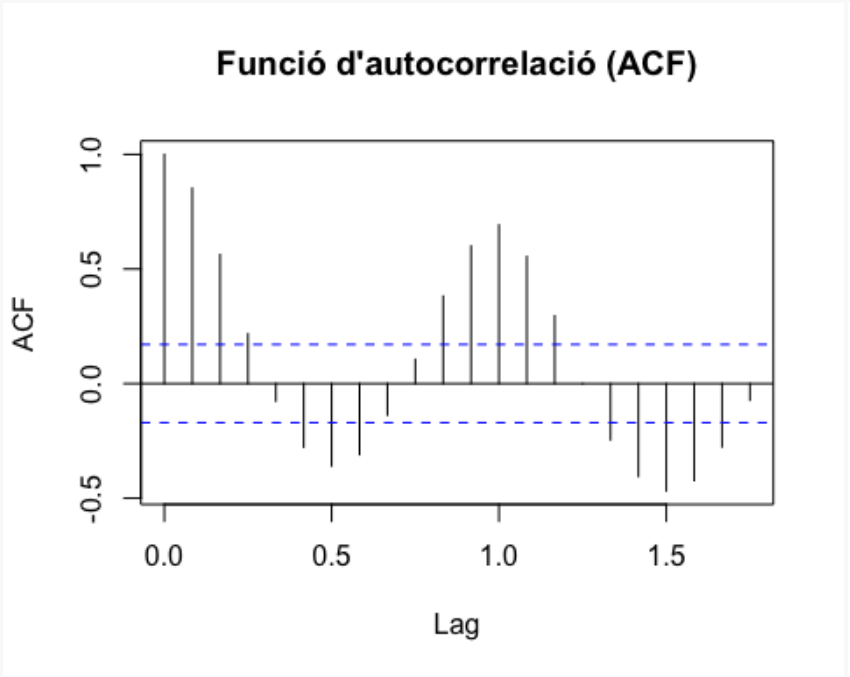

```
##########################
# 6. SÈRIE TEMPORAL
##########################

# Definir principi de la sèrie
inicio <- c(datos$Any[1], datos$mes[1])

# Sèrie temporal mensual
serie_ts <- ts(datos$Pern,
               start = inicio,
               frequency = 12)
# Resultat
plot(serie_ts,
    main = "Sèrie temporal mensual",
    ylab = "Pernoctacions",
    xlab = "Temps")
```



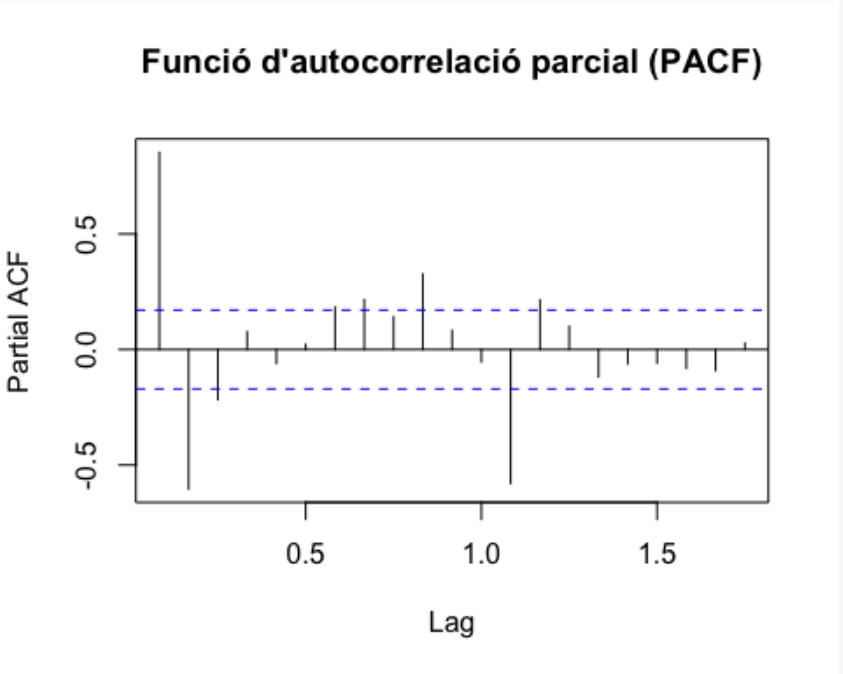

```r
datos <- datos[order(datos$Any, datos$mes), ]

########################
# 7. DESCOMPOSICIÓ
########################

descomp <- decompose(serie_ts, type = "additive")
plot(descomp)
```



```
#########################
# 8. ESTACIONARIETAT
#########################

resultat_adf <- adf.test(serie_ts)

## Warning in adf.test(serie_ts): p-value smaller than printed p-value

print(resultat_adf)

##
##  Augmented Dickey-Fuller Test
##
## data:  serie_ts
## Dickey-Fuller = -4.3357, Lag order = 5, p-value = 0.01
## alternative hypothesis: stationary

# El test Dickey-Fuller augmentat rebutja la hipòtesi nul·la d'arrel unitària amb un nivell de significació inferior a l'1%, indicant que la sèrie és estacionària.

#########################
# 9. ACF I PACF
#########################
acf(serie_ts, main = "Funció d'autocorrelació (ACF)")
```



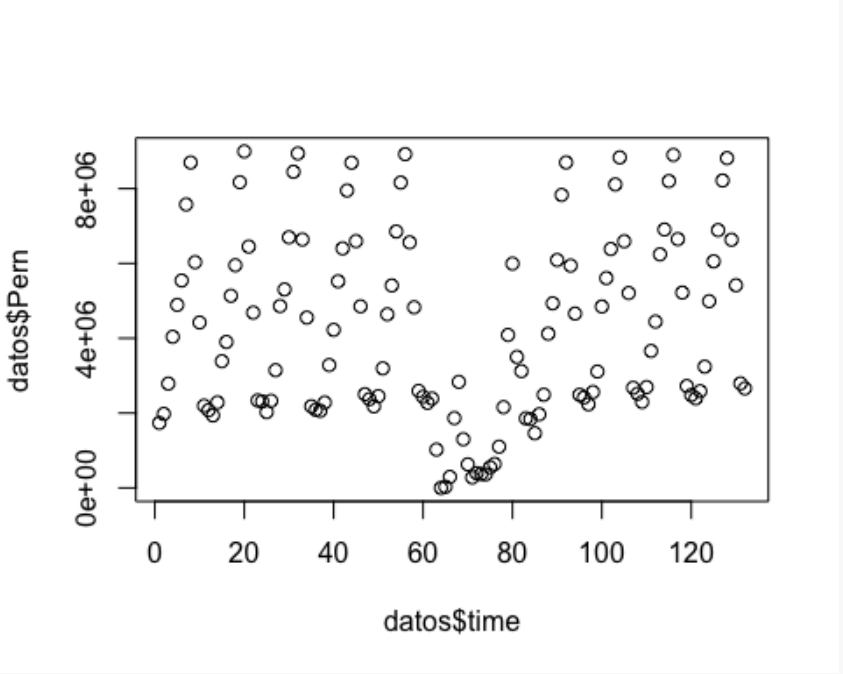
```
pacf(serie_ts, main = "Funció d'autocorrelació parcial (PACF)")
```




```
#########################
# 10. MODELS ECONOMÈTRICS
#########################

# MODEL 1: Tendència temporal
ModelT1 <- lm(Pern ~ time, data = datos)
ModelT1<-lm(datos$Pern~datos$time)
SModelT1<-summary(ModelT1)
summary(ModelT1)

##
## Call:
## lm(formula = datos$Pern ~ datos$time)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -4171496 -1842781  -644567  1914581  4944852
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3993780     432955   9.224 6.71e-16 ***
```

```
## datos$time      2777       5649   0.492    0.624
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2473000 on 130 degrees of freedom
## Multiple R-squared:  0.001855,   Adjusted R-squared:  -0.005823
## F-statistic: 0.2416 on 1 and 130 DF,  p-value: 0.6239
```

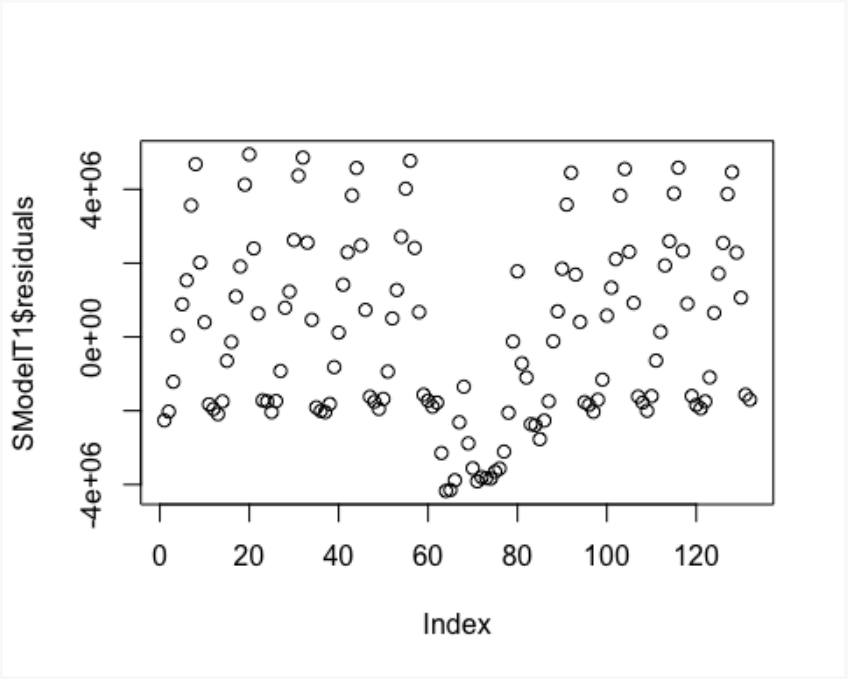
```
plot(datos$time,datos$Pern)
```




```
names(SModelT1)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelT1$residuals)
```




```
library(stargazer)

##
## Please cite as:

##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

stargazer(ModelT1, type="text")

##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                                Pern
## ------------------------------------------------
## time                          2,776.820
##                             (5,648.969)
##
## Constant                   3,993,780.000***
##                            (432,954.600)
##
## ------------------------------------------------
## Observations                      132
```

```
## R2                          0.002
## Adjusted R2                -0.006
## Residual Std. Error  2,473,016.000 (df = 130)
## F Statistic           0.242 (df = 1; 130)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

```
# MODEL 2: Tendència + estacionalitat mensual
ModelT2<-lm(datos$Pern~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+
datos$dm7+datos$dm8+datos$dm9+
            datos$dm10+datos$dm11+datos$dm12)
SModelT2<-summary(ModelT2)


summary(ModelT2)
```

```
##
## Call:
## lm(formula = datos$Pern ~ datos$time + datos$dm2 + datos$dm3 +
##     datos$dm4 + datos$dm5 + datos$dm6 + datos$dm7 + datos$dm8 +
##     datos$dm9 + datos$dm10 + datos$dm11 + datos$dm12)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -5281846   170144   380219   812333  1592621
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1800417     505826   3.559 0.000535 ***
## datos$time      1739       3490   0.498 0.619246
## datos$dm2     263673     648894   0.406 0.685221
## datos$dm3     806388     648922   1.243 0.216437
## datos$dm4    1789339     648969   2.757 0.006749 **
## datos$dm5    2653866     649035   4.089 7.91e-05 ***
## datos$dm6    3555287     649119   5.477 2.45e-07 ***
## datos$dm7    5231040     649223   8.057 6.80e-13 ***
## datos$dm8    6110434     649345   9.410 4.67e-16 ***
## datos$dm9    3798081     649485   5.848 4.46e-08 ***
## datos$dm10   2405214     649645   3.702 0.000325 ***
```

```
## datos$dm11    315968     649823   0.486 0.627694
## datos$dm12    219230     650020   0.337 0.736511
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1522000 on 119 degrees of freedom
## Multiple R-squared:  0.654,  Adjusted R-squared:  0.6191
## F-statistic: 18.75 on 12 and 119 DF,  p-value: < 2.2e-16
```

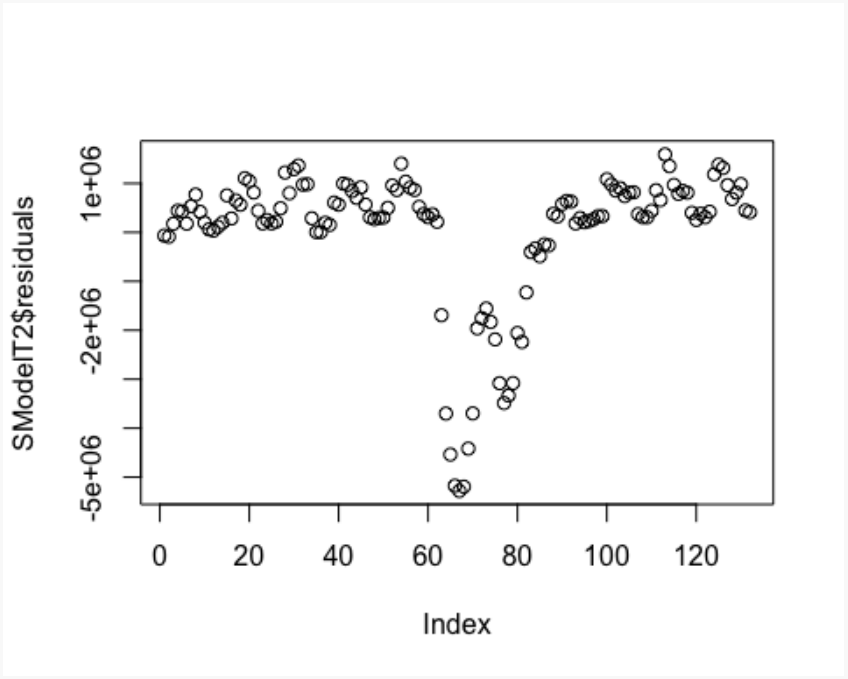
```
names(SModelT2)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
stargazer(ModelT1, ModelT2, type="text")
```

```
##
## ==========================================================================
##                              Dependent variable:
##                    ------------------------------------------------
##                                     Pern
##                          (1)                  (2)
## --------------------------------------------------------------------
## time                   2,776.820              1,739.022
##                      (5,648.969)            (3,490.444)
##
## dm2                                          263,673.400
##                                            (648,894.100)
##
## dm3                                          806,388.300
##                                            (648,922.300)
##
## dm4                                          1,789,339.000***
##                                            (648,969.200)
##
## dm5                                          2,653,866.000***
##                                            (649,034.900)
##
```

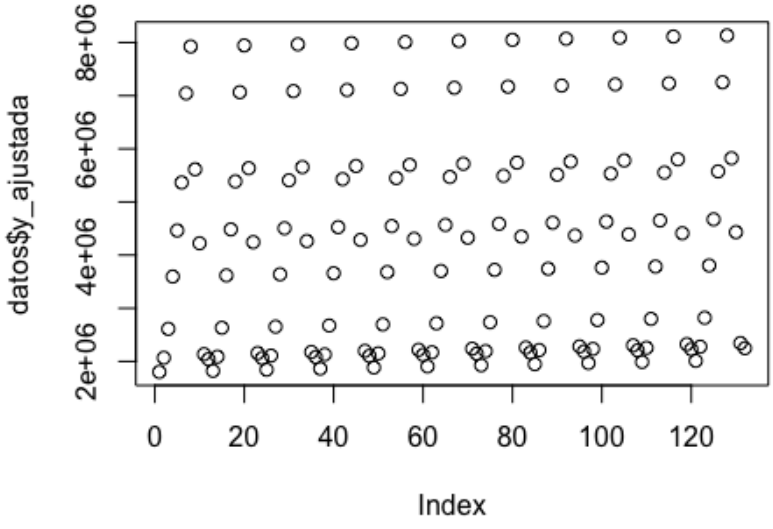
```
## dm6                                3,555,287.000***
##                                   (649,119.400)
##
## dm7                                5,231,040.000***
##                                   (649,222.600)
##
## dm8                                6,110,433.000***
##                                   (649,344.600)
##
## dm9                                3,798,081.000***
##                                   (649,485.300)
##
## dm10                               2,405,214.000***
##                                   (649,644.700)
##
## dm11                                315,967.500
##                                   (649,822.800)
##
## dm12                                219,229.500
##                                   (650,019.700)
##
## Constant              3,993,780.000***        1,800,417.000***
##                      (432,954.600)          (505,825.500)
##
## ------------------------------------------------------------------
## Observations               132                 132
## R2                        0.002               0.654
## Adjusted R2               -0.006              0.619
## Residual Std. Error 2,473,016.000 (df = 130) 1,521,770.000 (df = 119)
## F Statistic          0.242 (df = 1; 130)    18.747*** (df = 12; 119)
## ==================================================================
## Note:                                *p<0.1; **p<0.05; ***p<0.01

plot(SModelT2$residuals)
```




```
y_ajustada <- fitted(ModelT2)

y_ajustada

##      1      2      3      4      5      6      7      8      9     10
## 1802156 2067568 2612022 3596712 4462978 5366138 7043630 7924762 5614148 4223021
##     11     12     13     14     15     16     17     18     19     20
## 2135513 2040514 1823024 2088436 2632890 3617580 4483846 5387006 7064498 7945630
##     21     22     23     24     25     26     27     28     29     30
## 5635017 4243889 2156382 2061383 1843892 2109305 2653758 3638448 4504714 5407874
##     31     32     33     34     35     36     37     38     39     40
## 7085366 7966499 5655885 4264758 2177250 2082251 1864760 2130173 2674627 3659317
##     41     42     43     44     45     46     47     48     49     50
## 4525582 5428743 7106235 7987367 5676753 4285626 2198118 2103119 1885629 2151041
##     51     52     53     54     55     56     57     58     59     60
## 2695495 3680185 4546451 5449611 7127103 8008235 5697622 4306494 2218986 2123987
##     61     62     63     64     65     66     67     68     69     70
## 1906497 2171909 2716363 3701053 4567319 5470479 7147971 8029104 5718490 4327362
##     71     72     73     74     75     76     77     78     79     80
## 2239855 2144856 1927365 2192778 2737232 3721922 4588187 5491347 7168840 8049972
##     81     82     83     84     85     86     87     88     89     90
## 5739358 4348231 2260723 2165724 1948233 2213646 2758100 3742790 4609056 5512216
##     91     92     93     94     95     96     97     98     99    100
```

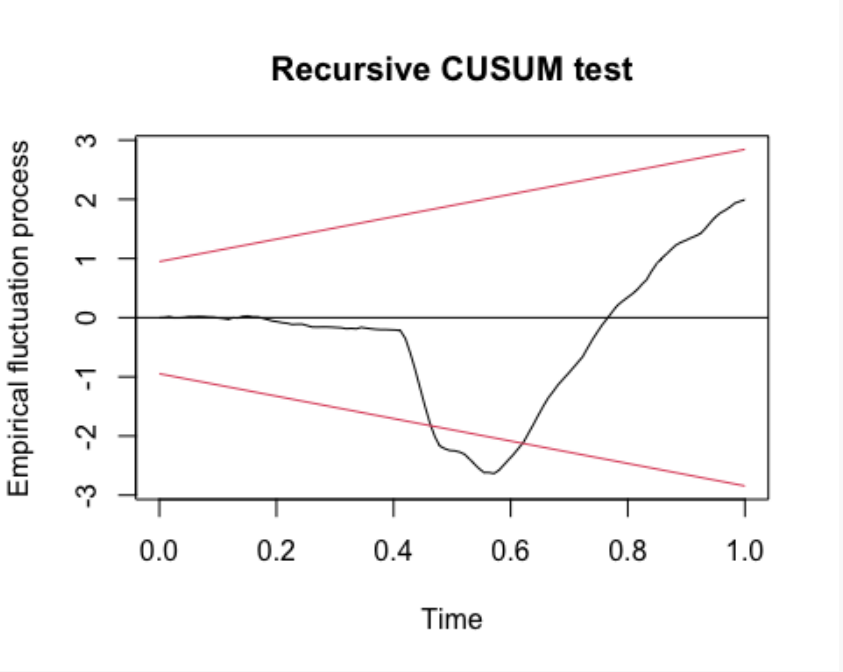
```
## 7189708 8070840 5760226 4369099 2281591 2186592 1969102 2234514 2778968 3763658
##     101     102     103     104     105     106     107     108     109     110
## 4629924 5533084 7210576 8091708 5781095 4389967 2302459 2207460 1989970 2255382
##     111     112     113     114     115     116     117     118     119     120
## 2799836 3784526 4650792 5553952 7231444 8112577 5801963 4410836 2323328 2228329
##     121     122     123     124     125     126     127     128     129     130
## 2010838 2276251 2820705 3805395 4671660 5574820 7252313 8133445 5822831 4431704
##     131     132
## 2344196 2249197
```

```
datos$y_ajustada<-fitted(ModelT2)

plot(datos$y_ajustada )
```




```
Data_1 <- datos

library(strucchange)

##
## Attaching package: 'strucchange'

## The following object is masked from 'package:stringr':
##
##     boundary
```

```
efp_model <-
efp(datos$Pern~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+datos$dm7
+datos$dm8+datos$dm9+
        datos$dm10+datos$dm11+datos$dm12, type="Rec-CUSUM")

plot (efp_model)
```




```
sctest (efp_model)

##
##  Recursive CUSUM test
##
## data:  efp_model
## S = 1.2406, p-value = 0.003984

# MODEL 3: Tendència + estacionalitat + CÒVID
ModelT3<-lm(datos$Pern~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+
datos$dm7+datos$dm8+datos$dm9+
        datos$dm10+datos$dm11+datos$dm12+datos$Covid)
SModelT3<-summary(ModelT3)

names(ModelT3)
```

```
##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "xlevels"       "call"          "terms"         "model"
```

```
stargazer(ModelT1, ModelT2,ModelT3, type="text")
```

```
##
##
==========================================================================
=======================
##                                  Dependent variable:
##                  ---------------------------------------------------------------------
##                                        Pern
##                        (1)                (2)                 (3)
## ---------------------------------------------------------------------------------------
## time                  2,776.820          1,739.022           3,887.225***
##                     (5,648.969)        (3,490.444)          (1,461.721)
##
## dm2                                    263,673.400          261,525.200
##                                      (648,894.100)        (271,221.000)
##
## dm3                                    806,388.300        1,145,804.000***
##                                      (648,922.300)        (271,609.600)
##
## dm4                                   1,789,339.000***     2,126,607.000***
##                                      (648,969.200)        (271,624.500)
##
## dm5                                   2,653,866.000***     2,988,986.000***
##                                      (649,034.900)        (271,647.200)
##
## dm6                                   3,555,287.000***     3,888,259.000***
##                                      (649,119.400)        (271,677.700)
##
## dm7                                   5,231,040.000***     5,561,864.000***
##                                      (649,222.600)        (271,716.200)
##
## dm8                                   6,110,433.000***     6,439,109.000***
##                                      (649,344.600)        (271,762.500)
```

```
##
## dm9                                3,798,081.000***        4,124,608.000***
##                                   (649,485.300)           (271,816.600)
##
## dm10                               2,405,214.000***        2,385,881.000***
##                                   (649,644.700)           (271,536.000)
##
## dm11                                315,967.500             294,485.500
##                                   (649,822.800)           (271,610.700)
##
## dm12                                219,229.500             195,599.200
##                                   (650,019.700)           (271,693.300)
##
## Covid                                                      -3,780,837.000***
##                                                           (159,320.900)
##
## Constant                3,993,780.000***        1,800,417.000***        2,013,089.000***
##                        (432,954.600)           (505,825.500)           (211,611.900)
##
## ------------------------------------------------------------------------------------------
## Observations                 132                     132                     132
## R2                         0.002                   0.654                   0.940
## Adjusted R2                -0.006                  0.619                   0.933
## Residual Std. Error 2,473,016.000 (df = 130) 1,521,770.000 (df = 119)  636,060.500 (df = 118)
## F Statistic          0.242 (df = 1; 130)    18.747*** (df = 12; 119) 142.371*** (df = 13; 118)
##
==========================================================================
======================
## Note:                                           *p<0.1; **p<0.05; ***p<0.01
```

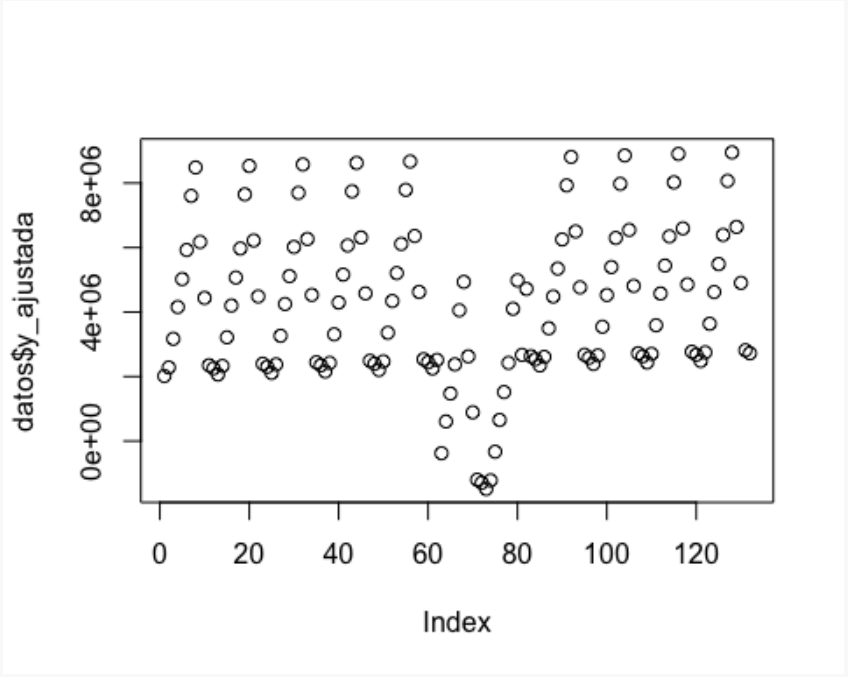
```
y_ajustada <- fitted(ModelT3)
```

```
y_ajustada
```

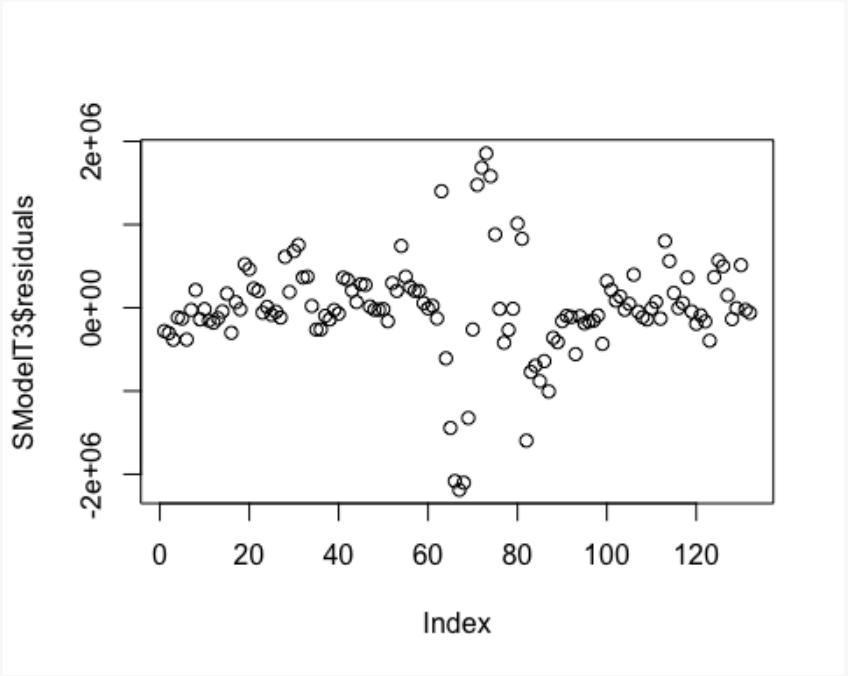
```
##         1         2         3         4         5         6         7
##  2016975.9  2282388.3  3170554.8  4155244.8  5021510.5  5924670.6  7602162.8
##         8         9        10        11        12        13        14
##  8483295.0  6172681.3  4437841.4  2350333.6  2255334.6  2063622.6  2329035.0
```

```
##       15       16       17       18       19       20       21
## 3217201.5 4201891.5 5068157.2 5971317.3 7648809.5 8529941.7 6219328.0
##       22       23       24       25       26       27       28
## 4484488.1 2396980.3 2301981.3 2110269.3 2375681.8 3263848.2 4248538.2
##       29       30       31       32       33       34       35
## 5114803.9 6017964.0 7695456.2 8576588.4 6265974.7 4531134.8 2443627.0
##       36       37       38       39       40       41       42
## 2348628.0 2156916.0 2422328.5 3310494.9 4295184.9 5161450.6 6064610.7
##       43       44       45       46       47       48       49
## 7742102.9 8623235.1 6312621.4 4577781.5 2490273.7 2395274.7 2203562.7
##       50       51       52       53       54       55       56
## 2468975.2 3357141.6 4341831.6 5208097.3 6111257.4 7788749.6 8669881.8
##       57       58       59       60       61       62       63
## 6359268.1 4624428.2 2536920.4 2441921.4 2250209.4 2515621.9 -377049.2
##       64       65       66       67       68       69       70
##  607640.8 1473906.5 2377066.6 4054558.8 4935691.1 2625077.3  890237.5
##       71       72       73       74       75       76       77
## -1197270.4 -1292269.4 -1483981.4 -1218568.9 -330402.5  654287.5 1520553.2
##       78       79       80       81       82       83       84
## 2423713.3 4101205.5 4982337.8 2671724.0 4717721.7 2630213.8 2535214.8
##       85       86       87       88       89       90       91
## 2343502.8 2608915.3 3497081.7 4481771.7 5348037.4 6251197.5 7928689.7
##       92       93       94       95       96       97       98
## 8809822.0 6499208.2 4764368.4 2676860.5 2581861.5 2390149.5 2655562.0
##       99      100      101      102      103      104      105
## 3543728.4 4528418.4 5394684.1 6297844.2 7975336.4 8856468.7 6545854.9
##      106      107      108      109      110      111      112
## 4811015.1 2723507.2 2628508.2 2436796.2 2702208.7 3590375.1 4575065.1
##      113      114      115      116      117      118      119
## 5441330.8 6344490.9 8021983.1 8903115.4 6592501.6 4857661.8 2770154.0
##      120      121      122      123      124      125      126
## 2675155.0 2483442.9 2748855.4 3637021.8 4621711.8 5487977.5 6391137.6
##      127      128      129      130      131      132
## 8068629.8 8949762.1 6639148.3 4904308.5 2816800.7 2721801.7
```

```
datos$y_ajustada<-fitted(ModelT3)

plot(datos$y_ajustada )
```




```
plot(SModelT3$residuals)
```




```
# Canvi de nom per poder fer la comparació
Data_1 <- datos
```

```
ModelAC<-lm(Data_1$Pern~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
            Data_1$dm10+Data_1$dm11+Data_1$dm12+Data_1$Covid)

stargazer(ModelAC, type="text")

##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                               Pern
## ------------------------------------------------
## time                       3,887.225***
##                           (1,461.721)
##
## dm2                         261,525.200
##                          (271,221.000)
##
## dm3                       1,145,804.000***
##                          (271,609.600)
##
## dm4                       2,126,607.000***
##                          (271,624.500)
##
## dm5                       2,988,986.000***
##                          (271,647.200)
##
## dm6                       3,888,259.000***
##                          (271,677.700)
##
## dm7                       5,561,864.000***
##                          (271,716.200)
##
## dm8                       6,439,109.000***
##                          (271,762.500)
##
## dm9                       4,124,608.000***
```

```
##                        (271,816.600)
##
## dm10                    2,385,881.000***
##                        (271,536.000)
##
## dm11                     294,485.500
##                        (271,610.700)
##
## dm12                     195,599.200
##                        (271,693.300)
##
## Covid                   -3,780,837.000***
##                        (159,320.900)
##
## Constant                2,013,089.000***
##                        (211,611.900)
##
## ------------------------------------------------
## Observations                  132
## R2                           0.940
## Adjusted R2                  0.933
## Residual Std. Error   636,060.500 (df = 118)
## F Statistic         142.371*** (df = 13; 118)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

```
#Model Estimat etapa Pre-còvid

Data_1 <- subset(datos, time< 61)



ModelAC1<-lm(Data_1$Pern~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
            Data_1$dm10+Data_1$dm11+Data_1$dm12)

stargazer(ModelAC1, type="text")
```

```
##
## ================================================
##                       Dependent variable:
##                   ---------------------------
##                             Pern
## ------------------------------------------------
## time                     9,515.358***
##                         (1,596.023)
##
## dm2                      267,969.600**
##                        (132,700.500)
##
## dm3                     1,153,032.000***
##                        (132,729.300)
##
## dm4                     2,316,429.000***
##                        (132,777.200)
##
## dm5                     3,226,954.000***
##                        (132,844.400)
##
## dm6                     4,253,999.000***
##                        (132,930.600)
##
## dm7                     6,016,961.000***
##                        (133,036.000)
##
## dm8                     6,794,244.000***
##                        (133,160.400)
##
## dm9                     4,392,394.000***
##                        (133,303.800)
##
## dm10                    2,594,236.000***
##                        (133,466.100)
##
## dm11                     280,068.000**
```

```
##                         (133,647.300)
##
## dm12                     164,080.900
##                         (133,847.300)
##
## Constant                1,750,321.000***
##                         (101,958.300)
##
## ------------------------------------------------
## Observations                  60
## R2                           0.993
## Adjusted R2                  0.992
## Residual Std. Error    209,802.700 (df = 47)
## F Statistic         586.408*** (df = 12; 47)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

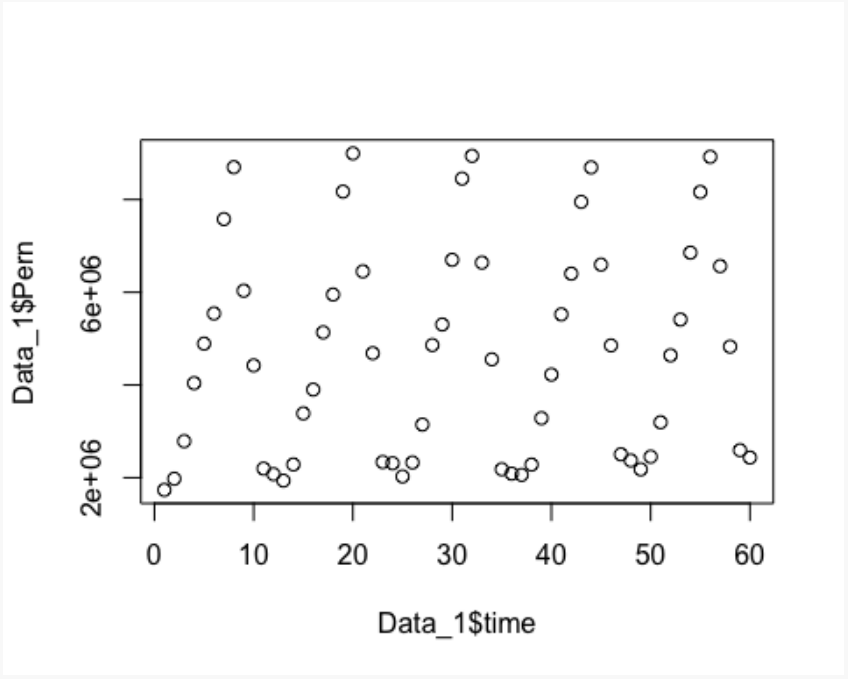
```
summary(ModelAC1)
```

```
##
## Call:
## lm(formula = Data_1$Pern ~ Data_1$time + Data_1$dm2 + Data_1$dm3 +
##     Data_1$dm4 + Data_1$dm5 + Data_1$dm6 + Data_1$dm7 + Data_1$dm8 +
##     Data_1$dm9 + Data_1$dm10 + Data_1$dm11 + Data_1$dm12)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -518101 -123216   -3141   80800  527800
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1750321     101958  17.167  < 2e-16 ***
## Data_1$time     9515       1596   5.962 3.07e-07 ***
## Data_1$dm2    267970     132700   2.019   0.0492 *
## Data_1$dm3   1153032     132729   8.687 2.45e-11 ***
## Data_1$dm4   2316429     132777  17.446  < 2e-16 ***
## Data_1$dm5   3226954     132844  24.291  < 2e-16 ***
## Data_1$dm6   4253999     132931  32.002  < 2e-16 ***
```

```
## Data_1$dm7   6016961     133036  45.228  < 2e-16 ***
## Data_1$dm8   6794244     133160  51.023  < 2e-16 ***
## Data_1$dm9   4392394     133304  32.950  < 2e-16 ***
## Data_1$dm10  2594236     133466  19.437  < 2e-16 ***
## Data_1$dm11   280068     133647   2.096   0.0415 *
## Data_1$dm12   164081     133847   1.226   0.2264
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 209800 on 47 degrees of freedom
## Multiple R-squared:  0.9934, Adjusted R-squared:  0.9917
## F-statistic: 586.4 on 12 and 47 DF,  p-value: < 2.2e-16
```

```
SModelAC1<-summary(ModelAC1)
```

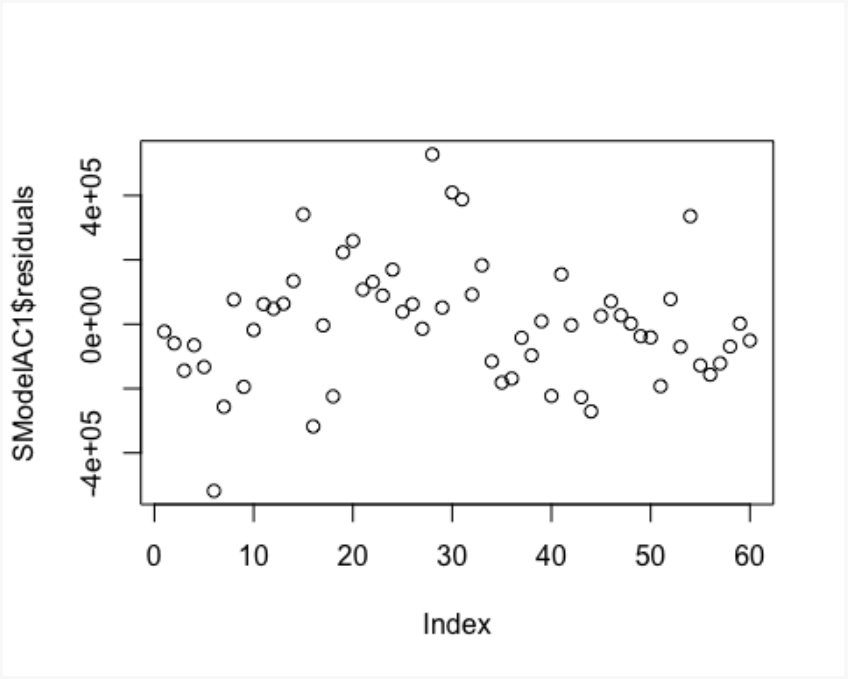
```
plot(Data_1$time,Data_1$Pern)
```




```
names(SModelAC1)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelAC1$residuals)
```




```
#Model Estimat etapa Post-còvid (suposem que comença a partir de l'observació 81)

Data_1 <- subset(datos, time>81)


ModelAC2<-lm(Data_1$Pern~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
        Data_1$dm10+Data_1$dm11+Data_1$dm12)

stargazer(ModelAC1,ModelAC2, type="text")

##
## ==========================================================================
##                              Dependent variable:
##                  ------------------------------------------------
##                                    Pern
##                        (1)                  (2)
## --------------------------------------------------------------------
## time                  9,515.358***          21,372.730***
##                      (1,596.023)           (2,956.740)
##
## dm2                   267,969.600**          335,192.300
##                     (132,700.500)          (217,295.300)
```

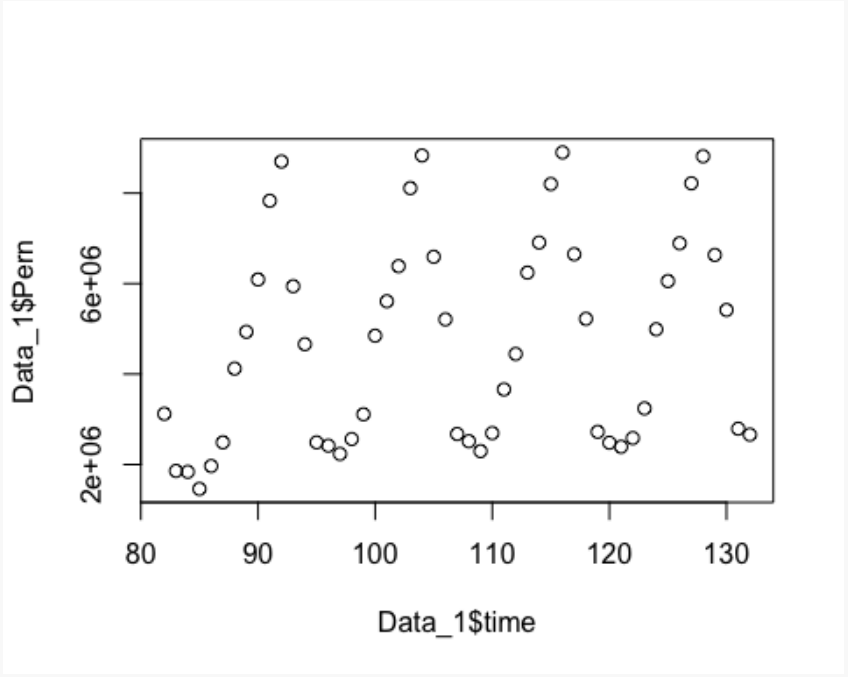
```
##
## dm3                1,153,032.000***        986,838.300***
##                   (132,729.300)          (217,355.600)
##
## dm4                2,316,429.000***        2,440,764.000***
##                   (132,777.200)          (217,456.200)
##
## dm5                3,226,954.000***        3,529,108.000***
##                   (132,844.400)          (217,596.800)
##
## dm6                4,253,999.000***        4,365,256.000***
##                   (132,930.600)          (217,777.500)
##
## dm7                6,016,961.000***        5,865,635.000***
##                   (133,036.000)          (217,998.200)
##
## dm8                6,794,244.000***        6,565,822.000***
##                   (133,160.400)          (218,258.700)
##
## dm9                4,392,394.000***        4,187,300.000***
##                   (133,303.800)          (218,558.900)
##
## dm10               2,594,236.000***        2,565,633.000***
##                   (133,466.100)          (206,316.100)
##
## dm11                280,068.000**           326,558.200
##                   (133,647.300)          (206,464.300)
##
## dm12                164,080.900            180,080.500
##                   (133,847.300)          (206,654.800)
##
## Constant           1,750,321.000***         -105,035.600
##                   (101,958.300)          (341,103.300)
##
## ------------------------------------------------------------------
## Observations              60                     51
## R2                       0.993                  0.986
```

```
## Adjusted R2                 0.992                    0.982
## Residual Std. Error  209,802.700 (df = 47)    307,273.500 (df = 38)
## F Statistic         586.408*** (df = 12; 47) 226.849*** (df = 12; 38)
## ==============================================================================
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

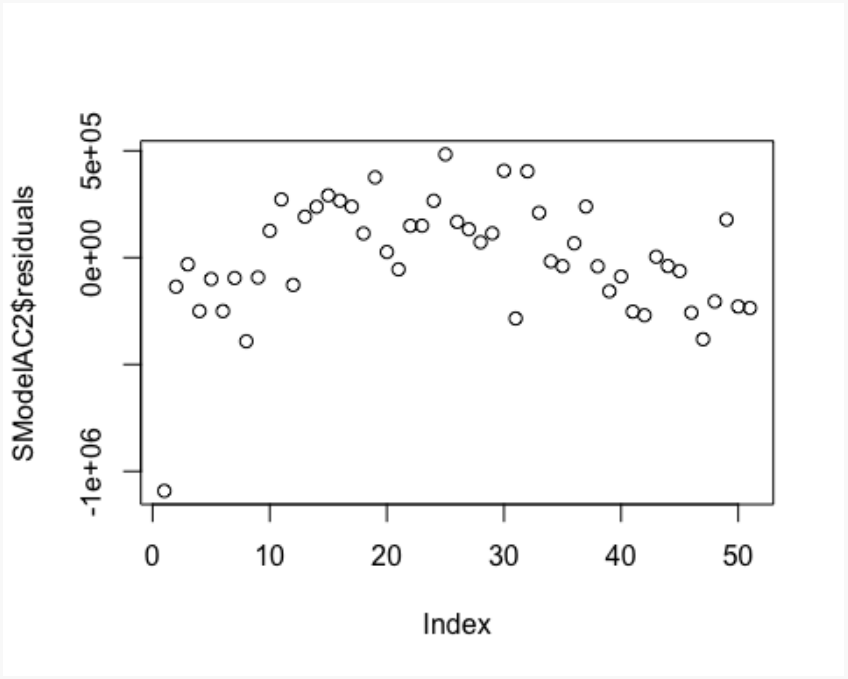
```
summary(ModelAC2)
```

```
##
## Call:
## lm(formula = Data_1$Pern ~ Data_1$time + Data_1$dm2 + Data_1$dm3 +
##     Data_1$dm4 + Data_1$dm5 + Data_1$dm6 + Data_1$dm7 + Data_1$dm8 +
##     Data_1$dm9 + Data_1$dm10 + Data_1$dm11 + Data_1$dm12)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1091632  -146595   -17277   184731   483063
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -105036     341103  -0.308    0.760
## Data_1$time    21373       2957   7.228 1.21e-08 ***
## Data_1$dm2    335192     217295   1.543    0.131
## Data_1$dm3    986838     217356   4.540 5.52e-05 ***
## Data_1$dm4   2440764     217456  11.224 1.25e-13 ***
## Data_1$dm5   3529108     217597  16.219  < 2e-16 ***
## Data_1$dm6   4365256     217778  20.045  < 2e-16 ***
## Data_1$dm7   5865635     217998  26.907  < 2e-16 ***
## Data_1$dm8   6565822     218259  30.083  < 2e-16 ***
## Data_1$dm9   4187300     218559  19.159  < 2e-16 ***
## Data_1$dm10  2565633     206316  12.435 5.75e-15 ***
## Data_1$dm11   326558     206464   1.582    0.122
## Data_1$dm12   180080     206655   0.871    0.389
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 307300 on 38 degrees of freedom
```

```
## Multiple R-squared:  0.9862, Adjusted R-squared:  0.9819
## F-statistic: 226.8 on 12 and 38 DF,  p-value: < 2.2e-16

SModelAC2<-summary(ModelAC2)


plot(Data_1$time,Data_1$Pern)
```




```
names(SModelAC2)

##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"

plot(SModelAC2$residuals)
```

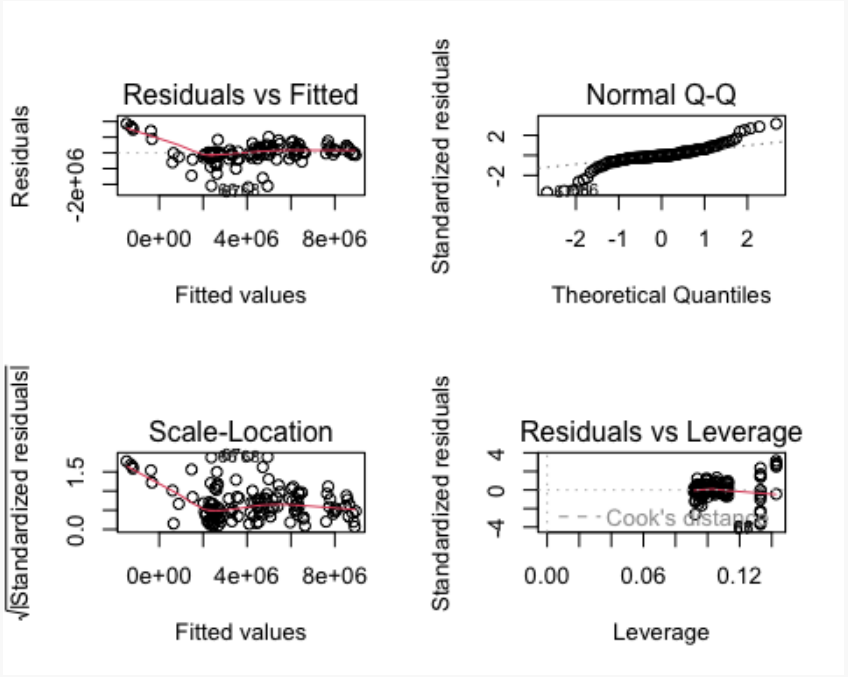



```
stargazer(ModelAC, ModelAC1,ModelAC2, type="text")

##
##
==========================================================================
=======================
##                                  Dependent variable:
##                 ---------------------------------------------------------------------
##                                        Pern
##                       (1)                (2)                (3)
## ---------------------------------------------------------------------------------------
## time                 3,887.225***          9,515.358***          21,372.730***
##                     (1,461.721)          (1,596.023)           (2,956.740)
##
## dm2                   261,525.200          267,969.600**          335,192.300
##                    (271,221.000)         (132,700.500)         (217,295.300)
##
## dm3                  1,145,804.000***        1,153,032.000***        986,838.300***
##                    (271,609.600)         (132,729.300)         (217,355.600)
##
## dm4                  2,126,607.000***        2,316,429.000***       2,440,764.000***
##                    (271,624.500)         (132,777.200)         (217,456.200)
##
```

```
## dm5                2,988,986.000***       3,226,954.000***       3,529,108.000***
##                   (271,647.200)          (132,844.400)          (217,596.800)
##
## dm6                3,888,259.000***       4,253,999.000***       4,365,256.000***
##                   (271,677.700)          (132,930.600)          (217,777.500)
##
## dm7                5,561,864.000***       6,016,961.000***       5,865,635.000***
##                   (271,716.200)          (133,036.000)          (217,998.200)
##
## dm8                6,439,109.000***       6,794,244.000***       6,565,822.000***
##                   (271,762.500)          (133,160.400)          (218,258.700)
##
## dm9                4,124,608.000***       4,392,394.000***       4,187,300.000***
##                   (271,816.600)          (133,303.800)          (218,558.900)
##
## dm10               2,385,881.000***       2,594,236.000***       2,565,633.000***
##                   (271,536.000)          (133,466.100)          (206,316.100)
##
## dm11                 294,485.500          280,068.000**          326,558.200
##                   (271,610.700)          (133,647.300)          (206,464.300)
##
## dm12                 195,599.200          164,080.900            180,080.500
##                   (271,693.300)          (133,847.300)          (206,654.800)
##
## Covid              -3,780,837.000***
##                   (159,320.900)
##
## Constant           2,013,089.000***       1,750,321.000***       -105,035.600
##                   (211,611.900)          (101,958.300)          (341,103.300)
##
## ------------------------------------------------------------------------------------------
## Observations              132                    60                    51
## R2                       0.940                  0.993                 0.986
## Adjusted R2              0.933                  0.992                 0.982
## Residual Std. Error  636,060.500 (df = 118)  209,802.700 (df = 47)  307,273.500 (df = 38)
## F Statistic        142.371*** (df = 13; 118) 586.408*** (df = 12; 47) 226.849*** (df = 12; 38)
##
```

```
================================================================================
======================
## Note:                                   *p<0.1; **p<0.05; ***p<0.01

#########################
# 10. RESULTATS
#########################
summary(ModelT3)

##
## Call:
## lm(formula = datos$Pern ~ datos$time + datos$dm2 + datos$dm3 +
##     datos$dm4 + datos$dm5 + datos$dm6 + datos$dm7 + datos$dm8 +
##     datos$dm9 + datos$dm10 + datos$dm11 + datos$dm12 + datos$Covid)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2188434  -160241   -22828   218658  1855138
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2013089     211612   9.513 2.85e-16 ***
## datos$time      3887       1462   2.659  0.00892 **
## datos$dm2     261525     271221   0.964  0.33689
## datos$dm3    1145804     271610   4.219 4.85e-05 ***
## datos$dm4    2126607     271624   7.829 2.35e-12 ***
## datos$dm5    2988986     271647  11.003  < 2e-16 ***
## datos$dm6    3888259     271678  14.312  < 2e-16 ***
## datos$dm7    5561864     271716  20.469  < 2e-16 ***
## datos$dm8    6439109     271762  23.694  < 2e-16 ***
## datos$dm9    4124608     271817  15.174  < 2e-16 ***
## datos$dm10   2385880     271536   8.787 1.46e-14 ***
## datos$dm11    294486     271611   1.084  0.28048
## datos$dm12    195599     271693   0.720  0.47299
## datos$Covid -3780838     159321 -23.731  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 636100 on 118 degrees of freedom
## Multiple R-squared:  0.9401, Adjusted R-squared:  0.9335
## F-statistic: 142.4 on 13 and 118 DF,  p-value: < 2.2e-16

#########################
# 11. RESIDUS
#########################

par(mfrow=c(2,2))
plot(ModelT3)
```




```
# Autocorrelació de residus (Un Durbin-Watson de 0.675 indica una forta autocorrelació positiva en los errores. Això viola la hipòtesi d'independència dels errors. Per això, encara que els coeficients puguin ser significatius, les inferències poden estar sesgades)
dwtest(ModelT3)

##
##  Durbin-Watson test
##
## data:  ModelT3
## DW = 0.63894, p-value = 1.262e-14
## alternative hypothesis: true autocorrelation is greater than 0
```

```
# Heterocedasticitat (no hi ha homocedasticitat, la variància dels errors no és constant)
bptest(ModelT3)

##
##  studentized Breusch-Pagan test
##
## data:  ModelT3
## BP = 60.353, df = 13, p-value = 4.543e-08

########################
# 12. ERRORS ROBUSTOS
########################

coeftest(ModelT3, vcov = vcovHC(ModelT3, type="HC1"))

##
## t test of coefficients:
##
##                Estimate  Std. Error  t value  Pr(>|t|)
## (Intercept)  2013088.66   193099.20  10.4252 < 2.2e-16 ***
## datos$time      3887.23      751.45   5.1730 9.544e-07 ***
## datos$dm2     261525.23   262696.39   0.9955 0.3215100
## datos$dm3    1145804.41   287375.18   3.9871 0.0001163 ***
## datos$dm4    2126607.19   239042.78   8.8963 8.067e-15 ***
## datos$dm5    2988985.69   277525.44  10.7701 < 2.2e-16 ***
## datos$dm6    3888258.55   315416.57  12.3274 < 2.2e-16 ***
## datos$dm7    5561863.51   309327.06  17.9805 < 2.2e-16 ***
## datos$dm8    6439108.56   308585.15  20.8666 < 2.2e-16 ***
## datos$dm9    4124607.61   267712.12  15.4069 < 2.2e-16 ***
## datos$dm10   2385880.52   266903.60   8.9391 6.406e-15 ***
## datos$dm11    294485.47   260744.47   1.1294 0.2610188
## datos$dm12    195599.25   271189.08   0.7213 0.4721730
## datos$Covid -3780837.49   327615.30 -11.5405 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

########################
# 13. VALORS AJUSTATS
########################
```

```r
datos$ajustada <- fitted(ModelT3)

ggplot(datos, aes(fecha)) +
 geom_line(aes(y=Pern), color="black") +
 geom_line(aes(y=ajustada), color="blue") +
 labs(title="Valores reales vs ajustados",
     y="Pernoctaciones")
```



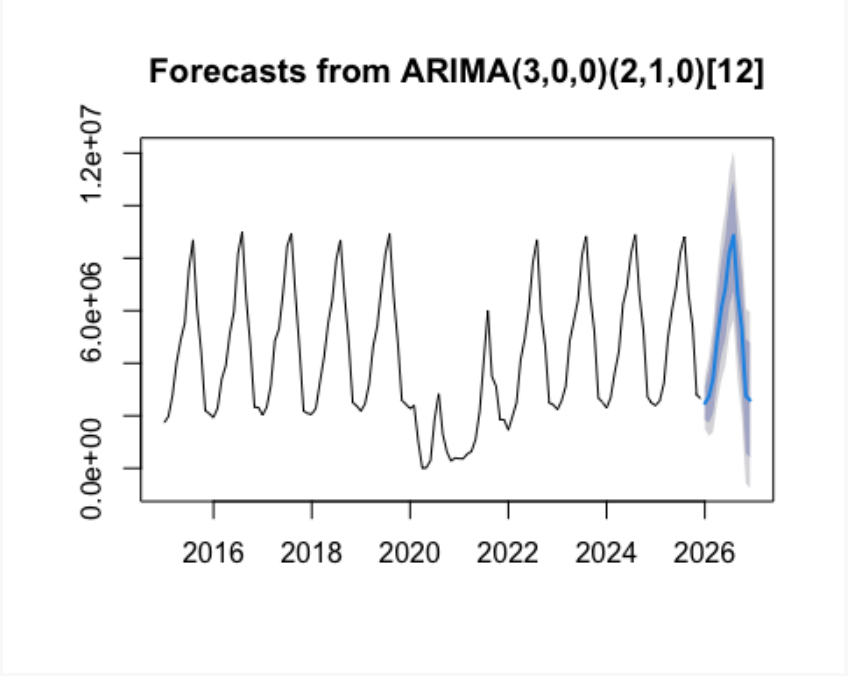

```r
#########################
# 14. SUBPERIÒDES
#########################

# PRECOVID
pre <- subset(datos, fecha < as.Date("2020-03-01"))

ModelPRE <- lm(Pern ~ time + mes_factor, data=pre)

# POSTCOVID / RECUPERACIÓN
post <- subset(datos, fecha > as.Date("2021-09-30"))

ModelPOST <- lm(Pern ~ time + mes_factor, data=post)
```

```r
stargazer(ModelPRE, ModelPOST,
        type="text",
        title="Comparación Pre y Post Covid")

##
## Comparación Pre y Post Covid
## ==========================================================================
##                              Dependent variable:
##                    ------------------------------------------------
##                                    Pern
##                          (1)                  (2)
## --------------------------------------------------------------------
## time                  9,087.906***          21,372.730***
##                      (1,487.723)           (2,956.740)
##
## mes_factor2            241,196.800**          335,192.300
##                     (119,768.700)          (217,295.300)
##
## mes_factor3           1,160,805.000***        986,838.300***
##                     (125,745.600)          (217,355.600)
##
## mes_factor4           2,324,629.000***       2,440,764.000***
##                     (125,684.000)          (217,456.200)
##
## mes_factor5           3,235,581.000***       3,529,108.000***
##                     (125,640.000)          (217,596.800)
##
## mes_factor6           4,263,054.000***       4,365,256.000***
##                     (125,613.600)          (217,777.500)
##
## mes_factor7           6,026,443.000***       5,865,635.000***
##                     (125,604.700)          (217,998.200)
##
## mes_factor8           6,804,154.000***       6,565,822.000***
##                     (125,613.600)          (218,258.700)
##
## mes_factor9           4,402,731.000***       4,187,300.000***
```

```
##                     (125,640.000)          (218,558.900)
##
## mes_factor10          2,605,001.000***       2,565,633.000***
##                     (125,684.000)          (206,316.100)
##
## mes_factor11           291,260.100**          326,558.200
##                     (125,745.600)          (206,464.300)
##
## mes_factor12           175,700.400            180,080.500
##                     (125,824.800)          (206,654.800)
##
## Constant             1,754,089.000***         -105,035.600
##                      (96,426.970)          (341,103.300)
##
## ------------------------------------------------------------------
## Observations                62                     51
## R2                        0.993                  0.986
## Adjusted R2               0.992                  0.982
## Residual Std. Error  207,429.400 (df = 49)    307,273.500 (df = 38)
## F Statistic        620.212*** (df = 12; 49) 226.849*** (df = 12; 38)
## ==================================================================
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

```
#########################
# 15. MODEL ARIMA
#########################

modelo_arima <- auto.arima(serie_ts)

summary(modelo_arima)

## Series: serie_ts
## ARIMA(3,0,0)(2,1,0)[12]
##
## Coefficients:
##         ar1     ar2      ar3     sar1     sar2
##       1.1665  0.0158  -0.2716  -0.4102  -0.1763
## s.e.  0.0938  0.1505   0.0912   0.0969   0.0916
```

```
##
## sigma^2 = 2.497e+11:  log likelihood = -1744.89
## AIC=3501.78   AICc=3502.52   BIC=3518.5
##
## Training set error measures:
##                    ME     RMSE      MAE  MPE MAPE      MASE       ACF1
## Training set 8905.928 466440.6 305646.7 -Inf  Inf 0.3160606 0.01862765

forecast_arima <- forecast(modelo_arima, h=12)


plot(forecast_arima)
```




```
############################################################
# FI
############################################################
```

Annex 2: Anàlisi dels viatgers a Catalunya

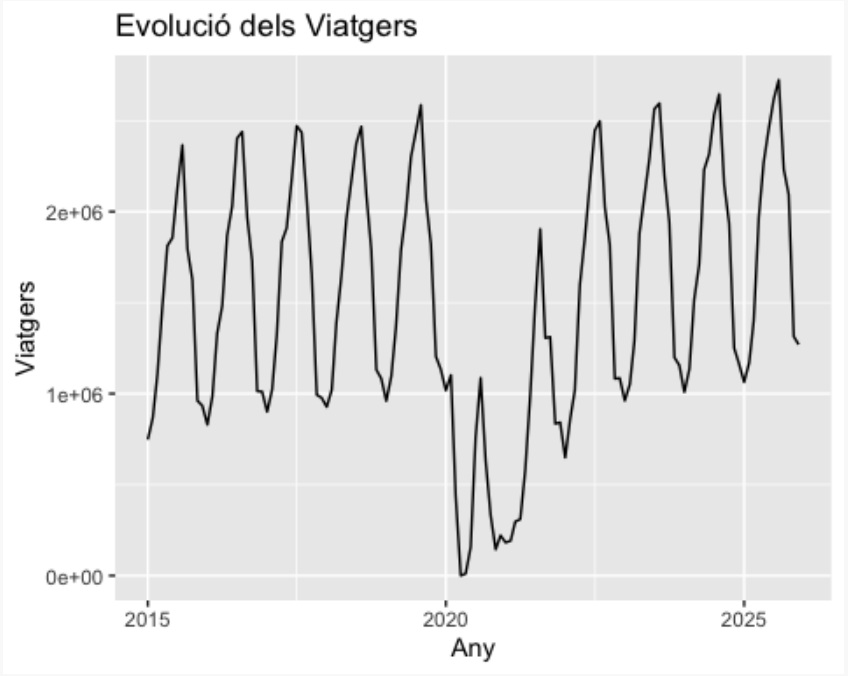
Introducció

Aquest document analitza l'evolució dels viatgers a Catalunya mitjançant tècniques de sèries temporals i regressió economètrica. L'objectiu és identificar la tendència estructural, l'estacionalitat mensual i l'impacte de la COVID-19 sobre l'activitat hotelera.

```
#########################
# 1. PAQUETS
#########################



library(readxl)
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.2.1     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.4     ✔ tibble    3.2.1
## ✔ lubridate 1.9.5     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(lubridate)
library(forecast)

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

library(tseries)
library(lmtest)

## Loading required package: zoo
##
```

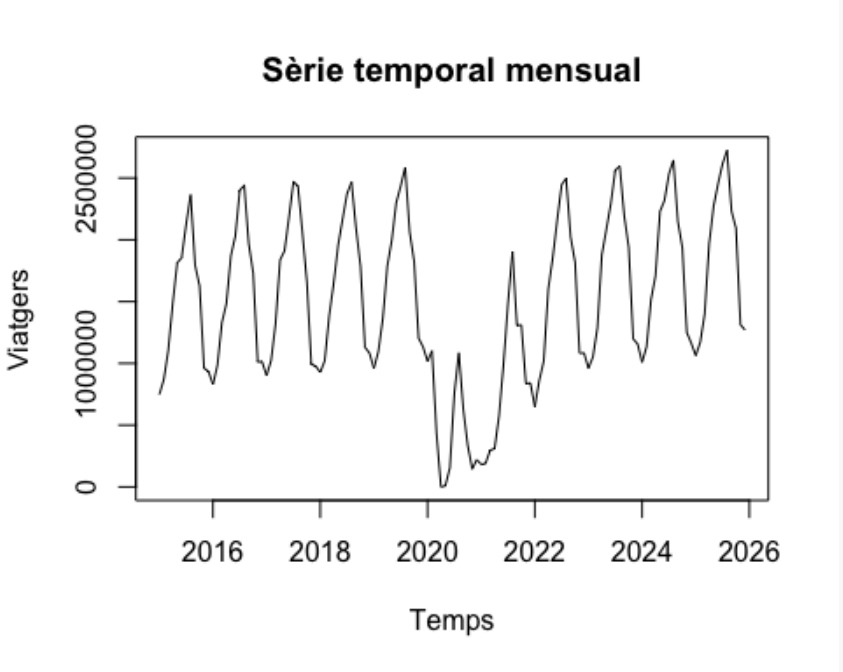
```
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric

library(sandwich)
library(modelsummary)

#########################
# 2. IMPORTAR DADES
#########################
datos <- read_excel("/Users/judit/Downloads/DataV_C.xlsx")

# Ordenar cronològicament
datos <- datos %>%
  arrange(Any, mes)

mean(datos$Viat)

## [1] 1499741

sd(datos$Viat)

## [1] 684006.4

min(datos$Viat)

## [1] 0

max(datos$Viat)

## [1] 2723647

# Data mensual
datos$fecha <- as.Date(paste(datos$Any, datos$mes, "01", sep="-"))

# Tendència temporal
datos$time <- 1:nrow(datos)

# Convertir mes en factor (gener com a base)
datos$mes_factor <- factor(datos$mes)
```

```
#########################
# 4. VARIABLE COVID
#########################

# Dummy COVID: març 2020 - setembre 2021
datos$Covid <- ifelse(datos$fecha >= as.Date("2020-03-01") &
                      datos$fecha <= as.Date("2021-09-30"), 1, 0)

#########################
# 5. ANÀLISI DESCRIPTIVA
#########################

ggplot(datos, aes(x = fecha, y = Viat)) +
  geom_line(color = "black") +
  labs(title = "Evolució dels Viatgers",
       x = "Any",
       y = "Viatgers")
```

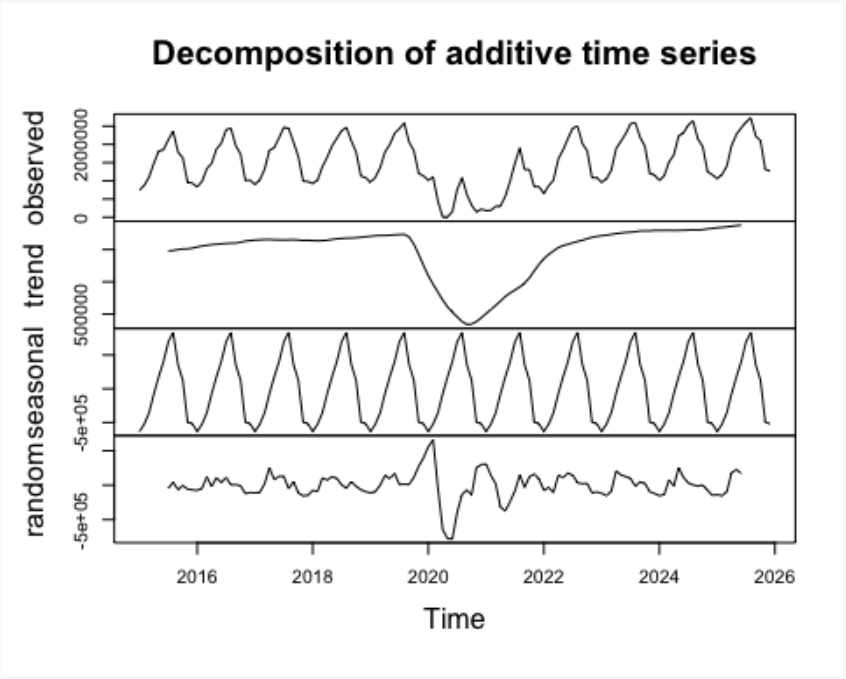



```
#########################
# 6. SÈRIE TEMPORAL
#########################
```

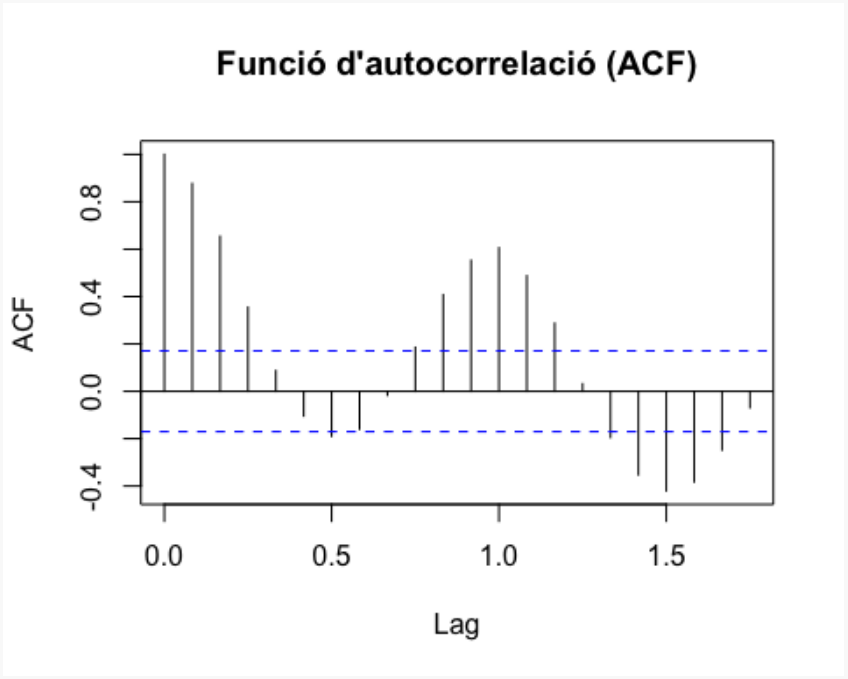
```r
# Definir principi de la sèrie
inicio <- c(datos$Any[1], datos$mes[1])

# Sèrie temporal mensual
serie_ts <- ts(datos$Viat,
               start = inicio,
               frequency = 12)
# Resultat
plot(serie_ts,
    main = "Sèrie temporal mensual",
    ylab = "Viatgers",
    xlab = "Temps")
```



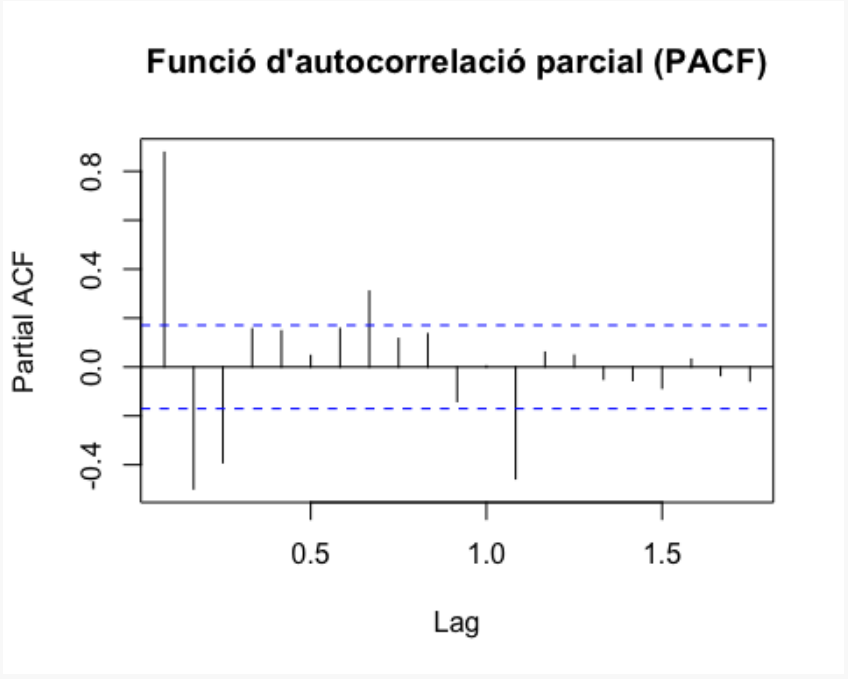

```r
datos <- datos[order(datos$Any, datos$mes), ]

##########################
# 7. DESCOMPOSICIÓ
##########################

descomp <- decompose(serie_ts, type = "additive")
plot(descomp)
```




```
##########################
# 8. ESTACIONARIETAT
##########################

resultat_adf <- adf.test(serie_ts)
print(resultat_adf)

##
##  Augmented Dickey-Fuller Test
##
## data:  serie_ts
## Dickey-Fuller = -3.27, Lag order = 5, p-value = 0.07932
## alternative hypothesis: stationary

# El test Dickey-Fuller augmentat no rebutja la hipòtesi nul·la (nivell de significació major a 0.01), indicant que la sèrie no és estacionària.

##########################
# 9. ACF I PACF
##########################
acf(serie_ts, main = "Funció d'autocorrelació (ACF)")
```

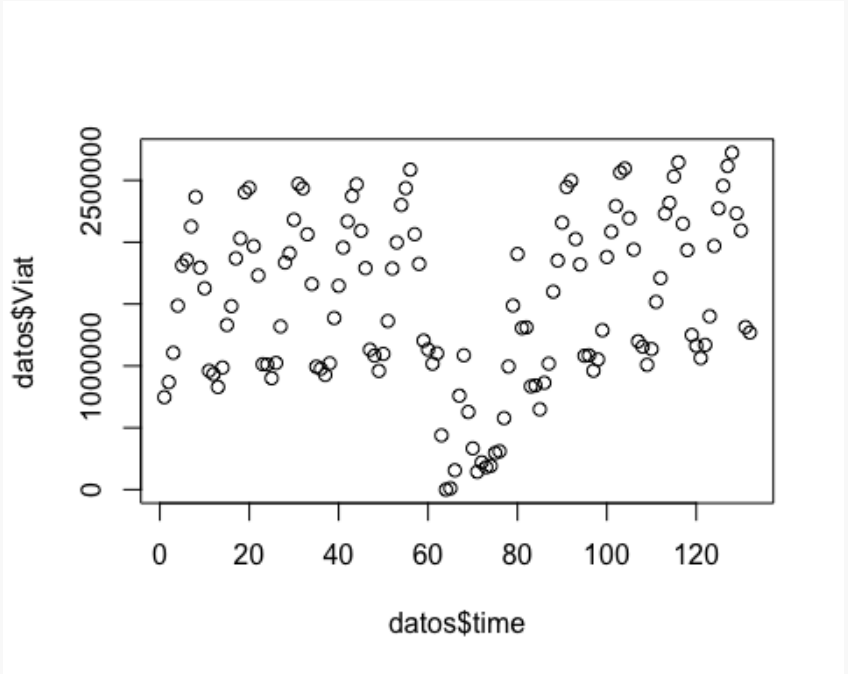



```
pacf(serie_ts, main = "Funció d'autocorrelació parcial (PACF)")
```




```
#########################
# 10. MODELS ECONOMÈTRICS
#########################

# MODEL 1: Tendència temporal
ModelT1 <- lm(Viat ~ time, data = datos)
ModelT1<-lm(datos$Viat~datos$time)
```

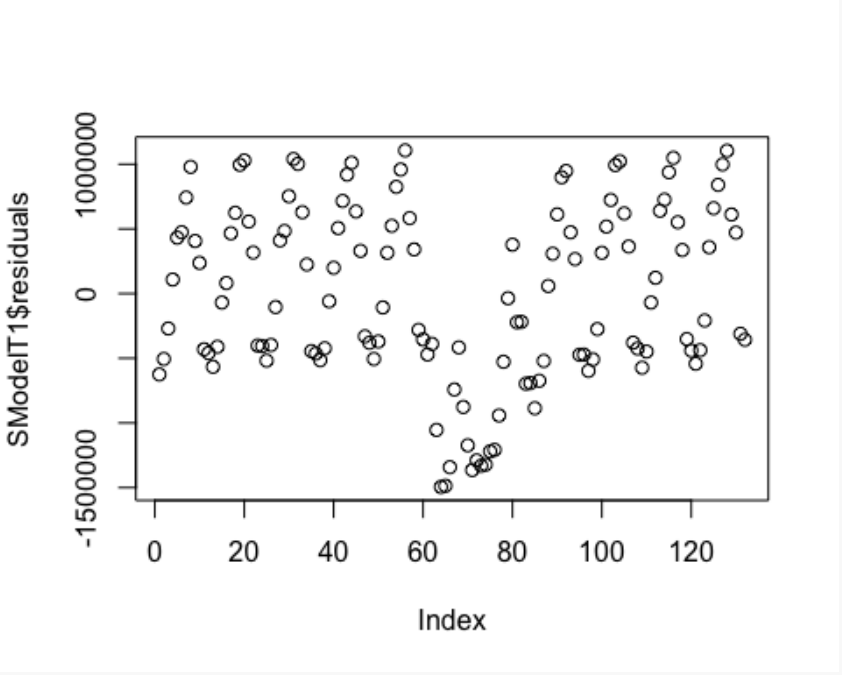
```
SModelT1<-summary(ModelT1)
summary(ModelT1)

##
## Call:
## lm(formula = datos$Viat ~ datos$time)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1494884  -465371   -64351   563342  1106758
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1370553     119498  11.469   <2e-16 ***
## datos$time      1943       1559   1.246    0.215
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 682600 on 130 degrees of freedom
## Multiple R-squared:  0.0118, Adjusted R-squared:  0.0042
## F-statistic: 1.552 on 1 and 130 DF,  p-value: 0.215

plot(datos$time,datos$Viat)
```



```
names(SModelT1)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelT1$residuals)
```




```
library(stargazer)
```

```
##
## Please cite as:
```

```
##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
stargazer(ModelT1, type="text")
```

```
##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                                Viat
## ------------------------------------------------
## time                          1,942.677
```

```
##                          (1,559.152)
##
## Constant                 1,370,553.000***
##                         (119,498.300)
##
## ------------------------------------------------
## Observations                   132
## R2                            0.012
## Adjusted R2                   0.004
## Residual Std. Error   682,568.600 (df = 130)
## F Statistic           1.552 (df = 1; 130)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

```
# MODEL 2: Tendència + estacionalitat mensual
ModelT2<-lm(datos$Viat~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+d
atos$dm7+datos$dm8+datos$dm9+
            datos$dm10+datos$dm11+datos$dm12)
SModelT2<-summary(ModelT2)

summary(ModelT2)
```

```
##
## Call:
## lm(formula = datos$Viat ~ datos$time + datos$dm2 + datos$dm3 +
##     datos$dm4 + datos$dm5 + datos$dm6 + datos$dm7 + datos$dm8 +
##     datos$dm9 + datos$dm10 + datos$dm11 + datos$dm12)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1744618    47676   142727   251600   484636
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   738964     163098   4.531 1.41e-05 ***
## datos$time      1665       1126   1.480 0.141597
## datos$dm2     113409     209230   0.542 0.588811
## datos$dm3     289290     209239   1.383 0.169381
```

```
## datos$dm4     582578     209254   2.784 0.006247 **
## datos$dm5     841270     209275   4.020 0.000103 ***
## datos$dm6    1052206     209302   5.027 1.78e-06 ***
## datos$dm7    1350820     209336   6.453 2.47e-09 ***
## datos$dm8    1488166     209375   7.108 9.43e-11 ***
## datos$dm9    1011338     209420   4.829 4.12e-06 ***
## datos$dm10    787153     209472   3.758 0.000267 ***
## datos$dm11    154468     209529   0.737 0.462442
## datos$dm12    129681     209592   0.619 0.537279
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 490700 on 119 degrees of freedom
## Multiple R-squared:  0.5325, Adjusted R-squared:  0.4854
## F-statistic:  11.3 on 12 and 119 DF,  p-value: 8.199e-15
```

```
names(SModelT2)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

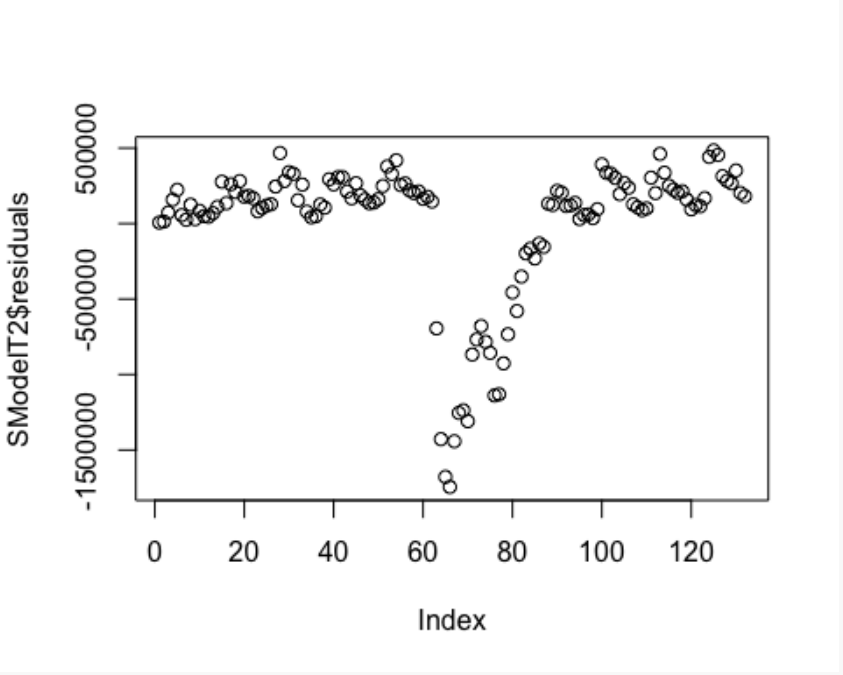
```
stargazer(ModelT1, ModelT2, type="text")
```

```
##
## ==========================================================================
##                              Dependent variable:
##                    ----------------------------------------------
##                                    Viat
##                          (1)                  (2)
## --------------------------------------------------------------------
## time                   1,942.677            1,665.338
##                      (1,559.152)          (1,125.460)
##
## dm2                                        113,408.800
##                                          (209,229.600)
##
## dm3                                        289,290.200
##                                          (209,238.700)
```

```
##
## dm4                                582,577.900***
##                                   (209,253.800)
##
## dm5                                841,269.500***
##                                   (209,275.000)
##
## dm6                               1,052,206.000***
##                                   (209,302.300)
##
## dm7                               1,350,820.000***
##                                   (209,335.500)
##
## dm8                               1,488,166.000***
##                                   (209,374.900)
##
## dm9                               1,011,338.000***
##                                   (209,420.200)
##
## dm10                               787,152.700***
##                                   (209,471.600)
##
## dm11                               154,468.200
##                                   (209,529.100)
##
## dm12                               129,680.600
##                                   (209,592.500)
##
## Constant              1,370,553.000***        738,964.200***
##                      (119,498.300)          (163,098.500)
##
## ------------------------------------------------------------------
## Observations               132                 132
## R2                        0.012               0.533
## Adjusted R2               0.004               0.485
## Residual Std. Error 682,568.600 (df = 130)  490,679.900 (df = 119)
## F Statistic         1.552 (df = 1; 130)   11.297*** (df = 12; 119)
```

```
## ================================================================================
## Note:                            *p<0.1; **p<0.05; ***p<0.01
```

```
plot(SModelT2$residuals)
```




```
y_ajustada <- fitted(ModelT2)
```

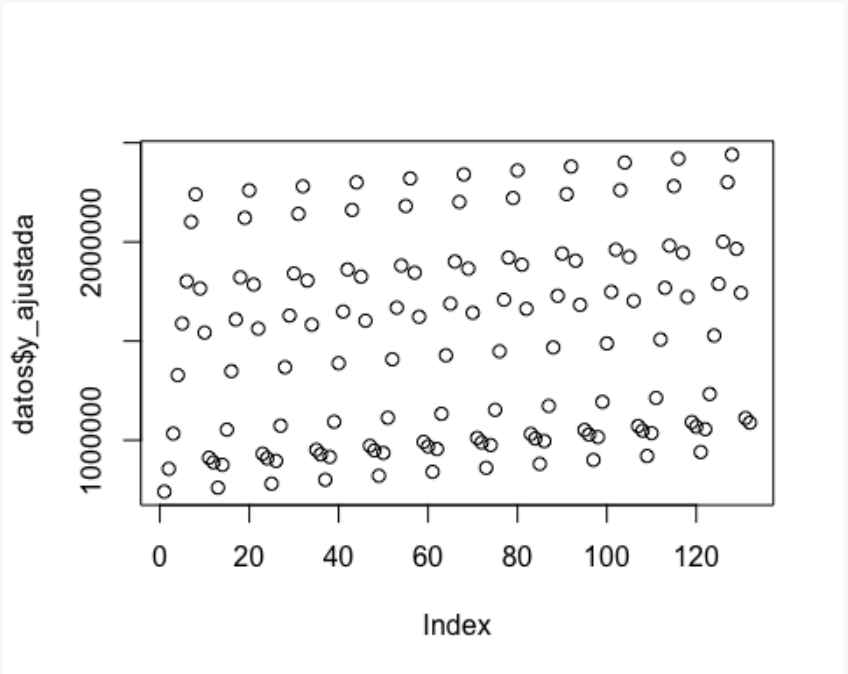
```
y_ajustada
```

```
##        1        2        3        4        5        6        7        8
##  740629.5  855703.7 1033250.4 1328203.4 1588560.3 1801161.8 2101441.7 2240452.7
##        9       10       11       12       13       14       15       16
## 1765290.5 1542770.3  911751.1  888628.9  760613.6  875687.8 1053234.5 1348187.5
##       17       18       19       20       21       22       23       24
## 1608544.4 1821145.9 2121425.8 2260436.8 1785274.6 1562754.3  931735.1  908613.0
##       25       26       27       28       29       30       31       32
##  780597.6  895671.8 1073218.6 1368171.6 1628528.5 1841129.9 2141409.8 2280420.8
##       33       34       35       36       37       38       39       40
## 1805258.6 1582738.4  951719.2  928597.0  800581.7  915655.9 1093202.6 1388155.6
##       41       42       43       44       45       46       47       48
## 1648512.5 1861114.0 2161393.9 2300404.9 1825242.7 1602722.4  971703.2  948581.1
##       49       50       51       52       53       54       55       56
##  820565.8  935639.9 1113186.7 1408139.7 1668496.6 1881098.0 2181377.9 2320388.9
##       57       58       59       60       61       62       63       64
```

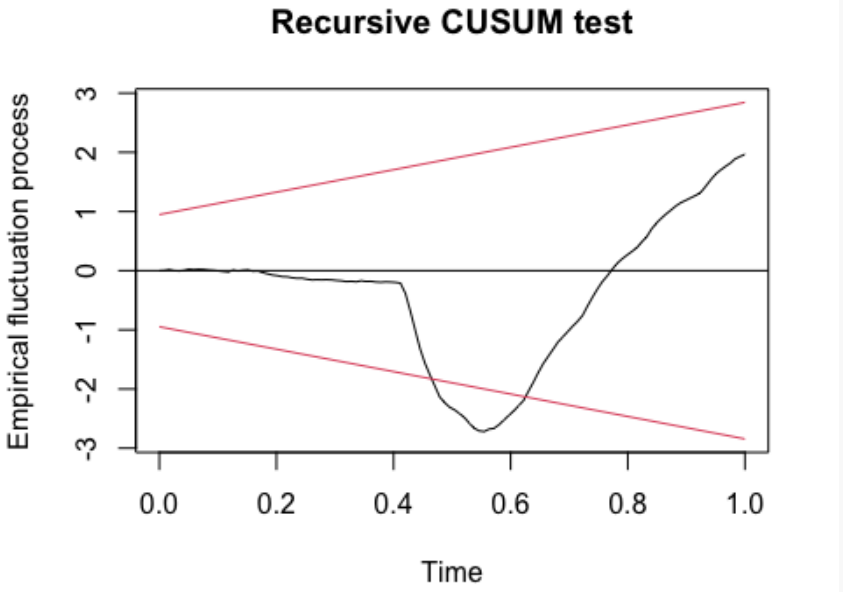
```
## 1845226.8 1622706.5  991687.3  968565.1  840549.8  955624.0 1133170.7 1428123.7
##       65       66       67       68       69       70       71       72
## 1688480.6 1901082.1 2201362.0 2340373.0 1865210.8 1642690.5 1011671.4  988549.2
##       73       74       75       76       77       78       79       80
##  860533.9  975608.1 1153154.8 1448107.8 1708464.7 1921066.1 2221346.1 2360357.1
##       81       82       83       84       85       86       87       88
## 1885194.9 1662674.6 1031655.4 1008533.2  880517.9  995592.1 1173138.8 1468091.8
##       89       90       91       92       93       94       95       96
## 1728448.8 1941050.2 2241330.1 2380341.1 1905178.9 1682658.7 1051639.5 1028517.3
##       97       98       99      100      101      102      103      104
##  900502.0 1015576.2 1193122.9 1488075.9 1748432.8 1961034.3 2261314.2 2400325.2
##      105      106      107      108      109      110      111      112
## 1925163.0 1702642.7 1071623.5 1048501.4  920486.0 1035560.2 1213107.0 1508060.0
##      113      114      115      116      117      118      119      120
## 1768416.9 1981018.3 2281298.2 2420309.2 1945147.0 1722626.8 1091607.6 1068485.4
##      121      122      123      124      125      126      127      128
##  940470.1 1055544.3 1233091.0 1528044.0 1788400.9 2001002.4 2301282.3 2440293.3
##      129      130      131      132
## 1965131.1 1742610.8 1111591.7 1088469.5
```

```
datos$y_ajustada<-fitted(ModelT2)
```

```
plot(datos$y_ajustada )
```



```
Data_1 <- datos

library(strucchange)

##
## Attaching package: 'strucchange'

## The following object is masked from 'package:stringr':
##
##     boundary

efp_model <-
efp(datos$Viat~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+datos$dm7+
datos$dm8+datos$dm9+
          datos$dm10+datos$dm11+datos$dm12, type="Rec-CUSUM")

plot (efp_model)
```



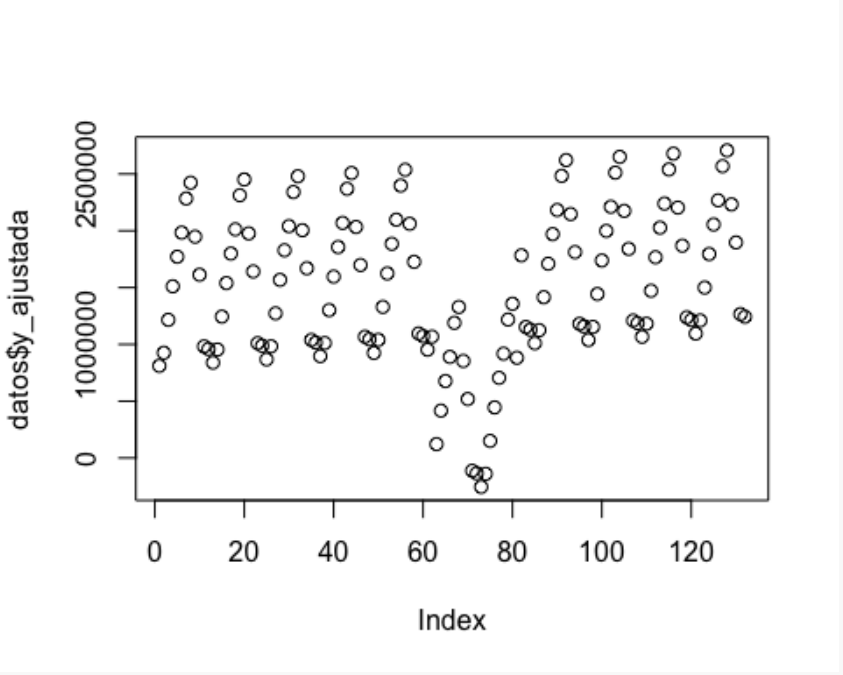

```
sctest (efp_model)

##
##  Recursive CUSUM test
##
## data:  efp_model
## S = 1.2938, p-value = 0.002334
```

```
# MODEL 3: Tendència + estacionalitat + CÒVID
ModelT3<-lm(datos$Viat~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+d
atos$dm7+datos$dm8+datos$dm9+
          datos$dm10+datos$dm11+datos$dm12+datos$Covid)
SModelT3<-summary(ModelT3)

names(ModelT3)

##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "xlevels"       "call"          "terms"         "model"

stargazer(ModelT1, ModelT2,ModelT3, type="text")

##
##
==========================================================================
====================
##                                  Dependent variable:
##                  ---------------------------------------------------------------
##                                        Viat
##                       (1)                (2)                 (3)
## ---------------------------------------------------------------------------------------
## time                  1,942.677          1,665.338           2,367.891***
##                     (1,559.152)        (1,125.460)           (437.794)
##
## dm2                                    113,408.800          112,706.300
##                                      (209,229.600)         (81,232.240)
##
## dm3                                    289,290.200          400,293.600***
##                                      (209,238.700)         (81,348.630)
##
## dm4                                    582,577.900***       692,878.700***
##                                      (209,253.800)         (81,353.070)
##
## dm5                                    841,269.500***       950,867.700***
##                                      (209,275.000)         (81,359.870)
##
```

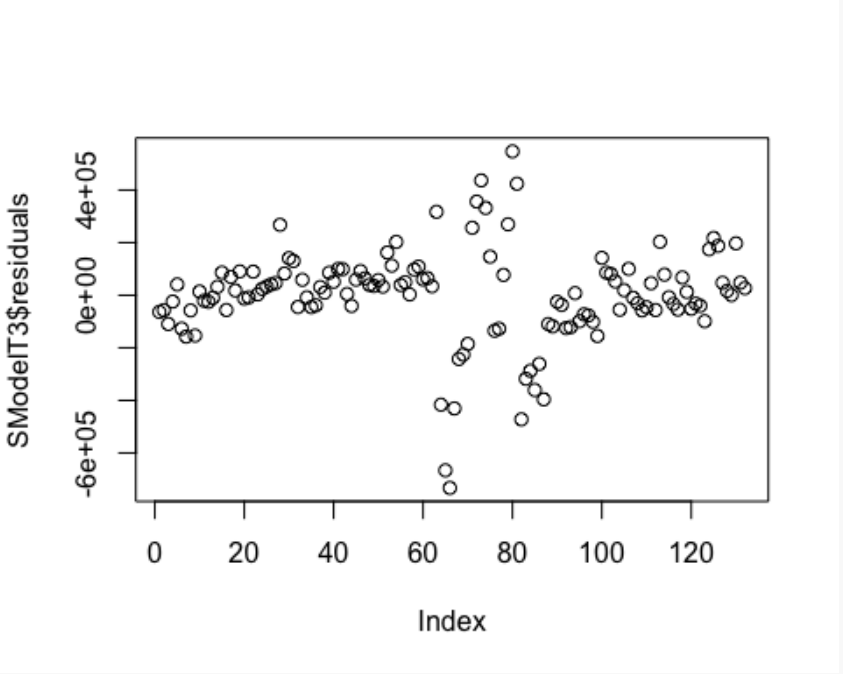
```
## dm6                                1,052,206.000***        1,161,101.000***
##                                   (209,302.300)           (81,369.030)
##
## dm7                                1,350,820.000***        1,459,013.000***
##                                   (209,335.500)           (81,380.540)
##
## dm8                                1,488,166.000***        1,595,656.000***
##                                   (209,374.900)           (81,394.400)
##
## dm9                                1,011,338.000***        1,118,126.000***
##                                   (209,420.200)           (81,410.610)
##
## dm10                                787,152.700***          780,829.700***
##                                   (209,471.600)           (81,326.560)
##
## dm11                                154,468.200             147,442.600*
##                                   (209,529.100)           (81,348.950)
##
## dm12                                129,680.600             121,952.600
##                                   (209,592.500)           (81,373.680)
##
## Covid                                                      -1,236,493.000***
##                                                           (47,717.520)
##
## Constant             1,370,553.000***        738,964.200***          808,516.900***
##                     (119,498.300)           (163,098.500)           (63,378.950)
##
## ------------------------------------------------------------------------------------
## Observations                132                     132                     132
## R2                        0.012                   0.533                   0.930
## Adjusted R2               0.004                   0.485                   0.922
## Residual Std. Error 682,568.600 (df = 130)  490,679.900 (df = 119)   190,503.700 (df = 118)
## F Statistic         1.552 (df = 1; 130)   11.297*** (df = 12; 119) 120.833*** (df = 13; 118)
##
==========================================================================
====================
## Note:                                          *p<0.1; **p<0.05; ***p<0.01
```

```
y_ajustada <- fitted(ModelT3)


y_ajustada

##        1        2        3        4        5        6        7        8
##  810884.8  925959.0 1215914.2 1510867.2 1771224.1 1983825.6 2284105.5 2423116.5
##        9       10       11       12       13       14       15       16
## 1947954.3 1613025.5  982006.4  958884.2  839299.5  954373.7 1244328.9 1539281.9
##       17       18       19       20       21       22       23       24
## 1799638.8 2012240.3 2312520.2 2451531.2 1976369.0 1641440.2 1010421.1  987298.9
##       25       26       27       28       29       30       31       32
##  867714.2  982788.4 1272743.6 1567696.6 1828053.5 2040654.9 2340934.9 2479945.9
##       33       34       35       36       37       38       39       40
## 2004783.7 1669854.9 1038835.8 1015713.6  896128.9 1011203.1 1301158.3 1596111.3
##       41       42       43       44       45       46       47       48
## 1856468.2 2069069.6 2369349.5 2508360.5 2033198.4 1698269.6 1067250.4 1044128.3
##       49       50       51       52       53       54       55       56
##  924543.6 1039617.8 1329573.0 1624526.0 1884882.9 2097484.3 2397764.2 2536775.2
##       57       58       59       60       61       62       63       64
## 2061613.1 1726684.3 1095665.1 1072543.0  952958.3 1068032.5  121494.5  416447.5
##       65       66       67       68       69       70       71       72
##  676804.4  889405.9 1189685.8 1328696.8  853534.6  518605.9 -112413.3 -135535.5
##       73       74       75       76       77       78       79       80
## -255120.2 -140046.0  149909.2  444862.2  705219.1  917820.6 1218100.5 1357111.5
##       81       82       83       84       85       86       87       88
##  881949.3 1783513.7 1152494.5 1129372.3 1009787.7 1124861.9 1414817.0 1709770.0
##       89       90       91       92       93       94       95       96
## 1970127.0 2182728.4 2483008.3 2622019.3 2146857.1 1811928.4 1180909.2 1157787.0
##       97       98       99      100      101      102      103      104
## 1038202.4 1153276.5 1443231.7 1738184.7 1998541.6 2211143.1 2511423.0 2650434.0
##      105      106      107      108      109      110      111      112
## 2175271.8 1840343.1 1209323.9 1186201.7 1066617.1 1181691.2 1471646.4 1766599.4
##      113      114      115      116      117      118      119      120
## 2026956.3 2239557.8 2539837.7 2678848.7 2203686.5 1868757.8 1237738.6 1214616.4
##      121      122      123      124      125      126      127      128
## 1095031.7 1210105.9 1500061.1 1795014.1 2055371.0 2267972.5 2568252.4 2707263.4
```

```
##      129      130      131      132
## 2232101.2 1897172.5 1266153.3 1243031.1

datos$y_ajustada<-fitted(ModelT3)


plot(datos$y_ajustada )
```




```
plot(SModelT3$residuals)
```



```
# Canvi de nom per poder fer la comparació
Data_1 <- datos

ModelAC<-lm(Data_1$Viat~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
            Data_1$dm10+Data_1$dm11+Data_1$dm12+Data_1$Covid)

stargazer(ModelAC, type="text")

##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                               Viat
## ------------------------------------------------
## time                        2,367.891***
##                            (437.794)
##
## dm2                          112,706.300
##                           (81,232.240)
##
## dm3                          400,293.600***
##                           (81,348.630)
##
## dm4                          692,878.700***
##                           (81,353.070)
##
## dm5                          950,867.700***
##                           (81,359.870)
##
## dm6                         1,161,101.000***
##                           (81,369.030)
##
## dm7                         1,459,013.000***
##                           (81,380.540)
##
## dm8                         1,595,656.000***
```

```
##                         (81,394.400)
##
## dm9                      1,118,126.000***
##                         (81,410.610)
##
## dm10                      780,829.700***
##                         (81,326.560)
##
## dm11                      147,442.600*
##                         (81,348.950)
##
## dm12                       121,952.600
##                         (81,373.680)
##
## Covid                    -1,236,493.000***
##                         (47,717.520)
##
## Constant                  808,516.900***
##                         (63,378.950)
##
## ------------------------------------------------
## Observations                    132
## R2                             0.930
## Adjusted R2                    0.922
## Residual Std. Error   190,503.700 (df = 118)
## F Statistic         120.833*** (df = 13; 118)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

```
#Model Estimat etapa Pre-còvid

Data_1 <- subset(datos, time< 61)


ModelAC1<-lm(Data_1$Viat~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
            Data_1$dm10+Data_1$dm11+Data_1$dm12)
```

```
stargazer(ModelAC1, type="text")

##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                               Viat
## ------------------------------------------------
## time                       4,918.767***
##                           (482.636)
##
## dm2                        121,568.800***
##                          (40,128.470)
##
## dm3                        418,236.900***
##                          (40,137.170)
##
## dm4                        760,265.700***
##                          (40,151.680)
##
## dm5                        1,016,582.000***
##                          (40,171.980)
##
## dm6                        1,209,886.000***
##                          (40,198.060)
##
## dm7                        1,459,841.000***
##                          (40,229.920)
##
## dm8                        1,551,263.000***
##                          (40,267.540)
##
## dm9                        1,084,349.000***
##                          (40,310.900)
##
## dm10                       809,675.900***
```

```
##                         (40,359.990)
##
## dm11                    138,766.500***
##                         (40,414.780)
##
## dm12                    100,675.400**
##                         (40,475.250)
##
## Constant                749,892.200***
##                         (30,832.060)
##
## ------------------------------------------------
## Observations                  60
## R2                          0.989
## Adjusted R2                 0.987
## Residual Std. Error    63,444.090 (df = 47)
## F Statistic         364.528*** (df = 12; 47)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

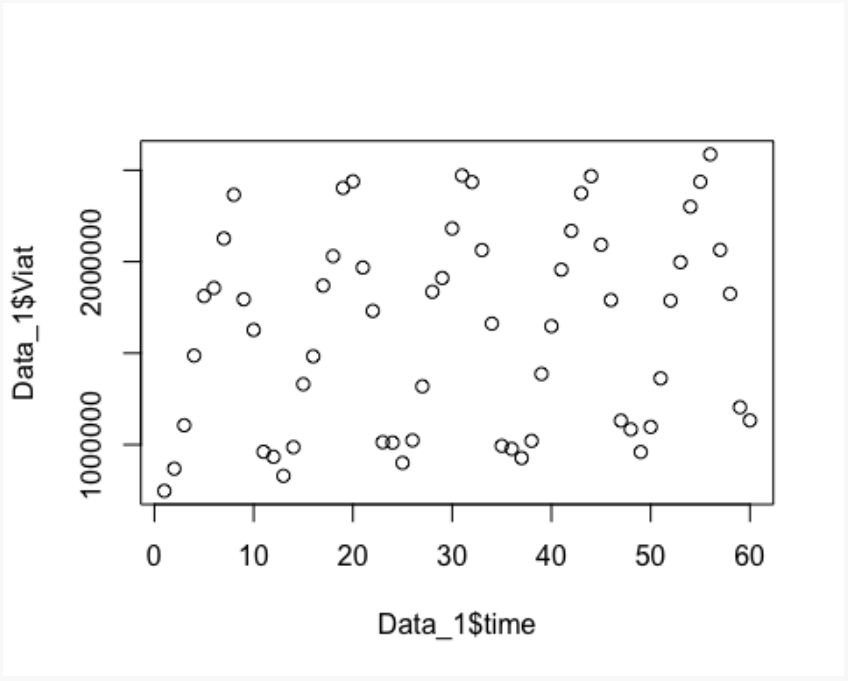
```
summary(ModelAC1)
```

```
##
## Call:
## lm(formula = Data_1$Viat ~ Data_1$time + Data_1$dm2 + Data_1$dm3 +
##     Data_1$dm4 + Data_1$dm5 + Data_1$dm6 + Data_1$dm7 + Data_1$dm8 +
##     Data_1$dm9 + Data_1$dm10 + Data_1$dm11 + Data_1$dm12)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -132701  -38804    2996   25845  187531
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  749892.2    30832.1  24.322  < 2e-16 ***
## Data_1$time    4918.8      482.6  10.191 1.73e-13 ***
## Data_1$dm2   121568.8    40128.5   3.029  0.00397 **
## Data_1$dm3   418236.9    40137.2  10.420 8.34e-14 ***
```

```
## Data_1$dm4   760265.7    40151.7  18.935  < 2e-16 ***
## Data_1$dm5  1016581.9    40172.0  25.306  < 2e-16 ***
## Data_1$dm6  1209885.8    40198.1  30.098  < 2e-16 ***
## Data_1$dm7  1459841.0    40229.9  36.287  < 2e-16 ***
## Data_1$dm8  1551263.0    40267.5  38.524  < 2e-16 ***
## Data_1$dm9  1084348.7    40310.9  26.900  < 2e-16 ***
## Data_1$dm10  809675.9    40360.0  20.061  < 2e-16 ***
## Data_1$dm11  138766.5    40414.8   3.434  0.00125 **
## Data_1$dm12  100675.4    40475.3   2.487  0.01648 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 63440 on 47 degrees of freedom
## Multiple R-squared:  0.9894, Adjusted R-squared:  0.9867
## F-statistic: 364.5 on 12 and 47 DF,  p-value: < 2.2e-16
```

```
SModelAC1<-summary(ModelAC1)
```

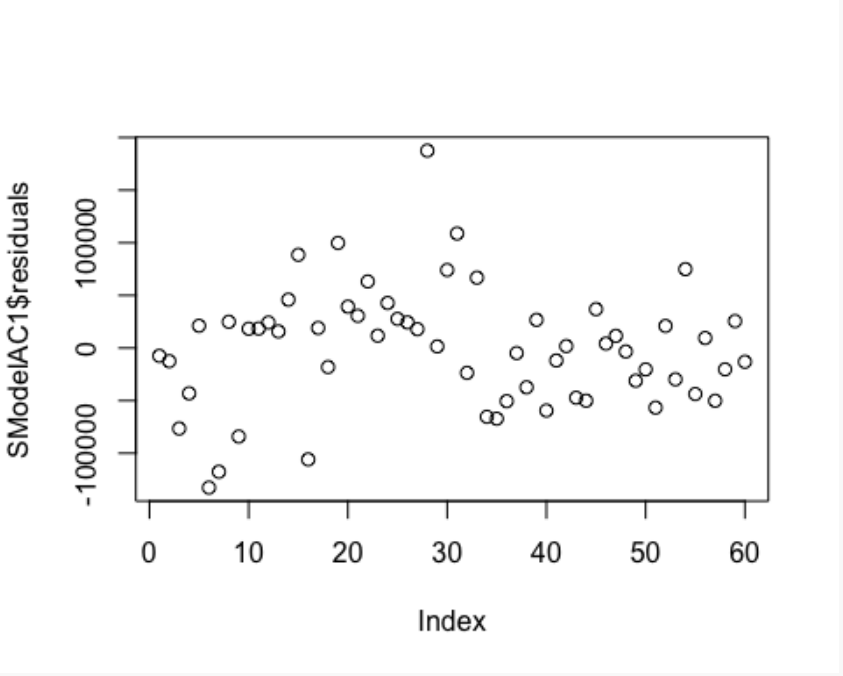
```
plot(Data_1$time,Data_1$Viat)
```




```
names(SModelAC1)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelAC1$residuals)
```




```
#Model Estimat etapa Post-còvid (suposem que comença a partir de l'observació 81)

Data_1 <- subset(datos, time>81)



ModelAC2<-lm(Data_1$Viat~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
            Data_1$dm10+Data_1$dm11+Data_1$dm12)

stargazer(ModelAC1,ModelAC2, type="text")

##
## ==========================================================================
##                              Dependent variable:
##                  ---------------------------------------------------
##                                    Viat
##                          (1)                  (2)
## ---------------------------------------------------------------------
```

```
## time                 4,918.767***          9,003.874***
##                     (482.636)             (897.641)
##
## dm2                  121,568.800***         125,729.400*
##                    (40,128.470)          (65,968.970)
##
## dm3                  418,236.900***         367,645.000***
##                    (40,137.170)          (65,987.290)
##
## dm4                  760,265.700***         842,372.600***
##                    (40,151.680)          (66,017.800)
##
## dm5                  1,016,582.000***       1,153,204.000***
##                    (40,171.980)          (66,060.510)
##
## dm6                  1,209,886.000***       1,340,296.000***
##                    (40,198.060)          (66,115.370)
##
## dm7                  1,459,841.000***       1,564,632.000***
##                    (40,229.920)          (66,182.370)
##
## dm8                  1,551,263.000***       1,631,693.000***
##                    (40,267.540)          (66,261.460)
##
## dm9                  1,084,349.000***       1,157,585.000***
##                    (40,310.900)          (66,352.600)
##
## dm10                 809,675.900***         872,802.400***
##                    (40,359.990)          (62,635.780)
##
## dm11                 138,766.500***         179,418.000***
##                    (40,414.780)          (62,680.790)
##
## dm12                 100,675.400**          137,406.100**
##                    (40,475.250)          (62,738.610)
##
## Constant             749,892.200***         -6,777.318
```

```
##                         (30,832.060)          (103,556.000)
##
## ------------------------------------------------------------------
## Observations                  60                    51
## R2                          0.989                 0.982
## Adjusted R2                 0.987                 0.976
## Residual Std. Error   63,444.090 (df = 47)    93,285.570 (df = 38)
## F Statistic        364.528*** (df = 12; 47) 174.029*** (df = 12; 38)
## ==================================================================
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

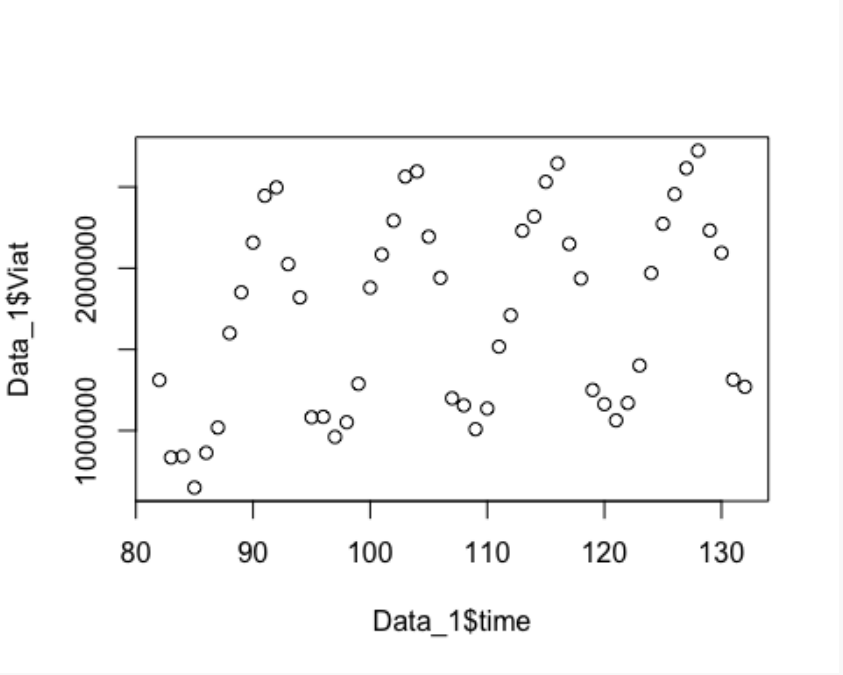
```
summary(ModelAC2)
```

```
##
## Call:
## lm(formula = Data_1$Viat ~ Data_1$time + Data_1$dm2 + Data_1$dm3 +
##     Data_1$dm4 + Data_1$dm5 + Data_1$dm6 + Data_1$dm7 + Data_1$dm8 +
##     Data_1$dm9 + Data_1$dm10 + Data_1$dm11 + Data_1$dm12)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -293082  -48776    7709   53265  156813
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -6777.3   103556.0  -0.065   0.9482
## Data_1$time    9003.9      897.6  10.031 3.14e-12 ***
## Data_1$dm2   125729.4    65969.0   1.906   0.0643 .
## Data_1$dm3   367645.0    65987.3   5.571 2.20e-06 ***
## Data_1$dm4   842372.6    66017.8  12.760 2.60e-15 ***
## Data_1$dm5  1153204.0    66060.5  17.457  < 2e-16 ***
## Data_1$dm6  1340295.9    66115.4  20.272  < 2e-16 ***
## Data_1$dm7  1564632.3    66182.4  23.641  < 2e-16 ***
## Data_1$dm8  1631693.4    66261.5  24.625  < 2e-16 ***
## Data_1$dm9  1157584.8    66352.6  17.446  < 2e-16 ***
## Data_1$dm10  872802.4    62635.8  13.935  < 2e-16 ***
## Data_1$dm11  179418.0    62680.8   2.862   0.0068 **
## Data_1$dm12  137406.1    62738.6   2.190   0.0347 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 93290 on 38 degrees of freedom
## Multiple R-squared:  0.9821, Adjusted R-squared:  0.9765
## F-statistic:   174 on 12 and 38 DF,  p-value: < 2.2e-16
```

```
SModelAC2<-summary(ModelAC2)
```

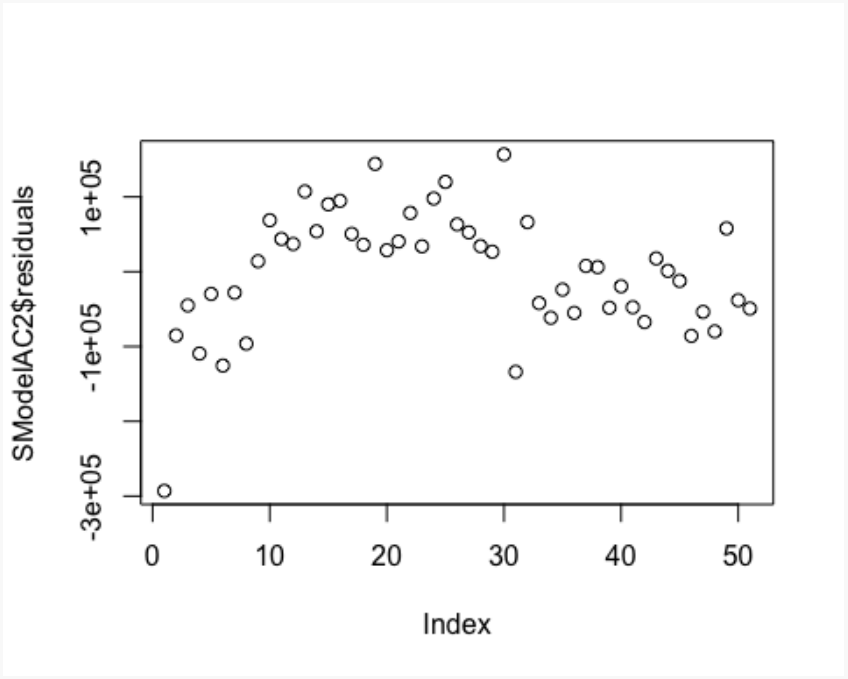
```
plot(Data_1$time,Data_1$Viat)
```




```
names(SModelAC2)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelAC2$residuals)
```

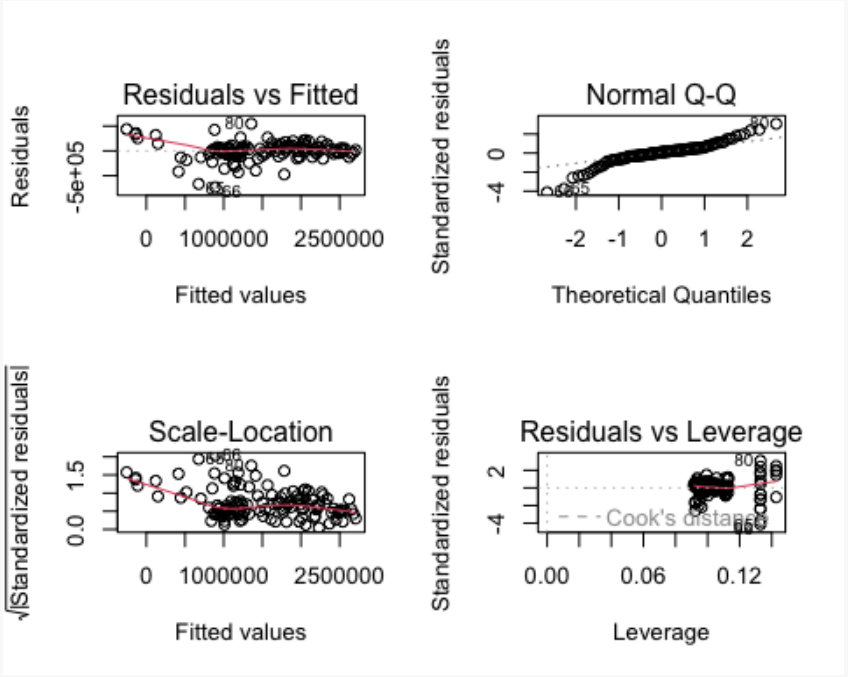



```
stargazer(ModelAC, ModelAC1,ModelAC2, type="text")

##
##
==========================================================================
=======================
##                                  Dependent variable:
##                 ---------------------------------------------------------------------
##                                        Viat
##                       (1)                (2)                (3)
## ---------------------------------------------------------------------------------------
## time                 2,367.891***        4,918.767***        9,003.874***
##                     (437.794)          (482.636)          (897.641)
##
## dm2                  112,706.300        121,568.800***       125,729.400*
##                    (81,232.240)        (40,128.470)        (65,968.970)
##
## dm3                  400,293.600***      418,236.900***       367,645.000***
##                    (81,348.630)        (40,137.170)        (65,987.290)
##
## dm4                  692,878.700***      760,265.700***       842,372.600***
##                    (81,353.070)        (40,151.680)        (66,017.800)
##
```

```
## dm5                 950,867.700***        1,016,582.000***        1,153,204.000***
##                    (81,359.870)          (40,171.980)          (66,060.510)
##
## dm6                 1,161,101.000***      1,209,886.000***        1,340,296.000***
##                    (81,369.030)          (40,198.060)          (66,115.370)
##
## dm7                 1,459,013.000***      1,459,841.000***        1,564,632.000***
##                    (81,380.540)          (40,229.920)          (66,182.370)
##
## dm8                 1,595,656.000***      1,551,263.000***        1,631,693.000***
##                    (81,394.400)          (40,267.540)          (66,261.460)
##
## dm9                 1,118,126.000***      1,084,349.000***        1,157,585.000***
##                    (81,410.610)          (40,310.900)          (66,352.600)
##
## dm10                780,829.700***        809,675.900***          872,802.400***
##                    (81,326.560)          (40,359.990)          (62,635.780)
##
## dm11                147,442.600*          138,766.500***          179,418.000***
##                    (81,348.950)          (40,414.780)          (62,680.790)
##
## dm12                121,952.600           100,675.400**           137,406.100**
##                    (81,373.680)          (40,475.250)          (62,738.610)
##
## Covid              -1,236,493.000***
##                    (47,717.520)
##
## Constant            808,516.900***        749,892.200***          -6,777.318
##                    (63,378.950)          (30,832.060)          (103,556.000)
##
## ------------------------------------------------------------------------------------------
## Observations               132                   60                   51
## R2                        0.930                 0.989                0.982
## Adjusted R2               0.922                 0.987                0.976
## Residual Std. Error  190,503.700 (df = 118)     63,444.090 (df = 47)     93,285.570 (df = 38)
## F Statistic        120.833*** (df = 13; 118) 364.528*** (df = 12; 47) 174.029*** (df = 12; 38)
##
```

```
================================================================================
======================
## Note:                                   *p<0.1; **p<0.05; ***p<0.01

#########################
# 10. RESULTATS
#########################
summary(ModelT3)

##
## Call:
## lm(formula = datos$Viat ~ datos$time + datos$dm2 + datos$dm3 +
##     datos$dm4 + datos$dm5 + datos$dm6 + datos$dm7 + datos$dm8 +
##     datos$dm9 + datos$dm10 + datos$dm11 + datos$dm12 + datos$Covid)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -732942  -56671   15192   81528  546890
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)   808516.9    63379.0  12.757  < 2e-16 ***
## datos$time      2367.9      437.8   5.409 3.37e-07 ***
## datos$dm2     112706.3    81232.2   1.387   0.1679
## datos$dm3     400293.6    81348.6   4.921 2.82e-06 ***
## datos$dm4     692878.7    81353.1   8.517 6.18e-14 ***
## datos$dm5     950867.7    81359.9  11.687  < 2e-16 ***
## datos$dm6    1161101.3    81369.0  14.270  < 2e-16 ***
## datos$dm7    1459013.3    81380.5  17.928  < 2e-16 ***
## datos$dm8    1595656.4    81394.4  19.604  < 2e-16 ***
## datos$dm9    1118126.3    81410.6  13.734  < 2e-16 ***
## datos$dm10    780829.7    81326.6   9.601  < 2e-16 ***
## datos$dm11    147442.6    81349.0   1.812   0.0725 .
## datos$dm12    121952.6    81373.7   1.499   0.1366
## datos$Covid -1236493.1    47717.5 -25.913  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

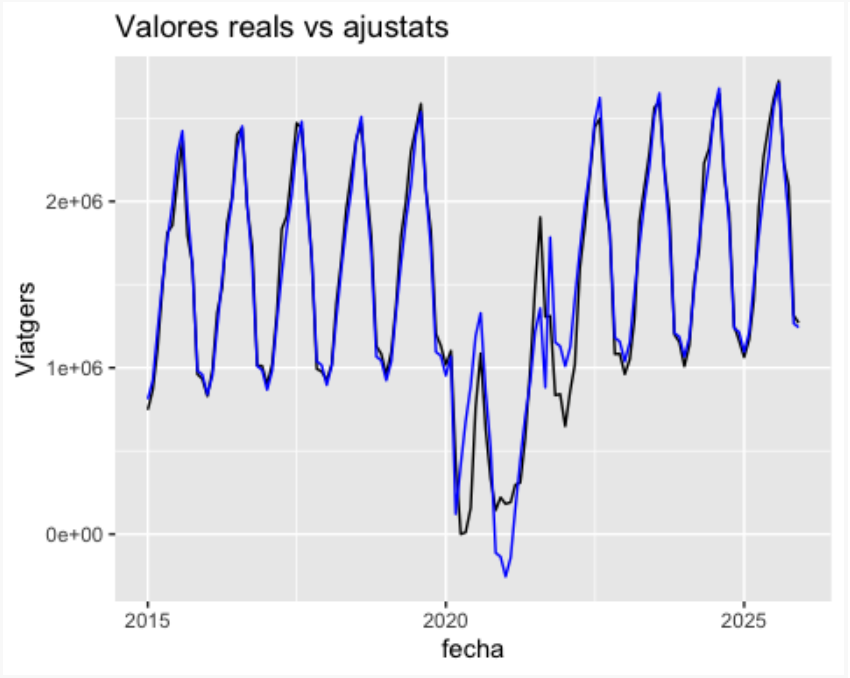
```
## Residual standard error: 190500 on 118 degrees of freedom
## Multiple R-squared:  0.9301, Adjusted R-squared:  0.9224
## F-statistic: 120.8 on 13 and 118 DF,  p-value: < 2.2e-16

#########################
# 11. RESIDUS
#########################

par(mfrow=c(2,2))
plot(ModelT3)
```




```
# Autocorrelación residuos
dwtest(ModelT3)

##
##  Durbin-Watson test
##
## data:  ModelT3
## DW = 0.73577, p-value = 6.934e-13
## alternative hypothesis: true autocorrelation is greater than 0

# Heterocedasticitat
bptest(ModelT3)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  ModelT3
## BP = 48.859, df = 13, p-value = 4.674e-06

########################
# 12. ERRORS ROBUSTOS
########################

coeftest(ModelT3, vcov = vcovHC(ModelT3, type="HC1"))

##
## t test of coefficients:
##
##               Estimate  Std. Error  t value  Pr(>|t|)
## (Intercept)   808516.93    54416.53  14.8579 < 2.2e-16 ***
## datos$time      2367.89      245.78   9.6340 < 2.2e-16 ***
## datos$dm2     112706.29    71031.45   1.5867   0.11525
## datos$dm3     400293.59    80875.74   4.9495 2.500e-06 ***
## datos$dm4     692878.70    81458.30   8.5059 6.557e-14 ***
## datos$dm5     950867.72    92887.91  10.2367 < 2.2e-16 ***
## datos$dm6    1161101.28    96156.81  12.0751 < 2.2e-16 ***
## datos$dm7    1459013.30    78438.17  18.6008 < 2.2e-16 ***
## datos$dm8    1595656.41    81796.81  19.5076 < 2.2e-16 ***
## datos$dm9    1118126.34    76379.17  14.6392 < 2.2e-16 ***
## datos$dm10    780829.71    79282.67   9.8487 < 2.2e-16 ***
## datos$dm11    147442.63    70971.44   2.0775   0.03993 *
## datos$dm12    121952.56    72941.01   1.6719   0.09719 .
## datos$Covid -1236493.13    91719.68 -13.4812 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

########################
# 13. VALORS AJUSTATS
########################

datos$ajustada <- fitted(ModelT3)
```

```
ggplot(datos, aes(fecha)) +
  geom_line(aes(y=Viat), color="black") +
  geom_line(aes(y=ajustada), color="blue") +
  labs(title="Valores reals vs ajustats",
       y="Viatgers")
```




```
#########################
# 14. SUBPERIÒDES
#########################

# PRECOVID
pre <- subset(datos, fecha < as.Date("2020-03-01"))

ModelPRE <- lm(Viat ~ time + mes_factor, data=pre)

# POSTCOVID / RECUPERACIÓN
post <- subset(datos, fecha > as.Date("2021-09-30"))

ModelPOST <- lm(Viat ~ time + mes_factor, data=post)

stargazer(ModelPRE, ModelPOST,
```

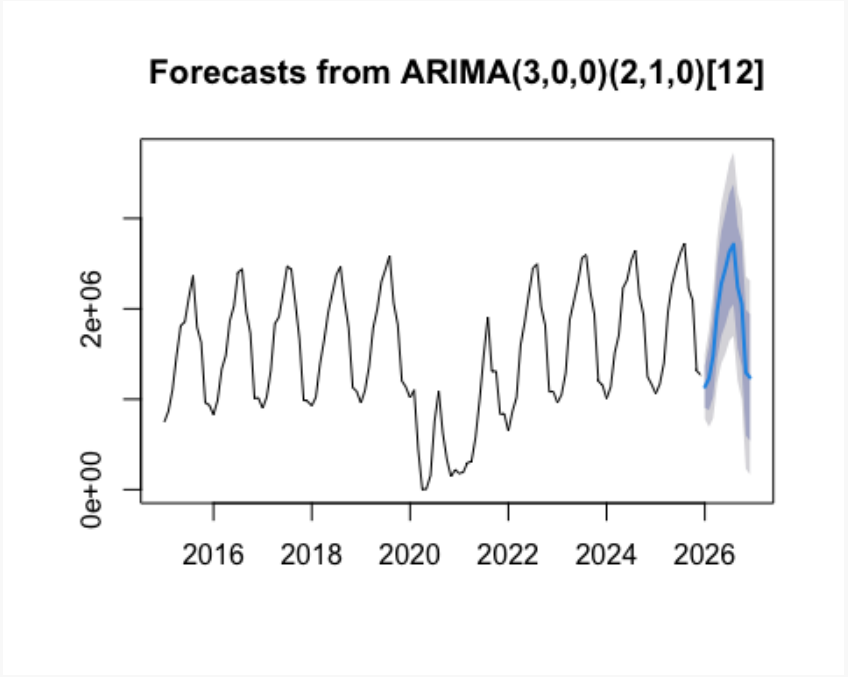
```
        type="text",
        title="Comparación Pre y Post Covid")

##
## Comparación Pre y Post Covid
## ==========================================================================
##                              Dependent variable:
##                    ------------------------------------------------
##                                    Viat
##                          (1)                  (2)
## ------------------------------------------------------------------
## time                  4,753.931***          9,003.874***
##                      (451.469)            (897.641)
##
## mes_factor2           114,676.400***        125,729.400*
##                     (36,345.360)          (65,968.970)
##
## mes_factor3           422,950.000***        367,645.000***
##                     (38,159.150)          (65,987.290)
##
## mes_factor4           765,143.700***        842,372.600***
##                     (38,140.450)          (66,017.800)
##
## mes_factor5           1,021,625.000***      1,153,204.000***
##                     (38,127.080)          (66,060.510)
##
## mes_factor6           1,215,093.000***      1,340,296.000***
##                     (38,119.070)          (66,115.370)
##
## mes_factor7           1,465,214.000***      1,564,632.000***
##                     (38,116.390)          (66,182.370)
##
## mes_factor8           1,556,800.000***      1,631,693.000***
##                     (38,119.070)          (66,261.460)
##
## mes_factor9           1,090,051.000***      1,157,585.000***
##                     (38,127.080)          (66,352.600)
```

```
##
## mes_factor10            815,542.900***          872,802.400***
##                        (38,140.450)            (62,635.780)
##
## mes_factor11            144,798.400***          179,418.000***
##                        (38,159.150)            (62,680.790)
##
## mes_factor12            106,872.000***          137,406.100**
##                        (38,183.180)            (62,738.610)
##
## Constant                749,629.600***           -6,777.318
##                        (29,262.020)            (103,556.000)
##
## ------------------------------------------------------------------
## Observations                62                      51
## R2                         0.989                   0.982
## Adjusted R2                0.987                   0.976
## Residual Std. Error   62,947.150 (df = 49)    93,285.570 (df = 38)
## F Statistic        383.166*** (df = 12; 49) 174.029*** (df = 12; 38)
## ==================================================================
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

```
#########################
# 15. MODEL ARIMA
#########################

modelo_arima <- auto.arima(serie_ts)

summary(modelo_arima)
```

```
## Series: serie_ts
## ARIMA(3,0,0)(2,1,0)[12]
##
## Coefficients:
##          ar1      ar2      ar3     sar1     sar2
##       1.1302  -0.0668  -0.1471  -0.4441  -0.2778
## s.e.  0.0949   0.1476   0.0935   0.0946   0.0879
##
```

```
## sigma^2 = 3.277e+10:  log likelihood = -1623.33
## AIC=3258.66   AICc=3259.4   BIC=3275.38
##
## Training set error measures:
##                    ME     RMSE      MAE  MPE MAPE      MASE       ACF1
## Training set 5191.017 168958.4 105616.2 -Inf  Inf 0.3149666 0.01568584

forecast_arima <- forecast(modelo_arima, h=12)


plot(forecast_arima)
```




```
############################################################
# FI
############################################################
```

Annex 3: Anàlisi de la despesa mitja diària de Catalunya

Introducció

Aquest document analitza l'evolució de la despesa mitjana diaria dels viatgers (en €) a Catalunya amb dades a partir de 2016 mitjançant tècniques de sèries temporals i regressió economètrica. L'objectiu és identificar la tendència estructural, l'estacionalitat mensual i l'impacte de la COVID-19 sobre l'activitat hotelera.

```
#########################
# 1. PAQUETS
#########################



library(readxl)
library(tidyverse)

## ── Attaching core tidyverse packages ──────────────────────────────── tidyverse 2.0.0 ──
## ✔ dplyr     1.2.1     ✔ readr     2.1.4
## ✔ forcats   1.0.0     ✔ stringr   1.5.0
## ✔ ggplot2   3.4.4     ✔ tibble    3.2.1
## ✔ lubridate 1.9.5     ✔ tidyr     1.3.0
## ✔ purrr     1.0.1
## ── Conflicts ────────────────────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors

library(lubridate)
library(forecast)

## Registered S3 method overwritten by 'quantmod':
##   method            from
##   as.zoo.data.frame zoo

library(tseries)
library(lmtest)
```

```
## Loading required package: zoo
##
## Attaching package: 'zoo'
##
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
library(sandwich)
library(modelsummary)
```

```
#########################
# 2. IMPORTAR DADES
#########################
datos <- read_excel("/Users/judit/Downloads/DadesDMD_C.xlsx")

# Ordenar cronològicament
datos <- datos %>%
  arrange(Any, mes)

mean(datos$Desp)
```

```
## [1] 180.4583
```

```
sd(datos$Desp)
```

```
## [1] 40.26569
```

```
min(datos$Desp)
```

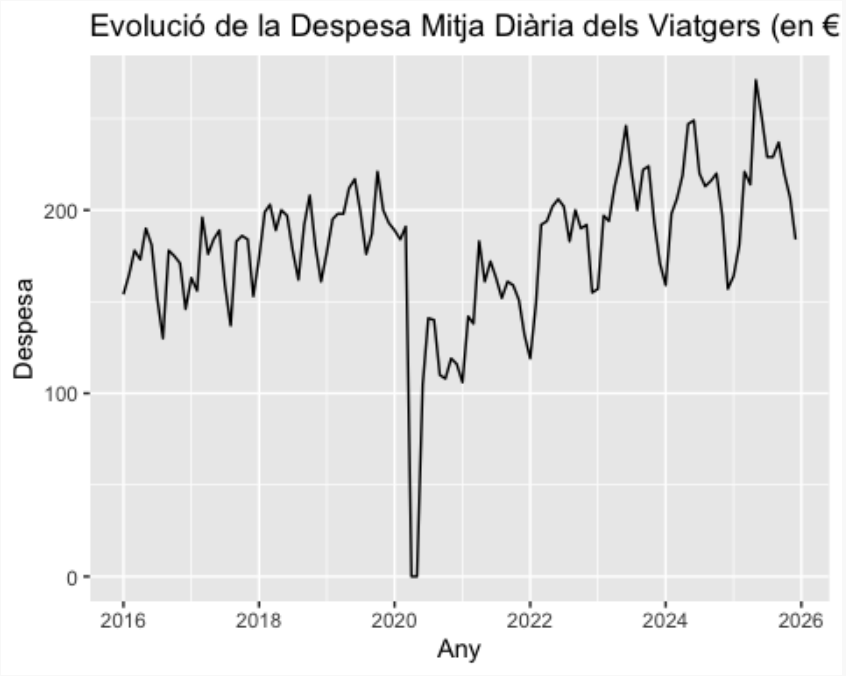
```
## [1] 0
```

```
max(datos$Desp)
```

```
## [1] 271
```

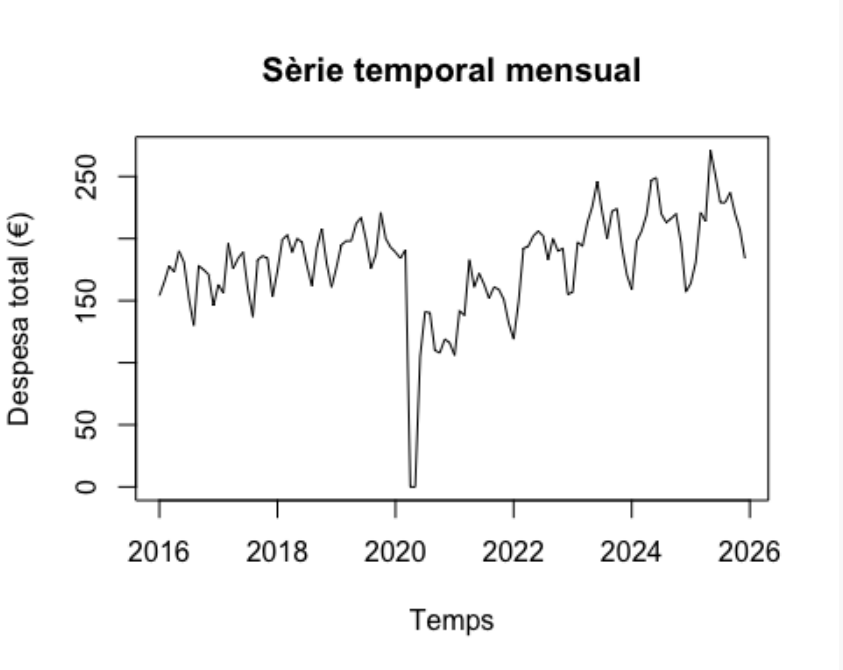
```
# Data mensual
datos$fecha <- as.Date(paste(datos$Any, datos$mes, "01", sep="-"))

# Tendència temporal
datos$time <- 1:nrow(datos)
```

```r
# Convertir mes en factor (gener com a base)
datos$mes_factor <- factor(datos$mes)

#########################
# 4. VARIABLE COVID
#########################

# Dummy COVID: març 2020 - setembre 2021
datos$Covid <- ifelse(datos$fecha >= as.Date("2020-03-01") &
                      datos$fecha <= as.Date("2021-09-30"), 1, 0)

#########################
# 5. ANÀLISI DESCRIPTIVA
#########################

ggplot(datos, aes(x = fecha, y = Desp)) +
  geom_line(color = "black") +
  labs(title = "Evolució de la Despesa Mitja Diària dels Viatgers (en €)",
       x = "Any",
       y = "Despesa")
```

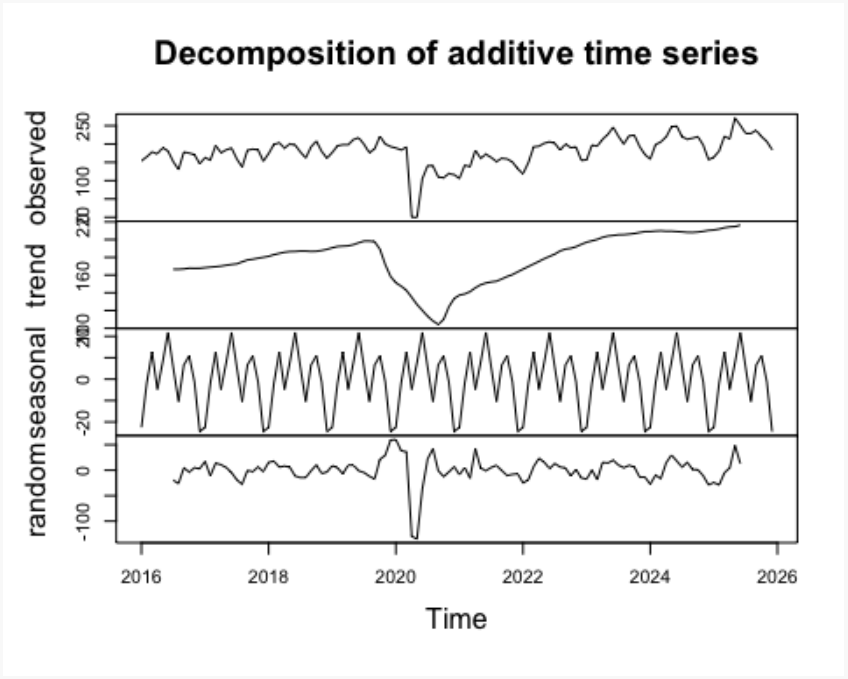



```r
#########################
# 6. SÈRIE TEMPORAL
```

```
#########################

# Definir principi de la sèrie
inicio <- c(datos$Any[1], datos$mes[1])

# Sèrie temporal mensual
serie_ts <- ts(datos$Desp,
               start = inicio,
               frequency = 12)
# Resultat
plot(serie_ts,
    main = "Sèrie temporal mensual",
    ylab = "Despesa total (€)",
    xlab = "Temps")
```




```
datos <- datos[order(datos$Any, datos$mes), ]

#########################
# 7. DESCOMPOSICIÓ
#########################

descomp <- decompose(serie_ts, type = "additive")
plot(descomp)
```




```
############################
# 8. ESTACIONARIETAT
############################

resultat_adf <- adf.test(serie_ts)
print(resultat_adf)

##
##  Augmented Dickey-Fuller Test
##
## data:  serie_ts
## Dickey-Fuller = -3.2645, Lag order = 4, p-value = 0.0805
## alternative hypothesis: stationary

# El test Dickey-Fuller augmentat no rebutja la hipòtesi nul·la (nivell de significació major a 0.05), indicant que la sèrie no és estacionària.
```

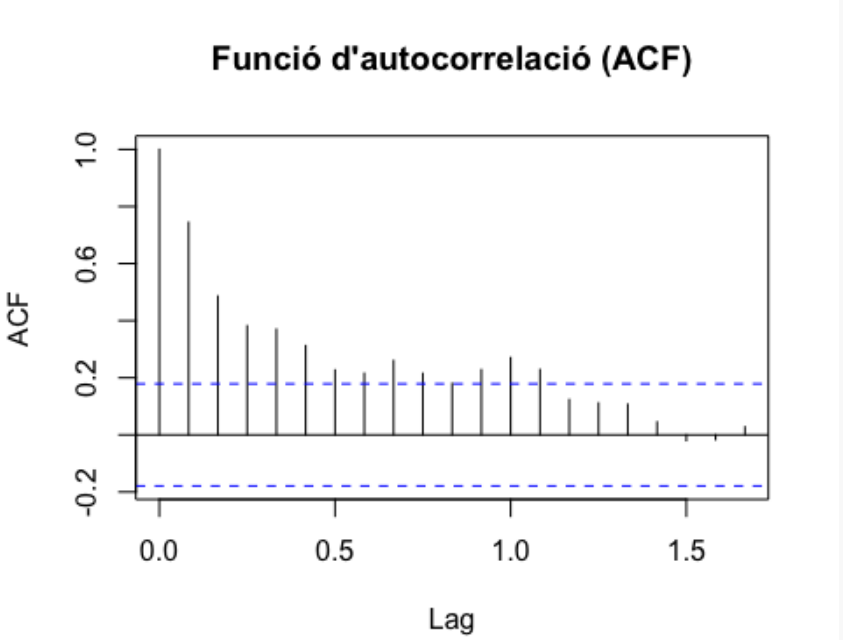
Interpretació:

El test Dickey-Fuller augmentat permet contrastar la presència d'arrel unitària.

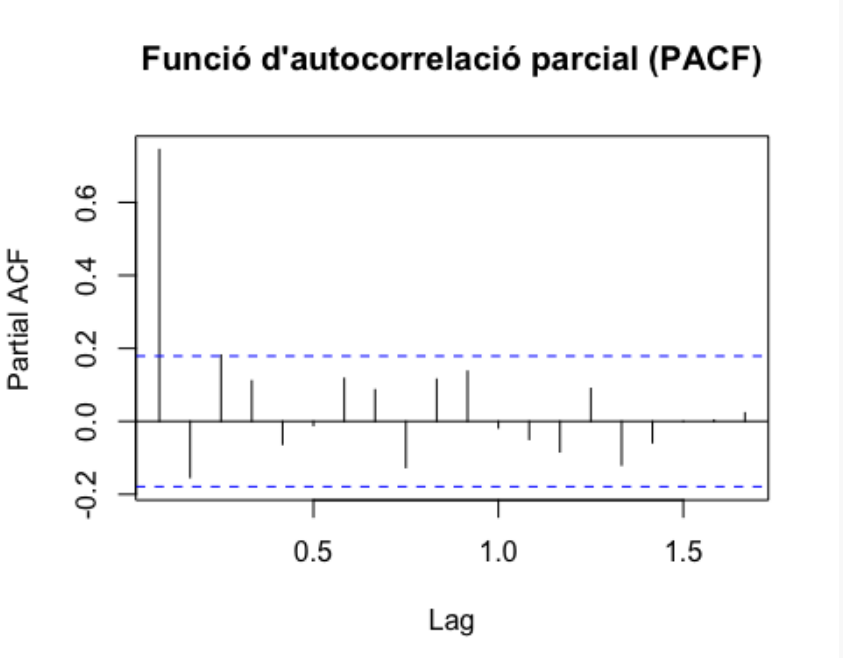
- H0: la sèrie no és estacionària
- H1: la sèrie és estacionària

Si p > 0.05 no es rebutja H0.

```
########################
# 9. ACF I PACF
########################
acf(serie_ts, main = "Funció d'autocorrelació (ACF)")
```

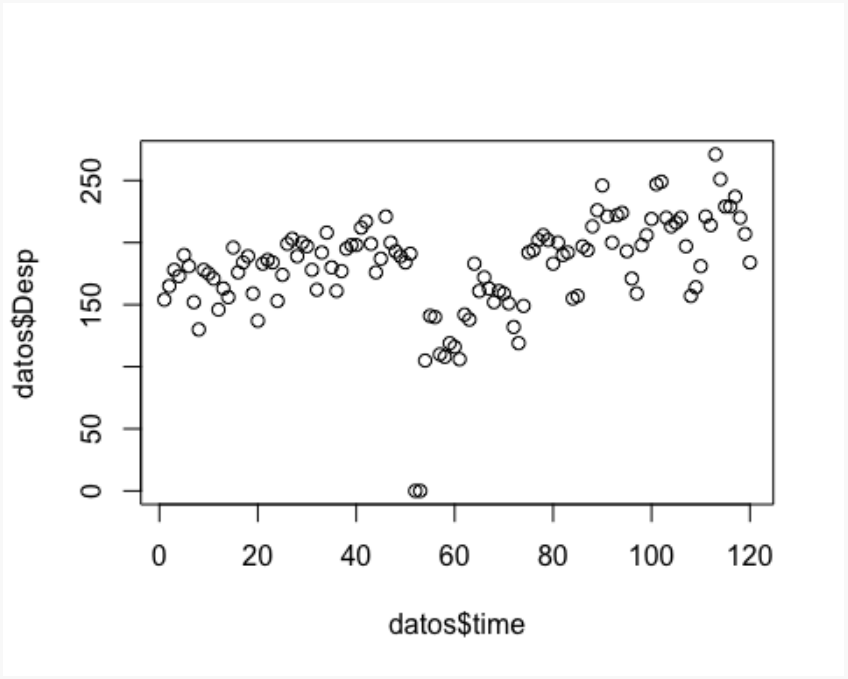



```
pacf(serie_ts, main = "Funció d'autocorrelació parcial (PACF)")
```




```
########################
# 10. MODELS ECONOMÈTRICS
```

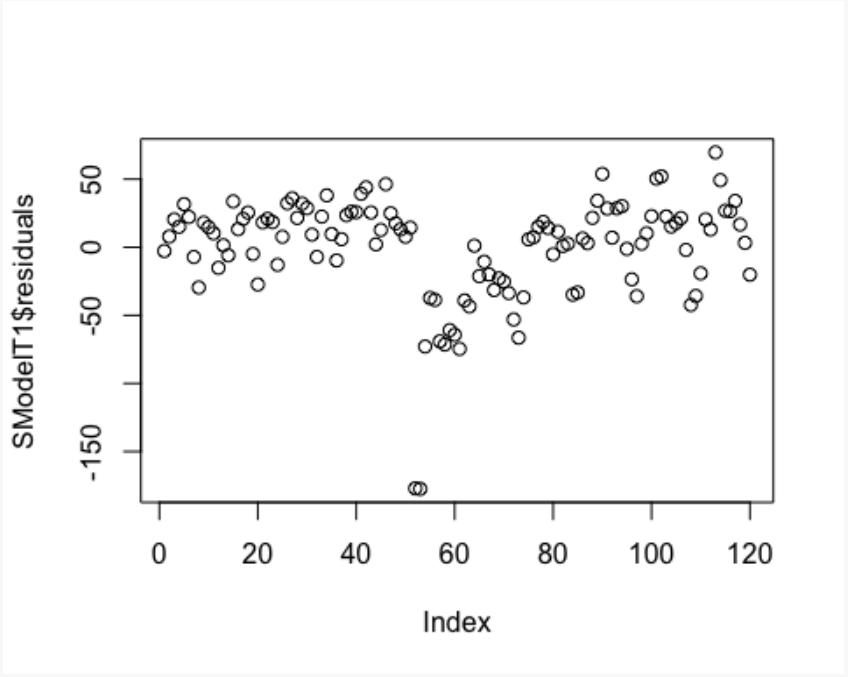
```
#########################

# MODEL 1: Tendència temporal
ModelT1 <- lm(Desp ~ time, data = datos)
ModelT1<-lm(datos$Desp~datos$time)
SModelT1<-summary(ModelT1)
summary(ModelT1)

##
## Call:
## lm(formula = datos$Desp ~ datos$time)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -177.472  -19.388    9.492   22.524   69.637
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 156.3682     6.9756  22.416  < 2e-16 ***
## datos$time    0.3982     0.1001   3.979  0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 37.97 on 118 degrees of freedom
## Multiple R-squared:  0.1183, Adjusted R-squared:  0.1109
## F-statistic: 15.84 on 1 and 118 DF,  p-value: 0.0001197

plot(datos$time,datos$Desp)
```




```
names(SModelT1)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelT1$residuals)
```




```
library(stargazer)
```

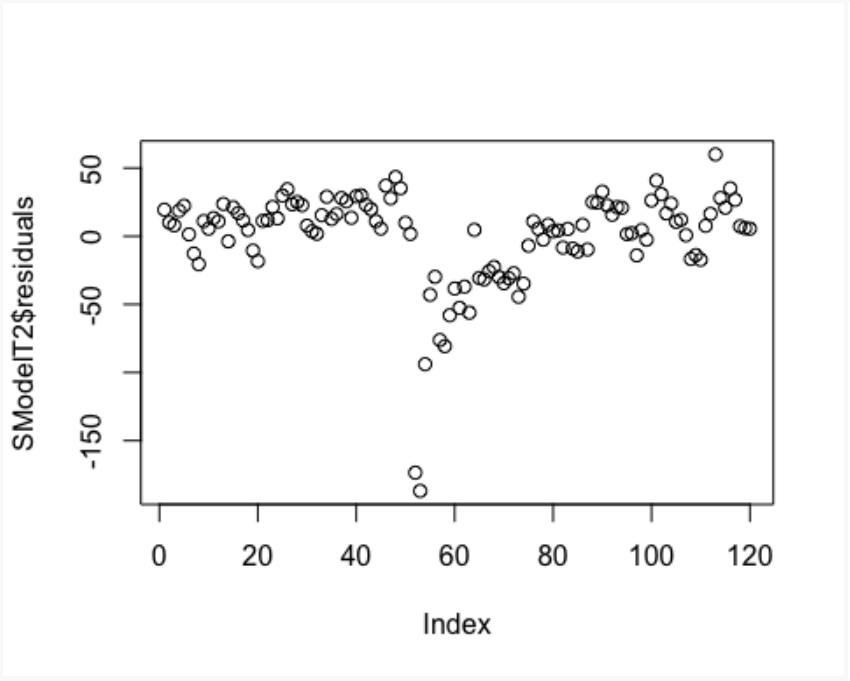
```
##
## Please cite as:

##  Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.

##  R package version 5.2.3. https://CRAN.R-project.org/package=stargazer

stargazer(ModelT1, type="text")

##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                                Desp
## ------------------------------------------------
## time                          0.398***
##                              (0.100)
##
## Constant                     156.368***
##                              (6.976)
##
## ------------------------------------------------
## Observations                    120
## R2                             0.118
## Adjusted R2                    0.111
## Residual Std. Error      37.968 (df = 118)
## F Statistic          15.836*** (df = 1; 118)
## ================================================
## Note:                *p<0.1; **p<0.05; ***p<0.01

# MODEL 2: Tendència + estacionalitat mensual
ModelT2<-lm(datos$Desp~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+
datos$dm7+datos$dm8+datos$dm9+
            datos$dm10+datos$dm11+datos$dm12)
SModelT2<-summary(ModelT2)


summary(ModelT2)

##
## Call:
```

```
## lm(formula = datos$Desp ~ datos$time + datos$dm2 + datos$dm3 +
##     datos$dm4 + datos$dm5 + datos$dm6 + datos$dm7 + datos$dm8 +
##     datos$dm9 + datos$dm10 + datos$dm11 + datos$dm12)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -186.892  -10.719    8.166   21.384   60.030
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 134.1282    12.9884  10.327  < 2e-16 ***
## datos$time    0.4013     0.0988   4.062 9.31e-05 ***
## datos$dm2    19.9987    16.6837   1.199   0.2333
## datos$dm3    34.6974    16.6846   2.080   0.0400 *
## datos$dm4    18.4961    16.6860   1.108   0.2701
## datos$dm5    31.4948    16.6881   1.887   0.0618 .
## datos$dm6    43.0935    16.6907   2.582   0.0112 *
## datos$dm7    27.7922    16.6939   1.665   0.0989 .
## datos$dm8    13.1909    16.6977   0.790   0.4313
## datos$dm9    29.1896    16.7021   1.748   0.0834 .
## datos$dm10   31.2883    16.7071   1.873   0.0638 .
## datos$dm11   19.1869    16.7126   1.148   0.2535
## datos$dm12   -3.8144    16.7188  -0.228   0.8200
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 37.31 on 107 degrees of freedom
## Multiple R-squared:  0.2282, Adjusted R-squared:  0.1416
## F-statistic: 2.636 on 12 and 107 DF,  p-value: 0.003981

names(SModelT2)

##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"

stargazer(ModelT1, ModelT2, type="text")
```

```
##
## ==========================================================================
##                          Dependent variable:
##                  ----------------------------------------------
##                                 Desp
##                        (1)                (2)
## ---------------------------------------------------------------
## time                  0.398***           0.401***
##                      (0.100)            (0.099)
##
## dm2                                      19.999
##                                         (16.684)
##
## dm3                                     34.697**
##                                         (16.685)
##
## dm4                                      18.496
##                                         (16.686)
##
## dm5                                      31.495*
##                                         (16.688)
##
## dm6                                     43.093**
##                                         (16.691)
##
## dm7                                      27.792*
##                                         (16.694)
##
## dm8                                      13.191
##                                         (16.698)
##
## dm9                                      29.190*
##                                         (16.702)
##
## dm10                                     31.288*
##                                         (16.707)
##
```

```
## dm11                                19.187
##                                    (16.713)
##
## dm12                                -3.814
##                                    (16.719)
##
## Constant              156.368***          134.128***
##                      (6.976)             (12.988)
##
## ------------------------------------------------------------------
## Observations              120                 120
## R2                       0.118               0.228
## Adjusted R2              0.111               0.142
## Residual Std. Error  37.968 (df = 118)    37.305 (df = 107)
## F Statistic       15.836*** (df = 1; 118) 2.636*** (df = 12; 107)
## ==================================================================
## Note:                              *p<0.1; **p<0.05; ***p<0.01
```

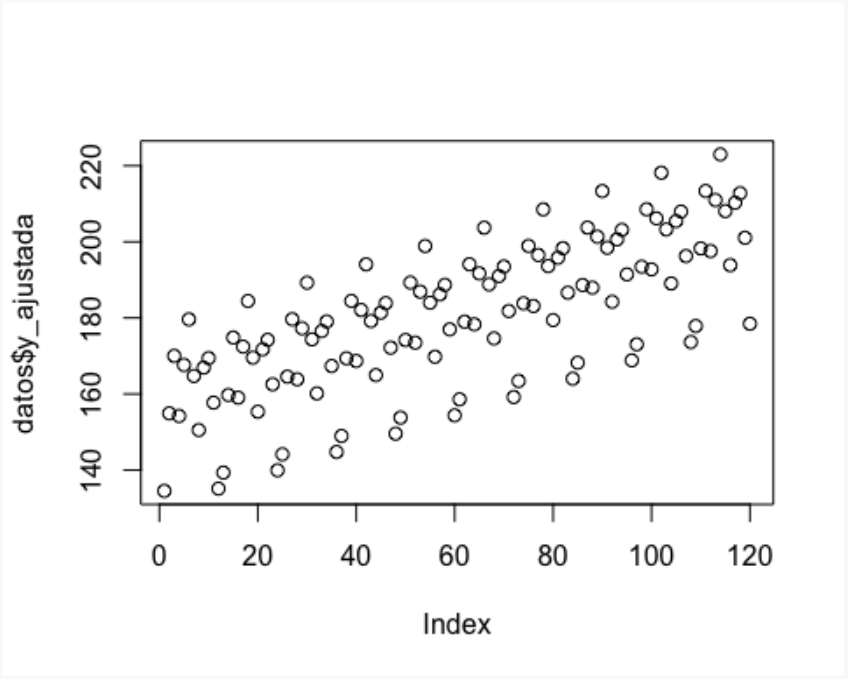
```
plot(SModelT2$residuals)
```



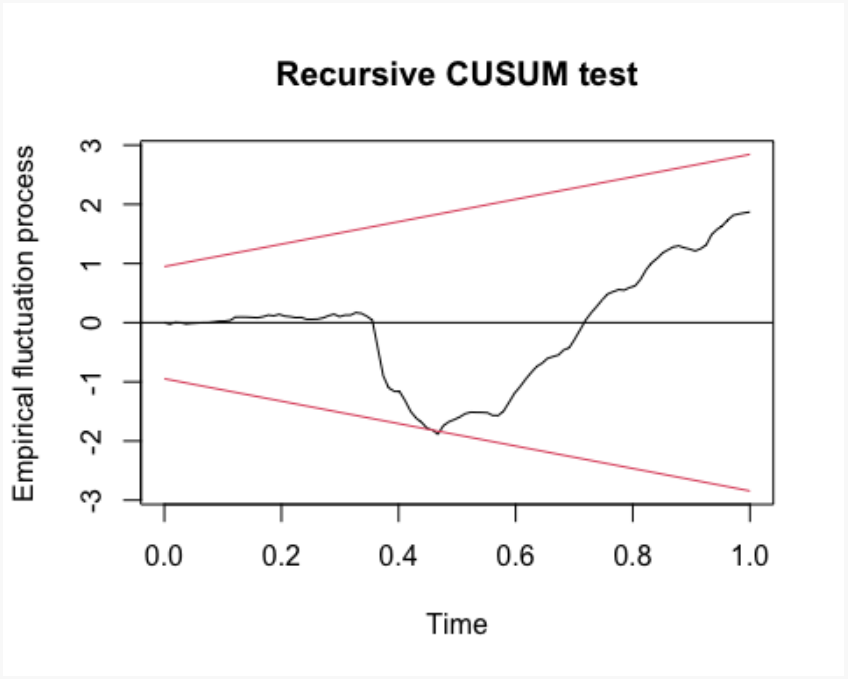

```
y_ajustada <- fitted(ModelT2)

y_ajustada
```

```
##       1       2       3       4       5       6       7       8
## 134.5295 154.9295 170.0295 154.2295 167.6295 179.6295 164.7295 150.5295
##       9      10      11      12      13      14      15      16
## 166.9295 169.4295 157.7295 135.1295 139.3452 159.7452 174.8452 159.0452
##      17      18      19      20      21      22      23      24
## 172.4452 184.4452 169.5452 155.3452 171.7452 174.2452 162.5452 139.9452
##      25      26      27      28      29      30      31      32
## 144.1609 164.5609 179.6609 163.8609 177.2609 189.2609 174.3609 160.1609
##      33      34      35      36      37      38      39      40
## 176.5609 179.0609 167.3609 144.7609 148.9765 169.3765 184.4765 168.6765
##      41      42      43      44      45      46      47      48
## 182.0765 194.0765 179.1765 164.9765 181.3765 183.8765 172.1765 149.5765
##      49      50      51      52      53      54      55      56
## 153.7922 174.1922 189.2922 173.4922 186.8922 198.8922 183.9922 169.7922
##      57      58      59      60      61      62      63      64
## 186.1922 188.6922 176.9922 154.3922 158.6078 179.0078 194.1078 178.3078
##      65      66      67      68      69      70      71      72
## 191.7078 203.7078 188.8078 174.6078 191.0078 193.5078 181.8078 159.2078
##      73      74      75      76      77      78      79      80
## 163.4235 183.8235 198.9235 183.1235 196.5235 208.5235 193.6235 179.4235
##      81      82      83      84      85      86      87      88
## 195.8235 198.3235 186.6235 164.0235 168.2391 188.6391 203.7391 187.9391
##      89      90      91      92      93      94      95      96
## 201.3391 213.3391 198.4391 184.2391 200.6391 203.1391 191.4391 168.8391
##      97      98      99     100     101     102     103     104
## 173.0548 193.4548 208.5548 192.7548 206.1548 218.1548 203.2548 189.0548
##     105     106     107     108     109     110     111     112
## 205.4548 207.9548 196.2548 173.6548 177.8705 198.2705 213.3705 197.5705
##     113     114     115     116     117     118     119     120
## 210.9705 222.9705 208.0705 193.8705 210.2705 212.7705 201.0705 178.4705
```

```
datos$y_ajustada<-fitted(ModelT2)

plot(datos$y_ajustada )
```




```
Data_1 <- datos

library(strucchange)

##
## Attaching package: 'strucchange'

## The following object is masked from 'package:stringr':
##
##     boundary

efp_model                                                                    <-
efp(datos$Desp~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+datos$dm7
+datos$dm8+datos$dm9+
            datos$dm10+datos$dm11+datos$dm12, type="Rec-CUSUM")

plot (efp_model)
```




```
sctest (efp_model)

##
##  Recursive CUSUM test
##
## data:  efp_model
## S = 0.97458, p-value = 0.04085

# MODEL 3: Tendència + estacionalitat + CÒVID
ModelT3<-lm(datos$Desp~datos$time+datos$dm2+datos$dm3+datos$dm4+datos$dm5+datos$dm6+
datos$dm7+datos$dm8+datos$dm9+
            datos$dm10+datos$dm11+datos$dm12+datos$Covid)
SModelT3<-summary(ModelT3)


names(ModelT3)

##  [1] "coefficients"  "residuals"     "effects"       "rank"
##  [5] "fitted.values" "assign"        "qr"            "df.residual"
##  [9] "xlevels"       "call"          "terms"         "model"

stargazer(ModelT1, ModelT2,ModelT3, type="text")

##
##
==========================================================================
```

```
===================
##                              Dependent variable:
##                 ---------------------------------------------------------------------
##                                        Desp
##                      (1)                (2)                (3)
## ---------------------------------------------------------------------------------------
## time                 0.398***           0.401***           0.398***
##                     (0.100)            (0.099)            (0.072)
##
## dm2                                     19.999             20.002
##                                        (16.684)           (12.075)
##
## dm3                                     34.697**           41.458***
##                                        (16.685)           (12.095)
##
## dm4                                     18.496             25.259**
##                                        (16.686)           (12.096)
##
## dm5                                     31.495*            38.261***
##                                        (16.688)           (12.098)
##
## dm6                                     43.093**           49.863***
##                                        (16.691)           (12.099)
##
## dm7                                     27.792*            34.564***
##                                        (16.694)           (12.102)
##
## dm8                                     13.191             19.966
##                                        (16.698)           (12.105)
##
## dm9                                     29.190*            35.967***
##                                        (16.702)           (12.108)
##
## dm10                                    31.288*            31.314**
##                                        (16.707)           (12.092)
##
## dm11                                    19.187             19.215
```

```
##                                 (16.713)            (12.096)
##
## dm12                             -3.814              -3.783
##                                 (16.719)            (12.100)
##
## Covid                                               -67.548***
##                                                     (6.814)
##
## Constant              156.368***          134.128***          141.039***
##                      (6.976)             (12.988)            (9.426)
##
## ------------------------------------------------------------------------------------------
## Observations              120                 120                 120
## R2                       0.118               0.228               0.599
## Adjusted R2              0.111               0.142               0.550
## Residual Std. Error   37.968 (df = 118)   37.305 (df = 107)   27.000 (df = 106)
## F Statistic        15.836*** (df = 1; 118) 2.636*** (df = 12; 107) 12.204*** (df = 13; 106)
##
==========================================================================
===================
## Note:                                           *p<0.1; **p<0.05; ***p<0.01
```

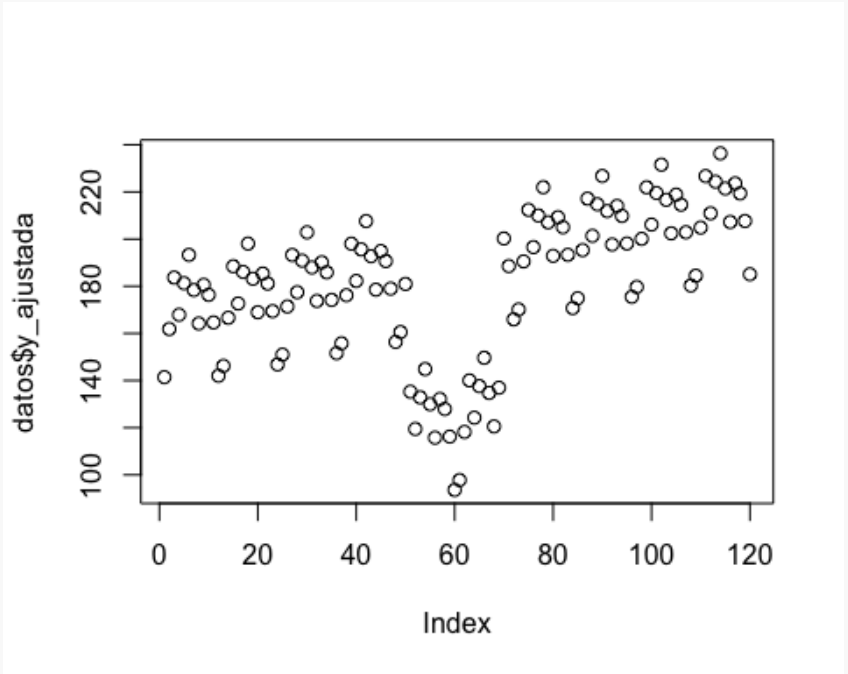
```
y_ajustada <- fitted(ModelT3)
```

```
y_ajustada
```

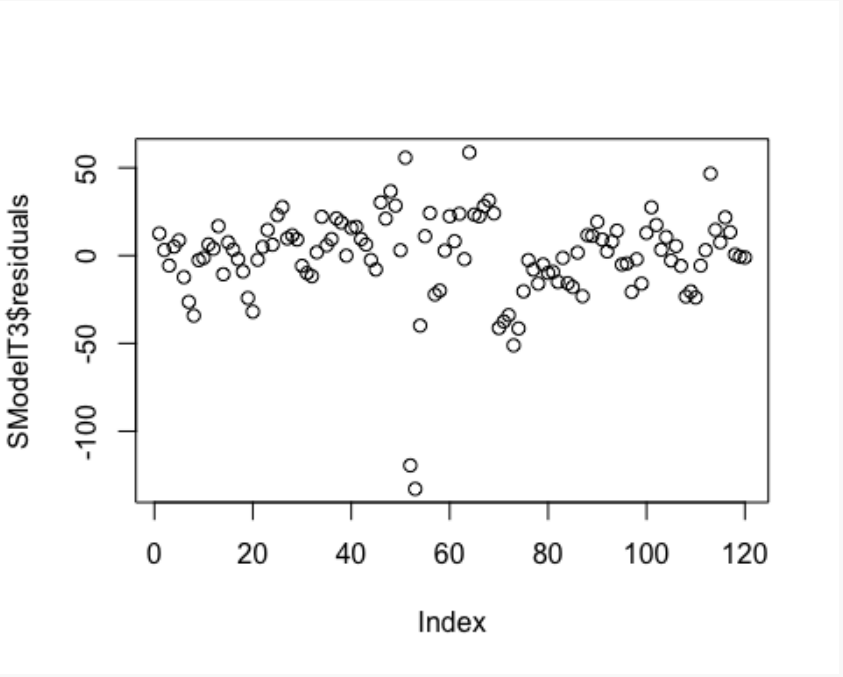
```
##        1        2        3        4        5        6        7        8
## 141.43790 161.83790 183.69274 167.89274 181.29274 193.29274 178.39274 164.19274
##        9       10       11       12       13       14       15       16
## 180.59274 176.33790 164.63790 142.03790 146.21944 166.61944 188.47428 172.67428
##       17       18       19       20       21       22       23       24
## 186.07428 198.07428 183.17428 168.97428 185.37428 181.11944 169.41944 146.81944
##       25       26       27       28       29       30       31       32
## 151.00098 171.40098 193.25582 177.45582 190.85582 202.85582 187.95582 173.75582
##       33       34       35       36       37       38       39       40
## 190.15582 185.90098 174.20098 151.60098 155.78252 176.18252 198.03736 182.23736
##       41       42       43       44       45       46       47       48
## 195.63736 207.63736 192.73736 178.53736 194.93736 190.68252 178.98252 156.38252
```

```
##       49       50       51       52       53       54       55       56
## 160.56407 180.96407 135.27054 119.47054 132.87054 144.87054 129.97054 115.77054
##       57       58       59       60       61       62       63       64
## 132.17054 127.91571 116.21571  93.61571  97.79725 118.19725 140.05208 124.25208
##       65       66       67       68       69       70       71       72
## 137.65208 149.65208 134.75208 120.55208 136.95208 200.24561 188.54561 165.94561
##       73       74       75       76       77       78       79       80
## 170.12715 190.52715 212.38198 196.58198 209.98198 221.98198 207.08198 192.88198
##       81       82       83       84       85       86       87       88
## 209.28198 205.02715 193.32715 170.72715 174.90869 195.30869 217.16352 201.36352
##       89       90       91       92       93       94       95       96
## 214.76352 226.76352 211.86352 197.66352 214.06352 209.80869 198.10869 175.50869
##       97       98       99      100      101      102      103      104
## 179.69023 200.09023 221.94507 206.14507 219.54507 231.54507 216.64507 202.44507
##      105      106      107      108      109      110      111      112
## 218.84507 214.59023 202.89023 180.29023 184.47177 204.87177 226.72661 210.92661
##      113      114      115      116      117      118      119      120
## 224.32661 236.32661 221.42661 207.22661 223.62661 219.37177 207.67177 185.07177
```

```
datos$y_ajustada<-fitted(ModelT3)

plot(datos$y_ajustada )
```



```
plot(SModelT3$residuals)
```




```
# Canvi de nom per poder fer la comparació
Data_1 <- datos

ModelAC<-lm(Data_1$Desp~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5+
Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
            Data_1$dm10+Data_1$dm11+Data_1$dm12+Data_1$Covid)

stargazer(ModelAC, type="text")

##
## ================================================
##                         Dependent variable:
##                     ---------------------------
##                               Desp
## ------------------------------------------------
## time                          0.398***
##                              (0.072)
##
## dm2                           20.002
##                              (12.075)
##
```

```
## dm3                 41.458***
##                    (12.095)
##
## dm4                 25.259**
##                    (12.096)
##
## dm5                 38.261***
##                    (12.098)
##
## dm6                 49.863***
##                    (12.099)
##
## dm7                 34.564***
##                    (12.102)
##
## dm8                  19.966
##                    (12.105)
##
## dm9                 35.967***
##                    (12.108)
##
## dm10                31.314**
##                    (12.092)
##
## dm11                 19.215
##                    (12.096)
##
## dm12                 -3.783
##                    (12.100)
##
## Covid               -67.548***
##                     (6.814)
##
## Constant            141.039***
##                     (9.426)
##
## ------------------------------------------------
```

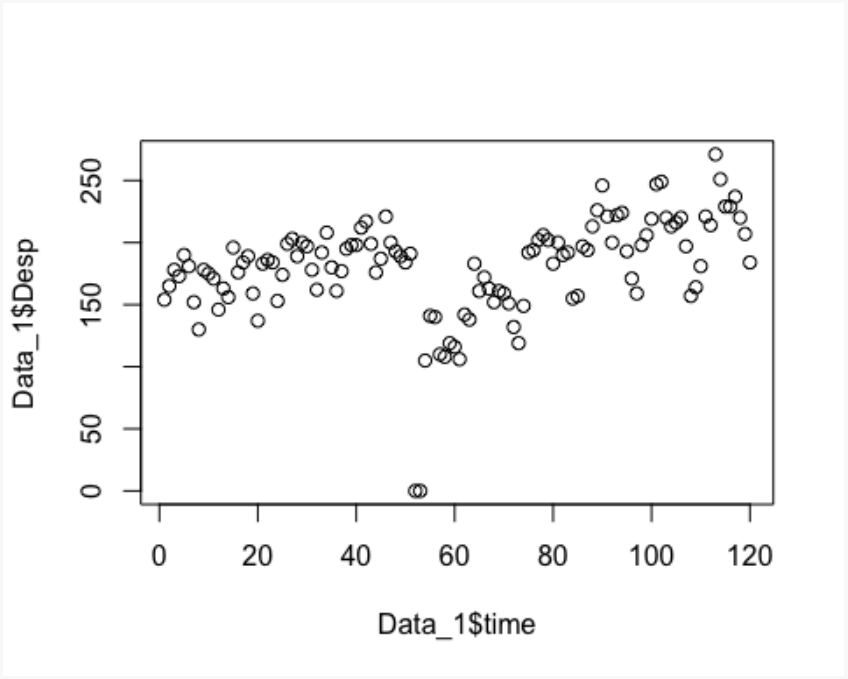
```
## Observations                120
## R2                        0.599
## Adjusted R2               0.550
## Residual Std. Error     27.000 (df = 106)
## F Statistic         12.204*** (df = 13; 106)
## ================================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

```
#Model Estimat etapa Pre-còvid


pre <- subset(datos, time< 51)



ModelAC1<-lm(Data_1$Desp~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5
+Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
           Data_1$dm10+Data_1$dm11+Data_1$dm12)

stargazer(ModelAC1, type="text")

##
## ================================================
##                       Dependent variable:
##                   ---------------------------
##                             Desp
## ------------------------------------------------
## time                       0.401***
##                           (0.099)
##
## dm2                         19.999
##                           (16.684)
##
## dm3                         34.697**
##                           (16.685)
##
## dm4                         18.496
##                           (16.686)
##
## dm5                         31.495*
```

```
##                         (16.688)
##
## dm6                     43.093**
##                         (16.691)
##
## dm7                     27.792*
##                         (16.694)
##
## dm8                     13.191
##                         (16.698)
##
## dm9                     29.190*
##                         (16.702)
##
## dm10                    31.288*
##                         (16.707)
##
## dm11                    19.187
##                         (16.713)
##
## dm12                    -3.814
##                         (16.719)
##
## Constant                134.128***
##                         (12.988)
##
## ------------------------------------------------
## Observations                  120
## R2                           0.228
## Adjusted R2                  0.142
## Residual Std. Error      37.305 (df = 107)
## F Statistic          2.636*** (df = 12; 107)
## ================================================
## Note:              *p<0.1; **p<0.05; ***p<0.01
```

```
summary(ModelAC1)
```

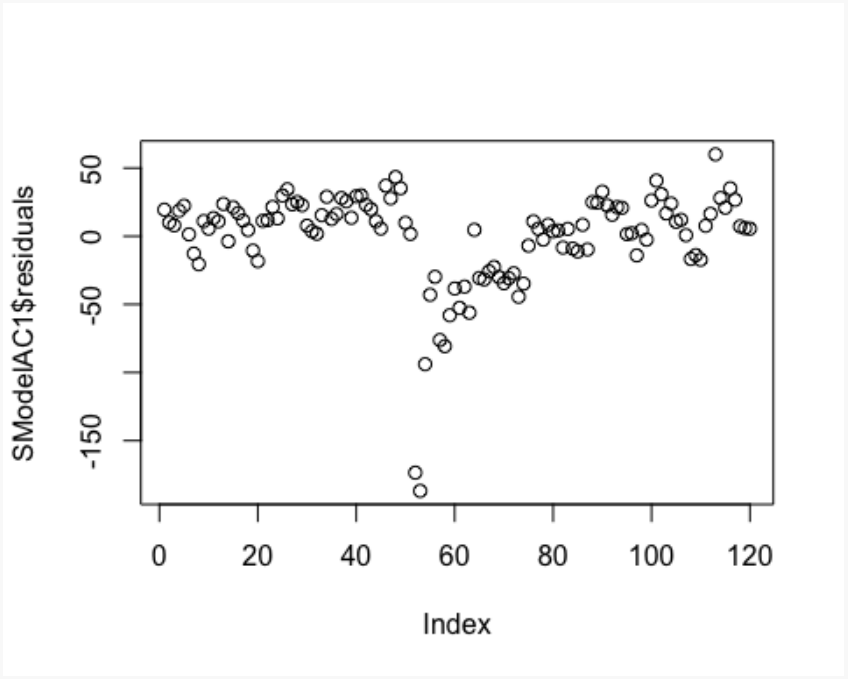
```
##
## Call:
## lm(formula = Data_1$Desp ~ Data_1$time + Data_1$dm2 + Data_1$dm3 +
##     Data_1$dm4 + Data_1$dm5 + Data_1$dm6 + Data_1$dm7 + Data_1$dm8 +
##     Data_1$dm9 + Data_1$dm10 + Data_1$dm11 + Data_1$dm12)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -186.892  -10.719    8.166   21.384   60.030
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 134.1282    12.9884  10.327  < 2e-16 ***
## Data_1$time   0.4013     0.0988   4.062 9.31e-05 ***
## Data_1$dm2   19.9987    16.6837   1.199   0.2333
## Data_1$dm3   34.6974    16.6846   2.080   0.0400 *
## Data_1$dm4   18.4961    16.6860   1.108   0.2701
## Data_1$dm5   31.4948    16.6881   1.887   0.0618 .
## Data_1$dm6   43.0935    16.6907   2.582   0.0112 *
## Data_1$dm7   27.7922    16.6939   1.665   0.0989 .
## Data_1$dm8   13.1909    16.6977   0.790   0.4313
## Data_1$dm9   29.1896    16.7021   1.748   0.0834 .
## Data_1$dm10  31.2883    16.7071   1.873   0.0638 .
## Data_1$dm11  19.1869    16.7126   1.148   0.2535
## Data_1$dm12  -3.8144    16.7188  -0.228   0.8200
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 37.31 on 107 degrees of freedom
## Multiple R-squared:  0.2282, Adjusted R-squared:  0.1416
## F-statistic: 2.636 on 12 and 107 DF,  p-value: 0.003981

SModelAC1<-summary(ModelAC1)


plot(Data_1$time,Data_1$Desp)
```




```
names(SModelAC1)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelAC1$residuals)
```




```
#Model Estimat etapa Post-còvid (suposem que comença a partir de l'observació 69)
post <- subset(datos, time>69)
```

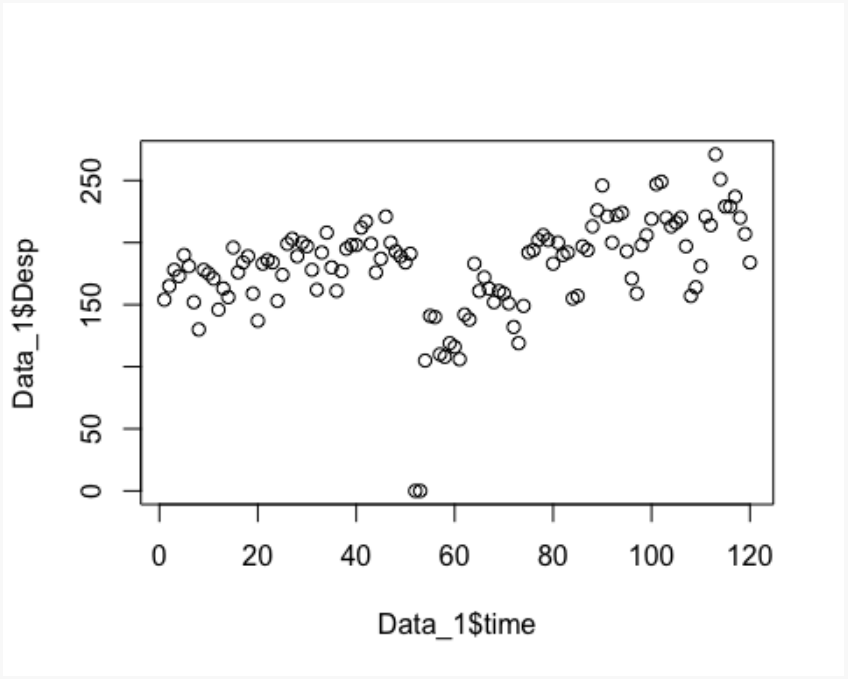
```
ModelAC2<-lm(Data_1$Desp~Data_1$time+Data_1$dm2+Data_1$dm3+Data_1$dm4+Data_1$dm5
+Data_1$dm6+Data_1$dm7+Data_1$dm8+Data_1$dm9+
          Data_1$dm10+Data_1$dm11+Data_1$dm12)

stargazer(ModelAC1,ModelAC2, type="text")

##
## ==============================================================
##                          Dependent variable:
##                        ----------------------------
##                                 Desp
##                           (1)          (2)
## --------------------------------------------------------------
## time                      0.401***     0.401***
##                          (0.099)      (0.099)
##
## dm2                        19.999       19.999
##                          (16.684)     (16.684)
##
## dm3                       34.697**     34.697**
##                          (16.685)     (16.685)
##
## dm4                        18.496       18.496
##                          (16.686)     (16.686)
##
## dm5                       31.495*      31.495*
##                          (16.688)     (16.688)
##
## dm6                       43.093**     43.093**
##                          (16.691)     (16.691)
##
## dm7                       27.792*      27.792*
##                          (16.694)     (16.694)
##
## dm8                        13.191       13.191
```

```
##                             (16.698)      (16.698)
##
## dm9                          29.190*       29.190*
##                             (16.702)      (16.702)
##
## dm10                         31.288*       31.288*
##                             (16.707)      (16.707)
##
## dm11                         19.187        19.187
##                             (16.713)      (16.713)
##
## dm12                         -3.814        -3.814
##                             (16.719)      (16.719)
##
## Constant                    134.128***    134.128***
##                             (12.988)      (12.988)
##
## ------------------------------------------------------------
## Observations                    120           120
## R2                            0.228         0.228
## Adjusted R2                   0.142         0.142
## Residual Std. Error (df = 107)  37.305        37.305
## F Statistic (df = 12; 107)     2.636***      2.636***
## ============================================================
## Note:                        *p<0.1; **p<0.05; ***p<0.01
```

```
summary(ModelAC2)
```

```
##
## Call:
## lm(formula = Data_1$Desp ~ Data_1$time + Data_1$dm2 + Data_1$dm3 +
##     Data_1$dm4 + Data_1$dm5 + Data_1$dm6 + Data_1$dm7 + Data_1$dm8 +
##     Data_1$dm9 + Data_1$dm10 + Data_1$dm11 + Data_1$dm12)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -186.892  -10.719    8.166   21.384   60.030
##
```

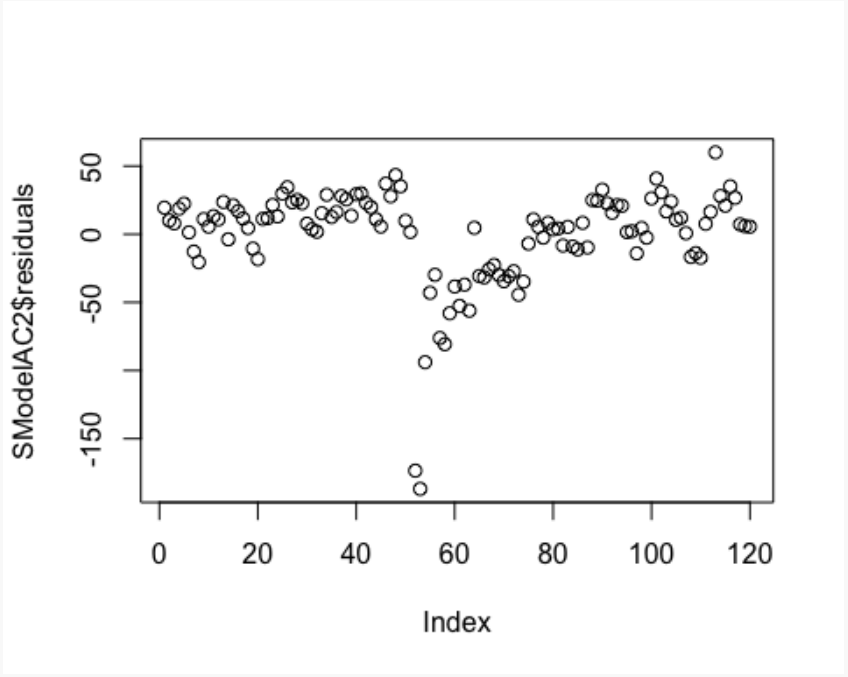
```
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 134.1282    12.9884  10.327  < 2e-16 ***
## Data_1$time   0.4013     0.0988   4.062 9.31e-05 ***
## Data_1$dm2   19.9987    16.6837   1.199   0.2333
## Data_1$dm3   34.6974    16.6846   2.080   0.0400 *
## Data_1$dm4   18.4961    16.6860   1.108   0.2701
## Data_1$dm5   31.4948    16.6881   1.887   0.0618 .
## Data_1$dm6   43.0935    16.6907   2.582   0.0112 *
## Data_1$dm7   27.7922    16.6939   1.665   0.0989 .
## Data_1$dm8   13.1909    16.6977   0.790   0.4313
## Data_1$dm9   29.1896    16.7021   1.748   0.0834 .
## Data_1$dm10  31.2883    16.7071   1.873   0.0638 .
## Data_1$dm11  19.1869    16.7126   1.148   0.2535
## Data_1$dm12  -3.8144    16.7188  -0.228   0.8200
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 37.31 on 107 degrees of freedom
## Multiple R-squared:  0.2282, Adjusted R-squared:  0.1416
## F-statistic: 2.636 on 12 and 107 DF,  p-value: 0.003981

SModelAC2<-summary(ModelAC2)


plot(Data_1$time,Data_1$Desp)
```




```
names(SModelAC2)
```

```
##  [1] "call"          "terms"         "residuals"     "coefficients"
##  [5] "aliased"       "sigma"         "df"            "r.squared"
##  [9] "adj.r.squared" "fstatistic"    "cov.unscaled"
```

```
plot(SModelAC2$residuals)
```




```
stargazer(ModelAC, ModelAC1,ModelAC2, type="text")
```

```
##
##
==========================================================================
===================
##                              Dependent variable:
##                 ---------------------------------------------------------------
##                                      Desp
##                      (1)              (2)              (3)
## ---------------------------------------------------------------------------------------
## time                 0.398***         0.401***         0.401***
##                     (0.072)          (0.099)          (0.099)
##
## dm2                   20.002           19.999           19.999
##                     (12.075)         (16.684)         (16.684)
##
## dm3                  41.458***        34.697**         34.697**
##                     (12.095)         (16.685)         (16.685)
##
## dm4                  25.259**         18.496           18.496
##                     (12.096)         (16.686)         (16.686)
##
## dm5                  38.261***        31.495*          31.495*
##                     (12.098)         (16.688)         (16.688)
##
## dm6                  49.863***        43.093**         43.093**
##                     (12.099)         (16.691)         (16.691)
##
## dm7                  34.564***        27.792*          27.792*
##                     (12.102)         (16.694)         (16.694)
##
## dm8                   19.966           13.191           13.191
##                     (12.105)         (16.698)         (16.698)
##
## dm9                  35.967***        29.190*          29.190*
##                     (12.108)         (16.702)         (16.702)
##
## dm10                 31.314**         31.288*          31.288*
```

```
##                    (12.092)          (16.707)          (16.707)
##
## dm11                 19.215            19.187            19.187
##                    (12.096)          (16.713)          (16.713)
##
## dm12                 -3.783            -3.814            -3.814
##                    (12.100)          (16.719)          (16.719)
##
## Covid               -67.548***
##                    (6.814)
##
## Constant             141.039***        134.128***        134.128***
##                    (9.426)           (12.988)          (12.988)
##
## ---------------------------------------------------------------------------------------
## Observations            120               120               120
## R2                    0.599             0.228             0.228
## Adjusted R2           0.550             0.142             0.142
## Residual Std. Error   27.000 (df = 106)     37.305 (df = 107)     37.305 (df = 107)
## F Statistic       12.204*** (df = 13; 106) 2.636*** (df = 12; 107) 2.636*** (df = 12; 107)
##
==========================================================================
==================
## Note:                                    *p<0.1; **p<0.05; ***p<0.01

taula <- data.frame(
  Periode = c("Pre-Covid", "Post-Covid"),
  Mitjana = c(mean(pre$Desp, na.rm=TRUE),
              mean(post$Desp, na.rm=TRUE)),
  Desviacio_tipica = c(sd(pre$Desp, na.rm=TRUE),
                       sd(post$Desp, na.rm=TRUE)),
  Minim = c(min(pre$Desp, na.rm=TRUE),
            min(post$Desp, na.rm=TRUE)),
  Maxim = c(max(pre$Desp, na.rm=TRUE),
            max(post$Desp, na.rm=TRUE)),
  Observacions = c(nrow(pre), nrow(post))
)
```

```
taula
```

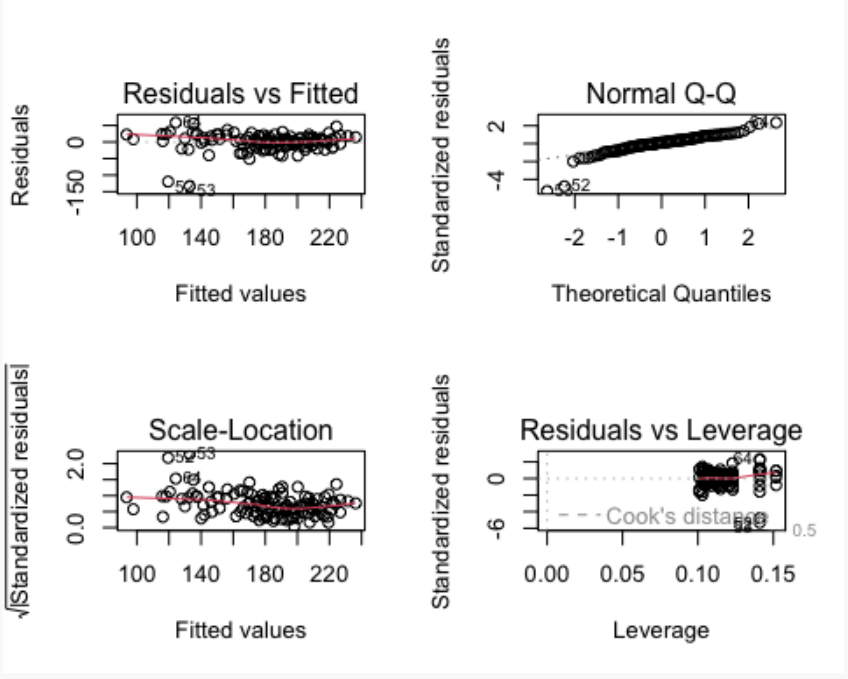
```
##      Periode  Mitjana Desviacio_tipica Minim Maxim Observacions
## 1  Pre-Covid 180.9600         19.81234   130   221           50
## 2 Post-Covid 199.9804         32.00093   119   271           51
```

```
#########################
# 10. RESULTATS
#########################
summary(ModelT3)
```

```
##
## Call:
## lm(formula = datos$Desp ~ datos$time + datos$dm2 + datos$dm3 +
##     datos$dm4 + datos$dm5 + datos$dm6 + datos$dm7 + datos$dm8 +
##     datos$dm9 + datos$dm10 + datos$dm11 + datos$dm12 + datos$Covid)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -132.871   -9.432    3.055   13.578   58.748
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) 141.03944    9.42637  14.962  < 2e-16 ***
## datos$time    0.39846    0.07151   5.572 1.93e-07 ***
## datos$dm2    20.00154   12.07508   1.656 0.100592
## datos$dm3    41.45791   12.09496   3.428 0.000868 ***
## datos$dm4    25.25945   12.09604   2.088 0.039172 *
## datos$dm5    38.26099   12.09753   3.163 0.002039 **
## datos$dm6    49.86253   12.09945   4.121 7.50e-05 ***
## datos$dm7    34.56407   12.10179   2.856 0.005161 **
## datos$dm8    19.96560   12.10455   1.649 0.102020
## datos$dm9    35.96714   12.10774   2.971 0.003677 **
## datos$dm10   31.31384   12.09201   2.590 0.010958 *
## datos$dm11   19.21538   12.09603   1.589 0.115136
## datos$dm12   -3.78308   12.10047  -0.313 0.755169
## datos$Covid -67.54836    6.81430  -9.913  < 2e-16 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 27 on 106 degrees of freedom
## Multiple R-squared:  0.5995, Adjusted R-squared:  0.5504
## F-statistic:  12.2 on 13 and 106 DF,  p-value: 8.176e-16
```

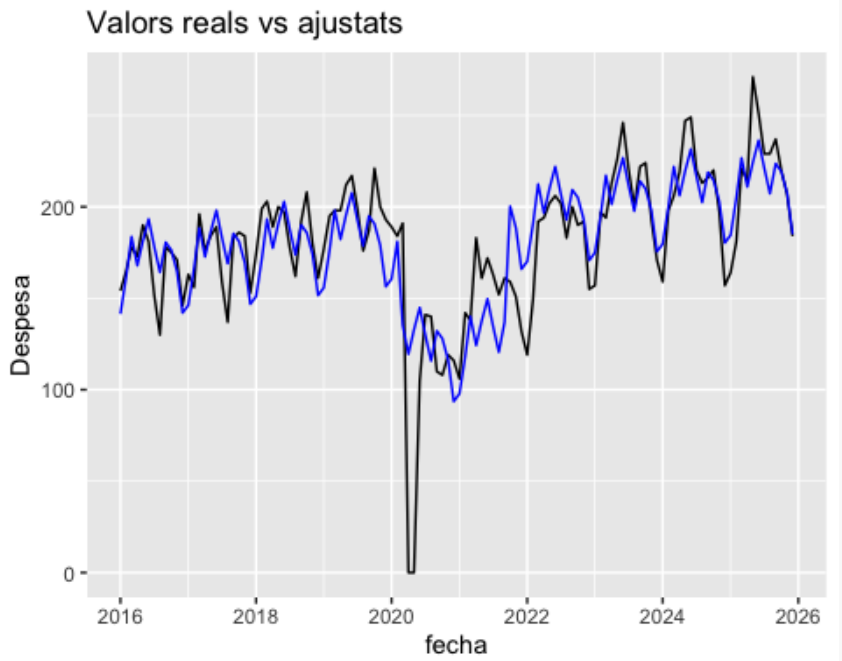
```
##########################
# 11. RESIDUS
##########################

par(mfrow=c(2,2))
plot(ModelT3)
```




```
# Autocorrelación residuos
dwtest(ModelT3)

##
##  Durbin-Watson test
##
## data:  ModelT3
## DW = 1.0344, p-value = 1.153e-07
## alternative hypothesis: true autocorrelation is greater than 0
```

```
# Heterocedasticitat
bptest(ModelT3)

##
##  studentized Breusch-Pagan test
##
## data:  ModelT3
## BP = 26.361, df = 13, p-value = 0.0152

########################
# 12. ERRORS ROBUSTOS
########################

coeftest(ModelT3, vcov = vcovHC(ModelT3, type="HC1"))

##
## t test of coefficients:
##
##               Estimate Std. Error t value  Pr(>|t|)
## (Intercept) 141.039439   8.490209 16.6120 < 2.2e-16 ***
## datos$time    0.398462   0.045325  8.7913 2.922e-14 ***
## datos$dm2    20.001538  10.743074  1.8618 0.0654003 .
## datos$dm3    41.457912  10.846472  3.8222 0.0002234 ***
## datos$dm4    25.259451  16.082343  1.5706 0.1192484
## datos$dm5    38.260989  17.195668  2.2250 0.0281972 *
## datos$dm6    49.862527  10.405438  4.7920 5.403e-06 ***
## datos$dm7    34.564065   9.913240  3.4867 0.0007130 ***
## datos$dm8    19.965603  11.041809  1.8082 0.0734136 .
## datos$dm9    35.967142   9.250173  3.8883 0.0001764 ***
## datos$dm10   31.313844  10.722775  2.9203 0.0042726 **
## datos$dm11   19.215382   9.742540  1.9723 0.0511782 .
## datos$dm12   -3.783079  10.701685 -0.3535 0.7244134
## datos$Covid -67.548359  11.837921 -5.7061 1.059e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

########################
# 13. VALORS AJUSTATS
########################
```

```r
datos$ajustada <- fitted(ModelT3)

ggplot(datos, aes(fecha)) +
 geom_line(aes(y=Desp), color="black") +
 geom_line(aes(y=ajustada), color="blue") +
 labs(title="Valors reals vs ajustats",
     y="Despesa")
```




```r
##########################
# 14. SUBPERIÒDES
##########################

# PRECOVID
pre <- subset(datos, fecha < as.Date("2020-03-01"))

ModelPRE <- lm(Desp ~ time + mes_factor, data=pre)

# POSTCOVID / RECUPERACIÓN
post <- subset(datos, fecha > as.Date("2021-09-30"))

ModelPOST <- lm(Desp ~ time + mes_factor, data=post)
```

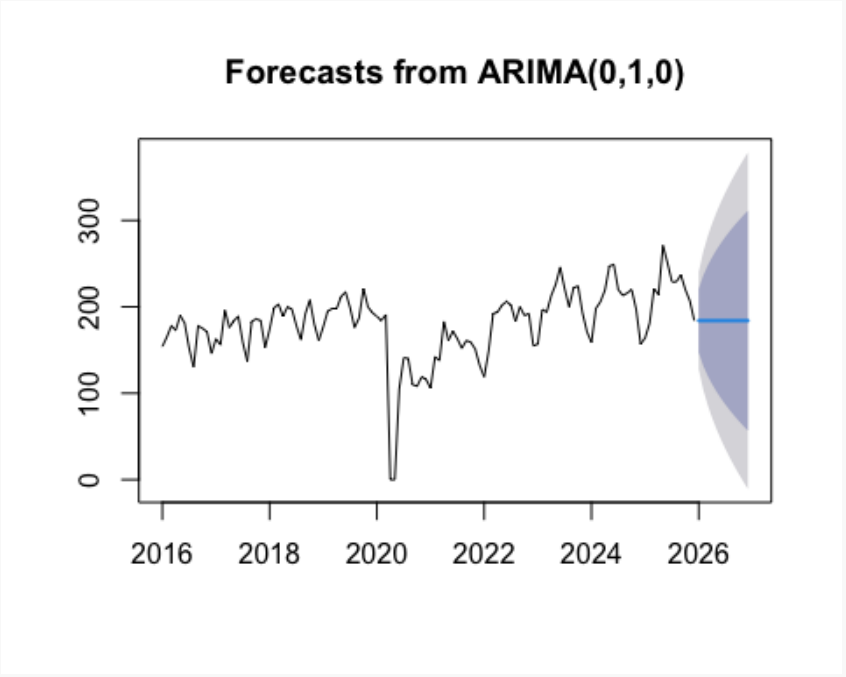
```r
stargazer(ModelPRE, ModelPOST,
        type="text",
        title="Comparación Pre y Post Covid")

##
## Comparación Pre y Post Covid
## ==========================================================================
##                              Dependent variable:
##                    ---------------------------------------------
##                                    Desp
##                          (1)                  (2)
## --------------------------------------------------------------------
## time                    0.849***             1.028***
##                        (0.084)              (0.111)
##
## mes_factor2               7.551              30.472***
##                        (5.352)              (8.141)
##
## mes_factor3              25.745***            51.443***
##                        (5.686)              (8.143)
##
## mes_factor4              15.146**             57.165***
##                        (5.681)              (8.147)
##
## mes_factor5              26.798***            82.637***
##                        (5.678)              (8.152)
##
## mes_factor6              25.449***            83.108***
##                        (5.676)              (8.159)
##
## mes_factor7               0.600              62.080***
##                        (5.676)              (8.167)
##
## mes_factor8             -20.999***            49.302***
##                        (5.676)              (8.177)
##
## mes_factor9              11.902**             60.773***
```

```
##                         (5.678)              (8.188)
##
## mes_factor10             23.554***            49.765***
##                         (5.681)              (7.730)
##
## mes_factor11              8.955               34.137***
##                         (5.686)              (7.735)
##
## mes_factor12             -12.394**             4.908
##                         (5.691)              (7.742)
##
## Constant                150.180***           56.172***
##                         (4.331)             (11.608)
##
## ------------------------------------------------------------------
## Observations               50                   51
## R2                       0.862                0.902
## Adjusted R2               0.818                0.871
## Residual Std. Error     8.461 (df = 37)      11.512 (df = 38)
## F Statistic         19.308*** (df = 12; 37) 29.030*** (df = 12; 38)
## ==================================================================
## Note:                              *p<0.1; **p<0.05; ***p<0.01
```

```
#########################
# 15. MODEL ARIMA
#########################

modelo_arima <- auto.arima(serie_ts)

summary(modelo_arima)

## Series: serie_ts
## ARIMA(0,1,0)
##
## sigma^2 = 822.6:  log likelihood = -568.24
## AIC=1138.49   AICc=1138.52   BIC=1141.27
##
## Training set error measures:
```

```
##                     ME     RMSE      MAE  MPE MAPE      MASE      ACF1
## Training set 0.2512833 28.56076 18.51795 -Inf  Inf 0.6461837 0.0082379
```

```
forecast_arima <- forecast(modelo_arima, h=12)


plot(forecast_arima)
```




```
############################################################
# FI
############################################################
```