RESUMEN DE VALORIZACION  
Este proyecto implementa un proceso de valorización de lecturas de medidores listo para producción (VALORIZACION), que convierte exportes crudos de EMRC en libros Excel auditables para informes y valorización. La solución mejora el rendimiento y la fiabilidad respecto al script legado, preservando la lógica de selección original para garantizar paridad de resultados, e incorporando detección tolerante de encabezados, optimizaciones de memoria y reglas determinísticas de clasificación (DIARIO/INTERMITENTE/WALKBY).

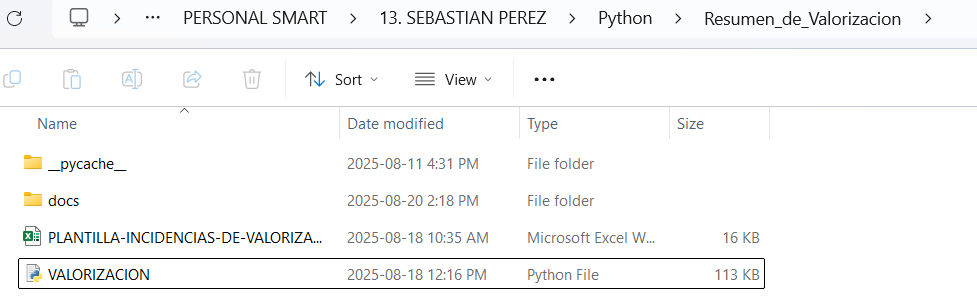
**Entregable principal:** VALORIZACION.py.

**Contribuciones clave**

* Mantener la función de selección heredada para paridad exacta y refactorizarla para mayor rendimiento y claridad.
* Añadir heurísticas flexibles de encabezados para manejar variaciones reales en archivos de entrada.
* Implementar optimizaciones de memoria (downcast de dtypes, conversión a category, procesamiento por chunks).
* Corregir errores por orden de operaciones y concurrencia (manejo de filas eliminadas y reclasificación DIARIO/WALKBY).
* Añadir manejo de incidencias y hoja de auditoría **ELIMINADOS** con motivos de eliminación.

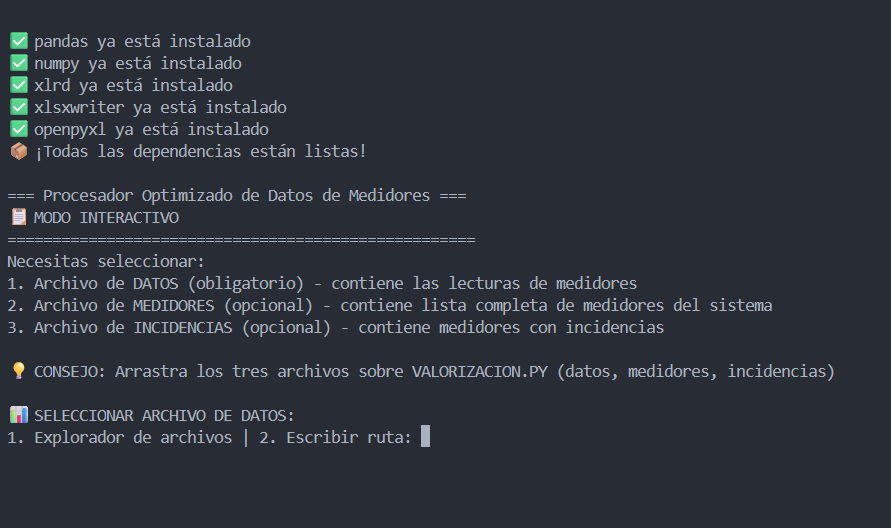
**Cómo funciona**  
Leer y concatenar hojas Excel → dividir Gateway vs WALKBY → aplicar reducciones por día/mes → crear tabla dinámica (medidor × fecha) → calcular *Número de Lecturas* y *Días Lecturados* → clasificar → exportar workbook con hojas **VALORIZABLE** y **ELIMINADOS**.

**TUTORIAL:**



En los archivos compartidos PERSONAL SMART -> SEBASTIAN PEREZ -> Python -> Resumen\_de\_Valorizacion se encuentra el script. Haga clic en el archivo para abrirlo.

2.

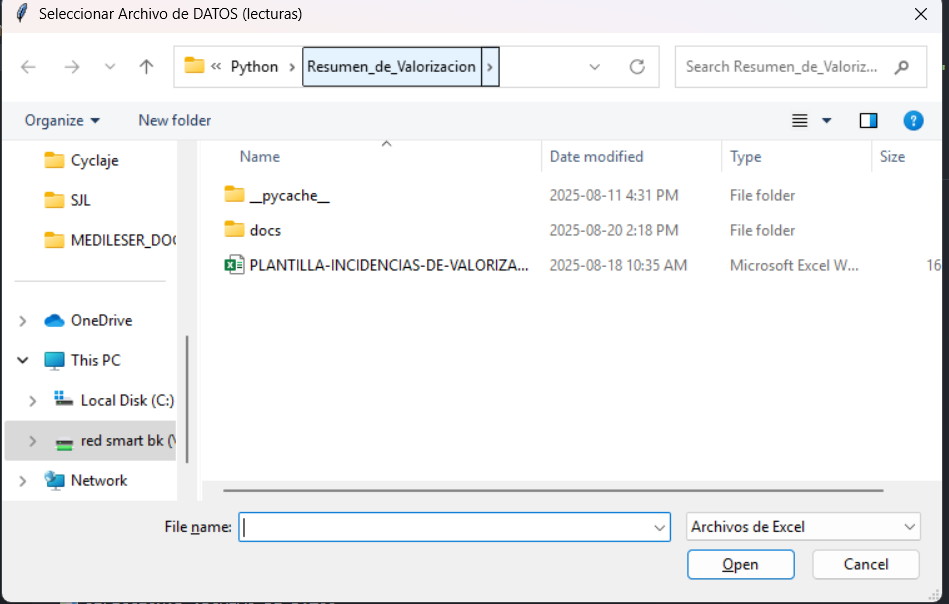


**Ahora que la interfaz gráfica (GUI) se ha abierto, se le presentan dos opciones:**

* Presione **‘1’** seguido de **Enter** para abrir el explorador de archivos y agregar los archivos.
* Presione **‘2’** seguido de **Enter** para escribir manualmente la ruta del archivo.

La **opción 1** se recomienda para la mayoría de los usuarios.

**Sugerencia:** si no ve marcas verdes de verificación en las bibliotecas instaladas, descargue **Python 3.13** desde Microsoft Store y ejecute su archivo dos veces para completar las instalaciones necesarias.

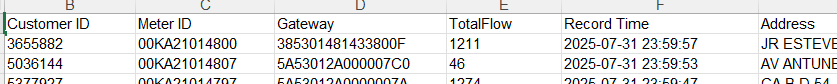


Selecciona el archivo apropiado sigue las instrucciones de la pantalla

**Archivo de entrada (1) Excel con datos de lecturas:**

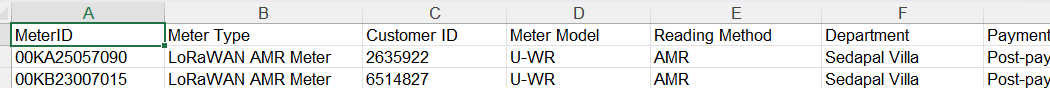
El primer archivo de entrada será el archivo de lecturas de medidores, descargado de EMRC con el nombre datainquiry-.

**Formato:**



El archivo de entrada descargado desde EMRC incluye la columna de fecha y hora combinada. No se debe separar esta columna, ya que el script no se ejecutará correctamente. No es necesario modificar los archivos descargados desde EMRC.

**Archivo de entrada (2)Excel con información de los medidores:**



El siguiente archivo de entrada será el archivo de lista de medidores. Es óptimo que esta lista se descargue directamente desde EMRC; sin embargo, si debe modificarse, como mínimo tanto la columna *‘meter ID’* como *‘customer ID’* deben encontrarse en las mismas columnas que en el ejemplo anterior. Las demás columnas pueden omitirse.

**Archivo de entrada (3) Plantilla Excel de INCIDENCIAS:  
**

El último archivo de entrada será una plantilla común en la que todos los usuarios deben actualizar la información de incidencias antes de añadir el archivo al script. Este archivo de Excel se encuentra en la misma carpeta que el script. Puede crear su propia copia para modificarla. Asegúrese de que todas las columnas estén completas para garantizar el correcto funcionamiento del script.

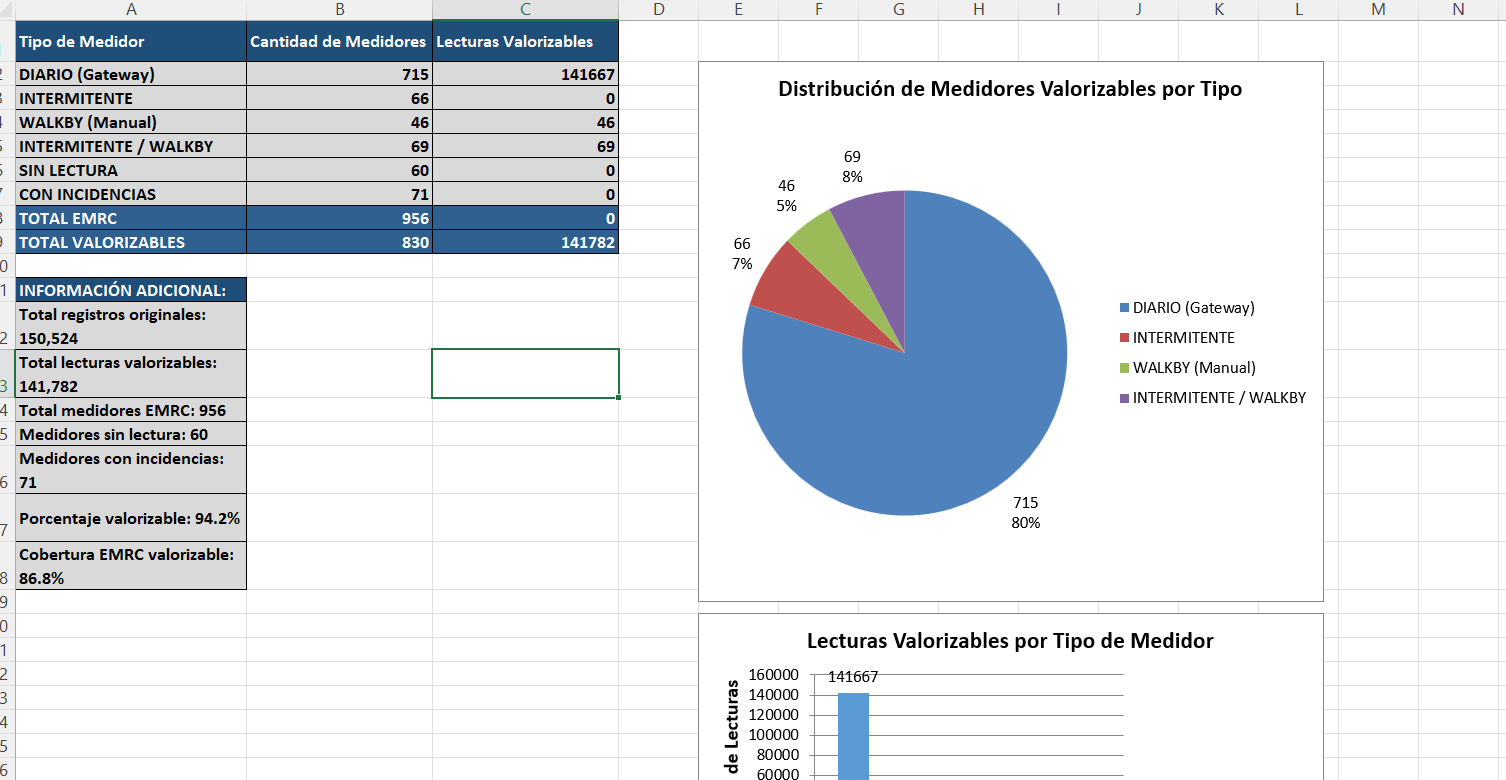
**AVISO**:

**Deje que el script se ejecute hasta que la terminal se cierre automáticamente para verificar una ejecución correcta.**  
El nuevo archivo de salida de Excel tendrá un nombre que comenzará con **dep\_-**.

**SU NUEVO ARCHIVO SE DESCARGARÁ EN LA MISMA CARPETA/UBICACIÓN QUE SU ARCHIVO DE LECTURAS DE MEDIDORES.**

**RESULTADO:**

**RESUMEN:**



Esta página contiene una tabla resumen y algunos gráficos para facilitar la interpretación de los datos presentes en la tabla dinámica final y en las lecturas valorizables.

**TOTAL:**

La hoja TOTAL contiene una lista completa de todas las lecturas de medidores combinadas en una sola pestaña. Se trata de los datos crudos sin ninguna modificación aplicada.

**GATEWAY DEPURADO:**

La hoja GATEWAY DEPURADO contiene una lista completa de todas las lecturas tipo “GATEWAY”, con las lecturas adicionales eliminadas. Ningún medidor en esta hoja tiene más de 7 lecturas por día.

Los criterios para eliminar lecturas son los siguientes:

1. Se dará prioridad a la eliminación de lecturas duplicadas.
2. Se eliminarán las lecturas demasiado cercanas entre sí, para maximizar el espaciado adecuado de lecturas a lo largo de un mismo día.
3. Se dará prioridad a eliminar lecturas más antiguas, siempre que sea posible, con el fin de conservar la información más reciente sobre la lectura del medidor.

**WALKBY:**

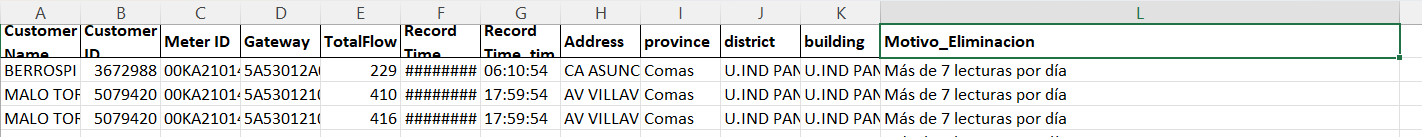
La hoja WALKBY contiene una lista de todas las lecturas de medidores tipo “WALKBY” sin alterar.

**VALORIZABLE GATEWAY:**

La hoja VALORIZABLE GATEWAY contiene una lista de todas las lecturas de medidores consideradas valorizables según los estándares de la empresa. Solo se mostrarán aquí las lecturas tipo DIARIO que no presenten incidencias.  
  
**VALORIZABLE WALKBY:**

La hoja VALORIZABLE WALKBY es similar a la anterior; sin embargo, contiene todas las lecturas tipo WALKBY que han sido filtradas y validadas.

**ELIMINADAS:**

****

# Tabla de Reglas de Eliminación / Auditoría

|  |  |
| --- | --- |
| Etiqueta (exacta) | Significado |
| Lectura WALKBY duplicada | Se encontraron múltiples lecturas WALKBY para el mismo medidor en el mismo día; solo se conservó la última lectura del día y las demás fueron eliminadas. |
| Lectura WALKBY adicional en el mes (solo primera lectura permitida) | Para medidores exclusivamente WALKBY (que no aparecen en datos de Gateway), solo se conserva la primera lectura de cada mes; las siguientes en el mismo mes se eliminan. |
| Medidor con incidencia | Las lecturas WALKBY fueron eliminadas porque el medidor aparece en el archivo de incidencias (medidor marcado con una incidencia). |
| Conflicto con lectura Gateway | Lecturas WALKBY eliminadas porque entraban en conflicto con lecturas Gateway en los mismos días para medidores reclasificados como DIARIO. Las lecturas Gateway tienen prioridad. |
| Más de 7 lecturas por día | Regla de Gateway: si hay más de 7 lecturas para el mismo medidor en el mismo día, las lecturas adicionales se eliminan, quedando un máximo de 7 por día. |

**TABLA DINAMICA:**

La hoja **Tabla Dinamica** muestra el estado final por medidor después de todo el procesamiento.  
Es la fuente principal para:

* La generación del resumen.
* La selección de lecturas valorizables.

Cada fila representa un medidor (**Meter ID**) y cada columna de fecha muestra la cantidad de lecturas registradas ese día (conteo). La tabla incluye además metadatos como cliente, clasificación, incidencias y totales por medidor.

**Columnas clave (orden típico)**

* **Meter ID**: Identificador del medidor (clave primaria por fila).
* **Customer ID**: Identificador o cliente asociado (si está disponible).
* **Clasificación**: Etiqueta de frecuencia del medidor. Posibles valores:
  + DIARIO
  + INTERMITENTE
  + WALKBY
  + DIARIO / WALKBY
  + INTERMITENTE / WALKBY
  + SIN LECTURA
* **Incidencia**: Puede ser CON INCIDENCIA o SIN INCIDENCIA. Indica si el medidor fue marcado en el archivo de incidencias.
* **Detalle de Incidencia**: Texto libre con el tipo o nota de la incidencia (si aplica).
* **Columnas de fecha** (ejemplo: 2025-07-01, 2025-07-02, …): cada celda contiene el número de lecturas registradas ese día para ese medidor.
* **Numero de Lecturas**: Suma de todas las columnas de fecha (total de lecturas encontradas).
* **Dias Lecturados**: Número de días con al menos una lectura (>0).

**Cómo se calculan los valores**

* La tabla se construye con un **pivot** sobre las lecturas Gateway (sumando las WALKBY cuando corresponda).
* Para las columnas de fecha, el cálculo usa aggfunc='count' sobre la columna de flujo/valor.
* **Numero de Lecturas** = suma de todos los conteos diarios para ese medidor.
* **Dias Lecturados** = número de columnas de fecha con valor > 0.

**Reglas de clasificación**

* Se agrupa por mes para determinar si un medidor es **DIARIO** (lecturas en todos los días de al menos un mes).
* Si además tenía WALKBY, puede ser DIARIO / WALKBY, y luego re-clasificado según conflicto.
* Si no cumple la condición de diario → INTERMITENTE (o INTERMITENTE / WALKBY si corresponde).
* Medidores que solo existen en WALKBY → WALKBY.
* Medidores sin lecturas → SIN LECTURA (ver sección *Medidores faltantes*).

**Manejo de incidencias**

* Si un medidor aparece en el archivo de incidencias:
  + **Incidencia = CON INCIDENCIA**.
  + **Detalle de Incidencia** se llena con la nota correspondiente.
* Para medidores con incidencia, las lecturas WALKBY pueden ser eliminadas o reemplazadas.
* Los cambios se reflejan en los totales y la clasificación.

**Conflictos entre Gateway y WALKBY**

* Si un medidor tiene lecturas tanto de Gateway como de WALKBY en el mismo día:
  + Se eliminan las lecturas WALKBY (Gateway tiene precedencia).
  + Los registros eliminados se documentan en la hoja **Eliminados** con motivo *Conflicto con lectura Gateway*.
* Después de resolver los conflictos, la clasificación puede transformarse (ejemplo: de DIARIO / WALKBY a DIARIO).

**Medidores faltantes**

Si se proporcionó un archivo con el listado completo de medidores, se agregan filas para aquellos que no aparecen en las lecturas:

* **Clasificación = SIN LECTURA**.
* **Incidencia = SIN INCIDENCIA** (o CON INCIDENCIA si aparecen en el archivo de incidencias).
* Columnas de fecha en 0.
* **Numero de Lecturas = 0**.
* **Dias Lecturados = 0**.

**Reglas de eliminación de lecturas antes del pivot**

**Gateway**

* Máximo 7 lecturas por día y medidor.
* Se priorizan lecturas espaciadas y más recientes.
* Lecturas eliminadas → motivo *Más de 7 lecturas por día*.

**WALKBY**

* Se conserva solo la última lectura del día si hay duplicados.
* WALKBY puro → solo se permite la primera lectura del mes. Las demás se eliminan con motivo *Lectura WALKBY adicional en el mes*.

**Incidencias y conflictos**

* Eliminaciones adicionales se documentan en **Eliminados** con su motivo.

**Lecturas valorizables**

* Se consideran valorizables principalmente los medidores con:
  + **Clasificación = DIARIO**.
  + **Incidencia = SIN INCIDENCIA**.
* También pueden incluirse algunos casos de WALKBY o INTERMITENTE / WALKBY según política.
* Las hojas **VALORIZABLE GATEWAY** y **VALORIZABLE WALKBY** se generan directamente a partir de estas condiciones.

**Uso práctico por el operador**

1. Verifique la fila de un medidor:
   * Si **Numero de Lecturas** es menor a lo esperado, consulte **Eliminados** para ver los motivos.
   * Si **Incidencia = CON INCIDENCIA**, revise el detalle y el archivo de incidencias.
   * Si aparece como **SIN LECTURA**, confirme que el archivo de entrada contenía ese *Meter ID*.
2. Para cambiar la política (ejemplo: permitir más de 7 lecturas/día), se debe ajustar la configuración del script y volver a ejecutar.

**Problemas comunes y verificación rápida**

* **Faltan columnas**: el pivot no se genera si no existen Meter ID, Record Time o la columna de valor en el archivo original.
* **Fechas desordenadas**: puede deberse a tipos de datos incorrectos; el script fuerza fechas como texto.
* **Diferencias entre hojas**: use **Eliminados**, **VALORIZABLE GATEWAY**, **VALORIZABLE WALKBY** y **TOTAL** para rastrear la causa.

**Ejemplo**

**Fila de ejemplo:**

* Meter ID: 12345
* Customer ID: C-987
* Clasificación: DIARIO
* Incidencia: SIN INCIDENCIA
* 2025-07-01: 1
* 2025-07-02: 1
* …
* Numero de Lecturas: 31
* Dias Lecturados: 31

**Interpretación**: medidor con lectura diaria, sin incidencias, candidato a valorización.

**Referencias cruzadas**

* **Eliminados**: lista todas las lecturas descartadas con su motivo.
* **VALORIZABLE GATEWAY / WALKBY**: lecturas elegidas para valorización.
* **RESUMEN**: cifras agregadas derivadas de la **Tabla Dinamica**.