

# Precio de la tortilla en México

Génesis Mendoza

Diplomado matemática en ciencia de datos

20th May 2024

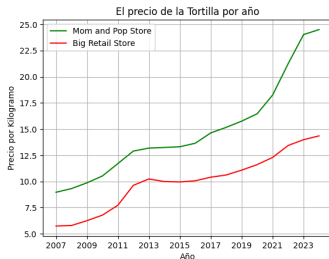
El análisis de el aumento de la tortilla através de los años es un problema nacional, ya que ha afectado negativamente la economía de la inmensa mayoría de las familias mexicanas, sobre todo en el de las más pobres. Por lo que siete de cada diez mexicanos creen que puede llegar a poner en riesgo la estabilidad social del país, el proposito es ver las razones de esta situación y como 1 kg de tortilla a aumentado através de los años esto gracias a la asignación de los recursos entre las empresas de la economía mexicana, es decir, cómo se distribuyen los factores de producción como el capital y el trabajo entre los productores.



Analizaremos :

- Como ha incrementado el precio de la tortilla através de los años 2007 a 2024.
- Cual es el comportamiento de precio por *kg* en diferentes estados.
- La comparación en precios en tipos de tiendas como familiares (que son tiendas pequeñas como abarrotes) y en tiendas minoristas (walmart, bodega aurrera,etc).

# Precio por año



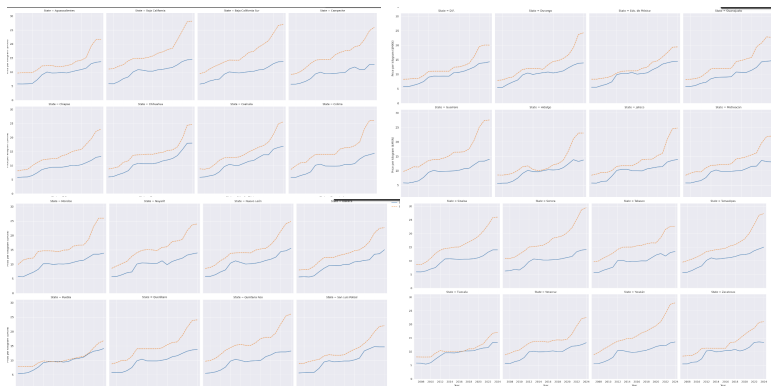
Año	Precio por kg
2007	8.966106
2008	9.318599
2009	9.867444
2010	10.523425
2011	11.707174
2012	12.902494
2013	13.18313
2014	13.246201
2015	13.312616
2016	13.647267
2017	14.63377
2018	15.167404
2019	15.758619
2020	16.461125
2021	18.253055
2022	21.234328
2023	24.02686
2024	24.50238

Se puede observar que en tiendas familiares el precio de la tortilla ha incrementado mas através del tiempo que en grandes tiendas minoristas, se puede observar que apartir del 2020 los precios para ambas tiendas casi se ha duplicado el doble.



Store type	Big Retail Store	Mom and Pop Store
Year		
2008	0.462831	3.340013
2009	6.726093	5.889783
2010	8.457905	6.647938
2011	13.977741	11.248697
2012	25.150417	10.073312
2013	6.539775	1.967108
2014	-2.813578	0.703717
2015	-0.643578	0.605959
2016	1.696219	2.172250
2017	2.835486	7.587014
2018	2.019938	3.641381
2019	4.797366	2.710275
2020	5.077862	0.187612
2021	5.094224	15.863070
2022	9.310361	17.416449
2023	4.017968	12.665753

# Analizando precios por año, estado y tipo de tienda.



# Analizando precios por año, estado y tipo de tienda.

Aunque en el último año el precio del maíz ha disminuido, el del kilo de tortilla no baja, sin embargo, esto se debe a factores como el desfase entre los ciclos de cosecha del grano, la diferencia de consumo entre regiones del país y las alzas en costos como el salario y la energía, entre otro.

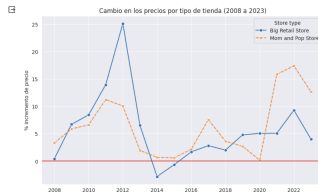
Year	Aguascalientes	Baja California	Baja California Sur	Campeche	Chiapas	Chihuahua	Coahuila	Colima	D.F.	Durango	Edo. de México	Guanajuato	Guerrero	Hidalgo	Jalisco	Michoacán	Morelos	Nayarit	Nuevo León	Oaxaca
2007	7.1771320717100	6.907270163200	7.101071700000	7.017770202610	6.900000110000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000	7.000171000000
2008	7.100000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2009	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2010	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2011	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2012	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2013	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2014	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2015	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2016	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2017	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2018	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2019	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2020	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2021	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2022	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2023	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000
2024	7.000000000000	6.900000000000	7.000000000000	7.000000000000	6.900000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000	7.000000000000



# Incremento año por año por tipo de tienda

Es importante tener en cuenta que las grandes empresas hacen compras de maíz por adelantado -unos cuatro meses-. “Apenas se están empezando a reflejar los menores precios”. El mercado se rige por la oferta y la demanda. Además conforme la inflación va subiendo los precios de materia prima también, en caso de las grandes empresas minoristas influye también el aumento en el precio del dolar y los precios internacionales del maíz

Year	Big Retail Store	Non and Pop Store
2008	9.402011289089407	3.58817318208380
2009	6.728087769069	6.08192881634811
2010	8.17604680518497	4.44747822201820
2011	10.87741701486512	11.2488019182143
2012	25.130410702025825	16.0731212869064
2013	6.558757020147115	1.8871471820386
2014	2.613017176320469	0.33317304738039
2015	0.803277027285442	0.88584648303186
2016	1.68621691181313	2.722592225557
2017	2.83458338400324	7.38321118028149
2018	2.14307042020647	3.64130820738714
2019	4.7072880780789	2.71827512108605
2020	8.01781288770263	8.18717088403135
2021	5.0442358844115	15.8518486071808
2022	8.71038042514857	17.4154482380563
2023	4.617086173816478	12.8657020181386



# Relación de precios

Conclusión: Teniendo en cuenta que el precio de la tortilla se rige de la oferta y demanda también hay otros puntos a considerar como lo que es la oferta y la competencia como podemos ver los estados del norte muestran precios significativamente más altos que el resto de los estados de México, pero otras regiones también sufren precios más altos como los de la Península de Yucatán, esto se debe a que en el norte consumen mas la tortilla de maiz.

State	Big Retail Store	Mini and Pop Store	Price ratio index
Aguascalientes	13.7	21.67	1.581753248173184
Baja California	14.372000000000001	20.140200000000002	1.4018020461773754
Baja California Sur	13.7688	27.9	1.949304811337885
Banquero	12.8482	26.8	2.054458583214735
Chiapas	13.270000000000001	23.06	1.731526479525582
Chihuahua	18.8184	24.087800000000004	1.280252787717131
Coahuila	18.862133333333336	25.580000000000008	1.341217023127044
Colima	14.237	26.8	1.882224483222485
CDP	14.27	20.0474	1.404867498848632
Durango	13.0885	24.28	1.793702182822131
Edo. de México	18.42	19.42	1.049236148222484
Guerrero	14.962787723000079	22.498887723000479	1.502384888903088
Guanajuato	13.975000000000001	27.51	1.9683141500388242
Hidalgo	13.1	22.8	1.7418621783382325
Jalisco	13.9	24.738	1.7807788704172682
Mérida	13.8572	22.8	1.6614788851308485
Morelia	13.8328	26.8	1.936448811333842
Nayarit	13.84	24.8	1.779888278862388
Nuevo León	15.532	24.557800000000008	1.58091158887758452
Oaxaca	15.3648	22.758800000000002	1.548702359582118
Puebla	14.078188888888889	18.888	1.386015252017022
Querétaro	13.7684	24.0444	1.7473288247358825
Quintana Roo	13.2588	26.1	1.932222452359848
San Luis Potosí	14.87	22.8	1.488959388337323
Sinaloa	14.072000000000001	28.8	1.957023183333333

# Comparación de Modelos de ML para predicción de precios



- Regresión lineal
- Random Forest
- Optimización con Grid Search

# Introducción

## Objetivos:

- Predecir el precio por Kilogramo utilizando datos historicos.
- Comparar la presición de diferentes modelos de maching learning.

## Importancia:

- Mejora la toma de decisiones en la fijación de precios.

# Descripción de Datos

## Análisis Exploratorio:

- Valores nulos en la columna 'Price per Kilogram'.
- Codificación de variables categoricas('state','City','Store Type')
- División del dataset en conjuntos de entrenamiento y prueba.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 278886 entries, 0 to 278885
Data columns (total 7 columns):
#   Column                Non-Null Count  Dtype
---  -
0   State                  278886 non-null object
1   City                   278886 non-null object
2   Year                   278886 non-null int64
3   Month                  278886 non-null int64
4   Day                    278886 non-null int64
5   Store type             278886 non-null object
6   Price per kilogram     272496 non-null float64
dtypes: float64(1), int64(3), object(3)
memory usage: 14.94 MB
```

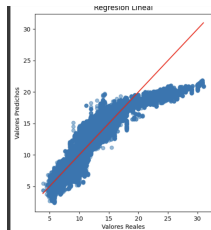
```
df = pd.read_csv("tortilla_prices.csv")
print(df.head())
```

	State	City	Year	Month	Day	Store type	Price per kilogram
0	Aguascalientes	Aguascalientes	2007	1	10	Non and Pop Store	9.9
1	Baja California	Mexicali	2007	1	10	Non and Pop Store	10.0
2	Baja California	Tijuana	2007	1	10	Non and Pop Store	10.0
3	Baja California Sur	La Paz	2007	1	10	Non and Pop Store	10.0
4	Campeche	Campeche	2007	1	10	Non and Pop Store	10.0

# Regresión lineal

Modelo de regresion simple que asume una relación lineal con las variables predictoras y la variable objetivo.

```
print(f'Regresión Lineal - MSE: {mse_lr}, R2: {r2_lr}')  
Regresión Lineal - MSE: 3.1550065215408702, R2: 0.8316115556503456
```

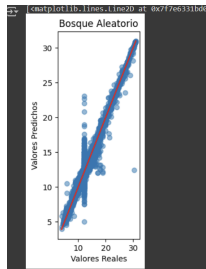


Muestra la dispersión significativamente alrededor de la línea ideal.

# Random forest

Modelo basado en múltiples árboles de decisión, reduce el riesgo de sobreajuste.

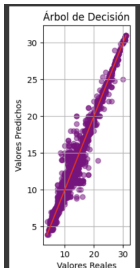
```
print(f'Bosque Aleatorio - MSE: {mse_rf}, R2: {r2_rf}')  
  
Bosque Aleatorio - MSE: 0.04091942447255191, R2: 0.99781
```



Mejor alineación de las predicciones con los valores reales comparado con la regresión lineal.

# Decision Tree

Puede manejar datos numéricos y categóricos y puede capturar relaciones no lineales entre características y la variable objetivo.



```
print(f'DecisionTree - MSE: {mse_dt}, R2: {r2_dt}')
```

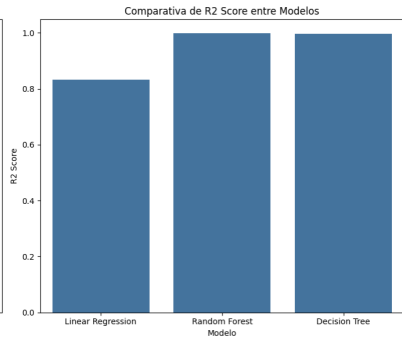
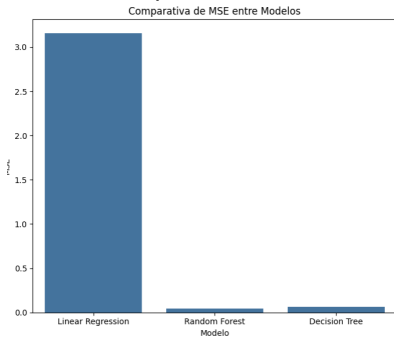
```
DecisionTree - MSE: 0.0652355373086163, R2: 0.996518260558667
```

Las predicciones muestran una mayor dispersión en comparación con Random Forest, lo que sugiere una capacidad predictiva inferior.



# Comparación visual de Modelos

## Grafico comparativo:



	Model	MSE	R2 Score
0	Linear Regression	3.155007	0.831612
1	Random Forest	0.040919	0.997816
2	Decision Tree	0.065236	0.996518

## Conclusiones:

Basándonos en la comparación visual y considerando la dispersión de las predicciones y su cercanía a la línea ideal, podemos concluir que Random Forest es el modelo más óptimo para este conjunto de datos. Esto se debe a su capacidad para capturar relaciones no lineales y su mejor rendimiento en la predicción de precios por kilogramo en comparación con la regresión lineal y el árbol de decisiones.

# !Gracias!