

Actividad Manejo de bases de datos con Python - Indicaciones para realizar la actividad

A continuación, se indican las instrucciones para realizar las tareas establecidas:

1. Cargar la base de datos Epidural.xlsx en Spyder y asignarle el nombre "epidural_practica".

Antes de nada, crea una carpeta en la que guardarás la base de datos **Epidural.xlsx** y el archivo de Python (que podrás denominar **actividad2.py**) con el que procesarás dicho conjunto de datos. Esto es muy importante, ya que, si la BBDD no se encuentra en el mismo directorio que el archivo de Python, esta no podrá importarse a Spyder.

Para cargar la BBDD en Spyder, escribe, primero, el siguiente código y ejecútalo:

```
import pandas as pd
```

El código "import pandas as pd" se utiliza para importar el paquete Pandas, que es una biblioteca de software de análisis y manipulación de datos.

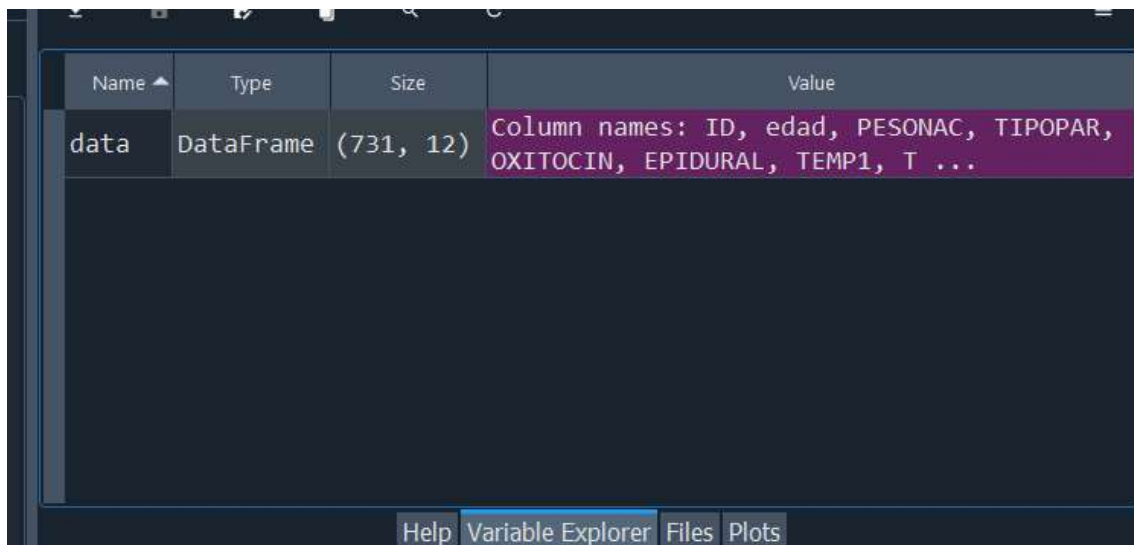
La sintaxis "as pd" se utiliza para renombrar el paquete de Pandas como "pd" en el código, lo que significa que se puede referenciar al paquete utilizando el alias "pd" en lugar del nombre completo "pandas".

Después, escribe el siguiente código y ejecútalo:

```
data = pd.read_excel("Epidural.xlsx")
```

2. Visualiza la base de datos.

En la consola de la parte superior derecha, clic en la pestaña "Variable Explorer" y, a continuación, sobre la fila que aparece en la consola. Al hacerlo, podrás visualizar la base de datos completa en una nueva pantalla que se abrirá al instante.



No olvides aportar en la resolución de la actividad un volcado de imagen de esta base de datos.

3. Mostrar el nombre de las variables.

Para ello, utiliza el siguiente código y ejecútalo. En la consola inferior derecha podrás visualizar los nombres de las variables. Aporta una captura de imagen para la respuesta en tu documento de la actividad.

```
print(data.columns)
```

4. Solicita los estadísticos descriptivos de la variable “edad”.

```
print(data[["edad"]].describe())
```

Este código te devuelve un resumen de estadísticas descriptivas que incluye el recuento, la media, la desviación estándar, el valor mínimo, el percentil 25, el percentil 50 (mediana), el percentil 75 y el valor máximo. Podrás ver los estadísticos en la consola inferior derecha.

5. Seleccionar (filtrar) a las mujeres que recibieron epidural y guardar el nuevo conjunto de datos con el nombre “epidural_si”. Ten en cuenta que el “sí” está codificado como 1.

```
epidural_si = data.loc[data["EPIDURAL"] == 1]
```

En la consola superior derecha, podrás visualizar este nuevo conjunto de datos.

Aporta una imagen de la consola para verificar que lo has hecho correctamente.

6. Recodifica la variable “edad” en tres categorías creando una nueva variable llamada “edad_recodificada”. Después, obtén una tabla de frecuencias.

- < 25 años llamada “por debajo”
- = 25 años llamada “por igual”
- > 25 años llamada “por encima”

Introduce el siguiente código y ejecútalo:

```
data["edad_recodificada"] = pd.cut(data['edad'], bins=[-float('inf'), 24, 25, float('inf')], labels=['por debajo', 'por igual', 'por encima'])
```

Para establecer los límites de los rangos se utiliza la función “bins”, así como “-float(‘inf’)” o “float(‘inf’)”, que representan el infinito negativo o positivo posible.

Para verificar que se ha recodificado la variable correctamente, acude a la base de datos en la consola, ábrela y busca la última columna. Vuelca una imagen en tu actividad.

A continuación, origina una tabla de frecuencias de la variable de “edad_recodificada”

```
frequency_table = data['edad_recodificada'].value_counts()
```

Pega la tabla que podrás visualizar en la consola.

7. Crear una nueva variable, llamada temp_media, como la media aritmética de las dos temperaturas.

Para calcular la media:

```
mean = data[['TEMP1', 'TEMP2']].mean(axis=1)
```

Para añadir la nueva variable en la base de datos:

```
data['TEMP_MEDIA'] = mean
```

Comprueba que se ha añadido esta nueva variable en tu base de datos.

Debes enviar a tu tutor tu respuesta en un solo fichero de pdf, en el que deberás incluir todos los elementos que estimes oportunos: texto explicativo, imágenes, volcados de pantalla, código de Python, etc., para demostrar que has resuelto estas cuestiones.