



Universidad Nacional Autónoma de México  
Facultad de Ciencias

Modelos de Supervivencia y de Series de Tiempo

Proyecto Series de Tiempo

Integrantes:

Casto Gómez Pedro Pablo  
Domínguez Ramos Aby  
López Salazar Erika Samantha  
Martínez Herrera Tania Melisa  
Mondragón Miranda Néstor Yair  
Pimentel Bolívar Luis Emanuel

# Objetivo

En este proyecto se hará un análisis de la base de datos “House Property Sales”, para ajustar una serie de tiempo, con el objetivo de poder responder a las siguientes preguntas:

- ¿En qué época del año es más conveniente comprar/vender una casa?
- ¿Cuál será el precio promedio de las casas para el siguiente año?

## Introducción

En estos días los bienes inmuebles representan lo que es una parte muy importante de la riqueza de la mayoría de las personas, una vivienda también puede convertirse en una excelente oportunidad de inversión, ya que se trata de un bien que con el tiempo aumentará su valor.

El término plusvalía se refiere a la capacidad que un inmueble tiene de incrementar su valor a través del tiempo, es decir, aumentar su precio comercial o de venta. Este fenómeno se logra a través de algunos factores externos que benefician la calidad de vida de quienes habitan el inmueble, pero también tiene que ver con el aumento generalizado y sostenido de los precios de bienes y servicios en un país, lo que se conoce como inflación.

Un estudio de mercado inmobiliario es un análisis de los valores de mercado actuales de las propiedades que ayuda al vendedor a elegir un precio de cotización y a los compradores a ver si el precio de venta es demasiado alto, bajo o razonable. Esta información proporciona conocimiento sobre cómo invertir y comprar de manera segura, ya que, entre otras cosas, permite:

- Saber si se debe invertir en una zona en lugar de otra.
- Conocer aspectos demográficos que ayudan a determinar la evolución de la zona donde se quiere invertir y conocer si a futuro tendrán un desarrollo que permitan elevar el valor de las propiedades.
- Aprovechar un precio más bajo y evitar el pago excesivo del valor de mercado de una propiedad.

Hay diferentes factores en el sector inmobiliario que influyen en el valor de una propiedad, los más importantes a considerar son:

- La demografía del lugar: Describe la composición de la población (edad, raza, género, ingresos, patrones de migración y crecimiento de la población).
- Tasas de interés: Los cambios en las tasas de interés pueden influir en la capacidad de una persona para comprar una propiedad residencial, ya que si disminuyen las tasas de interés, menor será el costo de obtener una hipoteca para comprar una casa, creando una mayor demanda de bienes raíces, lo que eleva los precios.
- La economía de un país: Mide la capacidad de un cliente particular para invertir en una propiedad.
- Políticas gubernamentales: Los créditos fiscales, las deducciones y los subsidios son algunas de las formas en que el gobierno puede aumentar temporalmente la demanda de bienes inmuebles.
- Comodidades: Cuando la propiedad está en buenas condiciones, con una mejor infraestructura y cuenta con servicios públicos y comunitarios (suministro de agua y electricidad).
- Ubicación: Es lo que más influyen a la hora de calcular el precio de una casa, ya que de esta misma se desprenden características importantes que influyen en la calidad de vida de las personas (los establecimientos de los que se rodea, la cercanía a zonas comerciales y centros de salud, el acceso a transporte público, las áreas verdes, los espacios recreativos, lugares de entretenimiento, seguridad de la zona, entre otras).
- Tamaño de la propiedad: El número de metros cuadrados es de las primeras cosas que se estudian y valoran. Cuanto mayor sea el tamaño de una propiedad, mayor será su valor en el mercado.

- Tipo de vivienda: Hay distintas razones para inclinarse entre diferentes tipos de propiedad, ya sea por, espacio, privacidad, seguridad, mantenimiento o precio, la elección se basa en los gustos, necesidades y presupuesto de cada persona.
- Época del año: Las condiciones bajo las cuales los inmuebles son puestos en venta cambian respecto a la época del año. Por ejemplo, se espera que en los últimos y primeros meses del año se tenga una mayor solvencia económica, producto de remuneraciones extra del trabajo o de los ahorros a lo largo del año.

## **La crisis financiera internacional de 2008**

Un evento de gran importancia dentro del sector inmobiliario es la crisis financiera internacional de 2008, que tuvo su origen en el problema de las hipotecas subprime de EE.UU. y se extendió después, afectando no sólo a toda la economía de ese país, sino también a las de otras naciones, de manera principal, miembros de la Unión Europea. (Juárez , Sánchez Daza, & Zurita Gonzáles , 2015)

El incidente presentó antecedentes en el 2002, cuando el entonces presidente de EE.UU., George W. Bush, dijo que parte de cumplir el sueño americano era tener una casa propia. Así, el presidente pidió ayuda al sector privado para que el mercado de capitales facilitará el financiamiento hipotecario a personas de ingresos más bajos. Este crecimiento de créditos alcanzó a personas a quienes era peligroso que alcanzara: la clase desfavorecida, sin ingresos, activos ni empleo.

Enseguida, las hipotecas dirigidas a un sector de mayor riesgo fueron bautizadas como “hipotecas subprime” y en realidad eran solo activos con alta probabilidad de no pago. La creación de estos activos ocasionó que las familias se dedicaran a especular ampliamente con los precios de los inmuebles y, sobre todo, con los costos de las hipotecas.

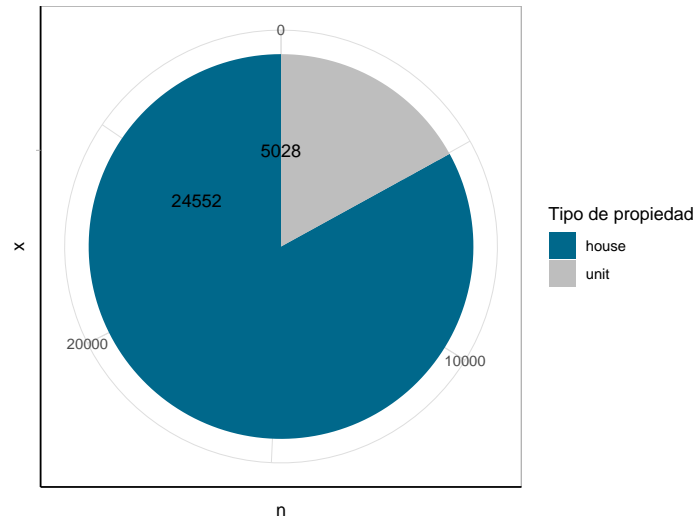
Posteriormente, la especulación llevó al desarrollo de una burbuja inmobiliaria que infló considerablemente los precios. Aunado a esto, los activos tóxicos comenzaron a ser “empaquetados” con otros, creando una gran variedad de nuevos productos financieros. Fue así como los bancos de inversión empezaron a repartir sus paquetes de activos por todo el mundo, aunque el contenido de estos era desconocido.

En consecuencia, la burbuja hipotecaria estalló y, gracias a ello, los precios de los activos se desplomaron. Al ver esto, las familias y propietarios de inmuebles endeudados optaron por devolver sus casas al banco y ocasionaron grandes pérdidas a los dueños de sus hipotecas. A su vez, esta situación generó naturalmente un colapso en los portafolios de los bancos de inversión a nivel mundial, debido a que estaban plagados de las hipotecas tóxicas subprime.

A raíz de todo lo acontecido, el lunes 15 de septiembre del 2008 fue que Lehman Brothers (compañía global de servicios financieros de Estados Unidos) se declaró en quiebra. El colapso de esta entidad, que era considerada “demasiado grande para caer”, fomentó el pánico general. (GESTIÓN, 2013)

## Análisis Descriptivo

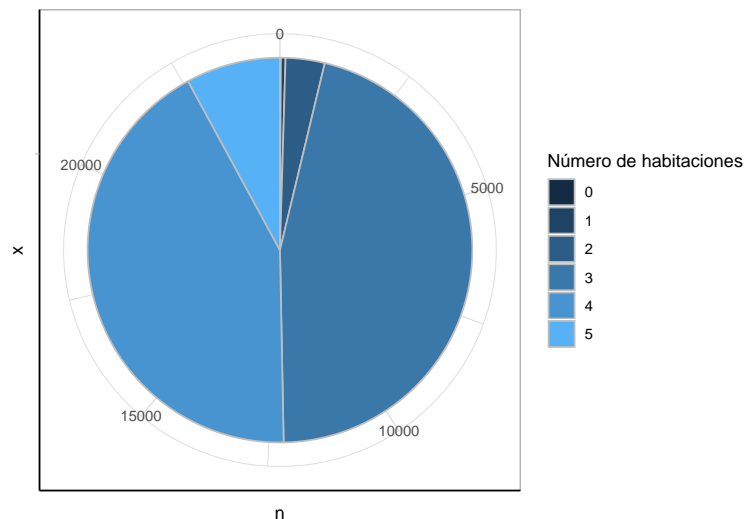
Se comienza el análisis con una gráfica de pastel, para conocer la proporción de viviendas de acuerdo al tipo de propiedad.



Se puede notar que la mayoría de las observaciones corresponden a las ventas de viviendas de tipo “house”; y que en promedio, el precio de estas es mayor que las del tipo “unit”.

Para fines de este proyecto se tomarán en cuenta únicamente el tipo de vivienda “house”, ya que los objetivos se centran en conocer el precios de las casas.

Ahora se mostrarán las observaciones de las casas segmentados por el número de habitaciones.

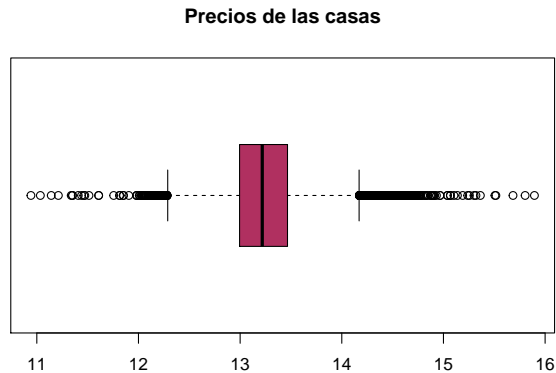


Es posible destacar que existe una menor cantidad de casas con menos de 2 habitaciones, por lo que se verán cuántas observaciones de este tipo hay y si representan una variable significativa para el modelo.

.	Total
Total de viviendas	29580
Viviendas con menos de 2 habitaciones	1657

Este tipo de viviendas representan a penas un poco más del 5% de nuestros datos, por lo que no serán consideradas para nuestro modelo.

En la siguiente gráfica de caja se observa claramente en qué región se encuentran concentrados los precios de las casas. Para una mejor visualización, a dichos precios se les aplicó logaritmo natural.

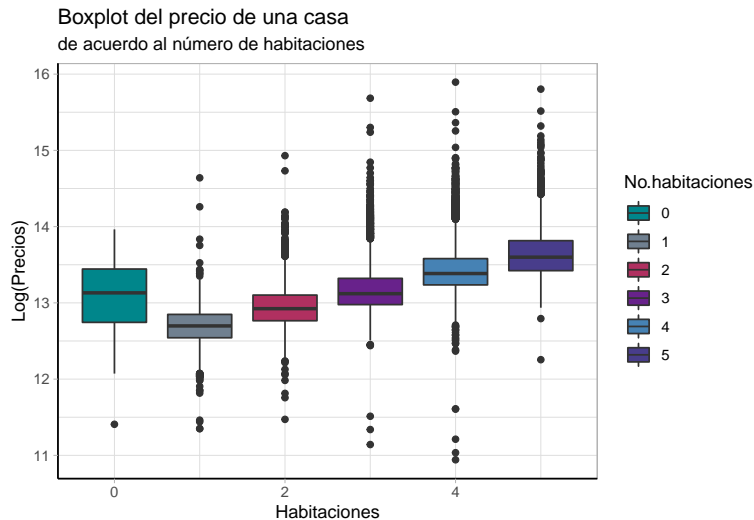


Existen pocos valores extremos a la izquierda de la gráfica. Para ver donde se encuentra la mayoría de ellos y saber cuáles se podrían descartar, se muestran los percentiles de las observaciones.

99.9%	99.91%	99.92%	99.93%	99.94%	99.95%	99.96%	99.97%	99.98%	99.99%	100%
2748420	2884472	2933420	2978841	3160000	3421050	3646050	4158841	4516840	5517310	8000000

Puede notarse que tan solo cerca del 0.03% de las observaciones son mayores que al rededor de 4,000,000 y algunas llegan hasta 8 millones; considerando estas como valores atípicos. Por lo que decidimos excluirlas de nuestro modelo.

Ahora, de forma general se describe la composición de los datos mediante una gráfica de caja con segmentación de acuerdo al número de habitaciones.



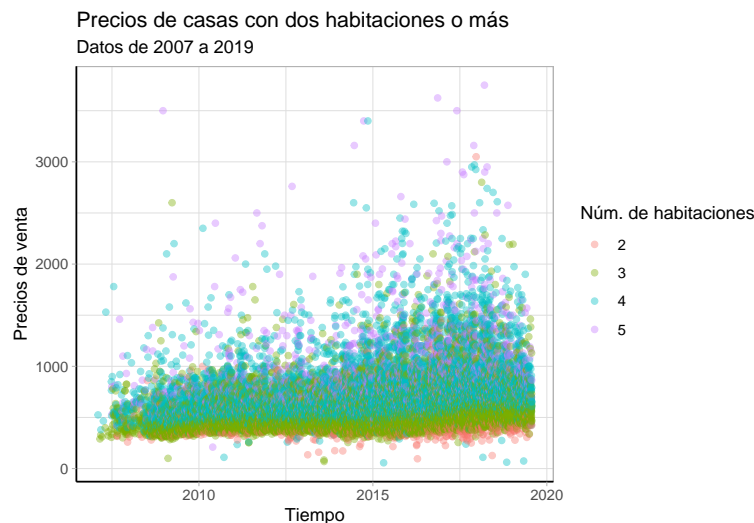
Se puede ver, que las viviendas con mayor número de habitaciones son las que en promedio tienen precios

más elevados. La excepción para este comportamiento son las casas con 0 habitaciones, pero como ya se mencionó anteriormente, no serán tomadas en cuenta para el modelo.

Para un mejor tratamiento de los datos, se consideró propicio tomar el precio promedio de casas para cada mes. De esta manera, se tiene un mejor resumen de la información, sin perder la interpretabilidad del modelo propuesto para la serie de tiempo.

Aunado a lo anterior, se decidió quitar aquellos datos cuyo comportamiento era demasiado alejado de los demás (precios arriba de \$3,900,000); dado que, estos datos no representan una proporción significativa del total.

En la siguiente gráfica se muestran los precios de casas segmentadas por el número de habitaciones, únicamente con las restricciones que se habían considerado previamente.



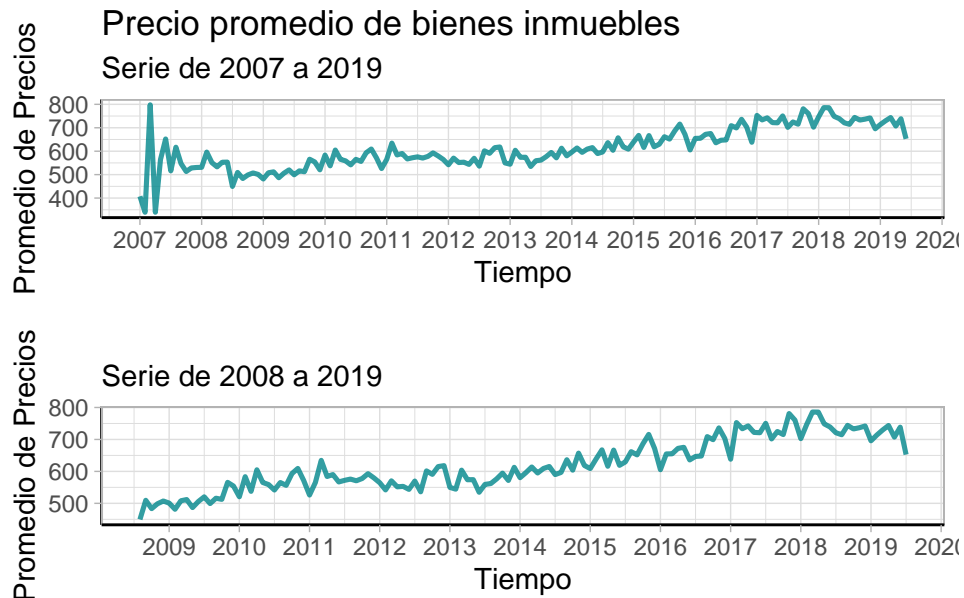
Al igual que el análisis previo, generalmente, el precio aumenta de acuerdo con el número de habitaciones, sin embargo, existe una mayor variación dentro de los precios de las casas con cuatro habitaciones. Lo cual tiene sentido debido a que el número de habitaciones está relacionado con el tamaño de la casa, aunque no lo determina completamente, por lo que no resulta anormal que una casa con dos habitaciones valga más que una con cinco.

Por otro lado, Se decidió tomar el promedio mensual de los precios de venta ya que de esta forma, se incluye la información de todas las observaciones y se puede hablar de cómo se comporta el mercado inmobiliario general.

Es a partir de estos datos que se construye la serie de tiempo. Gracias a esto será posible comprender mejor el comportamiento y así, ajustar un modelo que permita predecir posibles datos futuros de esta estadística y conocer la probable situación de este mercado.

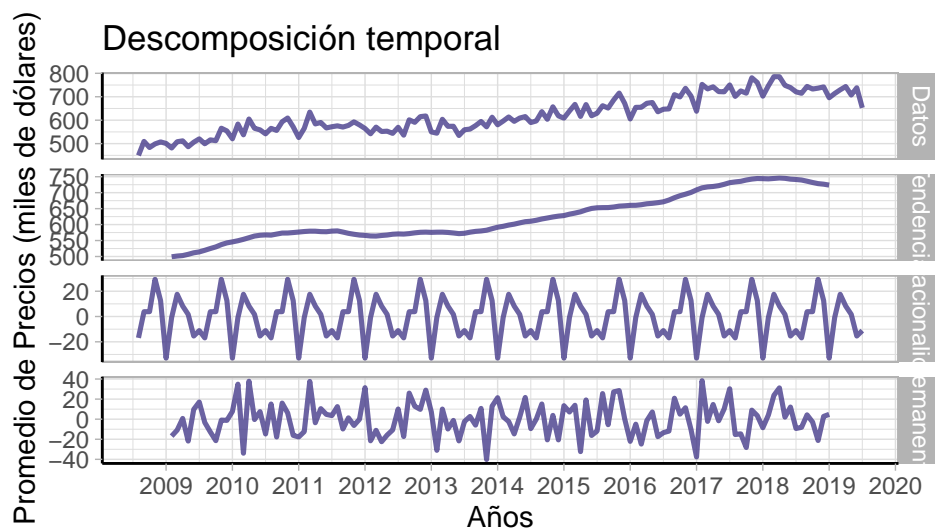
## Serie de Tiempo

Se inicia construyendo la serie de tiempo del promedio mensual de los precios de las casas. A modo comparativo, se ajusta para las fechas del 2007 al 2019 y del 2008 al 2019 respectivamente.



En ambas gráficas se puede apreciar que el promedio de los precios inmobiliarios ha aumentado a través del tiempo; no obstante, la primera gráfica muestra gran varianza en los datos del periodo 2007-2008. Este comportamiento anormal en los datos se puede adjudicar a los eventos que tuvieron lugar por la crisis financiera del 2008, y pueden ser considerados como eventos extraordinarios. Por esto, se decidió excluirlos del modelo para obtener un mejor ajuste, iniciando así la serie de tiempo en agosto del 2008. Dicha serie corresponde a la segunda gráfica.

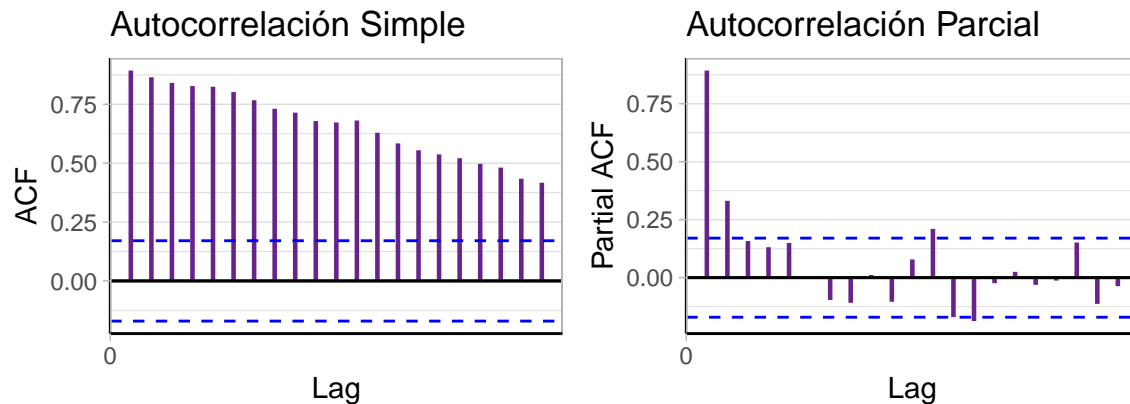
Se muestra a continuación la descomposición de la serie. Se utiliza el método de medias móviles para generar un modelo aditivo.



Aunque al inicio del comportamiento de la serie se ven datos alternos, se nota que al paso del tiempo el promedio del precio va aumentando gradualmente. Lo anterior se ve reflejado en la tendencia, que va creciendo drásticamente en los últimos años. Por otro lado, para finales del 2019 se puede apreciar el descenso del promedio mensual de los precios.

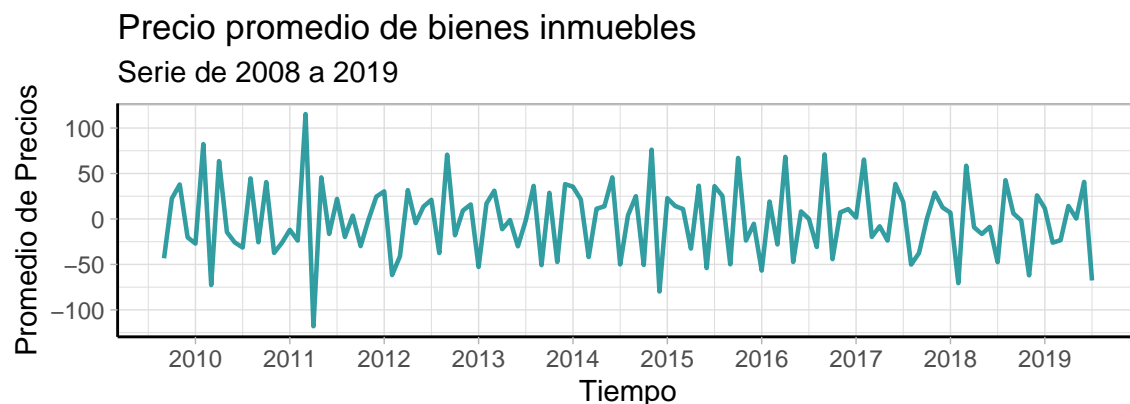
De igual manera, se ve una clara estacionalidad al inicio y a la mitad del año, donde hay una notoria recaída en el promedio de los precios; lo cual reafirma la importancia de la época del año en el comportamiento de estos. Finalmente, parece ser que el comportamiento se asemeja al de ruido blanco.

A continuación se muestran las gráficas de la autocorrelación (ACF) y autocorrelación parcial (Partial-ACF) de la serie.

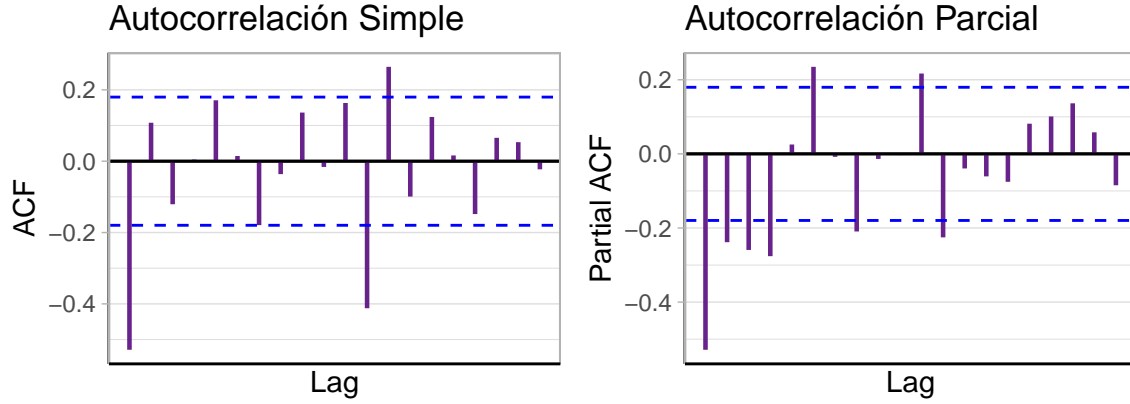


En la gráfica de autocorrelación simple, a pesar del decaimiento conforme avanzan los rezagos, se puede ver que cada 12 hay un aumento significativo en los retrasos, lo que indica cierta estacionalidad. Gracias a la tendencia presentada en la serie de tiempo, es posible saber que es no estacionaria, en busca de la estacionariedad de la serie, se consideró oportuno aplicar una diferencia estacional y una diferencia no estacional.

La serie con las diferencias aplicadas ya no presenta tendencia, así mismo, la media es constante a lo largo del tiempo, manteniéndose en 0, por último, la varianza no presenta grandes alteraciones. Debido a estas condiciones, no fue considerado aplicar alguna transformación a la serie.







Al aplicar ambas diferencias se ve que ya no hay ningún decaimiento y no se ven valores que sobresalgan periódicamente. En la gráfica del ACF, sobresalen significativamente los rezagos 1 y 12, lo que en primera instancia dice que el modelo que se ajusta a la serie de tiempo podría contener un componente de medias móviles estacional y otro no estacional. Por otro lado, en la gráfica del PACF hay varios rezagos que también sobresalen de tal manera que podría reafirmar la presencia de medias móviles, o tomarse como un indicio para la inclusión de componentes autorregresivos en el modelo.

Teniendo en cuenta el análisis anterior, se ajustaron los siguientes modelos:

Modelos	AIC
SARIMA(0,1,1)(0,1,1)[12]	2764.31
SARIMA(1,1,1)(0,1,1)[12]	2762.63
SARIMA(4,1,0)(2,1,0)[12]	2778.58

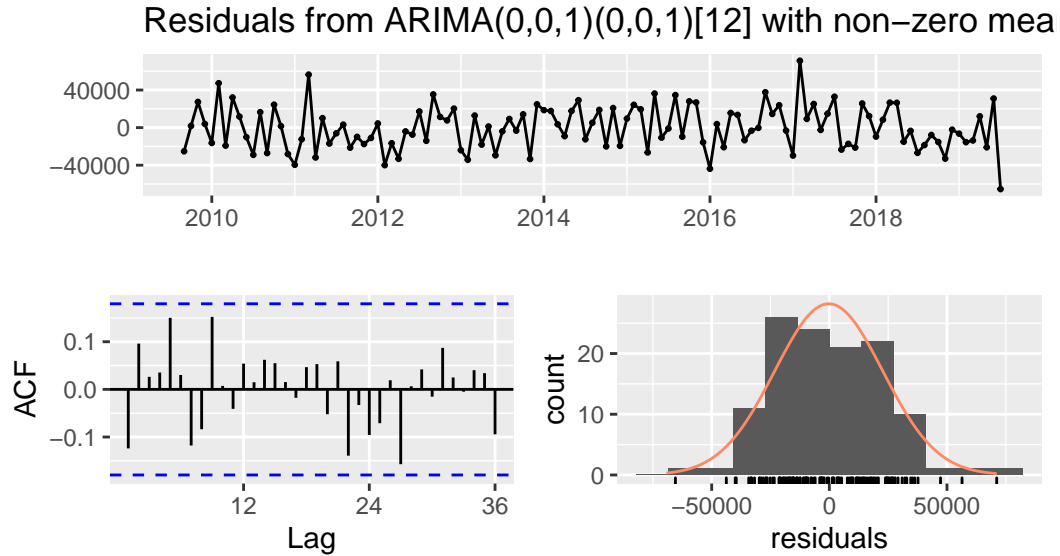
Entre los modelos ajustados, los dos primeros presentan similitudes. El segundo modelo tiene un mejor AIC, pero su diferencia con el primero (1.68) es mínima y se consideró que puede despreciarse para seleccionar el modelo con el que se trabajará. El primer modelo podría considerarse más adecuado que el segundo debido a que presenta menos componentes, lo que facilita su tratamiento e interpretación. El último fue generado mediante métodos numéricos (utilizando la función `auto.arima`) y es tomado únicamente como referencia, ya que resulta tener el peor AIC y presentar la mayor cantidad de componentes.

Complementando el análisis para estos modelos, se realizaron las siguientes pruebas de hipótesis:

Modelos	Ljung-Box	Dickey-Fuller	KPSS
SARIMA(0,1,1)(0,1,1)[12]	0.5428	0.1781	2.2e-16
SARIMA(1,1,1)(0,1,1)[12]	0.5041	0.1485	2.2e-16
SARIMA(4,1,0)(2,1,0)[12]	0.466	0.1513	2.2e-16

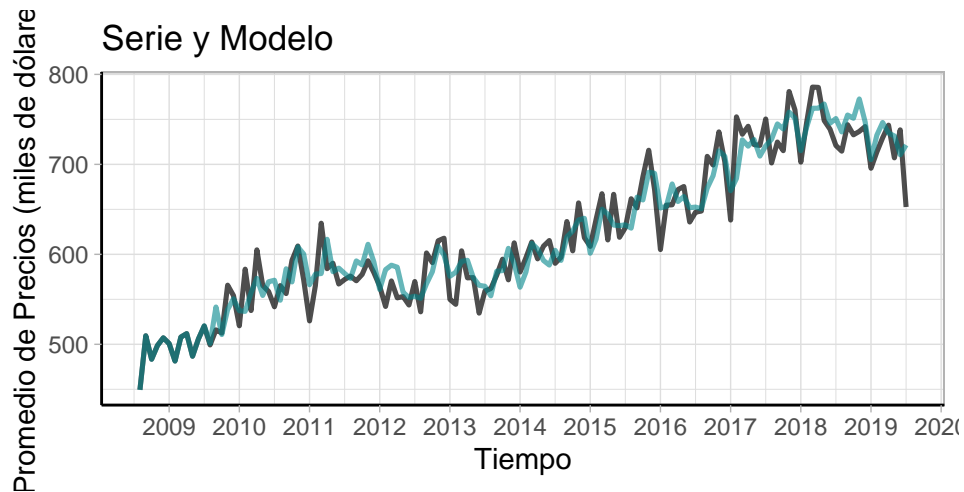
Todos los modelos fueron aceptados por las tres pruebas de hipótesis, donde se puede confirmar que todos ellos tienen residuales independientes (Ljung-Box), sus residuales son estacionarios (Dickey-Fuller) y la serie es estacionaria (KPSS). Debido a que los tres modelos pasaron las pruebas de hipótesis y tienen un AIC semejante, se consideró como mejor opción el primer modelo, gracias a su simplicidad con respecto a los otros.

El comportamiento de los residuales del modelo se puede ver mejor a través de la siguiente gráfica:



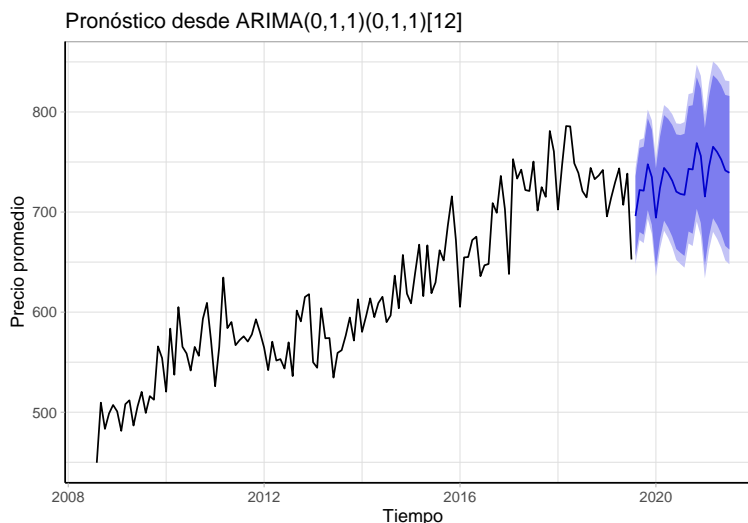
El modelo propuesto, como se vio anteriormente, parece no tener una varianza en el promedio de los precios al paso del tiempo. Igualmente, se nota que la autocorrelación simple, no tienen ningún valor sobresaliente por lo que no parece haber ninguna diferencia o variable necesaria para mejorar este modelo. Finalmente, se muestra que los residuales siguen un comportamiento parecido al del ruido blanco.

A continuación, en la siguiente gráfica se muestra la serie de tiempo (negro) y el modelo ajustado (turquesa).



Se puede apreciar que el modelo es similar a la serie de tiempo. Por lo cual, junto con las pruebas antes realizadas, se concluye que es un buen modelo.

Por último, se presentan las predicciones realizadas hasta mitad del año 2021, junto con sus respectivos intervalos de predicción para 90% y 95% de confianza.



Con respecto a la precisión de las predicciones, cabe destacar algunos puntos. En primera instancia, la obtención de un modelo con pocos componentes (parámetros) es una forma de evitar sobreajustar el modelo a los datos. Gracias a esto, tenemos que no se espera una repercusión negativa del “overfitting” en las predicciones. Adicionalmente, se comprobó mediante pruebas estadísticas la normalidad de los residuales (comportamiento de ruido blanco), siendo este comportamiento un supuesto para la construcción de los intervalos de predicción. Lo anterior indica que dichos intervalos fueron correctamente calculados.

En caso de ser necesarias pruebas más exhaustivas con respecto a las estimaciones puntuales y sus respectivos intervalos, se recomienda seguir un procedimiento diferente para la obtención de un modelo. Este procedimiento consiste en extraer una porción baja de los datos (alrededor de 20%) y dejarlo fuera del proceso de ajuste del modelo. Posteriormente, se procedería a realizar estimaciones (con sus intervalos de predicción) para las fechas de los datos previamente extraídos y se compararía con los datos reales para conocer la precisión del modelo. (Hyndman, R.J., & Athanasopoulos, G., 2018)

Debe ser considerado que el objetivo del estudio es conocer el comportamiento general del mercado inmobiliario, enfocándonos en las casas, tomando en cuenta los datos disponibles. Dado lo anterior y que se decidió trabajar con el precio promedio de los bienes inmuebles, no fueron precisados datos exactos para su interpretación, sino un vistazo de la tendencia que seguirá el sector en cuanto a posibles aumentos o disminuciones en los precios en general.

Por estas razones, consideramos adecuadas las predicciones y sus respectivos intervalos de predicción para su interpretación, misma que nos llevará a contestar los cuestionamientos planteados al inicio de este trabajo.

La predicción nos indica que el promedio de los precios se mantendrá muy parecido para los años 2020 y 2021 (sin tomar en cuenta una contingencia mundial). Aunque se destaca que hubo una disminución de los precios en los últimos años, se espera que esto no continúe y se mantenga para estos años. Por otro lado, si se es dueño de bienes inmuebles que entren en la categoría de los estudiados en este informe, no se recomienda tomar medidas como la venta de los mismos ya que su valor tenderá a mantenerse con el tiempo. Además, los bienes inmuebles pueden generar una renta para el propietario, lo que genera estabilidad financiera y se tiene la posibilidad de un aumento de posterior de precios. Esto refuerza la idea de que no es recomendable vender este tipo de bienes y muestra que comprar una casa es una buena opción.

Finalmente, con respecto a nuestros objetivos, se observó que el mejor momento para comprar una casa sería al final del año, pues se ve una disminución en los precios para este periodo; en cambio, para vender una el mejor periodo sería a principio del último trimestre, pues los precios alcanzan su punto cumbre en estas fechas. No obstante, cabe mencionar que la información disponible hasta el momento no indica una

tendencia establecida, por lo que sería recomendable esperar a tener información más actualizada antes de realizar movimientos arriesgados en el sector inmobiliario.

## Referencias

- Hyndman, R.J., & Athanasopoulos, G. (2018) *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2. Consultado en 25/01/2021.
- Quiero Casa. (8 de enero de 2019). *¿Por qué el valor de los bienes inmuebles crece con el tiempo?* Obtenido de El Blog de Quiero Casa: <https://blog.quierocasa.com.mx/por-que-valor-bienes-inmuebles-crece-tiempo>
- VIVANUNCIOS. (8 de agosto de 2019). *¿Ubicación o tamaño? Qué influye más en el valor de una casa.* El Financiero.
- VIVANUNCIOS. (s.f.). *Cuál es la mejor época del año para comprar una casa.* Obtenido de VIVANUNCIOS Web site: <https://blog.vivanuncios.com.mx/bienes-raices/comprar-casa/mejor-epoca-del-ano-para-comprar-una-casa/>
- International Monetary Fund. (2007, mayo). *Índices de los precios inmobiliarios.* IMF eLIBRARY.
- Parra, A. (2020, 14 septiembre). *¿Qué es un estudio de mercado inmobiliario?* QuestionPro.
- Brown, A. (2020, 7 diciembre). *5 factores que influyen en el sector inmobiliario.* Preemo.
- Ricardo, C. (2020, 3 julio). *4 factores clave que impulsan el mercado inmobiliario.* ExoNegocios.
- Juárez, G., Sánchez, A., & Zurita, J. (2015). *La crisis financiera internacional de 2008 y algunos de sus efectos económicos sobre México.* Contaduría y administración, 60(Supl. 2), 128-146.
- Redacción Gestión. (2013, 15 septiembre). *¿Cómo se originó la peor crisis financiera de la historia?* Gestión. <https://gestion.pe/economia/origino-peor-crisis-financiera-historia-48101-noticia/?ref=gesr>