

## Laboratorio 2. Series de Tiempo.

### INSTRUCCIONES:

Se le provee el conjunto de datos: GlobalLandTemperaturesByCountry.csv

Esta hoja de trabajo se realizará en cuartetos (**grupos de 4 personas**). Para que se pueda calificar su laboratorio debe estar inscrito en algún grupo de Canvas.

### DESCRIPCIÓN DEL DATASET

El conjunto de datos de importación de combustibles contiene 577,462 observaciones de 4 variables. Representa las temperaturas promedio mensuales, desde 1750, para muchos lugares del mundo. El documento de donde se extrajo la información lo puede encontrar en el siguiente vínculo: <http://berkeleyearth.org/data/>

Más información en:

- <http://berkeleyearth.org/about/>

### EJERCICIOS

1. Haga un análisis exploratorio de los datos que se le presentan (AverageTemperature), se sugiere explorar el comportamiento de la variable y si está distribuidas normalmente, en caso de ser continuas. Meses en los que más/menos temperatura hay, comportamiento en los últimos x años, etc.
2. Haga una serie univariante por la variable AverageTemperature:
  - a. Especifique Inicio, fin, y frecuencia.
  - b. Haga un gráfico de la serie y explique qué información puede obtener a primera vista.
  - c. Descomponga la serie en sus componentes. Teniendo en cuenta el diagrama de la serie y sus componentes discuta si es posible hablar de estacionalidad.
  - d. Determine si es necesario transformar la serie. Explique.
  - e. Una vez analizada la serie, elija los parámetros p, q y d del modelo ARIMA o SARIMA
  - f. Haga varios modelos SARIMA, y diga cuál es el mejor de ellos para estimar los datos de la serie.
  - g. Haga un modelo usando el algoritmo Prophet de Facebook. Compárelo con los modelos del inciso anterior. ¿Cuál funcionó mejor?
3. Haga una predicción de los valores de la serie para los últimos 3 años. Explique qué tan buena es su predicción.
4. Haga una predicción, tres años al futuro, usando los años anteriores y determine qué tan apegada es a la realidad teniendo en cuenta los meses pasados del año actual.
5. Basado en los resultados obtenidos, diría Ud que hay evidencia para corroborar la tan discutida aseveración de que el Cambio Climático y el Calentamiento Global es real

6. Repita los 5 pasos anteriores pero esta vez solo utilizando los datos para Guatemala.

### EVALUACIÓN

**NOTA:** La evaluación de cada integrante del grupo será de acuerdo con sus contribuciones al trabajo grupal

**(25 puntos)** Análisis exploratorio:

- Se elaboró un análisis exploratorio, hay gráficos explicativos y análisis que permiten comprender el conjunto de datos.
- Se crearon las series de tiempo correspondientes a los datos de la variable que representan la temperatura promedio.
- Para cada una de las series se informa inicio, fin y frecuencia.

**(30 puntos)** Análisis de la serie de tiempo

- Para cada una de las series creadas se analiza:
  - o El gráfico de la serie y su descomposición en componentes.
  - o Si la serie presenta estacionalidad o no y que implica que sí tenga.
  - o Si la serie presenta tendencia o no y esto que significa.

**(25 puntos)** Generación de modelos

- Para cada una de las series creadas:
  - o Se determinan los valores de los parámetros p, q, y d.
  - o Se explica la elección de los parámetros y de los modelos en caso de que se propongan varios modelos. Se deben explicar los parámetros, aunque sean propuestos de forma automática por R (en caso de usar este lenguaje).
  - o Se genera un modelo con el algoritmo Prophet.

**(20 puntos)** Predicción con los modelos generados.

- Para cada una de las series creadas:
  - o Se crean los conjuntos de entrenamiento y prueba siguiendo las instrucciones.
  - o Se explica que tan bueno es el modelo prediciendo las temperaturas promedio para el conjunto de prueba.
  - o Se comparan los modelos generados con ambos algoritmos.
  - o Se discute si los resultados apoyan la tesis del Cambio Climático y Calentamiento Global

### MATERIAL A ENTREGAR

- Archivo .pdf con el informe que contenga, los resultados de los análisis y las explicaciones.
- Link de Google drive donde trabajó el grupo.
- Script de R (.r o .rmd) o de Python que utilizó para hacer su análisis exploratorio y predicciones.
- Link del repositorio usado para versionar el código.