

Universidad del Valle de Guatemala  
Departamento de Matemática  
Licenciatura en Matemática Aplicada

**Estudiante:** Rudik Roberto Rompich  
**Correo:** rom19857@uvg.edu.gt  
**Carné:** 19857

CC3066 - Data Science I - Catedrático: Luis Furlan  
27 de septiembre de 2021

---

## Laboratorio 7 - Series de tiempo

### Instrucciones:

Para este ejercicio, trabajarán con 4 conjuntos de datos:

1. daily-total-female-births.csv
2. shampoo.csv
3. monthly-car-sales.csv
4. monthly-mean-temp.csv

Cada uno de estos conjuntos representan series de tiempo, pero muestran diferentes características relacionadas con la tendencias, estacionalidad, período, etc. (ver presentación de clase)

Hemos visto 5 métodos diferentes para predecir con series de tiempo, a decir:

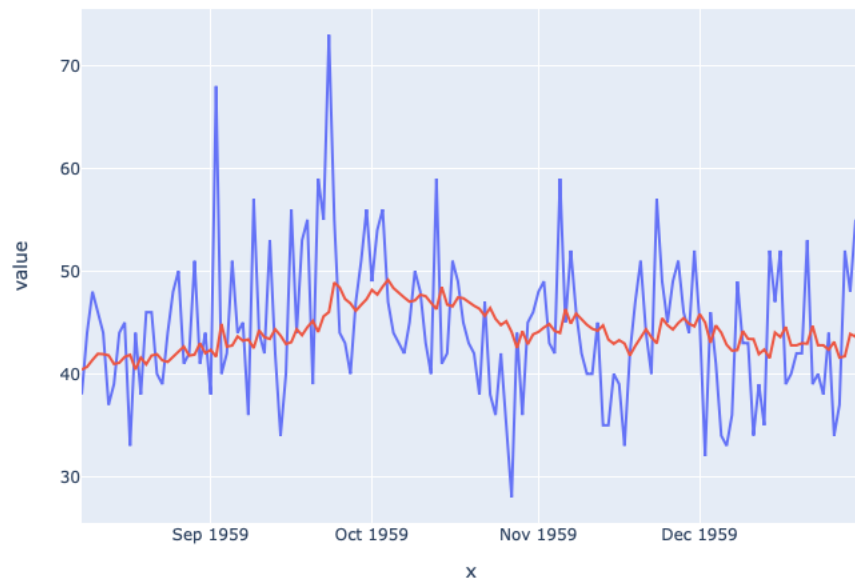
1. Promedio (para usar como base de referencia)
2. SARIMAX
3. Alisamiento exponencial (Winter-Holt)
4. Red Neuronal
5. FB Prophet

En este laboratorio, deben ejecutar los 5 métodos para cada uno de los conjuntos de datos arriba indicados.

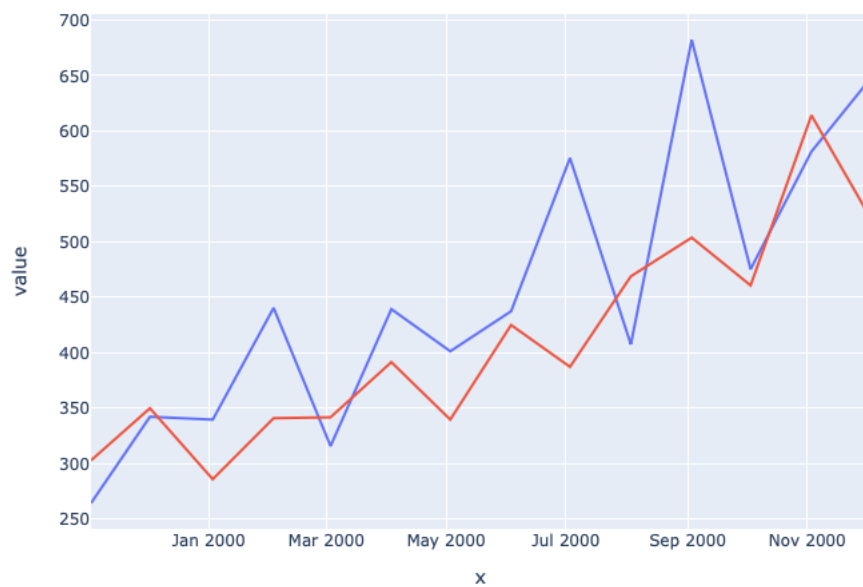
**Problema 1.** *Utilizando la prueba del RMSE (raíz cuadrada de la media de los errores al cuadrado). Generen gráficas para el caso que dio el mejor resultado de cada método en cada conjunto de datos.*

**Solución.** Tenemos los siguientes casos:

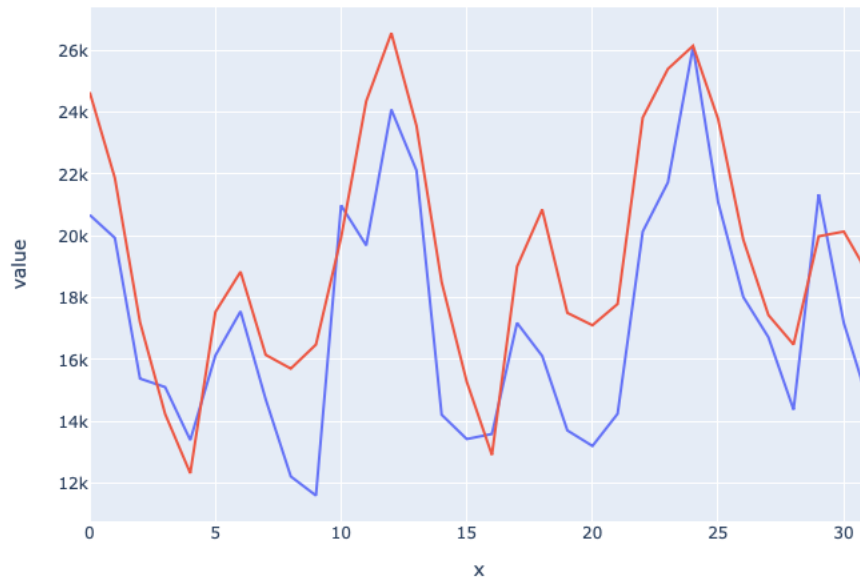
### 1. daily-total-female-births: SARIMAX.



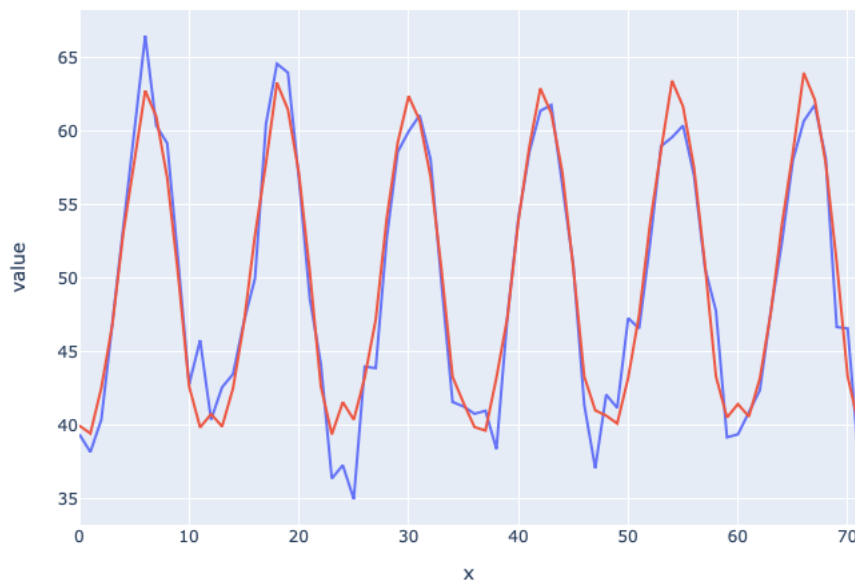
### 2. shampo: SARIMAX.



### 3. monthly-car-sales: Prophet.



#### 4. monthly-mean-temps: Prophet.



□

**Problema 2.** *¿Qué métodos logran captar mejor las tendencias y las variaciones estacionales?*

**Solución.** Eso dependerá mucho de la base de datos proporcionada. En resumen, los métodos que mejor capturan las tendencias y variaciones estacionales son: Prophet y redes neuronales. Aunque en algunos casos SARIMAX y Winter-Holt, también resultan muy efectivos. □

**Problema 3.** *Generen una tabla comparativa que muestre los RMSE más bajos para cada método con cada conjunto de datos. En la tabla indiquen qué método fue mejor para cada caso de datos. Por ejemplo:*

	Ninguna tendencia	Tendencia	Estacionalidad	Tendencia y estacionalidad
Método/Datos	Nacimiento	Shampoo	Temperatura	Carros
Promedio	6.37	113.15	5.14	3647
SARIMA	6.16 (Ganador)	62.83 (Ganador)	2.27	2600
Holt-Winter	6.74	97	4.58	3635
Red Neuronal	6.7	115	2.20 (Ganador)	2091 (Ganador)

**Solución.** Los resultados obtenidos son los siguientes:

Cuadro 1: Comparación de RMSE

Conjunto de datos/ Método	daily-total-female-births	shampoo	monthly-car-sales	monthly-mean-temp
Promedio	7.51	188.54	3676.92	5.91
SARIMAX	6.98	88.27	3134.12	4.39
Winter-Holt	7.21	101.75	3706.02	5.33
Red Neuronal	8.05	257.61	3481.97	5.063
Prophet	7.46	304	2839.13	2.19

□

**Problema 4.** Ahora apliquen el FB Prophet con cada uno de los conjuntos de datos. Compare los resultados con los de los métodos anteriores. ¿Hay algún ganador claro, entre todos los métodos?

**Solución.** Nuevamente, la respuesta dependerá del conjunto de datos de por sí; en la base de datos de carros y de temperaturas, Prophet es un claro vencedor. Sin embargo, en las primeras dos bases de datos de cumpleaños de mujeres y de shampoo, Prophet se queda bastante rezagado. Así que los mejores algoritmos están entre SARIMAX y Prophet. □

**Problema 5.** Escriban sus conclusiones sobre lo aprendido en este módulo sobre series de tiempo. ¿Cuál es el mejor procedimiento para resolver un problema de predicción de series de tiempo?

**Solución.** Algunas conclusiones:

1. No existe un método totalmente dominante, ya que todos tienen fluctuaciones y principalmente sus rendimientos se basan en la base de datos proporcionada.
2. Para comprender mejor que algoritmo es mejor es necesario tener un conocimiento teórico del tema, ya que existen diversidad de parámetros que poseen los métodos que podrían ser utilizados y resultar bastante efectivos.
3. El modelo de SARIMAX pareciera ser bastante efectivo al colocarle sus parámetros correctos por medio del Grid Search, por lo cual es de suma utilidad implementar este método.

□