

Minería de datos: Tarea 4

Profesor Guillermo Herrera

El objetivo de esta tarea es practicar lo aprendido en clase sobre dataframes en Python. Cómo hacer consultas, estadísticas y gráficas a partir de una base de datos.

La fecha de envío es el lunes 19 de septiembre antes de las 11:59 am con los equipos formados previamente.

Solo deben enviar un reporte en formato .pdf y un script .py ambos en una carpeta comprimida. El reporte debe cumplir con las especificaciones que hemos platicado en clase. No deben reenviar la base de datos.

Actividad: Análisis de datos sísmicos

En el correo encontrarán una base de datos con información de sismos en México. Los datos son reales y fueron descargados de la página <http://www.ssn.unam.mx>

No deben descargar más datos, todos deben utilizar exactamente la misma base de datos.

Ejercicio 1. (2 puntos)

Cargar la base de datos en Python y crear un Dataframe con toda la información disponible (sin contar las primeras y últimas líneas). Para cargar la base de datos deben utilizar los comandos de lectura de texto como se vieron en clase. NO deben cargar la base directamente como un csv. Aunque sí es posible leerlo como un csv en Python ese no es el objetivo de la tarea.

A partir del dataframe van a obtener información.

Ejercicio 2. (1 puntos)

Deben buscar los 10 sismos con magnitud mayor dentro de la base de datos y desplegar toda la información disponible de ellos.

Ejercicio 3. (1 puntos)

Para cada año en la base de datos calcular el promedio de las magnitudes de los sismos que se presentan. Después deben crear una gráfica por año con dichos valores.

Ejercicio 4. (1 puntos)

Generar un histograma por meses utilizando la base de datos completa. ¿Es cierto que en septiembre hay más temblores?

Ejercicio 5. (1 puntos)

Generar una gráfica de dispersión con todos los sismos utilizando su ubicación geográfica. Los sismos que tengan una magnitud mayor o igual a 5.0 deben ser marcados con color rojo. Los que tiene magnitud menor a 5.0 deben ser marcados con verde. Todos en la misma gráfica. Deben utilizar parámetros que hagan ilustrativa la gráfica.

Ejercicio 6. (1 puntos)

Generar una gráfica de dispersión con todos los sismos del año 2021 utilizando su ubicación geográfica. Los sismos que tengan una magnitud mayor o igual a 5.0 deben ser marcados con color rojo. Los que tiene magnitud menor a 5.0 deben ser marcados con verde. Todos en la misma gráfica. Deben utilizar parámetros que hagan ilustrativa la gráfica.

Ejercicio 7. (1 puntos)

Si el 11 de marzo de 2020 la OMS declaró la pandemia de COVID-19. ¿Cuántos sismos de magnitud mayor o igual 4.0 han ocurrido desde esa fecha hasta el final de la base de datos?

Finalmente deberán hacer un reporte en el que hablen sobre los sismos en México

Ejercicio 8. (4 puntos)

En el reporte deben tener 3 secciones, 1) introducción al tema 2) análisis de datos, donde describirán la base de datos y además debe: Agregar la información del ejercicio 2. La gráfica del ejercicio 3. Responder el ejercicio 4. Mostrar las gráficas de los ejercicios 5 y 6. Y finalmente la respuesta del ejercicio 7. Todo esto en un texto de lectura amena, sin cambios drásticos entre un párrafo y otro. 3) Una conclusión a partir de su análisis.

La calificación es sobre 10 puntos, aunque es posible obtener una suma de puntos mayor.

Para citar los datos háganlo de esta manera:

México. Universidad Nacional Autónoma de México, I. d. G., Servicio Sismológico Nacional. (2022). Catálogo de sismos. Extraído de <http://www2.ssn.unam.mx:8080/catalogo/>

El Identificador de Objeto Digital o DOI por sus siglas en inglés (Digital Object Identifier), es un código estandarizado que permite identificar contenido digital específico de manera única, permanente e independiente de las direcciones electrónicas. Su uso es común en las citas de publicaciones especializadas. El DOI para los datos extraídos de este catálogo de sismos es el siguiente:

[10.21766/SSNMX/EC/MX](https://doi.org/10.21766/SSNMX/EC/MX)

¿Cómo citar el uso del catálogo de sismos?

En cualquier publicación, el uso de datos del catálogo de sismos del Servicio Sismológico Nacional debe ser citado en el formato que cada caso requiera:

APA

México. Universidad Nacional Autónoma de México, I. d. G., Servicio Sismológico Nacional. (2022). Catálogo de sismos. Extraído de <http://www2.ssn.unam.mx:8080/catalogo/>

Copiar

IEEE

MLA

AER

(American Economic Review)

JGR

(Journal of Geophysical Research, AGU)

Los agradecimientos se deben redactar de la siguiente manera:

"El catálogo de sismos del Servicio Sismológico Nacional (México) es posible gracias a todo su personal y producto de los cálculos realizados por su área de Análisis e Interpretación de Datos Sísmicos."