

# Análisis estadístico de los datos de los productos

2022-07-04

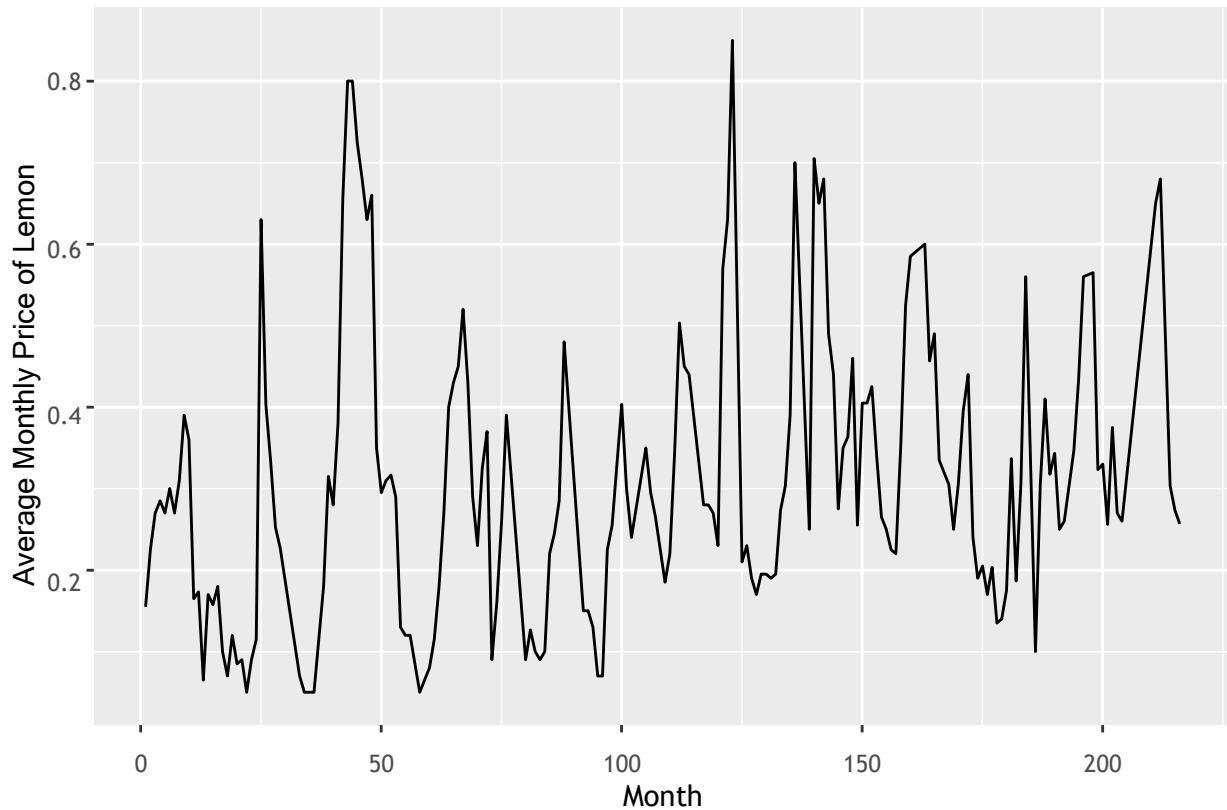
## Análisis

El conjunto de datos contiene los precios mínimos, máximos y medios aplicados a 20 productos del grupo de cítricos y frutales. El conjunto de datos mensuales comienza en enero de 2000 y termina en enero de 2022, pero no había valores de precios antes de 2005. Por lo tanto, los datos principales utilizados para el análisis fueron de enero de 2005 a enero de 2022, lo que se traduce en 216 meses, pero incluso este conjunto de datos contenía muchos valores perdidos. Dado que los precios máximos y mínimos pueden sesgar los datos, decidí analizar el precio medio cobrado, ya que las medias representan el precio típico cobrado por el producto. Obsérvese que para muchos de los productos, especialmente el limón y la almendra, hay muchos valores para el precio medio del producto para algunos de los meses. Por lo tanto, se obtuvo la media de los valores para obtener el precio medio pagado por el producto en ese mes. Hay 20 productos en el conjunto de datos, pero como se ha dicho, muchos de estos productos no tienen datos para algunos de los meses, y el problema de los datos faltantes era grave para algunos de los productos. Por ejemplo, el producto Cal sólo tiene 1 punto de datos de un total de 216 puntos de datos; por lo tanto, no se puede hacer ningún análisis serio para este producto.

En general, 16 de los 20 productos del conjunto de datos tienen un problema grave de falta de datos, por lo que no se puede realizar un análisis estadístico serio para ellos. Algunos de los productos que tienen un problema grave de valores perdidos son los jinjoles (faltan 203 puntos de datos), los albaricoques (faltan 164 puntos de datos), las aceitunas (faltan 181 puntos de datos) y las cerezas (faltan 193 puntos de datos), etc. Los únicos productos con suficientes puntos de datos para realizar un análisis estadístico son el limón, la almendra, la mandarina y la naranja. El limón sólo tenía 26 meses sin datos, la almendra 7, la naranja 40 y la mandarina 78. Por lo tanto, la visualización de los datos y el análisis estadístico se hicieron sólo para estos cuatro productos.

## Statistical Visualization/Analysis

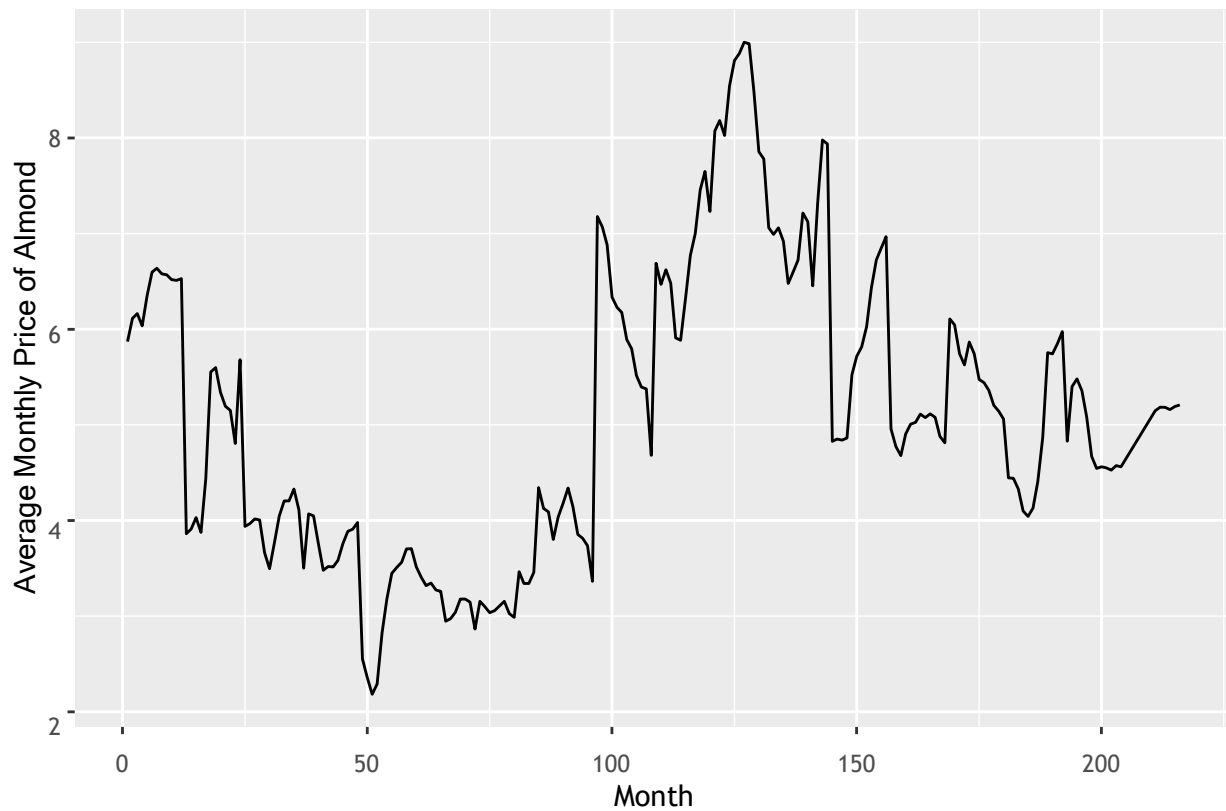
### Average Monthly Lemon Price from 2005–2022



El gráfico del precio medio mensual del limón desde 2005 hasta 2022 (arriba) muestra que el precio medio del limón ha tenido una tendencia general al alza cada mes: el aumento medio mensual del precio del limón resulta ser del 0,06477%, lo que se traduce en un aumento anual del 0,780015%. También parece haber un patrón estacional en los datos, aunque es difícil afirmar de forma concluyente que exista un patrón estacional, ya que los picos y los mínimos no se repiten de forma uniforme. En general, lo principal que se puede extraer del gráfico es que el precio medio mensual de los limones ha sido ascendente a lo largo del periodo analizado.

## [1] 7

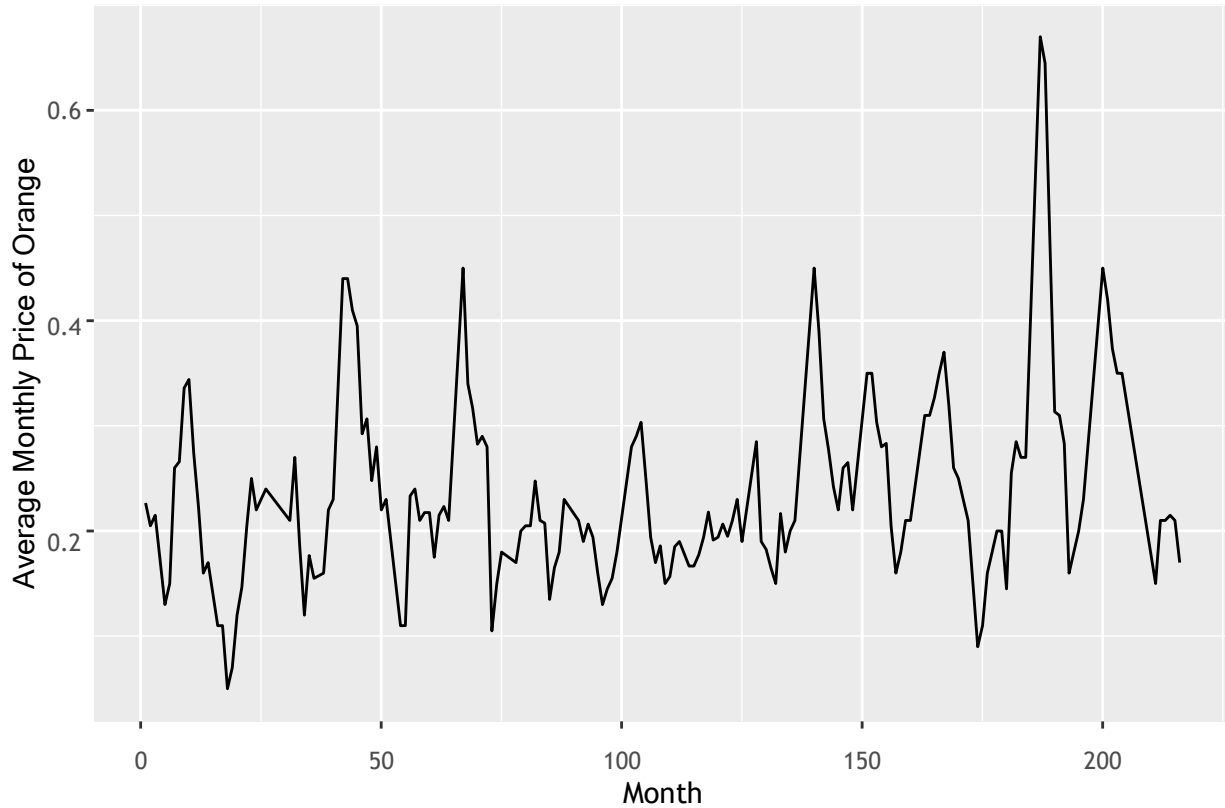
Average Monthly Almond Price from 2005–2022



En el gráfico del precio medio mensual de la almendra entre 2005 y 2022 (arriba), es evidente que no hay una tendencia al alza ni a la baja en los datos. Aunque no hay una tendencia clara en el conjunto de datos, existe una tendencia no lineal (cúbica) que una prueba estadística ha confirmado. El patrón de la gráfica es un poco aleatorio, lo que indica que no hay estacionalidad en el conjunto de datos.

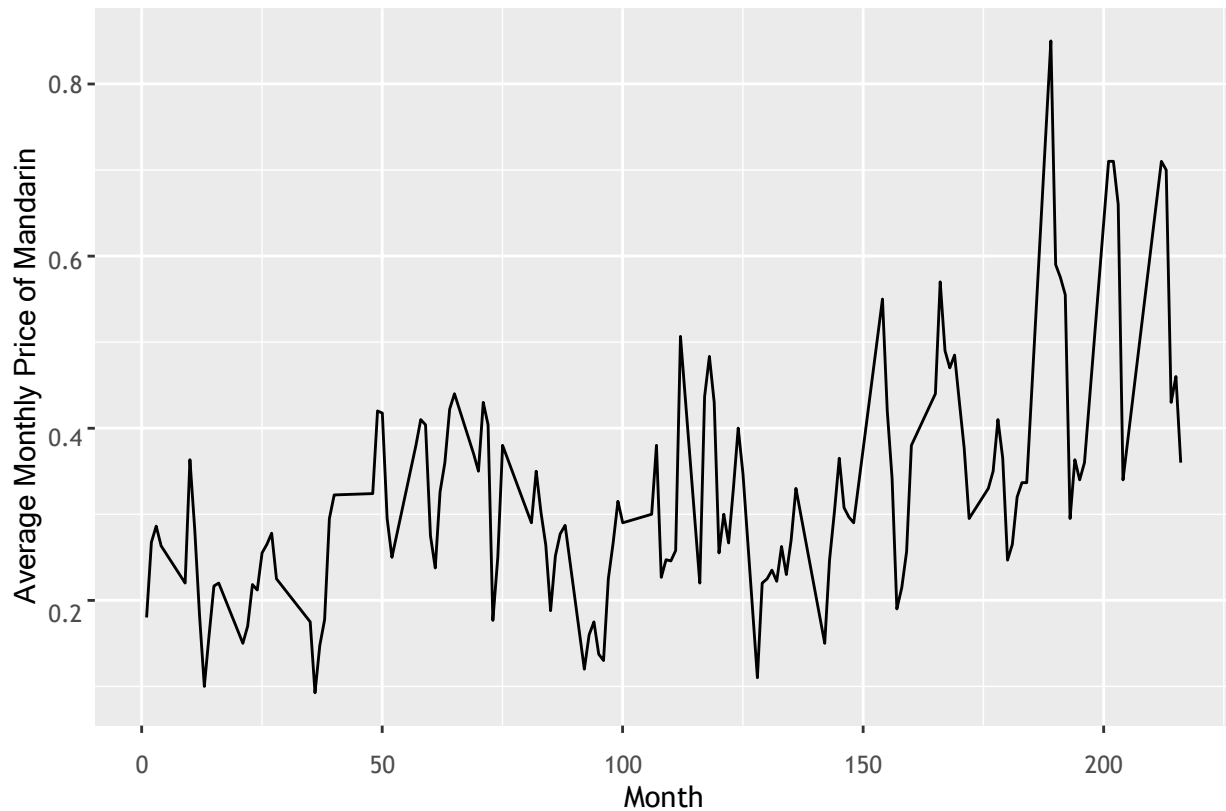
## [1] 40

Average Monthly Orange Price from 2005–2022



El gráfico del precio medio mensual de la naranja que se muestra arriba muestra una pequeña pero significativa tendencia al alza en los datos (la tendencia es realmente significativa según una prueba de hipótesis de los datos). La tendencia obtenida, que es de 0,0003505, nos indica que el aumento mensual del precio medio de la naranja es del 0,03505%. Esto equivale a una subida anual del 0,4214%. También parece que hay estacionalidad en el conjunto de datos de la naranja, aunque esto aún está por confirmar mediante un análisis estadístico.

Average Monthly Mandarin Price from 2005–2022



El gráfico del precio medio mensual de las mandarinas muestra una tendencia significativa al alza en los datos (la tendencia es efectivamente significativa según una prueba de hipótesis de los datos, aunque parece que se trata de una tendencia cuadrática/cúbica y no lineal). También parece que hay estacionalidad en el conjunto de datos de las naranjas, aunque esto aún está por confirmar mediante un análisis estadístico.

#### Reservas

En la vida real existe una estacionalidad significativa en los precios de la mayoría de los productos agrícolas, pero no pude ver ninguna estacionalidad clara en el conjunto de datos analizado. En mi opinión, esto podría deberse a los datos que faltan (que se han indicado anteriormente) y al promedio realizado al preprocesar los datos (algunos de los meses tienen más de un punto de datos para algunos de los productos, lo que hace inevitable el promedio). No se puede hacer un buen análisis de muchos productos a los que les faltan muchos puntos de datos. Los productos que tenían un conjunto de datos suficiente para ser factibles son el limón, la mandarina, la almendra y la naranja. Si la actividad final era la previsión de los precios de estos productos utilizando varios métodos estadísticos, el problema de los datos que faltan para estos cuatro conjuntos de datos no es un gran problema.