



Series Temporales

**ANÁLISIS PASAJEROS
DEL SUBTE CABA**

Ariadna Garmendia

Trabajo Práctico

01

Elegir y "Armar" el dataset

Pre-procesamiento,
gráfico y descripción

02

Des. aditiva y tests ESTACIONARIEDAD

Modelo de 4 componentes, Adfuller,
KPSS, Autocorrelación

03

PREPROCESAMIENTO

Box-Cox, media y mediana móvil,
diferenciación

04

ANÁLISIS DE TENDENCIA

Tendencia cíclica por
cuadrados mínimos

05

06

07

08

TRANSFORMACIONES Y MODELOS

Modelado con Sarima,
selección del mejor modelo

ANÁLISIS DE RESÍDUOS

¿Es bueno el modelo?

ANÁLISIS ESPECTRAL

Periodograma

MODELADO CON LSTM

Modelado con Red Neuronal LSTM y
evaluación de resultados

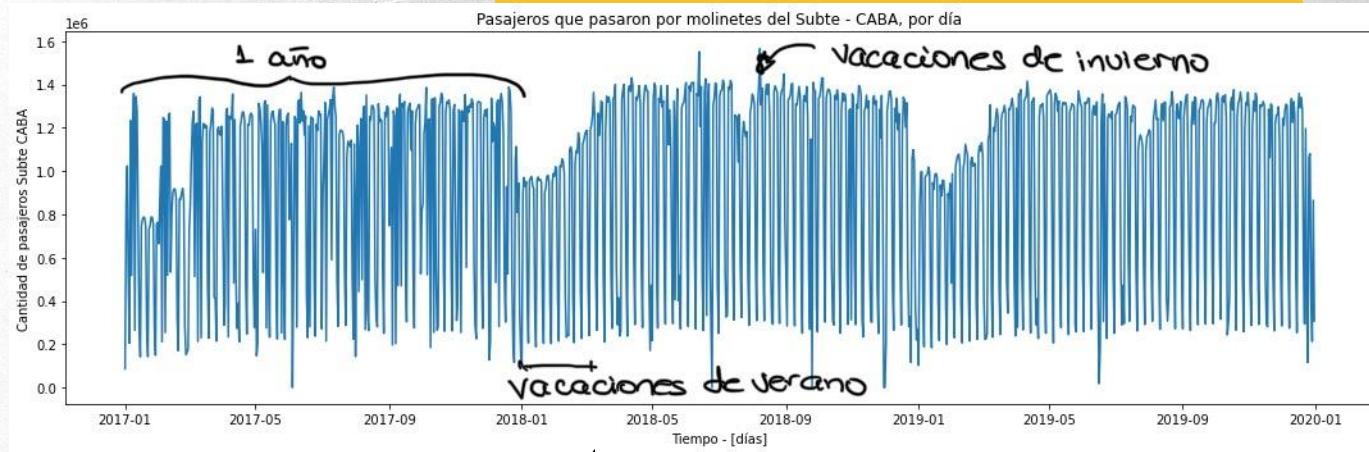
01

Exploración inicial

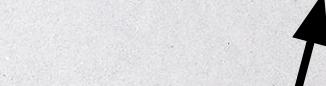
Armado de la serie temporal y análisis visual/descripción inicial



Total pasajeros de subte por día

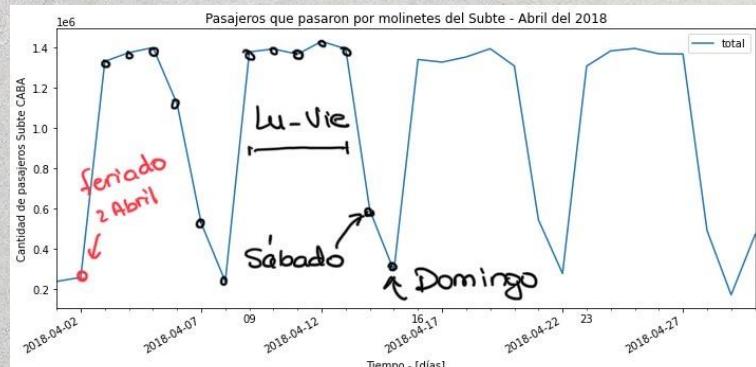


1 MES DE DATOS



3 AÑOS DE DATOS - 2017 A 2019

Obtenido combinando datos de 3 datasets, y
tomando totales diarios



B

H

<https://data.buenosaires.gob.ar/dataset/subte-viajes-molinete>



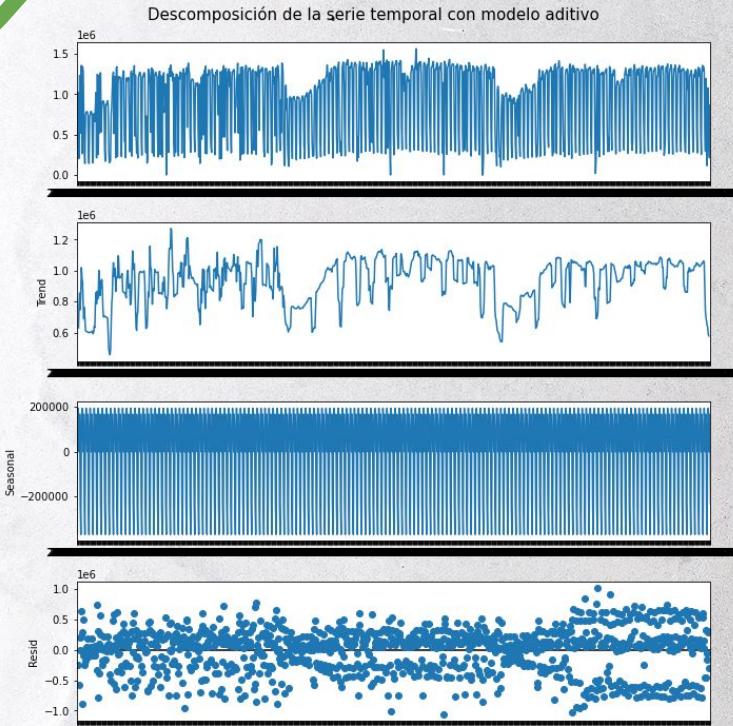
02

DESCOMPOSICIÓN ADITIVA Y ESTACIONARIEDAD

**Modelo de 4 componentes, Adfuller, KPSS,
Autocorrelación**

DESCOMPOSICIÓN ADITIVA Y TESTS DE ESTACIONARIEDAD

D



DICKEY-FULLER

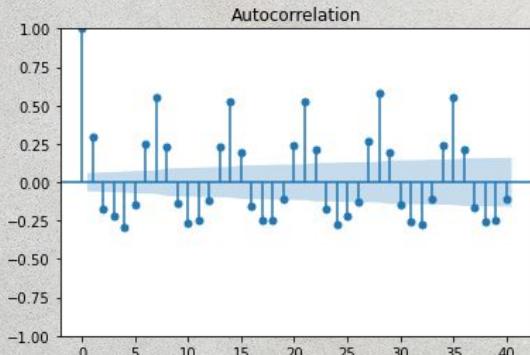
P_valor = 0.00008 (rechazo)

KPSS

P_valor = 0.1 (no rechazo H0)

Indican que es estacionaria, pero porque miran si hay caminante aleatorio

Tiene forma cíclica y
valores no nulos para
lags grandes



NO ES ESTACIONARIA

B

03 - 04

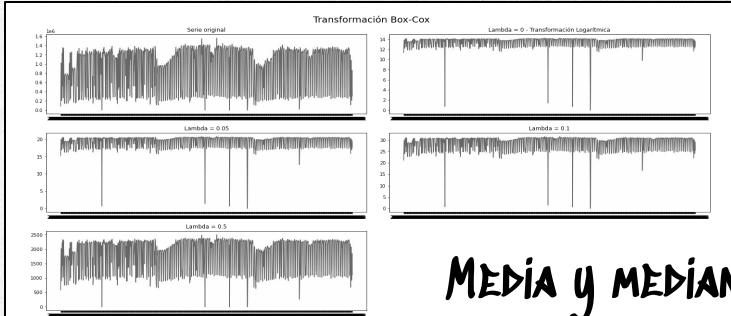
Pre-procesamiento y análisis de tendencia

**Box-Cox, media y mediana móvil,
diferenciación. Tendencia cíclica por
cuadrados mínimos**

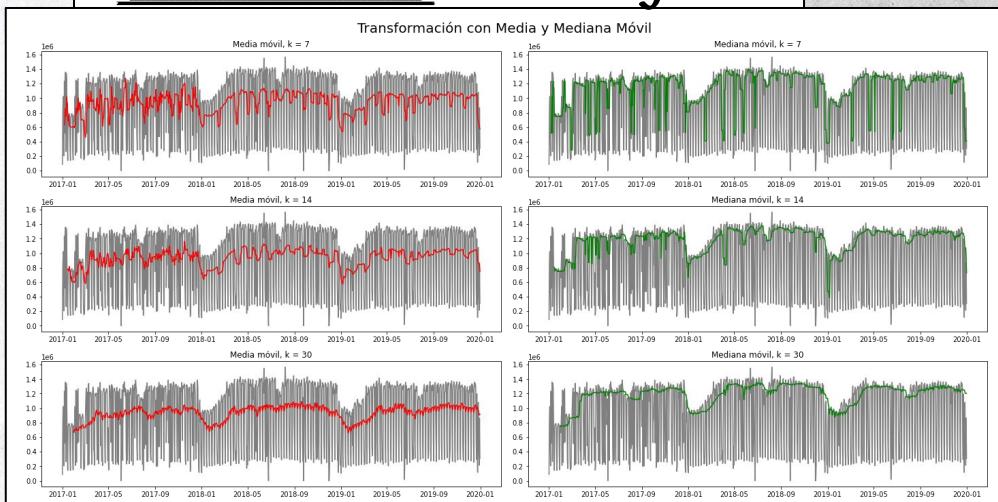


TRANSFORMACIONES

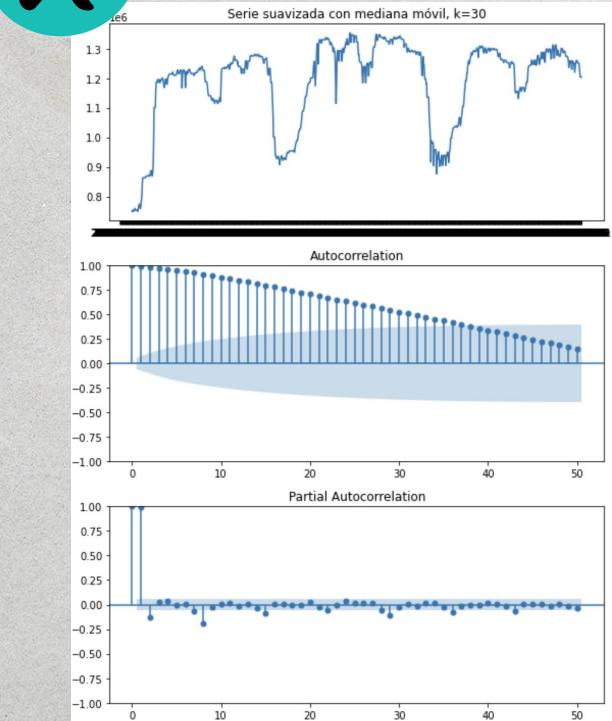
BOX-COX



MÉDIA Y MEDIANA MÓVIL

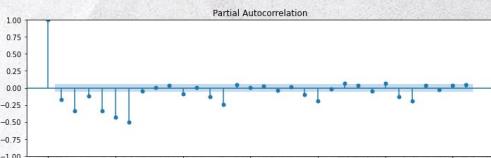
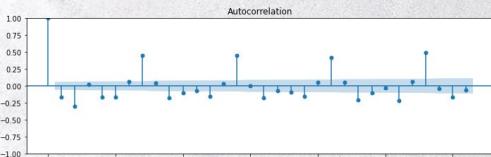
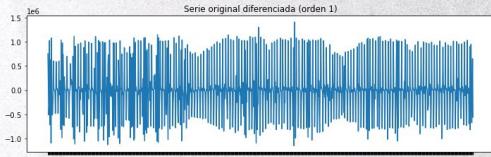


MEDIANA MÓVIL, $k=30$

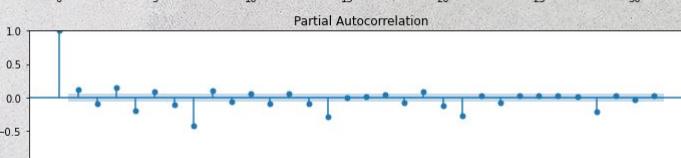
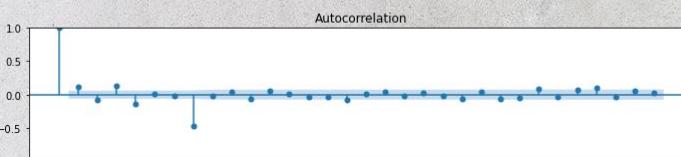
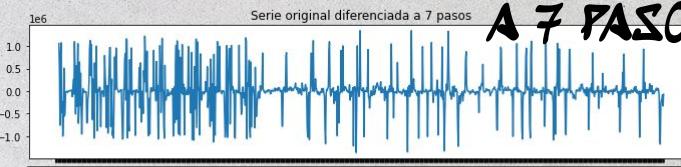


DIFERENCIACIÓN

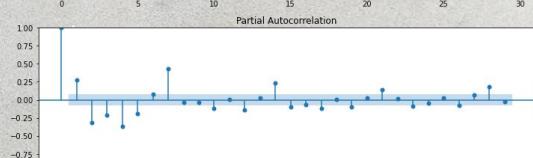
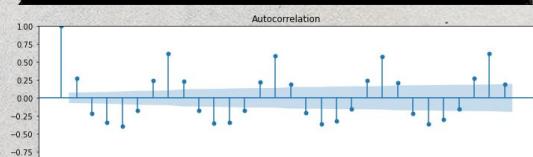
A 1 PASO



A 7 PASOS

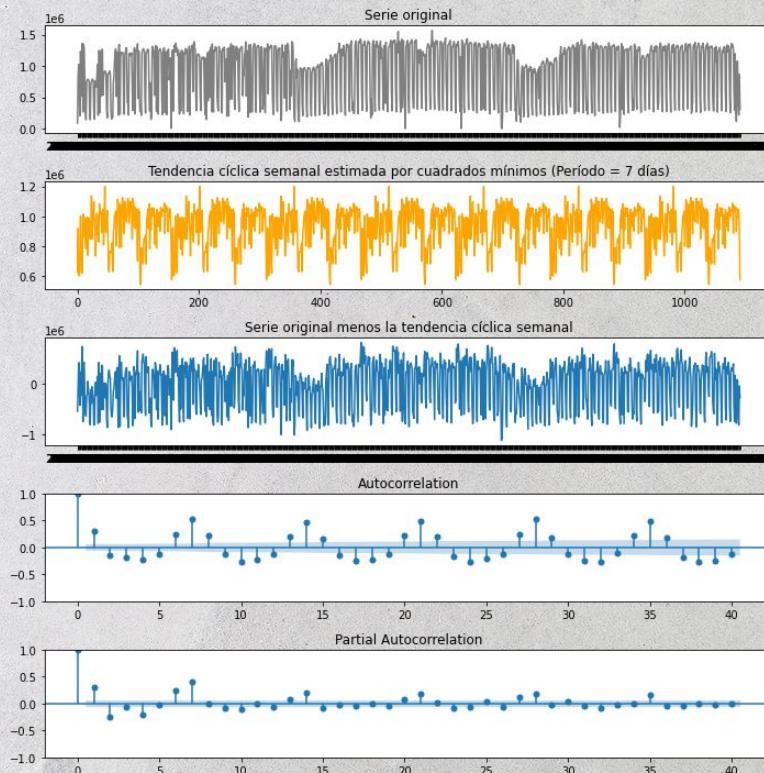


A 365 PASOS



B

TENDENCIA CÍCLICA POR CUADRADOS MÍNIMOS



B

05 - 06

Transformaciones y modelos. Análisis de residuos

Modelado con Sarima, selección del mejor
modelo



MODELOS

D

AIC

BIC

SARIMA
 $(p,q,d)=(1,1,1)$
 $(P,D,Q,s)=(1,1,0,7)$

23320.177

23340.019

SARIMA
 $(p,q,d)=(1,1,1)$
 $(P,D,Q,s)=(3,1,0,7)$

23230.016

23259.778

SARIMA
 $(p,q,d)=(1,0,1)$
 $(P,D,Q,s)=(1,0,0,7)$

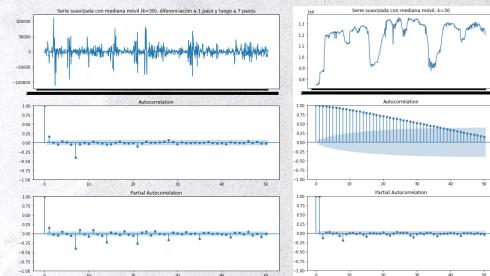
23138.011

23157.883

2

3

1



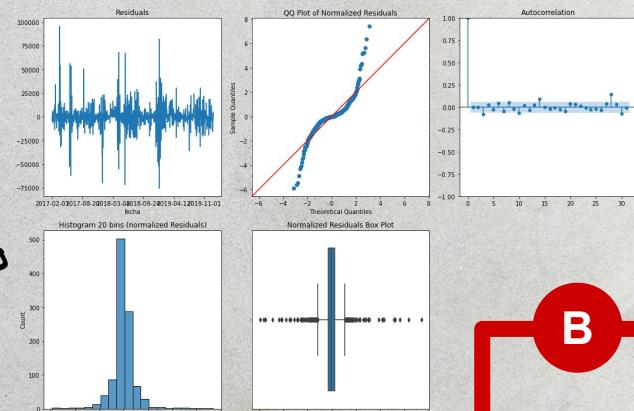
¿ES BUENO EL MODELO?



	coef	std err	z	P> z	[0.025	0.975]
ar.L1	0.9997	0.001	1616.628	0.000	0.999	1.001
ma.L1	0.2022	0.013	15.287	0.000	0.176	0.228
ar.S.L7	0.3452	0.018	19.174	0.000	0.310	0.380
sigma2	1.88e+08	1.58e-11	1.19e+19	0.000	1.88e+08	1.88e+08
<hr/>						
Ljung-Box (L1) (Q):			0.01	Jarque-Bera (JB):		6509.82
Prob(Q):			0.92	Prob(JB):		0.00
Heteroskedasticity (H):			0.51	Skew:		0.49
Prob(H) (two-sided):			0.00	Kurtosis:		15.09
<hr/>						



EL RUIDO
NO ES
GAUSSIANO



B



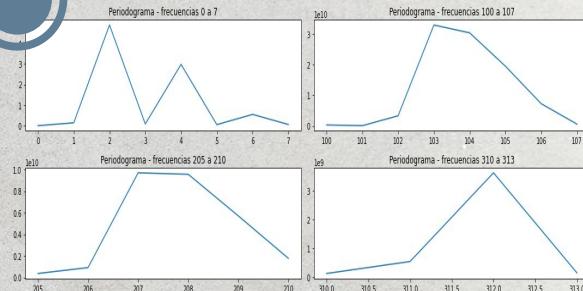
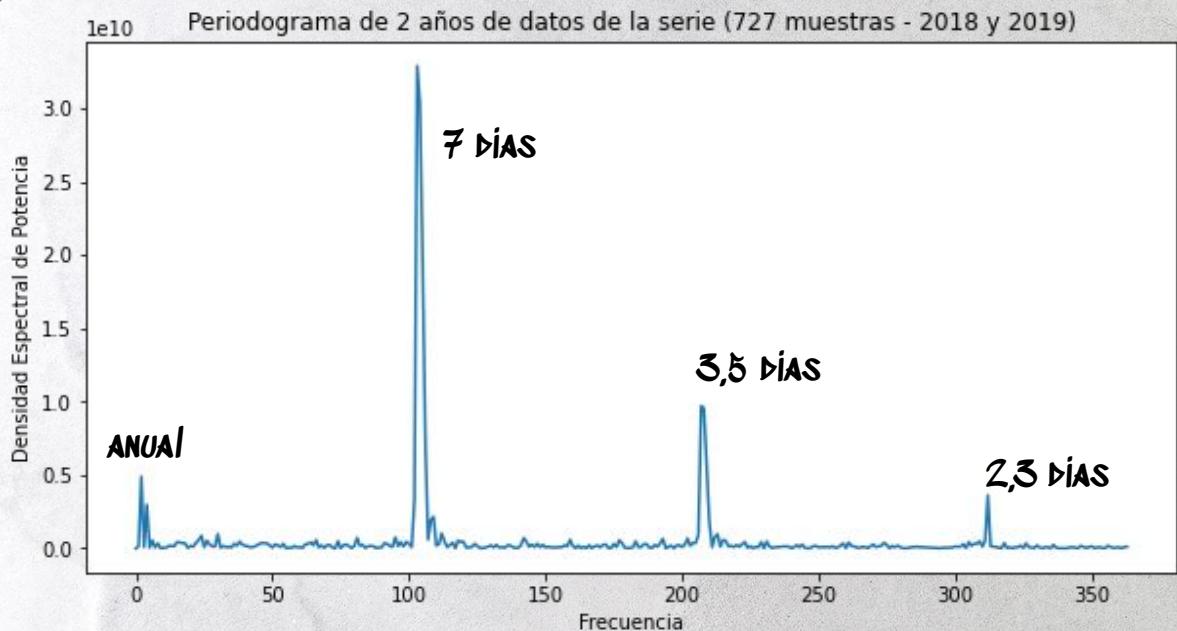
07

Análisis espectral

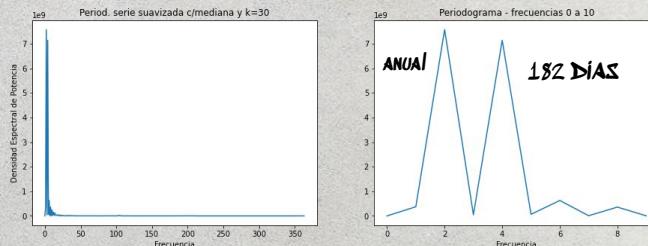
Periodograma

Periodograma

D



SERIE SUAVIZADA



E

08

Modelado con LSTM

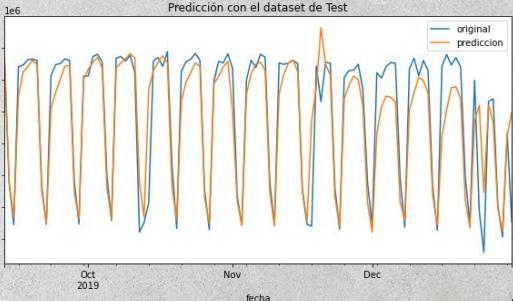
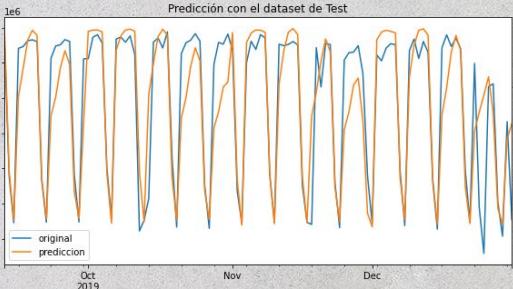
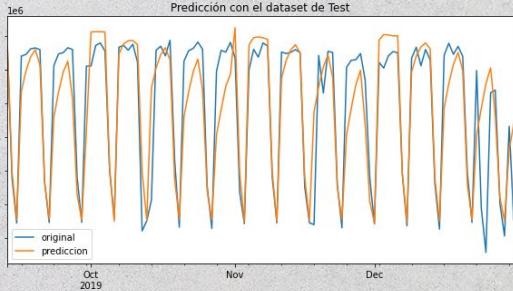
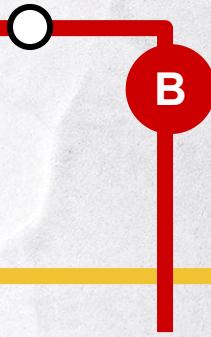
Modelado con Red Neuronal LSTM y evaluación de resultados



Con red neuronal

TRAIN / VALID / TEST
80 / 10 / 10%
872 / 110 / 110

LSTM



Sin encoding especial

Mean Absolute Error: 138394.19
Root Mean Squared Error: 212893.15
R^2 Score: 0.759

Con feriados

Mean Absolute Error: 135925.77
Root Mean Squared Error: 207553.5
R^2 Score: 0.771

Con feriados y encoding cíclico

Mean Absolute Error: 116624.1
Root Mean Squared Error: 187330.8
R^2 Score: 0.813

H

Conclusiones

- Para el análisis de estacionariedad, es importante utilizar varias herramientas y entender realmente qué es lo que busca cada test.
- El tiempo y análisis que lleva el modelado con Sarima es significativo en comparación con el modelo LSTM.
- Los resultados con LSTM son superiores, y en términos de forecast no tienen la desventaja de los procesos Arima, en los que el error crece indefinidamente.
- En contraste, implementar y entrenar el modelo Sarima es extremadamente sencillo y rápido.
- El periodograma es una herramienta muy útil para validar la inspección visual.



Gracias!

Preguntas?



7

CREDITS: This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#), and infographics & images by [Freepik](#)