

UNIVERSIDAD DE SONORA

Lic. Fisico-Matematico



Análisis Exploratorio de Datos en Python (EDA, Exploratory Data Analysis)

Fisica Computacional

Alumno: Gerardo A. Carreón Castro

Maestro: Carlos Lizarraga Celaya

4to Semestre

05 de Febrero de 2021

1 Introducción a la actividad y primeras impresiones

El objetivo principal de esta actividad es continuar con el uso de la biblioteca Pandas de Python describiendo y analizando los datos proporcionados por la estación meteorológica seleccionada en la actividad 1 pero ahora de manera más exhaustiva ya que buscamos conocer la estructura, contenidos y características del conjunto de datos que se desean analizar. Entre algunas características que se desean analizar son:

- Estructuras/patrones en los datos
- Número de datos faltantes
- Detección de datos anómalos y valores extremos
- Extraer y seleccionar variables importantes
- Probar teorías de comportamiento de datos

Espero que durante esta actividad aprendamos a graficar todo tipo de gráficas que nos ayuden a analizar los datos seleccionados y además poder manipular y diseñar dichas gráficas a nuestro criterio para poder elaborar gráficas más específicas y adecuadas para nuestros objetivos.

2 Bibliotecas

Las bibliotecas utilizadas en esta ocasión fueron las ya vistas anteriormente, Pandas, Numpy, Matplotlib e introducimos una nueva la cual es Seaborn biblioteca necesaria para poder graficar DataFrames, esto con la finalidad de poder manipular más fácilmente los datos contenidos en el archivo original.

De estas bibliotecas ya tenemos conocimientos y experiencia utilizandolas solo de la única que se nos dará una breve introducción en esta actividad es la de Seaborn sin embargo, ya que es una biblioteca derivada de Matplotlib está seguramente resulte muy fácil de dominar ya que las bases de Matplotlib ya las conocemos en su totalidad. Por lo tanto, no espero una dificultad en el uso de las bibliotecas de esta actividad.

3 Opinión

En mi opinión esta actividad aunque en su mayor parte fue sencilla no puedo negar que si se me complico en ciertas secciones ya que muchos de los comandos que se utilizaron para realizar las distintas gráficas no los conocía y me fueron difíciles memorizarlos por lo que fue gracias a los comentarios del profesor que pude, en su totalidad, comprender qué es lo que hacía cada comando y cómo debía construirse de acuerdo a nuestros datos. De ahí en adelante compilar las gráficas fue muy sencillo una vez se conocía la estructura de los comandos para realizar las distintas gráficas y los resultados obtenidos fueron sin duda muy satisfactorios dándonos gráficas muy detalladas y con el diseño adecuado para su comprensión.