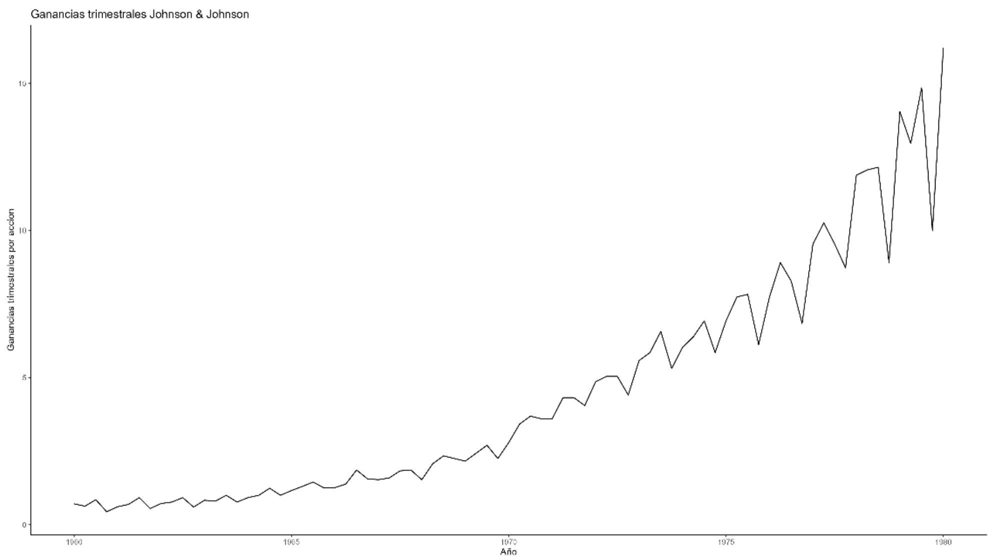
Profesora magistral: Juliana Helo
Profesores complementarios: Joaquín Rueda, Luis Felipe González, Carlos Guzmán y Jhon Vela Salcedo
Monitores: Camila Robayo, Juan José Gutiérrez, David Flórez y Nicolle Orjuela
Taller 5: contenido práctico
Miembros: José Ricardo Ricardo Hernández – 202113889
Andrés Serrano – 202116783
Laura Rodríguez – 202110325

1. Parte I:
Busquen una serie de tiempo que les gustaría analizar. Describan el nombre de la serie, la fuente de la cual la obtienen, su periodicidad, las fechas para las cuales tienen datos y una pequeña motivación para analizar esta serie de tiempo. Carguen los datos a R y guárdenlos como un objeto ts.
Nota: No es permitido usar las series del precio del petróleo BRENT o del par USD/COP, ya que fueron utilizadas en la clase complementaria. La periodicidad de la serie utilizada debe ser semanal, mensual o trimestral (puede descargar una serie en otra periodicidad y hacer una conversión a una periodicidad permitida, pero debe explicarla en su documento, e.g., una serie diaria se puede pasar a mensual con el promedio de los datos o el dato de cierre). Finalmente, independientemente de su periodicidad, utilice un mínimo de 3 años y un máximo de 6 años para sus datos.
La serie de tiempo que elegimos para analizar incluye los datos de los ingresos trimestrales de la empresa Johnson & Johnson. Esta serie fue extraída al utilizar la librería astsa (Time series and its aplications). Esta serie comienza desde el primer trimestre de 1960, hasta el último trimestre de 1980, por lo que hay 84 trimestres en total.
La motivación para analizar esta serie recae en que queremos evaluar las ganancias de la empresa Johnson & Johnson para así determinar el ritmo de crecimiento e identificar si la serie de tiempo es estacionaria, dado que en caso de que lo sea, se podría concluir que los ingresos que se tienen en la empresa se dan de manera constante.

2. Grafique su serie de tiempo y describa los comportamientos que observe en esta.
La descripción no debe ser solo visual, sino que debe complementarse con fuentes que otorguen un contexto a las descripciones.



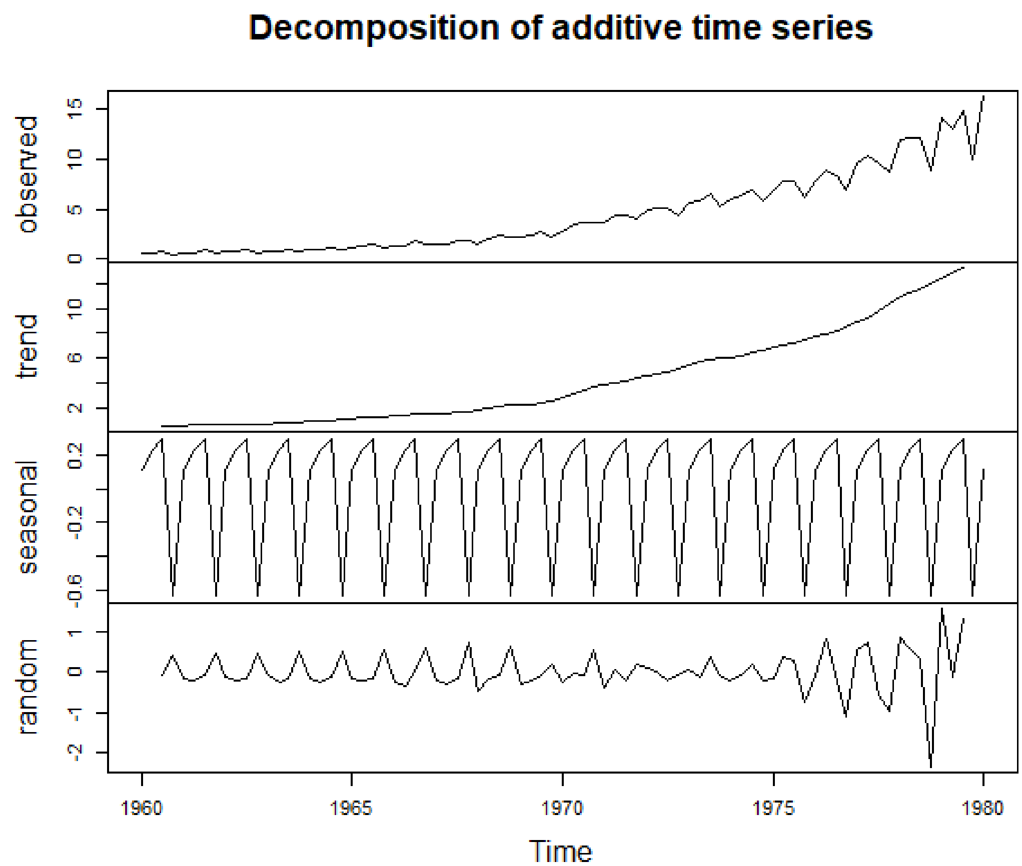
Gráfica 1. Gráfica de la serie de tiempo donde se la relación entre el año y la ganancia trimestral por acción.

La gráfica muestra la serie de tiempo de las ganancias trimestrales de la empresa Jonhson & Johnson, donde se evidencian las ganancias trimestrales por acción desde el año 1960 a 1980. En esta gráfica es posible evidenciar que hay una tendencia creciente en las ganancias, lo cual se puede entender como un crecimiento sostenido de la empresa durante el periodo, donde también se encuentran algunas fluctuaciones estacionales que se repiten de manera regular en cada año.

Este crecimiento constante pudo ser motivado también por la expansión del sector farmacéutico que se dio entre los años 60 y 70, en donde esta empresa logró posicionarse como una de las principales empresas del sector. También se puede entender que influyó el gran crecimiento económico que tenía Estados Unidos durante los 60 y los 70, que, a pesar de experimentar algunas crisis, la empresa logró mantener el crecimiento constante que había tenido durante los años analizados.

En la gráfica también se puede evidenciar estacionalidad, donde se evidencian patrones continuos a lo largo de cada uno de los años analizados, lo que se puede relacionar con ciclos de ventas o factores estacionales que son directamente relacionados con el sector.

3. Ahora, describa los principales patrones de la serie de tiempo, de acuerdo con las definiciones vistas en clase: tendencia, estacionalidad y ciclos. Si no encuentra alguno de estos patrones, menciónelo explícitamente.



Gráfica 2. Patrones de la serie de tiempo: tendencia, estacionalidad y ciclos.

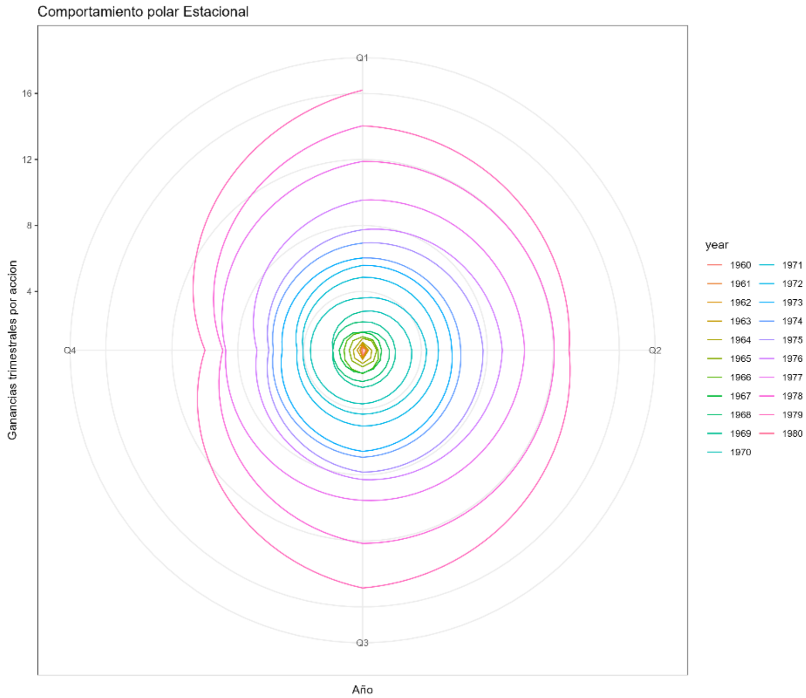
En la gráfica 2 es posible ver los patrones de la serie de tiempo, en donde se incluye la tendencia, la estacionalidad, los ciclos y el residual.

Al observar la tendencia se puede evidenciar que hay un crecimiento sostenido y acelerado a lo largo de los años analizados, y este se aumenta de manera significativa a partir del año 1970. De acuerdo con esto, se puede entender que adicional a las fluctuaciones estacionales de cada año, hay una mejora significativa en las utilidades de la empresa, las cuales pueden ser motivadas por un crecimiento en la empresa dado por expansión, innovación y mayor demanda de los productos farmacéuticos en los años establecidos, así como se mencionó en el punto anterior.

Para la estacionalidad, se puede observar en la serie que hay un patrón estacional evidente y repetitivo. En cada año, las ganancias tienen una fluctuación regular, donde también hay aumentos y caídas sistemáticas, lo que se puede entender como estacionalidad. Esto puede ocurrir por los ciclos de ventas que se tienen en cada uno de los años.

En cuanto a ciclos, no es posible observar ciclos económicos claros que vayan más allá de la estacionalidad dado que no se encuentran patrones cíclicos que sean regulares o prolongados.

4. Hagan un gráfico que les permita analizar la estacionalidad de la serie de tiempo y comenten lo que ven en este gráfico.



Gráfica 3. Comportamiento Polar Estacional

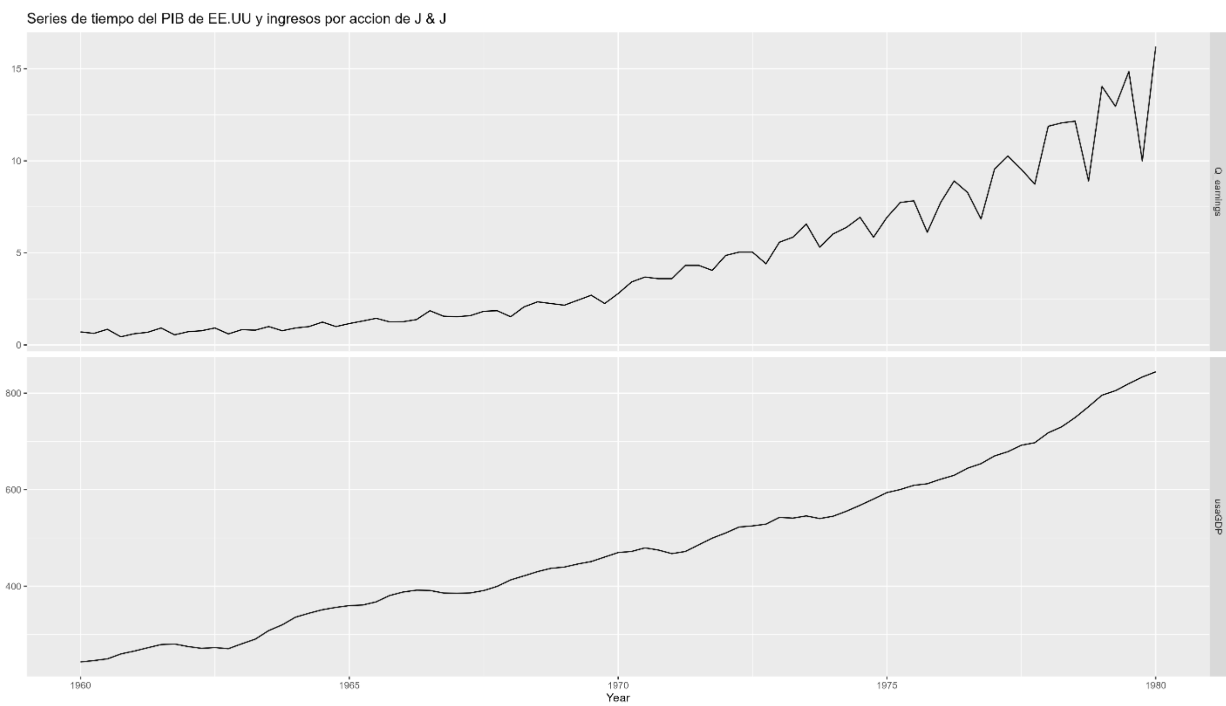
Con la gráfica mostrada, se puede observar un comportamiento estacional consistente en las ganancias trimestrales por acción entre los 60's y 80's. A lo largo de los años, es posible observar que hay un patrón repetitivo tanto de aumentos, como de caídas en algunos trimestres. Esto se puede evidenciar, por ejemplo, en las ganancias que tienen en los últimos trimestres del año (Q3 y Q4) las cuales tienden a aumentar.

También es posible observar que a medida que pasan los años, cada una de las curvas se va alejando del centro de forma progresiva, lo cual indica que hay un aumento continuo y sostenido de las ganancias en Johnson & Johnson. Adicionalmente, se puede observar que no hay algún tipo de ruptura del patrón dentro de las curvas y estas son simétricas la mayoría de los trimestres analizados, por lo que se entiende que hay estacionalidad en la serie de tiempo elegida.

5. Cargue otra serie de tiempo que tenga la misma periodicidad que su serie inicial. En este caso no hay restricciones (puede utilizar el precio del petróleo referencia Brent o el par USD/COP). Documente, de acuerdo con fuentes, por qué podría haber alguna correlación entre ambas series.

La segunda serie de tiempo utilizada corresponde al PIB de Estados Unidos, representado de manera trimestral, durante 1960 y 1980.
Como bien se sabe, un crecimiento en el PIB de un país indica un crecimiento económico general, incluyendo también el consumo de este. Por lo que se puede entender que, al haber un crecimiento, se dio también un alza en la demanda de productos y por ende mayor consumo. De acuerdo con un informe realizado por la OECD, un país que tiene un alto PIB, al tener un aumento de un punto porcentual en su crecimiento económico, experimenta un crecimiento de 0,75 puntos porcentuales en el sector de la salud, por lo que también puede ayudar a comprender la relación que se da entre las series de tiempo mostradas, entendiendo que a medida que crece una economía también aumenta el consumo en diferentes factores relevantes, como elementos médicos que provee J&J.

6. Grafiquen las dos series de tiempo seleccionadas en un gráfico de líneas y comenten lo que observan.

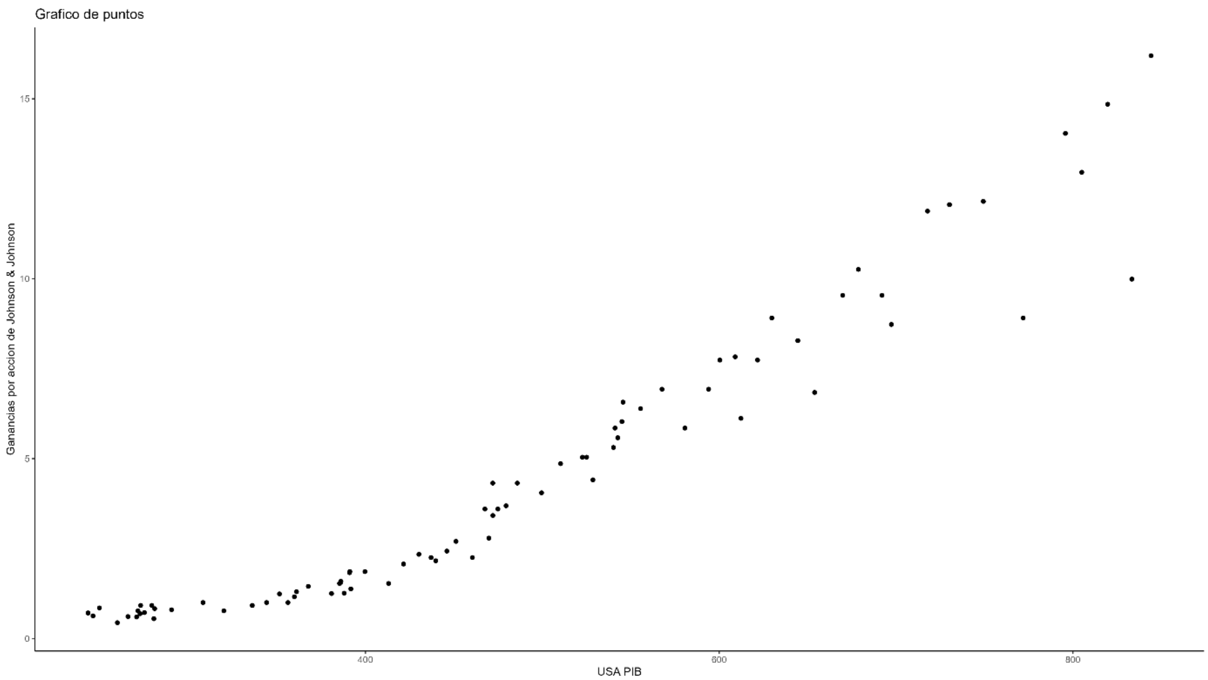


Gráfica 4. Serie de tiempo del PIB de Estados Unidos e ingresos por acción de Johnson y Johnson.

La gráfica 4 permite visualizar las dos series de tiempo que están siendo usadas entre 1960 y 1980, siendo las ganancias por acciones de J&J y el PIB de Estados Unidos. En esta gráfica es posible observar que ambas líneas muestran una tendencia ascendente entre los años mencionados, por lo que es posible mencionar que ambos factores crecieron de manera sostenida durante el periodo analizado. También es posible observar en la gráfica que hay una relación entre ambas variables, cuando el PIB crece, también lo hacen las ganancias de Johnson & Johnson, lo que se puede entender como una correlación positiva entre estos datos.

Por otro lado, es posible evidenciar que en la serie de tiempo que muestra las ganancias de J&J, hay fluctuaciones más marcadas, principalmente a partir de 1970. Al contrario de esto, el PIB es mucho más estable y mantiene un crecimiento mucho más estable, por lo que se puede entender que esta variable tiene estas variaciones dado a que es una serie macroeconómica y no un indicador empresarial.

7. Grafiquen las dos series de tiempo seleccionadas en un gráfico de puntos (scatter) y comenten lo que observan. Calcule el coeficiente de correlación y relacione el resultado con lo antes comentado.

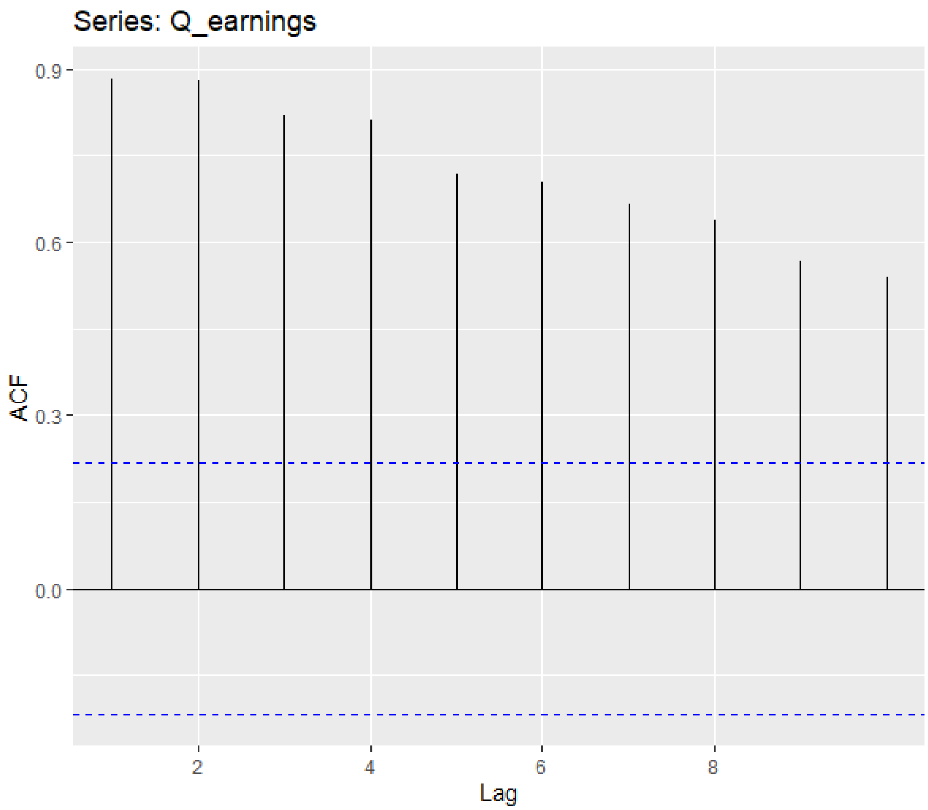


Gráfica 5. Serie de tiempo del PIB de Estados Unidos e ingresos por acción de Johnson y Johnson en un gráfico de puntos.

El gráfico 5 permite observar las dos series de tiempo elegidas en un gráfico de puntos. En este es posible observar que hay una relación positiva fuerte, dado que a medida que aumenta el PIB de estados Unidos, las ganancias por acción de Johnson y Johnson también aumentan. Así mismo, también es posible observar que no hay mucha dispersión en los datos, por lo que la relación que se evidencia entre las series no contiene gran cantidad de valores atípicos y en sí la mayoría son cercanos entre ellos mismos.

Por otro lado, el coeficiente de correlación calculado dio 0.9586944. Este coeficiente es muy cercano a 1, por lo que se entiende que hay una correlación positiva fuerte y significativa entre las variables de las ganancias por acciones de Johnson & Johnson y el PIB de Estados Unidos. Esto se puede argumentar con lo mencionado en puntos anteriores, donde un crecimiento económico general permite que haya más consumo e inversión en diferentes sectores, como lo es la salud, por lo que esto también influye en las ganancias recibidas por parte de privados como J&J.

8. Calcule y reporte los coeficientes de autocorrelación para un rezago de hasta 10 períodos. Luego, haga un gráfico de autocorrelación y relacione los resultados. Comente sobre la significancia de los coeficientes de autocorrelación.



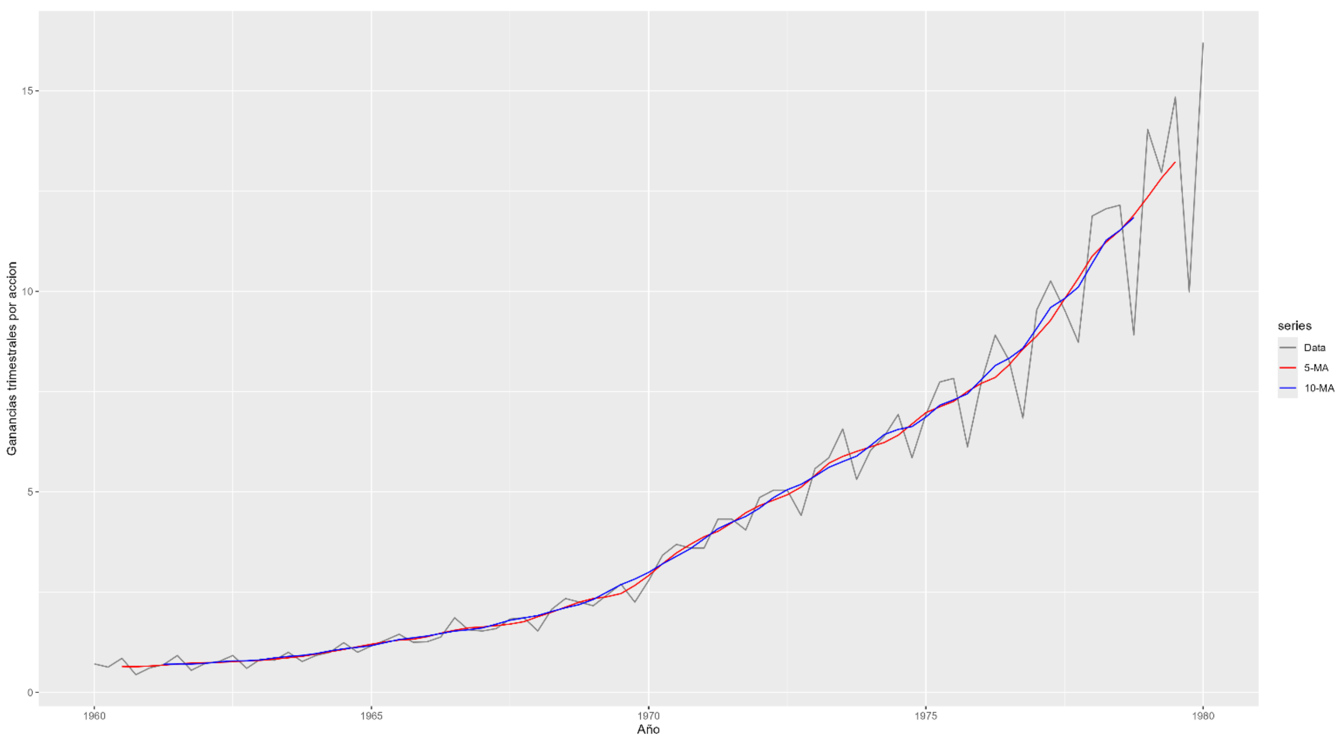
Gráfica 6. Gráfico de autocorrelación para un rezago de 10 periodos.

V1	V2
0	1
0,25	0,882692
0,5	0,879565
0,75	0,819073
1	0,812313
1,25	0,719366
1,5	0,703548
1,75	0,664954
2	0,638379
2,25	0,565995
2,5	0,538682

Tabla 1. Coeficientes de autocorrelación para un rezago de 10 periodos

Los coeficientes de correlación se representan en la Gráfica 6, en donde es posible evidenciar que todos los coeficientes están por encima del umbral de significancia, lo que indica que todas las correlaciones sin estadísticamente significativas. Adicionalmente, es posible identificar una dependencia temporal entre los valores actuales y los pasados, con coeficientes mayores a 0.5.
Los valores de la correlación son altos y decrecientes, lo que se puede entender como una estructura temporal que puede ser modelada por procesos autorregresivos. Esto es importante para construir modelos ARIMA ya que indica si los valores pasados funcionan como buenos predictores del futuro. En caso de que los coeficientes estuvieran dentro del intervalo de confianza, se podría afirmar que la serie es ruido blanco.

9. Muestre gráficamente el suavizamiento de la serie por promedio móvil. Pruebe distintos valores para el número de periodos considerados en el promedio. Explique con detalle qué se está calculando con la metodología y comente los resultados observados.

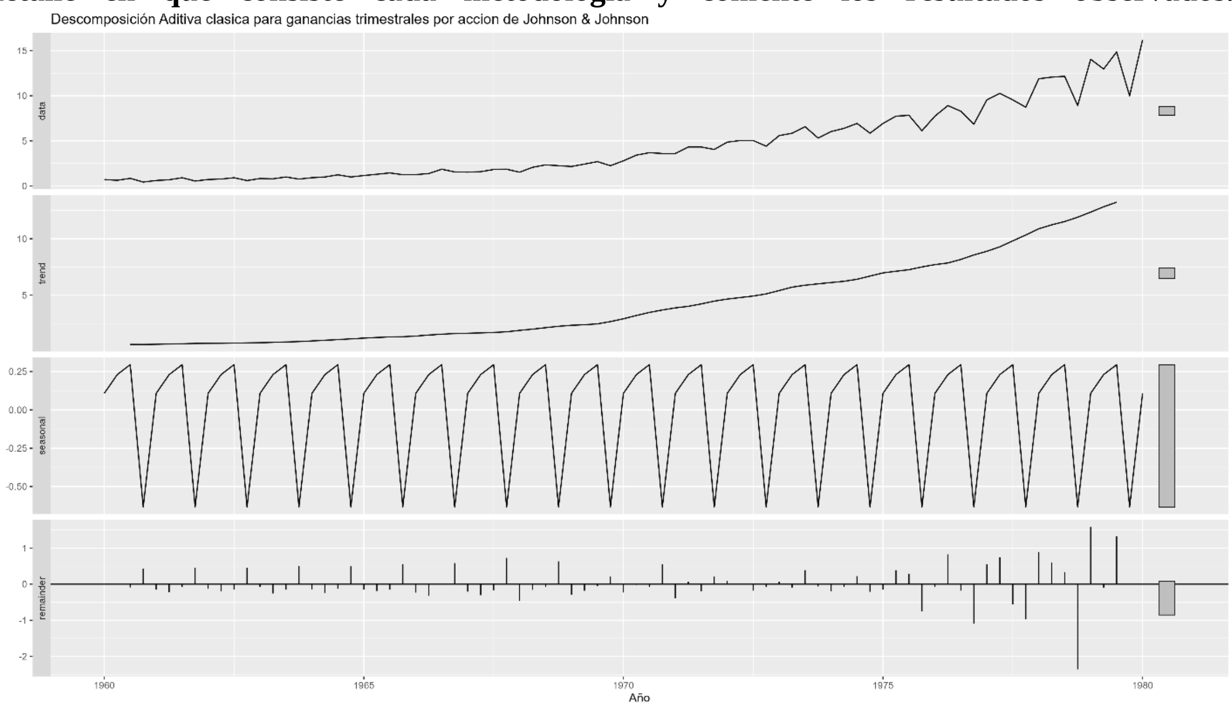


Gráfica 7. Suavizamiento de la serie por promedio móvil.

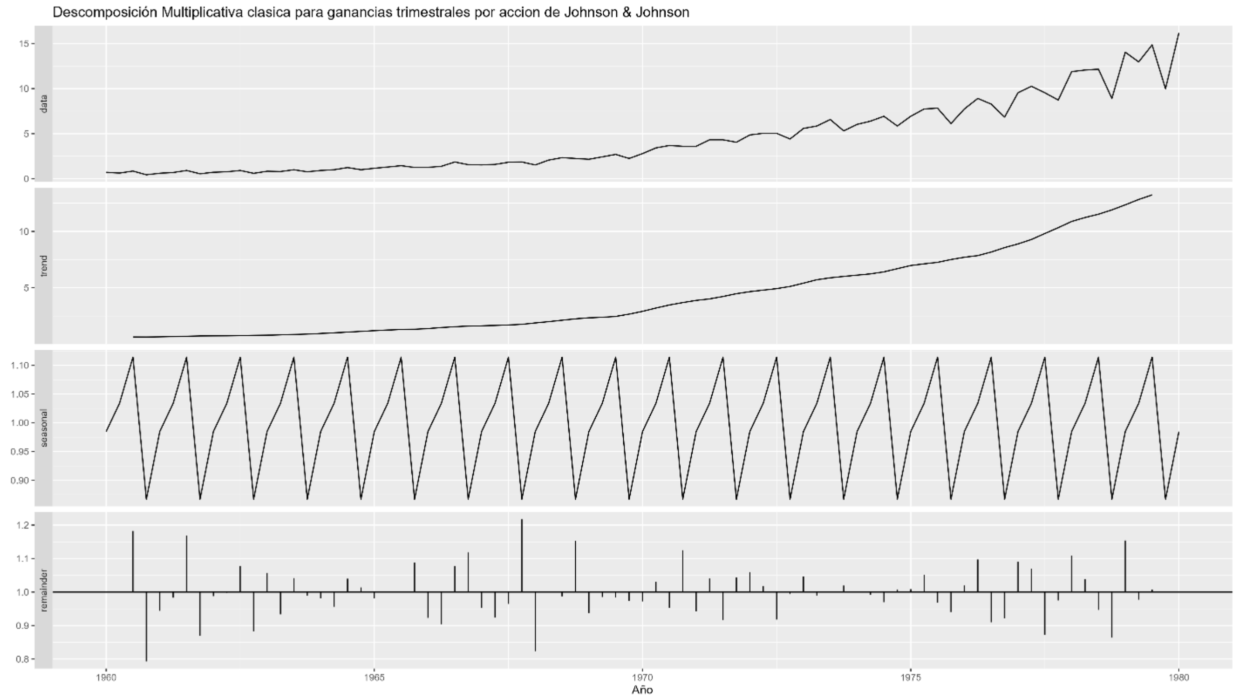
En la gráfica 7 es posible evidenciar el suavizamiento de la serie por promedio móvil esto permite que se reduzca la volatilidad de una serie de tiempo. En este caso, se usa esta metodología con un promedio móvil de 5 periodos (5-MA) y un promedio móvil de 10 periodos (10-MA). Al usar esta metodología, se permite observar mejor la tendencia de una serie de tiempo eliminando fluctuaciones de corto plazo, esto se hace al sustituir cada observación por el promedio de un número fijo de valores consecutivos anteriores, eliminando así el "ruido" que se tiene en corto plazo.

En la gráfica es posible evidenciar la serie original, la cual muestra las ganancias por acción con todas las variaciones, incluido el ruido que se pretende eliminar. Para la línea de 5 periodos, es posible evidenciar que se logra suavizar la serie de manera moderada, esto dado que se ve cierta variabilidad a lo largo de los años. Por otro lado, para la línea de 10 periodos se puede observar que se logra suavizar mucho más la serie, por lo que hay una tendencia más estable y se entiende que se eliminan casi todas las fluctuaciones que hay en corto plazo.

10. Muestre gráficamente el suavizamiento de la serie por el método de descomposición clásica. Realice tanto una descomposición aditiva como una multiplicativa. Explique con detalle en qué consiste cada metodología y comente los resultados observados.



Gráfica 8. Suavizamiento de la serie por medio de descomposición aditiva.



Gráfica 9. Suavizamiento de la serie por medio de descomposición multiplicativa.

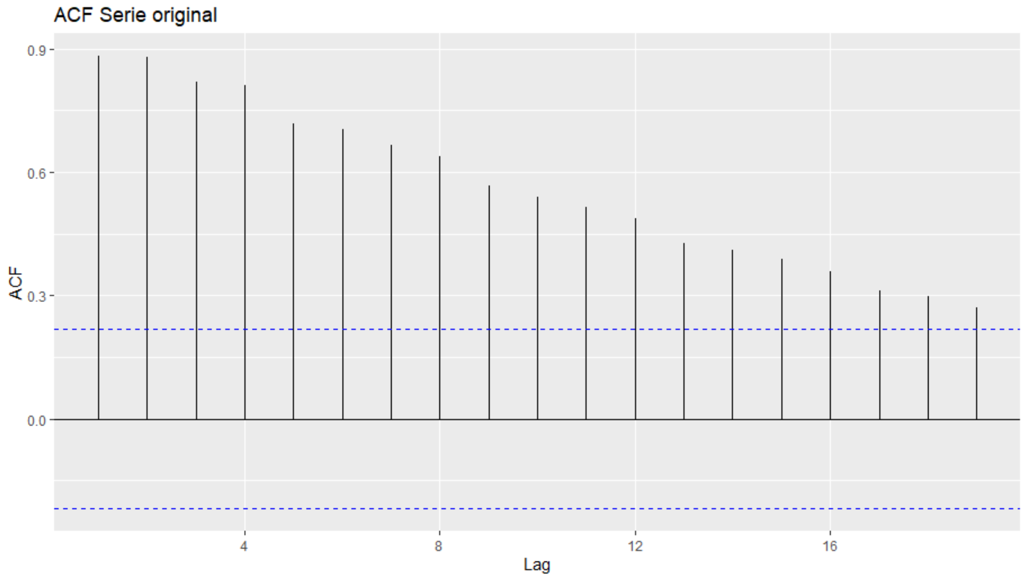
La suavización de la serie por medio de descomposición clásica consiste en separar la serie en tendencia (T), estacionalidad (S) y ruido (R), la cual permite tener una mejor comprensión de los patrones y previene posibles tendencias futuras. La descomposición aditiva se utiliza cuando los componentes son independientes del nivel de la serie. En este caso, la serie está basada en el modelo Yt=Tt+St+Rt, en donde los componentes se suman, por lo que es útil cuando las fluctuaciones estacionales y el ruido tienen una magnitud que es constante en el tiempo. En cuanto a la descomposición multiplicativa, se usa cuando las fluctuaciones estacionales y el ruido dependen del nivel de la serie. Este método se basa en el modelo Yt=Tt×St×Et, lo cual permite que este sea adecuado cuando las variaciones crecen o disminuyen proporcionalmente con la magnitud de la serie.

En la gráfica 8 se evidencian los resultados usando el método de descomposición aditiva. En este gráfico es posible determinar que la tendencia muestra un crecimiento sostenido en las ganancias de Johnson & Johnson. Por otro lado, la estacionalidad tiene una forma regular, por lo que se entiende que hay un patrón repetitivo similar en cada uno de los años analizados. En cuanto al ruido, estos son más pequeños y oscilan alrededor de cero, lo que demuestra que la descomposición captura de manera adecuada la estructura de la serie.

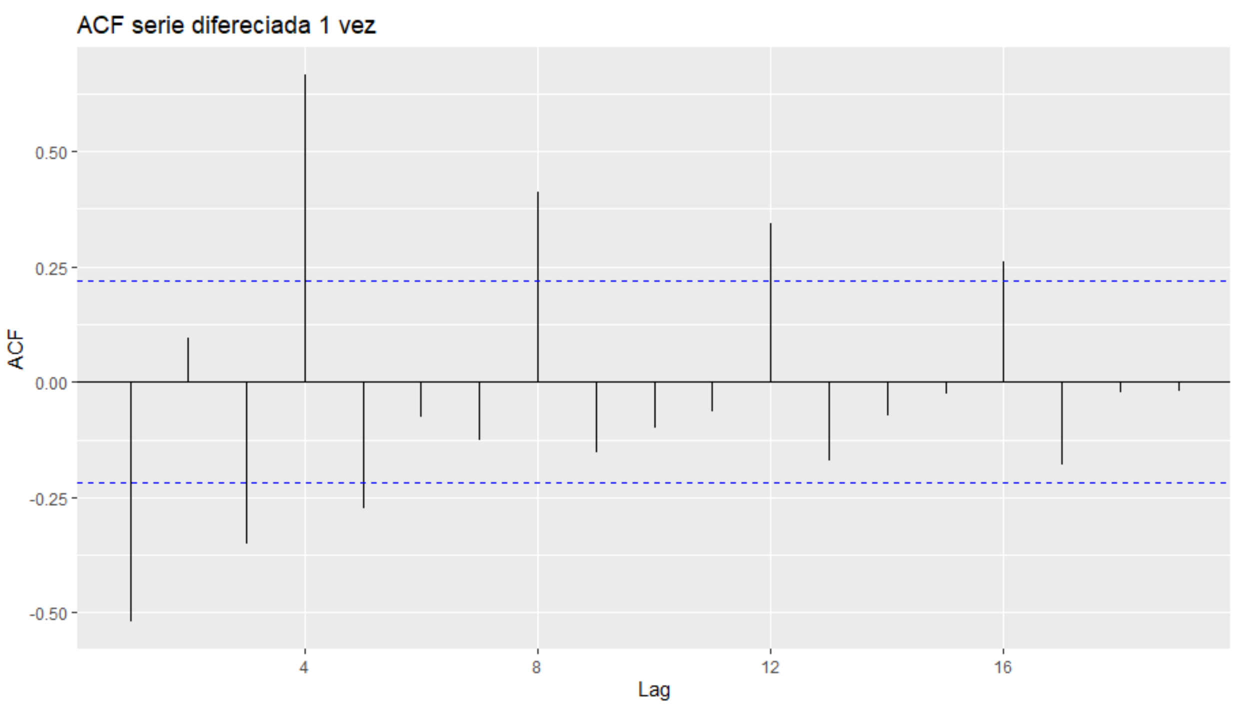
En la gráfica 9 es posible observar los resultados usando el método de descomposición multiplicativa. En esta, la tendencia muestra también un crecimiento sostenido de las ganancias de la empresa. En cuanto a la estacionalidad, esta oscila alrededor de 1 y su amplitud parece aumentar con el nivel de la serie. En cuanto a los residuales, estos son proporcionales al nivel de la serie.

2. Parte II

11. Muestre un gráfico de la función de autocorrelación (FAC) para su serie original y para su serie diferenciada una vez. Realice un test con el estadístico de Ljung-Box para su serie original y para su serie diferenciada una vez. Explique y comente los resultados.



Gráfica 10. Función de autocorrelación para la serie original.



Gráfica 11. Función de autocorrelación para la serie diferenciada 1 vez.

```
	Box-Ljung test

data:  Q_earnings
X-squared = 65.477, df = 1, p-value =
5.551e-16
```

Imagen 1. Test Box-Ljung para serie original.

```
	Box-Ljung test

data:  diff_Q_earnings
X-squared = 22.407, df = 1, p-value =
2.205e-06
```

Imagen 2. Test Box-Ljung para serie Diferenciada 1 vez.

De acuerdo con la gráfica 10, es posible observar la autocorrelación para la serie original. En esta se evidencia que hay una autocorrelación positiva y significativa para diferentes rezagos a lo largo del periodo analizado, lo que puede indicar que hay una dependencia temporal y por ende la serie no es estacionaria. Es posible ver también que la autocorrelación disminuye de forma lenta, siendo algo característico en una serie con tendencia. Por otro lado, la gráfica 11 permite ver la autocorrelación de la serie al ser diferenciada una vez. En este caso la autocorrelación ha disminuido de manera considerable al compararla con la gráfica anterior. Adicional, solo algunos rezagos muestran valores levemente significativos, lo que se puede entender como un efecto de la diferenciación al ayudar a eliminar la dependencia temporal, acercando así a que la serie sea estacionaria.

En cuanto a la prueba de Ljung-Box, el resultado que se tiene con la serie original presenta un p-value pequeño, por lo que se rechaza la hipótesis nula que establece que no hay autocorrelación, lo cual indica que la serie originar presenta autocorrelación significativa y no es ruido blanco. En cuanto a los resultados para la serie siendo diferenciada una vez, se puede observar que el p-value sigue siendo pequeño, y menor al que se evidenció en la serie original, por lo que nuevamente se rechaza la hipótesis nula de que no hay autocorrelación. Este resultado permite entender que aún hay correlaciones existentes pero la dependencia se ha reducido.

12. Realice un test de raíz unitaria por medio de la prueba de Kwiatkowski-Phillips Schmidt-Shin (KPSS). Comente sus resultados. En caso de requerir realizar diferenciaciones, calcule cuántas deben realizar a partir del comando ndiffs().

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 0.9846

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Imagen 3. Test de raíz unitaria KPSS

De acuerdo con el resultado en la imagen 3, se puede observar que el estadístico de prueba equivale a 0.9846 y los valores críticos 0.347 (10pct), 0.463 (5pct), 0.574 (2.5 pct) y 0.739 (1 pct). En este caso se entiende que a hipótesis nula equivale a que la serie es estacionaria, por lo que no tiene raíz unitaria; y la hipótesis alternativa se entiende como que la serie no es estacionaria, por lo que tiene raíz unitaria.

Al observar el estadístico, se entiende que al ser este mayor que todos los valores críticos, se rechaza la hipótesis nula. Esto indica que la serie no es estacionaria y debe ser diferenciada. De acuerdo con el comando de ndiffs(), esta serie debe ser diferenciada dos veces para que sea estacionaria.

13. Realice y describa las transformaciones y diferenciaciones a su serie original y muestre, por medio del test de KPSS y del test de Ljung-Box, cuál cuenta con una serie estacionaria

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 0.2324

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

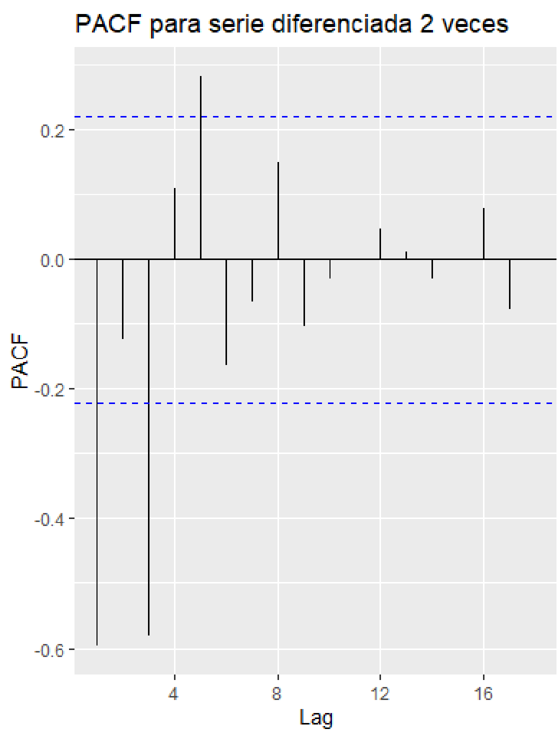
Imagen 4. Test de raíz unitaria para serie diferenciada 2 veces

```
        Box-Ljung test

data:  secondDiff
X-squared = 29.236, df = 1, p-value =
6.406e-08
```

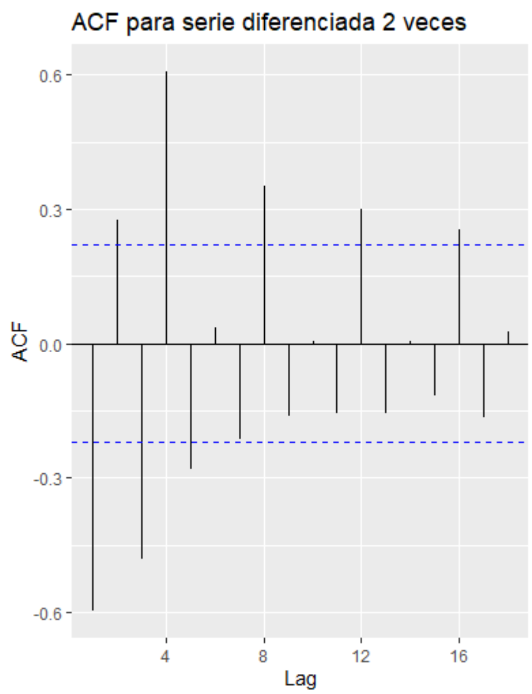
Imagen 5. Test Box-Ljung para serie diferenciada 2 veces

Los resultados indican que no hay evidencia estadísticamente significativa de que la serie sea estacionaria, por lo que la diferenciamos por segunda vez. Cuando se realiza esto, el resultado para el test de KPSS indica que el p-value es menor a los valores críticos, por lo que no se rechaza la hipótesis nula que indica que la serie es estacionaria. De acuerdo con este test, al diferenciar 2 veces la serie, esta se vuelve estacionaria.
En cuanto al test de Ljung-Box, el p-valor que muestra es muy pequeño, por lo que se rechaza la hipótesis nula que establece que no hay autocorrelación, por lo que se puede entender que aún hay autocorrelación significativa a pesar de que la segunda diferenciación haga que la serie sea estacionaria.

14. Realicen los gráficos de la función de autocorrelación (ACF) y de la función de autocorrelación parcial (PACF). Comenten lo que observan en los gráficos y propongan al menos 3 posibles modelos para realizar los pronósticos de la serie (modelos ARIMA no estacionales). Comenten cuál modelo escogerían entre los propuestos y expliquen el criterio de selección.



Gráfica 12. Función de autocorrelación Parcial



Gráfica 13. Función de autocorrelación

```
#planteamiento de modelos
m1 <- arima(secondDiff, order = c(1,2,0))
m2 <- arima(secondDiff, order = c(1,1,0))
m3 <- arima(secondDiff, order = c(1,2,2))
m4 <- arima(secondDiff, order = c(2,2,5))
m5 <- arima(secondDiff, order = c(0,2,6))
m6 <- arima(secondDiff, order = c(2,2,0))
m7 <- arima(secondDiff, order = c(0,0,3))
m8 <- arima(secondDiff, order = c(0,1,0))
m9 <- arima(secondDiff, order = c(2,0,6))
```

Imagen 6. Planteamiento de modelos para realizar pronósticos de la serie.

En la gráfica 12 se puede observar la función de autocorrelación parcial para la serie diferenciada 2 veces. En el primer rezago se puede observar un valor significativamente distinto de cero, por lo que se entiende que hay un componente autorregresivo en la serie. A partir del segundo rezago, los valores de la función se encuentran dentro del intervalo de confianza, por lo que indica que no hay necesidad de incluir rezagos adicionales. De esta manera, el primer rezago es el único que tiene un efecto relevante en el valor actual de la serie.

En la gráfica 13, es posible ver la función de autocorrelación en donde la serie es diferenciada dos veces. En esta, el valor de la función muestra un valor alto en el primer rezago y luego se observa una caída rápida de estos valores pero siguen existiendo algunos valores ligeramente significativos hasta el tercer o cuarto rezago. Esto puede indicar que hay un componente de media móvil. Al haber una caída rápida y dentro del intervalo, se puede interpretar que no hay una autocorrelación persistente en los últimos rezagos, entendiendo que después de la segunda diferenciación la serie se volvió estacionaria.
En cuanto a los modelos, escogeríamos el M9. Gracias al PACF se sabe que se deben incluir 2 rezagos. También se sabe que la serie es estacionaria cuando es diferenciada 2 veces (d=2), pero como ya tenemos la serie diferenciada no es necesario incluir d=2 nuevamente. Adicional, dado el número de lags fuera del intervalo de confianza, escogemos que sea q=6.

15.Utilice la función auto.arima() y comente cuál es el modelo que selecciona esta función. Recuerde que busca un modelo ARIMA no estacional. Compare el resultado con los modelos propuestos en el literal anterior.

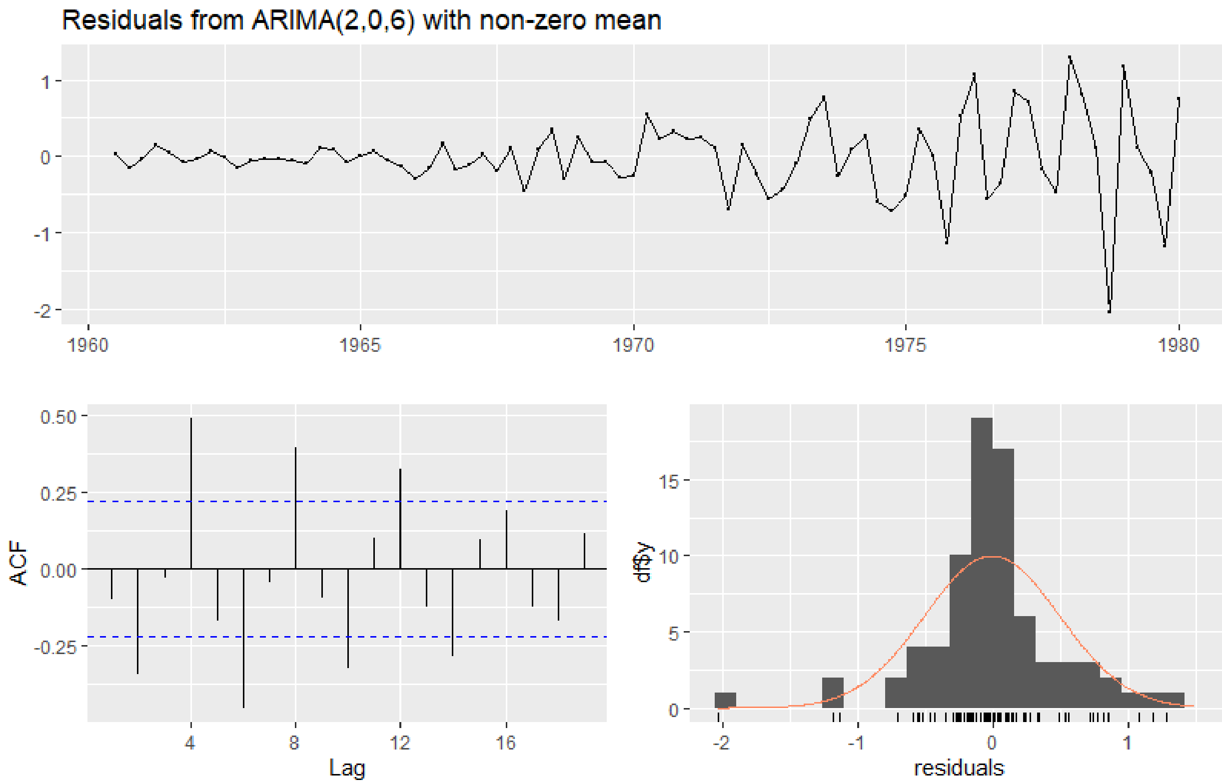
El modelo que selecciona la función auto.arima() es (0,0,3) y el modelo propuesto es de (2,0,6). El resultado de la función indica que no se aplicaron diferenciaciones, no hay ningún componente autorregresivo y se incorporaron 3 términos de media móvil. Este modelo sugiere que la serie ya era estacionaria en niveles, lo cual no concuerda con el análisis ya hecho anteriormente. El modelo que se propuso y se eligió anteriormente indica que no se debe hacer diferenciación dado que la serie ya estaba diferenciada dos veces, se deben incluir dos términos autorregresivos y 6 términos de media móvil. Este modelo escogido incluye más parámetros e incluye mayor riesgo de sobreajuste.

16. Tome el modelo que selecciona la función auto.arima() y alguno de sus modelos propuestos, que sea distinto al del auto.arima(). Compruebe que los residuales se comporten como ruido blanco y haga el test de raíz unitaria. Explique sus resultados y la importancia de que esto se cumpla para realizar el pronóstico.

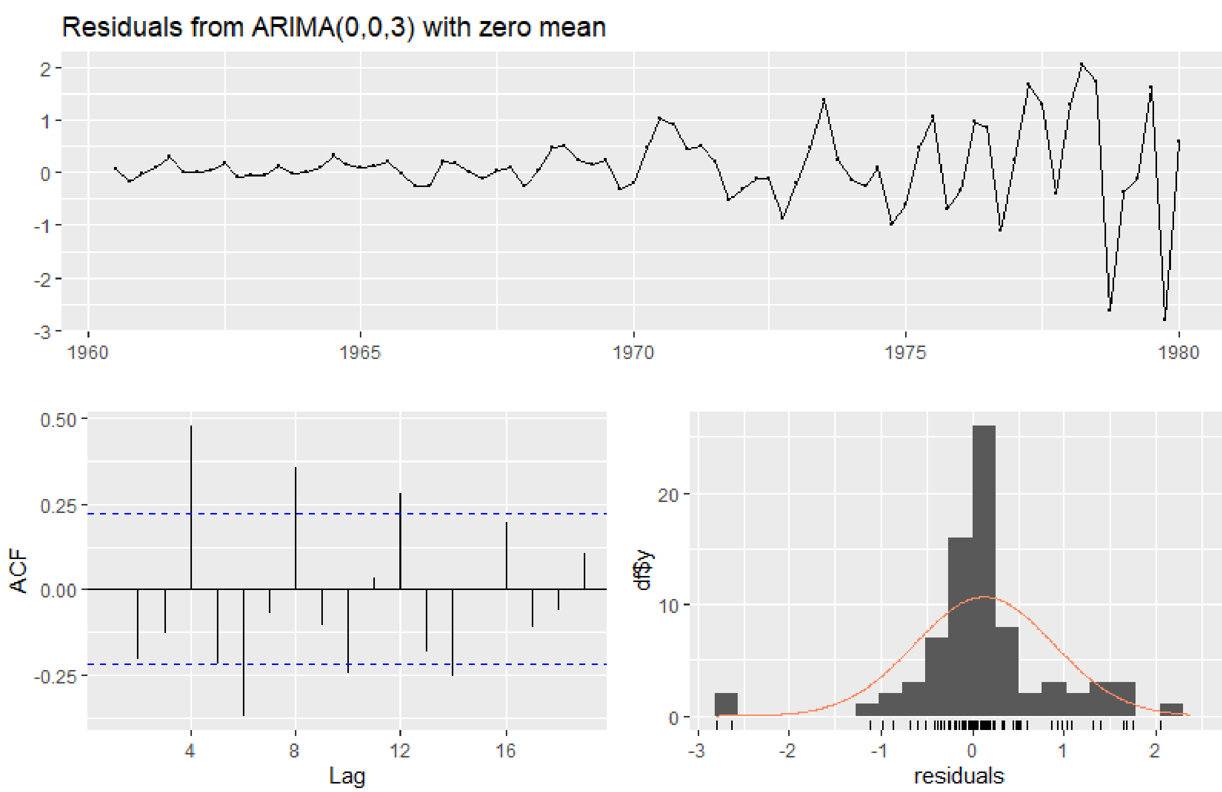
Modelo de auto.arima(0,0,3) y modelo propuesto (2,0,6)

```
#######################
# KPSS Unit Root Test #
#######################

Test is of type: mu with 3 lags.

Value of test-statistic is: 0.2324

Critical value for a significance level of:
                10pct  5pct 2.5pct  1pct
critical values 0.347 0.463  0.574 0.739
```

Imagen 7. Test de raíz unitaria



Gráfica 14. Residuales del modelo ARIMA (2,0,6)



Gráfica 15. Residuales del modelo ARIMA (0,0,3)

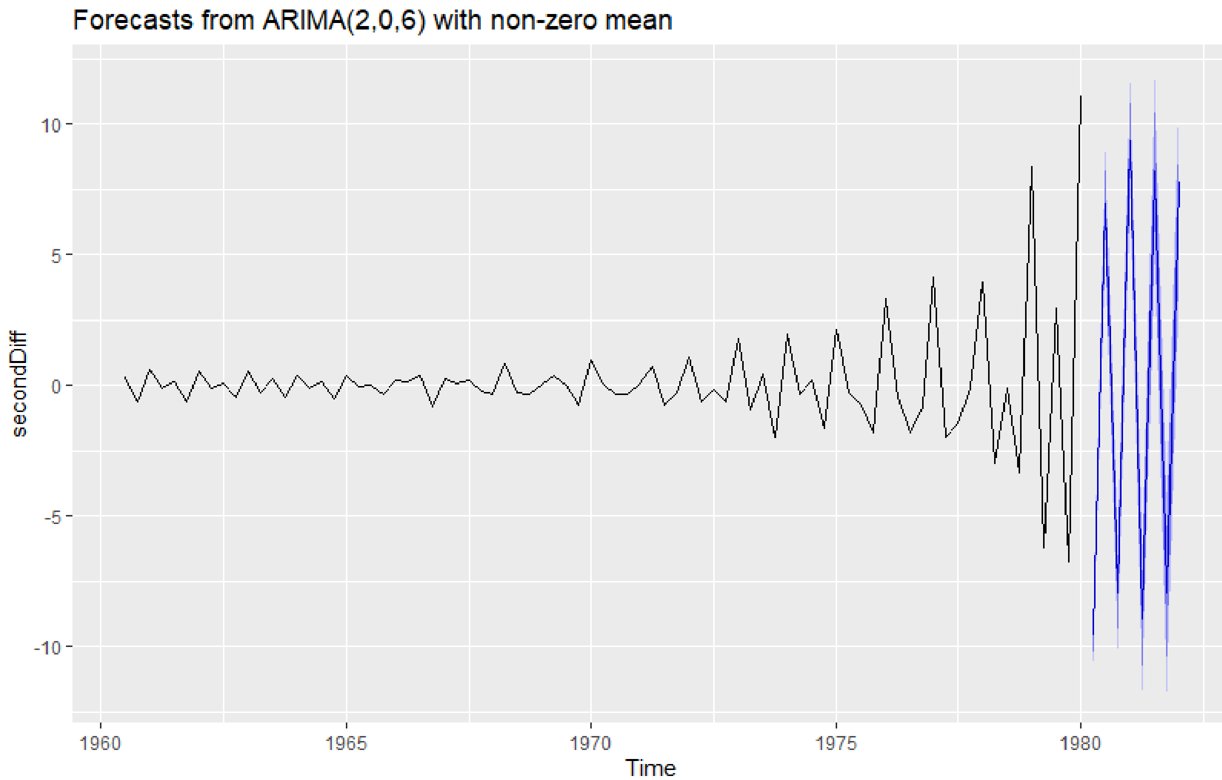
De acuerdo con el test de raíz unitaria, este tiene hipótesis nula de que la serie es estacionaria, yla alternativa indica que tiene raíz unitaria, que se entiende que la serie no es estacionaria. Al observar el p-.value, este es menor que los valores críticos, por lo que no se rechaza la hipótesis nula y se entiende que los residuos son estacionarios.

Al analizarlos residuales del modelos ARIMA (0,0,3) se evidencia que los residuos oscilan dentro de la media cero y no incluyen tendencias. Además, todos los rezagos están dentro del intervalo de confianza, indicando que no hay autocorrelación significativa. El histograma es simétrico, por lo que se asemeja a una distribución normal.

En cuanto al modelo ARIMA (2,0,6), los residuos también oscilan alrededor de la media pero tienen mayor variabilidad. Algunos de los rezagos están cerca del límite de significancia, por lo que se puede entender que hay una posible autocorrelación débil. El histograma es menos simétrico, por lo que se asume que los errores no se distribuyen de forma normal.

Al analizar ambos modelos, se evidencia que el modelo ARIMA (0,0,3) muestra residuos que se comportan más como ruido blanco, mientras que el otro modelo muestra cierta correlación y dispersión en sus errores. Es importante que los residuales de los modelos ARIMA sean estacionarios y se comporten como ruido blanco porque esto permite que se valide el hecho de que el modelo ha capturado toda la estructura temporal de la serie original, así mismo evita errores sistemáticos en los pronósticos y garantizan que no haya necesidad de incluir más términos autorregresivos o de media móvil al modelo.

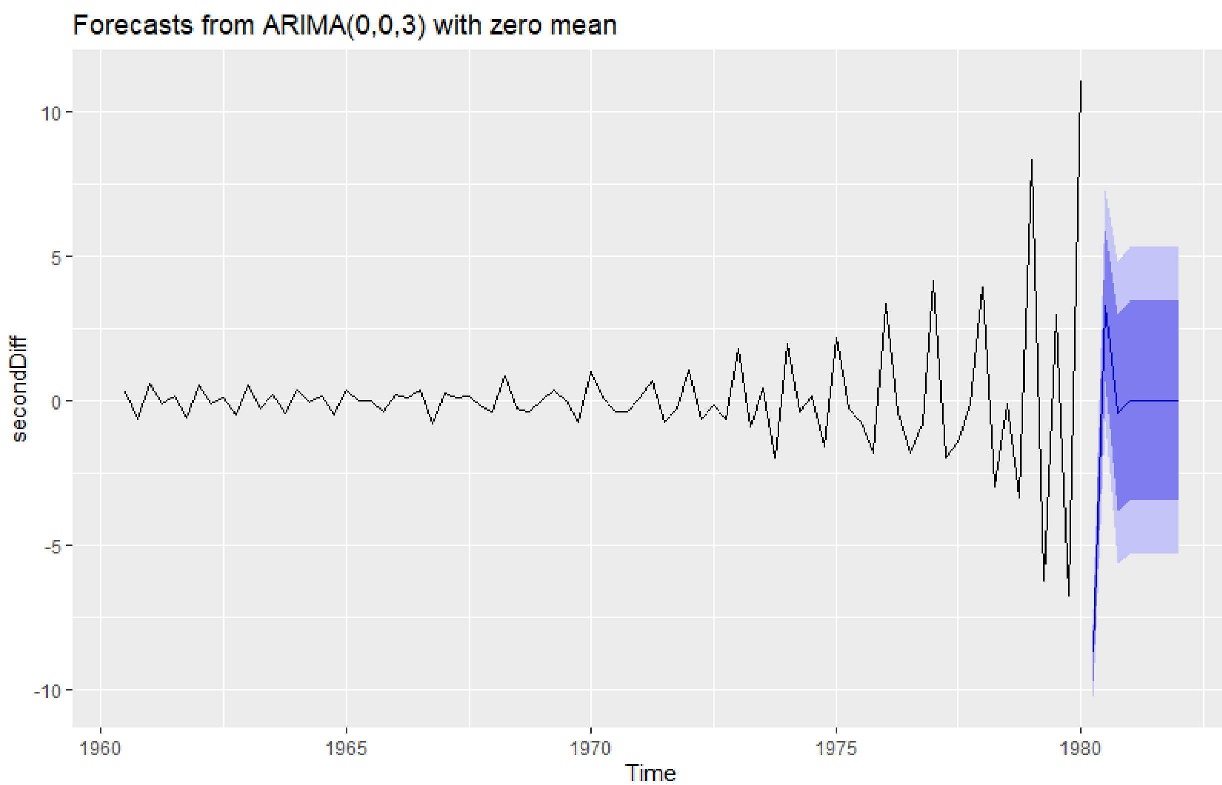
17. Realice un pronóstico con cada uno de los modelos seleccionados en el literal anterior (para un total de 2 pronósticos). Muestre sus resultados en una tabla, así como gráficamente, y compárelos. Si sus datos son trimestrales, pronostique 8 periodos; si son mensuales, pronostique 12 periodos; y si son semanales, pronostique 26 semanas.



Gráfica 16. Pronósticos del modelo ARIMA (2,0,6)

Point Foreca	Lo 90	Hi 90	Lo 95	Hi 95	Lo 99	Hi 99
-8.75365663	-10.0286368	-7.47867643	-10.2728892	-7.23442402	-10.7502667	-6.75704653
3.31603824	-0.03897451	6.67105098	-0.68170602	7.3137825	-1.93788831	8.56996479
-0.42682251	-4.7934022	3.93975718	-5.62992313	4.77627811	-7.26485583	6.41121082
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359

Tabla 2. Pronósticos del modelo ARIMA (2,0,6)



Gráfica 17. Pronósticos del modelo ARIMA (0,0,3)

Point Foreca	Lo 90	Hi 90	Lo 95	Hi 95	Lo 99	Hi 99
-8.75365663	-10.0286368	-7.47867643	-10.2728892	-7.23442402	-10.7502667	-6.75704653
3.31603824	-0.03897451	6.67105098	-0.68170602	7.3137825	-1.93788831	8.56996479
-0.42682251	-4.7934022	3.93975718	-5.62992313	4.77627811	-7.26485583	6.41121082
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359
0	-4.46271694	4.46271694	-5.31765522	5.31765522	-6.98858359	6.98858359

Tabla 3. Pronósticos del modelo ARIMA (0,0,3)

Al observar los pronósticos realizados para ambos modelos, se puede establecer que los primeros tres puntos pronosticados coinciden en ambos modelos, pero a partir del cuarto periodo se ve como ambos modelos tienden a cero, manteniendo los mismos intervalos de confianza.

El modelo ARIMA (2,0,6) muestra mayor volatilidad en los pronósticos, las bandas de confianza son más amplicas al inicio y el pronóstico se estabiliza en valores constantes después del tercer periodo. Para el modelo ARIMA (0,0,3) los intervalos de confianza se ven más estables y con menor dispersión, y ocurre algo similar que con el anterior modelo, donde tiende a convergir en 0. Por otro lado, en el primer modelo se evidencia que la media no es cero, mientras que en el segundo modelo la media es cero.

Anexo Codigo

```
#----------------- #PARTE 1#--------------------------------#


#limpiamos el ambiente
rm(list = ls())
cat("\014")
#cargamos las librerias
require(pacman)
p_load(astsa, tidyverse, forecast, ggplot2, tseries, rio, fpp2, lubricate, writexl)
# utilizamos la libreria astsa(Time series and its aplications).
# la cual tiene datos de los ingresos trimestrales de la empresa johnson and johnson
# La serie empieza desde el primer trimestre de 1960 hasta el ultimo trimestre de 1980
# Es decir, hay 84 trimestres en total.
# Motivacion: queremos evaluar las ganancias de la empresa  Johnson & Johnson
# para determinar el ritmo de crecimiento y saber si la serie es estacionaria.
# De ser estacionaria, se podria concluir que los ingresos de  Johnson & Johnson son constantes.



Q_earnings <- ts(jj, start = 1960, end = 1980, frequency = 4)


#2#


autoplot(Q_earnings)+
  ggtitle("Ganancias trimestrales Johnson & Johnson")+
  xlab("Año")+
  ylab("Ganancias trimestrales por accion")+
  theme_classic()
```

```
#guardamos

setwd("C:/Users/richa/OneDrive - Universidad de los Andes/Universidad/octavo/Econometria 2/talleres/talleres-econometria-2-en-r/taller5/Resultados/")

ggsave("Grafico1.png", plot  = last_plot(),width = 16, height = 9)



#Descomponemos en patrones

plot(decompose(Q_earnings))


#guardar

ggsave("Grafico2.png", plot  = last_plot(), width = 16, height = 9)



#4#


ggseasonplot(Q_earnings, polar = TRUE)+
  ggtitle("Comportamiento polar Estacional")+
  xlab("Año")+
  ylab("Ganancias trimestrales por accion")+
  theme_bw()
#
ggsave("Grafico4.png", plot  = last_plot(), width = 16, height = 9)



#5# cargar otra serie de tiempo
```

```
usaGDP <- ts(GDP, start = 1960, end = 1980, frequency =4)

# podria haber una correlacion entre el crecimiento economico general de estados\
#unidos con el alto y creciente sharevalue.
cor(Q_earnings,usaGDP)  = 0.9585944



#unimos las bases de datos
 jj_GDP <- cbind(Q_earnings, usaGDP)
#grafico de los 2
 autoplot(jj_GDP, facets = TRUE)+
   xlab("Year") + ylab("") +
   ggtitle("Series de tiempo del PIB de EE.UU y ingresos por accion de J & J")

 #guardamos
 ggsave("Grafico6.png", plot = last_plot(), width = 16, height = 9)

#7#

 as.data.frame(jj_GDP)%>%
   ggplot(aes(x = usaGDP, y = Q_earnings))+
   geom_point()+
   xlab("USA PIB") + ylab("Ganancias por accion de Johnson & Johnson") +
   ggtitle("Grafico de puntos")+
   theme_classic()
#guardamos
```

```
 ggsave("grafico de puntos.png", plot = last_plot(), width = 16, height = 9)

 #8#
 print(acf(Q_earnings, lag.max = 10, type = "correlation" ,plot = FALSE))
acfQE <- acf(Q_earnings, lag.max = 10, type = "correlation", plot = FALSE )
#guardamos
write_xlsx(as.data.frame(cbind(acfQE$lag, acfQE$acf)), "acf.xlsx")
#grafico
ggAcf(Q_earnings,lag.max = 10, type = "correlation")

#9#

Smooth_QEarnings5 <- ma(Q_earnings, order = 5)
Smooth_QEarnings10 <- ma(Q_earnings, order = 10)
#grafico
autoplot(Q_earnings, series = "Data")+
  autolayer(Smooth_QEarnings5, series = "5-MA")+
  autolayer(Smooth_QEarnings10, series = "10-MA")+
  xlab("Año")+
  ggtitle("Serie con Suavizamiento")+
  ylab("Ganancias trimestrales por accion")+
  scale_colour_manual( values = c("Data"="grey50","5-MA"="red", "10-MA"="blue"),
                       breaks = c("Data","5-MA", "10-MA"))
ggsave("Grafico9.png", plot=last_plot(), width = 16, height =9)

#10#
#aditiva
```

```
Q_earnings %>%
  decompose(type = "additive")%>%
  autoplot()+xlab("Año")+
  ggtitle("Descomposición Aditiva clasica para ganancias trimestrales por accion de Johnson & Johnson")


 ggsave("AdditiveDecomposition.png", plot = last_plot(), width = 16, height = 9)
#multiplicativa
 Q_earnings %>%
   decompose(type = "multiplicative")%>%
   autoplot()+xlab("Año")+
   ggtitle("Descomposición Multiplicativa clasica para ganancias trimestrales por accion de Johnson & Johnson"
 ggsave("MultiplicativeDecomposition.png", plot=last_plot(), width = 16,height = 9)


 #--------------- parte 2----------------------------#
 require(pacman)
 p_load(haven, tidyverse, forecast, ggplot2, tseries, readxl, writexl, dplyr,
        fable, fpp2, lubridate, mFilter, urca, astsa)
 #11#


 Q_earnings <- ts(jj, start = 1960, end = 1980, frequency = 4)


 diff_Q_earnings<- diff(Q_earnings)


 ggAcf(Q_earnings, main="ACF Serie original")
 ggAcf(diff_Q_earnings, main="ACF serie difereciada 1 vez")
```

```
 #test Ljung Box para serie original
 Box.test(Q_earnings, type= "Ljung-Box")

#test Ljung Box para serie diferenciada

 Box.test(diff_Q_earnings, type= "Ljung-Box")

 #12 y 13#
 #test de raiz unitaria
 diff_Q_earnings%>% ur.kpss()%>%summary()

 # los resultados indican que no hay evidencia estadisticamente significativa de
 # que la serie sea estacionaria
 # diferenciamos por segunda vez
 secondDiff <- diff(diff(Q_earnings))
secondDiff %>% ur.kpss() %>% summary()
# cuando diferenciamos 2 veces el resultado del test estadistico es menor a los
# valores criticos. Por lo tanto, la serie diferenciada 2 veces es estacionaria.
#verificamos estacionariedad para test Ljung-box
Box.test(secondDiff, type = "Ljung-Box")
# De acuerdo con el test, el p-valor menor a 0.01.Por lo cual, rechazamos la hipotesis nula.
# Entonces, existe una correlacion entre los datos. No son independientes.

#14 y 15#
ggAcf(secondDiff, main="ACF para serie diferenciada 2 veces")
ggPacf(secondDiff, main= "PACF para serie diferenciada 2 veces")
```

#planteamiento de modelos

m1 <- arima(secondDiff, order = c(1,2,0))

m2 <- arima(secondDiff, order = c(1,1,0))

m3 <- arima(secondDiff, order = c(1,2,2))

m4 <- arima(secondDiff, order = c(2,2,5))

m5 <- arima(secondDiff, order = c(0,2,6))

m6 <- arima(secondDiff, order = c(2,2,0))

m7 <- arima(secondDiff, order = c(0,0,3))

m8 <- arima(secondDiff, order = c(0,1,0))

m9 <- arima(secondDiff, order = c(2,0,6))

#Posible modelo que escogeriamos M4. Dado el pacf, sabemos que hay que incluir 2 rezagos

Sabemos que la serie estacionaria cuando diferenciamos 2 veces. Entonces, d=2 (ya la tenemos diferenciada 2 veces, no es necesario incluir d=2).

Por ultimo, dado el numero de lags fuera del intervalo de confianza, escogemos q=6.

#criterio de seleccion AIC(seleccionamos el AIC mas pequeno)

listAIC <- c(m1$aic,m2$aic,m3$aic,m4$aic,m5$aic,m6$aic, m7$aic, m8$aic, m9$aic)

which.min(listAIC) #= m9

fitauto<-auto.arima(secondDiff, seasonal = FALSE, stepwise = FALSE)

fitauto

#16 modelo de auto.arima(0,0,3) y modelo propuesto (2,0,6)

#comprobar que se comporten como ruido blanco y test de raiz unitaria

#(0,0,3):

checkresiduals(fitauto)

autoplot(fitauto)

```
#(2,0,6):

checkresiduals(m9)

autoplot(m9)+title("Test de raiz unitaria para ARIMA(2,0,6)")



#17#

#para el modelo (0,0,3)

autoplot(forecast(fitauto, h=8))

forecast(fitauto, h = 8, level = c(90, 95, 99))

write_xlsx(as.data.frame(forecast(fitauto,h = 8,level = c(90, 95, 99),... =
"resultadoss2(003).xlsx")))


#para el modelo (2,0,6)

autoplot(forecast(m9), h=8)

write_xlsx(as.data.frame(forecast(fitauto, h=8, level=c(90, 95, 99))), "resultados206.xlsx")
```

Referencias:

Aparicio Cabrera, A. (2014). *Historia económica mundial 1950–1990*. *Economía Informa*, (385), 70–83. https://doi.org/10.1016/S0185-0849(14)70420-7

Correa Martínez, M. A.(2021, 29 de septiembre). *Variables macroeconómicas que influyen sobre las empresas*. *El Economista*. https://www.eleconomista.com.mx/revistaimef/Variables-macroeconomicas-que-influyen-sobre-las-empresas-20210929-0054.html

Marino, A. et al. (2017), "Future trends in health care expenditure: A modelling framework for cross-country forecasts", *OECD Health Working Papers*, No. 95, OECD Publishing, Paris, https://doi.org/10.1787/247995bb-en.

Pharmaphorum. (s.f.). *A history of Johnson & Johnson*. Recuperado el 18 de mayo de 2025, de https://pharmaphorum.com/views-analysis-sales-marketing/a-history-of-johnson-johnson