

# Actividad 3

Thelma Mariana Vallejo Valdez

## 1 Introducción

Para esta actividad se realizo el manejo de una base de datos con ayuda de la biblioteca Pandas. Para poder desarrollar la actividad se tuvo que seguir ciertos pasos y asi preparar la informacion que se deseaba utilizar. Primero se define un DataFrame con los datos climatológicos de una estación climatologica. Una vez estudiados los datos y preguntas que se menciona en la actividad, pasamos a identificar los datos en el programa como lo es en que numero de renglos comienzan los datos y asi quedarnos con los datos que sean necesarios. Como tercer paso se pide que realizar ciertas instrucciones haciendo un analisis y tratamiento de los datos. Y para finalizar los datos manejados los convertimos a una variable que python comprenda, esto con una serie de herramientas que ofrece Pandas.

## 2 Pandas

Pandas es una biblioteca de software escrita como extensión de NumPy para manipulación y análisis de datos para el lenguaje de programación Python. En particular, ofrece estructuras de datos y operaciones para manipular tablas numéricas y series temporales.

Algunas características y funciones:

- Tipo de datos DataFrame para manipulación de datos con indexación integrada.
- Herramientas para leer y escribir datos entre estructuras de dato en-memoria y formatos de archivo variados.
- Alineación de dato y manejo integrado de datos faltantes.
- Reestructuración y segmentación de conjuntos de datos.
- Segmentación vertical basada en etiquetas, indexación elegante, y segmentación horizontal de grandes conjuntos de datos.
- Inserción y eliminación de columnas en estructuras de datos.
- Agrupación predefinida en la biblioteca lo que permite realizar cadenas de operaciones dividir-aplicar-combinar sobre conjuntos de datos.
- Mezcla y unión de datos.
- Indexación jerárquica de ejes para trabajar con datos de altas dimensiones en estructuras de datos de menor dimensión.
- Funcionalidad de series de tiempo: generación de rangos de fechas y conversión de frecuencias, desplazamiento de ventanas estadísticas y de regresiones lineales, desplazamiento de fechas y retrasos