

# Machine Learning para estructuras de Series de Tiempo usando Stata/Python

Franco A. Mansilla Ibáñez

# Agenda

1. Machine Learning.
2. Contexto.
3. Patrones temporales.
4. Estructuras de un modelo.
5. Pre-procesamiento.
6. Medidas de evaluación.
7. Ejercicio práctico.

# Machine Learning

- Es el sistema que entrega el aprendizaje a los sistemas inteligencia o máquinas.

reconocimiento voz

conducción automática

ChatGPT

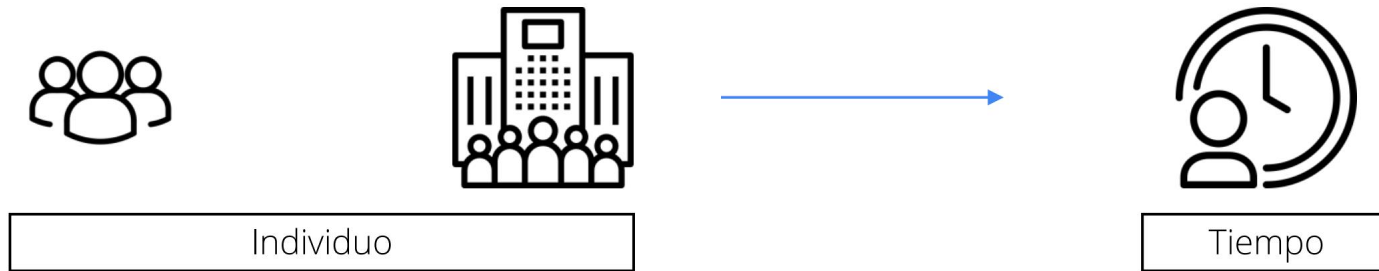
Examples	Capabilities	Limitations
"Explain quantum computing in simple terms" →	Remembers what user said earlier in the conversation	May occasionally generate incorrect information
"How do I make an HTTP request in Javascript?" →	Trained to decline inappropriate requests	May occasionally produce harmful instructions or biased content
		Limited knowledge of world and events after 2021

chat inteligentes

Machine Learning en Empresas

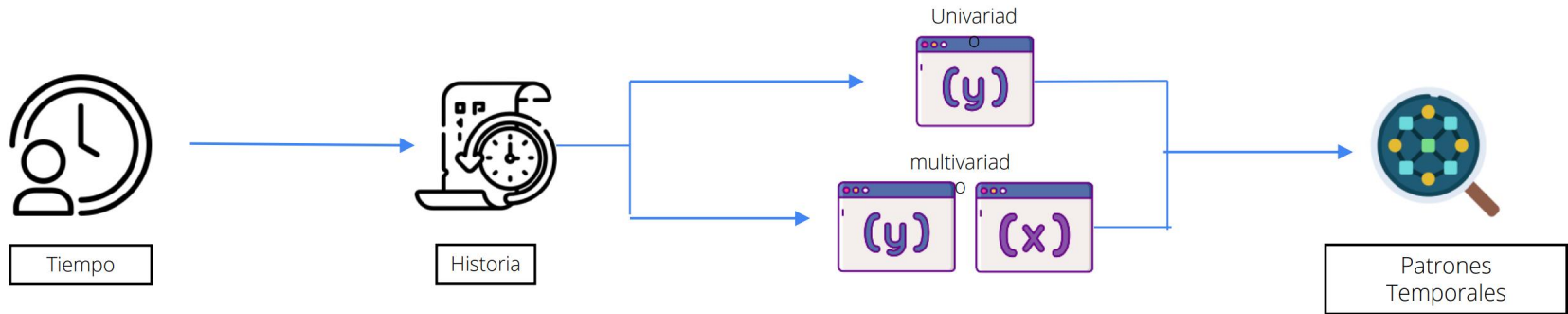
áreas comerciales, riesgo, operacional, planificación

# Contexto



- Las proyección de ventas de una empresa, es capturar las ventas *a lo largo del tiempo, ordenadas cronológicamente*.

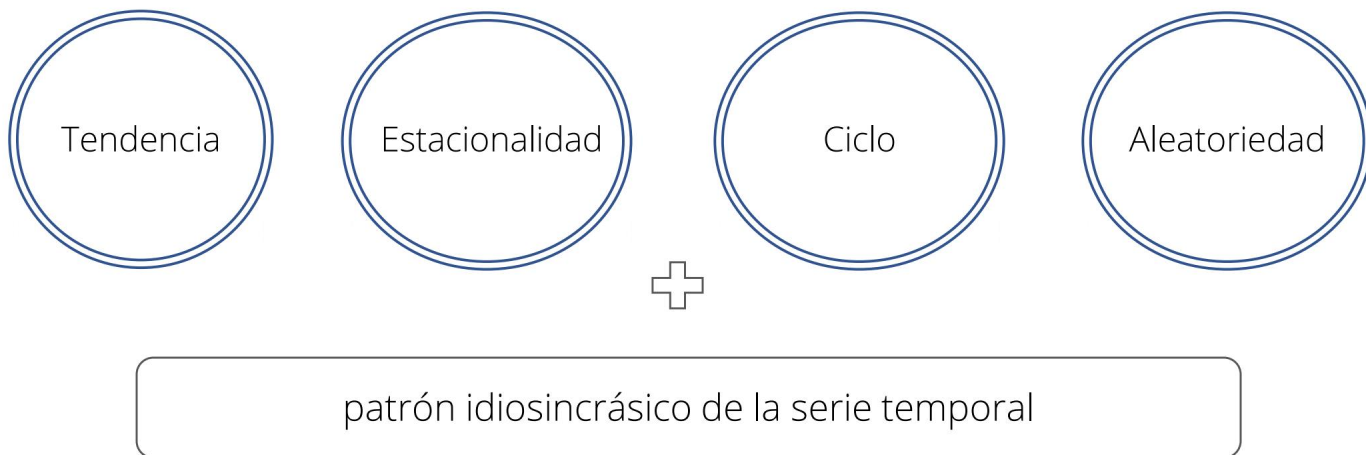
# Contexto



- Las variables **tienen memoria** en su historia para ser recogidas modelando con técnicas analíticas.

# Patrones Temporales

- Los patrones temporales son comportamientos que recogen la información contenida en la historia de la serie de tiempo.



# Estructura Modelo de Proyección (cont.)

- viene trabajando, con una estructura transversal al tiempo:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k$$

- En que  $i$  representa el individuo, empresa, etc.

- Desde ahora, se trabajará con una estructura temporal:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \dots + \beta_k Y_{T-t}$$

ó

Modelo Univariado

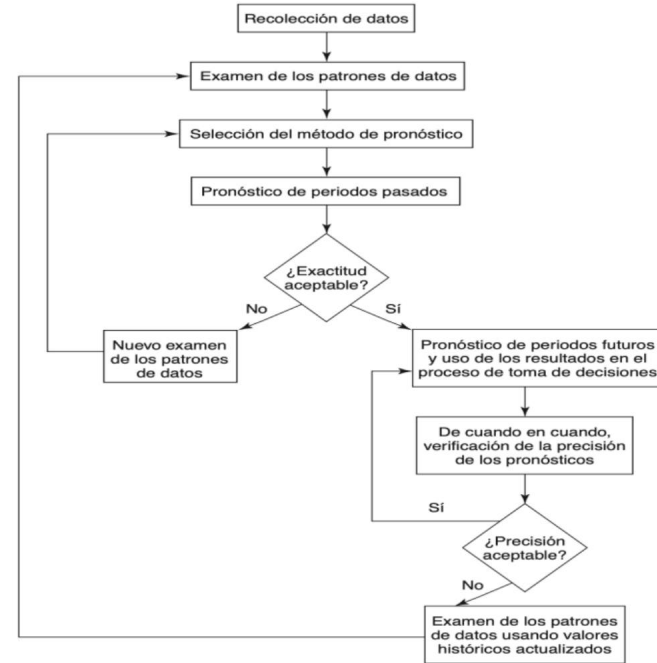
$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_{t-1} + \dots + \beta_k Y_{T-t} + \dots + \beta_k X_{T-t}$$

Modelo Multivariado

- En que  $t$  representa el tiempo.

*Objetivo:* lograr proyectar el futuro con el comportamiento histórico de la misma variable.

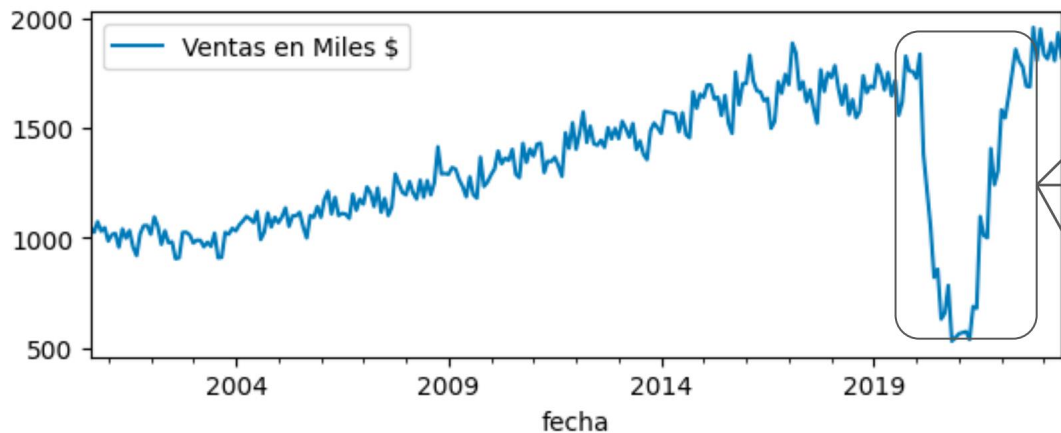
# Etapas de elaboración de un Modelo



Fuente: Pronóstico en los Negocios, 9º ed. Hanke, J & Wichern, D.



# pre-procesamiento.

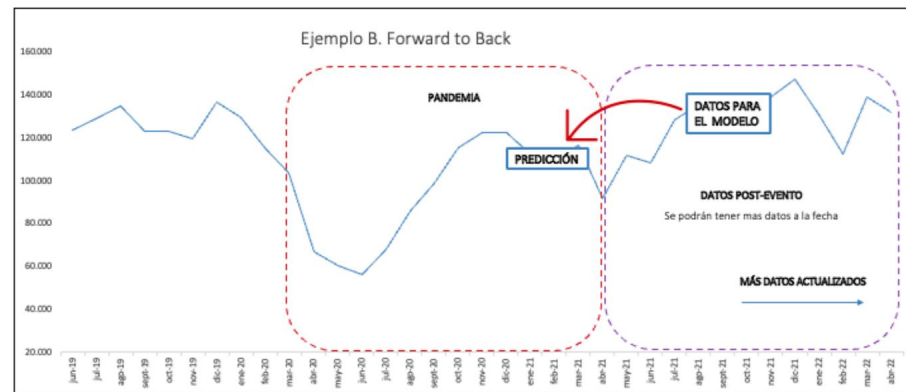
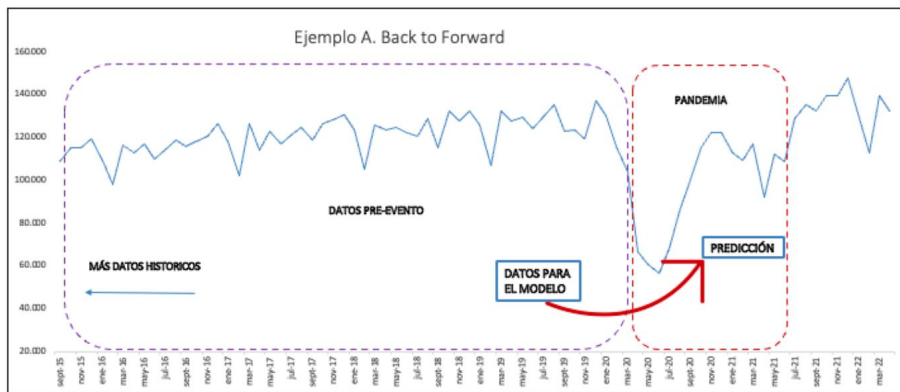


COVID19

evento  
exógenos

datos no  
representativos

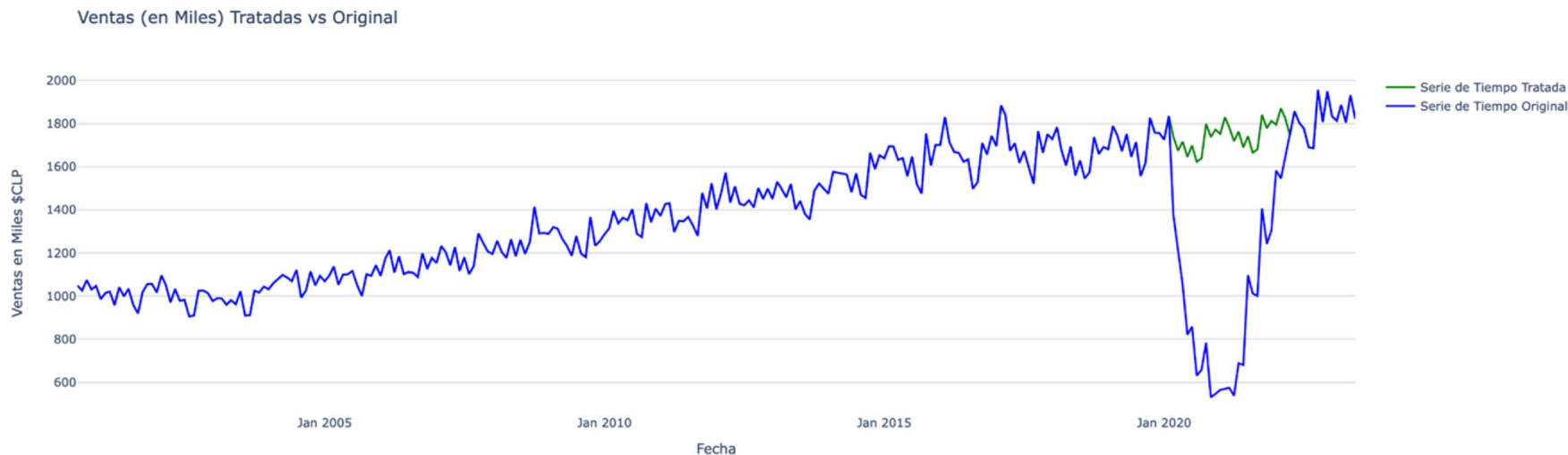
# pre-procesamiento.



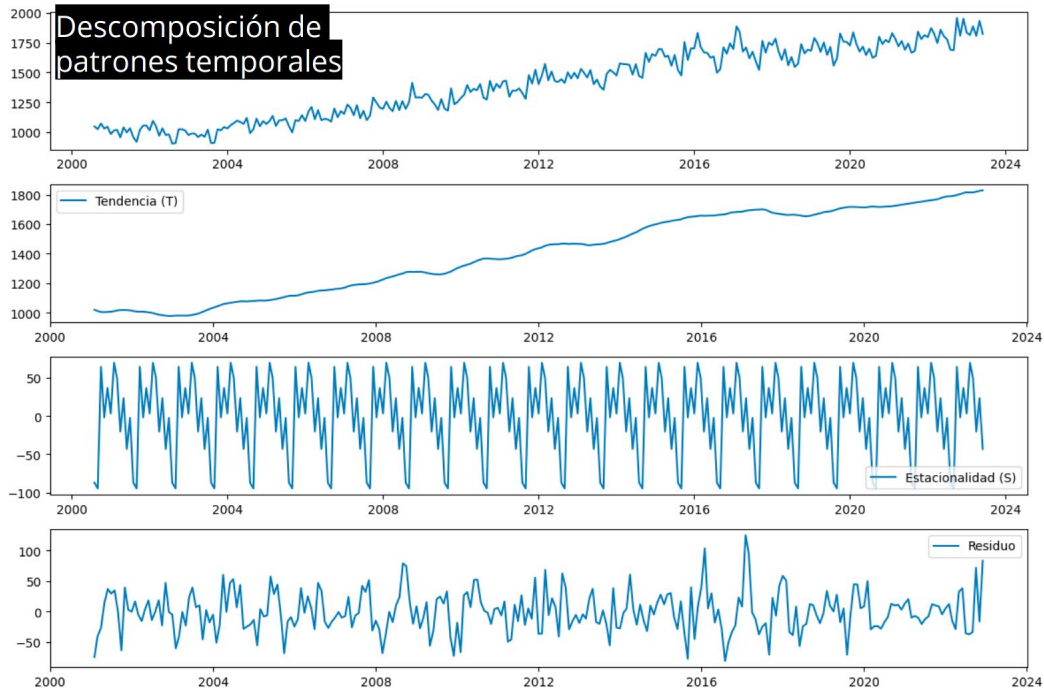
Tratamiento de Series Temporales por efecto COVID (u otro evento exógeno):

<https://fmansillaib.medium.com/>

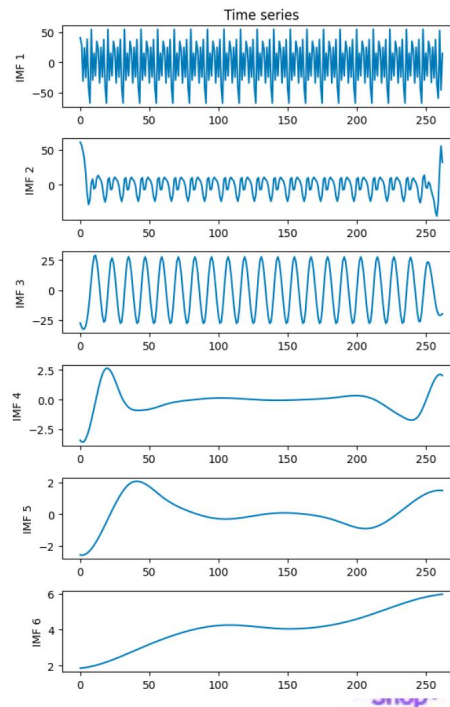
# pre-procesamiento.



# análisis previos



## Descomposición empírica de Modos (EMD)



# Técnicas de Pronósticos

Método	Patrón de datos	Horizonte de tiempo	Tipo del modelo	Datos mínimos requeridos	
				No estacionales	Estacionales
Simple	ST, T, S	S	TS	1	
Promedios simples	ST	S	TS	30	
Promedios móviles	ST	S	TS	4-20	
Suavizamiento exponencial	ST	S	TS	2	
Suavizamiento exponencial lineal	T	S	TS	3	
Suavizamiento exponencial cuadrático	T	S	TS	4	
Suavizamiento exponencial estacional	S	S	TS		2 × s
Filtración adaptativa	S	S	TS		
Regresión simple	T	I	C	10	
Regresión múltiple	C, S	I	C	10 × V	
Descomposición clásica	S	S	TS		5 × s
Modelos de tendencia exponencial	T	I, L	TS	10	
Ajuste de la curva S	T	I, L	TS	10	
Modelos de Gompertz	T	I, L	TS	10	
Curvas de crecimiento	T	I, L	TS	10	
Census X-12	S	S	TS		6 × s
Box-Jenkins	ST, T, C, S	S	TS	24	3 × s
Indicadores principales	C	S	C	24	
Modelos econométricos	C	S	C	30	
Regresión múltiple de series de tiempo	T, S	I, L	C		6 × s

Patrón de datos: ST, estacionario; T, de tendencia; S, estacional; C, cíclico

Horizonte de tiempo: S, corto plazo (menos de tres meses); I, mediano plazo; L, largo plazo

Tipo de modelo: TS, serie de tiempo; C, causal

Estacional: S, longitud de la estacionalidad

Variable: V, número de variables

Fuente: Pronóstico en los Negocios, 9° ed. Cap.4. Hanke. I & Wichern, D.

# Métricas de Evaluación

MAPE & RMSE

Sabiendo:

$$error\ total^2 = e_1^2 + e_2^2 + e_3^2 + \dots + e_i^2 = \sum e_i^2$$

Siendo:  $e_i = (y_i - \hat{y}_i)$

$$error\ total^2 = \sum e_i^2 = \sum (y_i - \hat{y}_i)^2$$

*Raiz Media Cuadratica (RMSE)*

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$$

Error Absoluto Porcentual Medio (MAPE)

$$MAPE = \frac{1}{n} \sum \frac{|y_i - \hat{y}_i|}{|y_i|}$$

# Métricas de Evaluación

Métrica de Negocio: SIGNO

Holt-Winter 6 meses

Ent: {'%MAPE': 3.82, '%sMAPE': 3.82, 'RMSE': 67.48, 'UTheil': 0.85, 'Ponderado': 15.66,  
Val: {'%MAPE': 2.61, '%sMAPE': 2.6, 'RMSE': 51.45, 'UTheil': 0.86, 'Ponderado': 11.85,

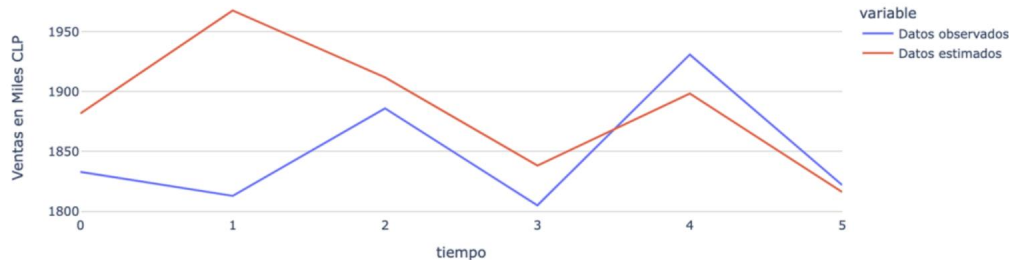
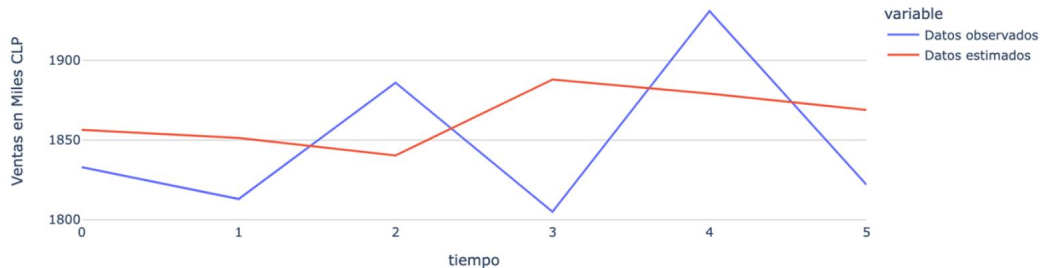
Holt-Winter 12 meses

Ent: {'%MAPE': 2.15, '%sMAPE': 2.15, 'RMSE': 37.95, 'UTheil': 0.42, 'Ponderado': 8.79,  
Val: {'%MAPE': 2.74, '%sMAPE': 2.67, 'RMSE': 69.76, 'UTheil': 0.91, 'Ponderado': 15.57,

# Métricas de Evaluación

Métrica de Negocio: SIGNO

Holt-Winter 6 meses



Holt-Winter 12 meses



# Métricas de Evaluación

Métrica de Negocio: SIGNO

Holt-Winter 6 meses

Val: {'%MAPE': 2.61, '%sMAPE': 2.6, 'RMSE': 51.45, 'UTheil': 0.86, 'Ponderado': 11.85, '% Acierto Signo': 40.0}

Holt-Winter 12 meses

Val: {'%MAPE': 2.74, '%sMAPE': 2.67, 'RMSE': 69.76, 'UTheil': 0.91, 'Ponderado': 15.57, '% Acierto Signo': 60.0}

# Consideraciones

estrategias uso - aprendizaje de largo y corto plazo



Conozca más sobre STATA escaneando el código QR.

