



UNIVERSIDAD DE SONORA

DEPARTAMENTO DE FÍSICA

Actividad 4

Análisis exploratorio de datos en python

Pacheco Rodriguez Sharon Estivalez

Expediente:

219221708

05 de febrero de 2021

1. Introducción

En esta cuarta actividad realizamos un análisis y visualización de los datos con las bibliotecas *Matplotlib* y *Seaborn*, además de las ya utilizadas en anteriores actividades como *Pandas* y *Numpy*.

Los datos fueron observados y analizados mediante

- Histogramas
- Funciones de densidad de probabilidad
- Gráficas de caja
- Gráficas de barra
- Gráficas de líneas

Todas estas herramientas en la actividad nos mostraron cómo visualizar de una mejor manera el comportamiento de un grupo de datos y poder utilizarlo en próximas actividades. Asimismo conocer y comprender poco a poco el lenguaje de *python* y su impacto, aterrizando de excelente manera sus usos.

2. Comentarios generales

En esta actividad fue la primer ocasión donde trabajé (y quizá la mayoría de mis compañeros también) con este tipo de gráficas de manera más formal con un lenguaje de programación. Esto lo hizo más interesante, el poder notar inmediatamente los resultados de lo que ingresabas en el código. Me pareció un poco más sencillo trabajar con *Seaborn* que con *Matplotlib*.

Primero realizamos un análisis de datos exploratorio, viendo su estructura, los datos que faltan, detectando anomalías y valores extremos, seleccionando las variables importantes, etc. Posteriormente creamos las gráficas mencionadas en la introducción con las variables importantes para poder entender el comportamiento de los datos del sitio.

De las gráficas y datos mostrados me llamó mucho la atención las temperaturas mínimas y máximas en Canatlán, ya que las cifras muestran que pueden variar de grados bajo cero hasta más de 40°C. Siendo los valores más frecuentes para la máxima entre 20°C y 30°C. Mientras que para la mínima son entre 5°C y 10°C.

A lo largo de los años también se encuentra en las cifras que la precipitación varía mucho, subiendo y bajando constantemente, cosa que también se ve en los meses, siendo de junio a octubre donde hay mayor precipitación en el año.

3. Retroalimentación

El tema de análisis exploratorio de datos me pareció muy interesante, nunca imaginé que se le invertiría tanto tiempo a esta parte del proceso; fue muy bueno conocer todas estas herramientas, me pareció muy útil e imagino que continuaremos utilizándolas en futuras actividades.

El reto de elaborar la actividad fue algo tedioso en el sentido de que eran muchas cosas nuevas, más instrucciones que en anteriores ocasiones; pero al igual que en actividades pasadas, el hecho de tener como apoyo el ejemplo del maestro en *Colab* facilitó todo de gran manera.

Lo que siento que más se me dificultó fue estar escribiendo y reteniendo tantos comandos o instrucciones nuevas, esto me hacía detenerme y por lo mismo dar más tiempo en la elaboración de la actividad.

No siento que algo haya sido aburrido, pero si recomendaría para la actividad detenernos un poco más en cada comando/instrucción para comprenderlo realmente y no memorizarlo de alguna manera. El grado de complejidad fue intermedio, considerando que teníamos toda la información en el archivo del profesor.