

# Actividad 6

Hannah Borrego

Febrero del 2021

Una breve introducción de la Actividad desarrollada y algunas características que hayas notado en el análisis inicial del sitio de estudio. Comentarios generales de la información analizada y realizada en Google Colab.

## 1 Introducción

En este trabajo continuamos trabajando con los datos de metodología, Se selecciona un periodo de datos donde no haya huecos. Apartir de los años 90 en adelante. Yo había estado trabajando con dos periodos de 30 años casi completos y el segundo periodo abarcaba desde 1980 en adelante. En esta ocasión decidí modificarlo y acotarlo para un periodo de 1990 en adelante, para este trabajo se necesitaban datos desde ese año en adelante así que dentro de ese periodo seleccioné otro más específico de 1990-1995, en donde no había huecos como si los había en los años 2000.

### 1.1 Desarrollo

El objetivo de esta actividad de el pronóstico de las series de tiempo, se necesita tener una serie de tiempo estacionaria. Lo que significa promedio y desviación estándar constantes o de mínima variabilidad. Para ello se emplean métodos para comprobar que la serie sea estacionaria o no. La prueba de Dickey-Fuller aumentada, es una prueba de hipótesis estadística, que nos ayudará a rechazar la hipótesis nula de si una serie de tiempo es no-estacionaria o no, en base al valor-p.

Se trabaja con algunos métodos para eliminar la Tendencia y la Estacionalidad, también transformar la serie de tiempo en una serie de tiempo estacionaria forma parte del proceso para completar la actividad.

### 1.2 Resultados de la actividad

¿Que puedes agregar de las series de temperatura que analizaste?

Las temperaturas máximas rondan entre los 12 y los 21 grados en el año 1990 y las temperaturas mínimas entre -1 y 3 grados en Santa Ana, Sonora. Después de hacer la serie estacionaria se entiende que no habrá variación o al menos muy poca durante el tiempo. Al menos entre 1990-1995 periodo elegido.

Respecto a la estacionaridad de tus series. ¿Que procedimientos seguiste para obtener una serie estacionaria?

Primero se usó La prueba de Dickey-Fuller aumentada, que es una prueba de hipótesis estadística, la cual nos ayudará a rechazar la hipótesis nula de si una serie de tiempo es no-estacionaria o no, en base al valor-p.

1.- Primero se selecciona el periodo de tiempo sobre el que se va a trabajar. Yo elegí 1990-1995.

2.- Despues se elige si es con temperatura máxima o mínima.

3.- El valor de p tiene que ser menor a 0.05, para que sea una serie estacionaria, en mi caso, en la temperatura máxima si lo era pero aún así decidí seguir con el método para confirmar que lo era. En la temperatura mínima no se podía asegurar que fuera estacionaria y de igual manera seguí con el proceso.

4.- Luego se selecciona la función log base 10 para cambiar la escala de la gráfica.

5.- Se calcula el promedio móvil de 365 días.

6.- Se hace una resta de funciones de la siguiente forma:

$$ts_{testlog_{movingave}}diff = ts_{testlog} - moving_{avg}$$

Es la diferencia entre la gráfica de la temperatura y el promedio móvil.

7.- Probamos ahora la estacionaridad de la nueva serie de tiempo escalada en Log.

8.- Como aún no se puede asegurar se repite el cálculo usando promedios móviles exponenciales. Con ayuda de función de pandas `pd.ewm()` y 30 datos.

9.- Se obtiene la diferencia entre el promedio exponencia y la serie de log de la temperatura máxima.

10.- Probamos ahora la estacionaridad de la nueva serie de tiempo escalada en Log y usando EWM.

11.- Ahora podemos rechazar la hipótesis nula de que la serie de tiempo no es estacionaria y afirmar que la serie de tiempo es estacionaria con 99

Tambien nos podemos ir por el camino del método de la descomposición. Graficando estos 3 componentes tendencia, estacionalidad y un residuo.

12.- Eliminando la Tendencia y Estacionalidad, probamos si la serie Residuo es estacionaria o no.

13.- Ahora ya Podemos afirmar que es una serie de tiempo estacionaria con mas de un 99

Respecto a las tendencias de Tmax y Tmin. ¿Qué se observa?

La P en la temperatura máxima si era menor a 0.05 lo cual ya nos daba para afirmar que es una serie estacionaria apesar que de igual manera se llevó todo el procedimiento para hacerla estacionaria. En la temperatura mínima no P en un inicio era mayor a 0.05.

¿Que puedes decir en general sobre los datos de la estación que estuviste analizando?

Las temperaturas no han variado mucho alrededor de los años, Santa Ana es un lugar 5 grados mas o menos más frío que Hermosillo. No sube ni baja de una manera exorbitante en ninguno de sus años. Tampoco abunda la lluvia.

¿Qué limitaciones encontraste en tus datos?, ¿Vacíos?, ¿Cuál fue el periodo más largo que pudiste encontrar?.

Al inicio pensé en seleccionar el periodo de años más cercano a la actualidad que se pudiera así que elegí un periodo entre el 2000 y el 2005 que es hasta donde se terminan los datos. Pero al intentar con ese había muchos datos vacíos y decidí tomar periodos más atrás hasta que encontré el que me funcionó mejor. De 1990 a 1995.

### **1.3 Conclusión**

El trabajo me pareció un poco complejo sobre todo para encontrar la utilidad, creo que tendría que hacer más ejemplos con otras series de tiempo para ver como se van transformando los datos despues de todo el análisis llevado. Recomendaría poder ver más el cambio ocurrido en el pasado, hacer predicciones hacia adelante donde ya existen datos, compararlos y ver si las predicciones se cumplen o no. Se me ocurre que podría ser algo interesante para analizar.