**Conversor de ficheros Excel/CSV**

Esta herramienta permite convertir ficheros Excel a una salida especificada extrayendo los datos de dicho fichero en base a un fichero de equivalencias.

Argumentos de llamada:

**equivalenceConverter.exe**

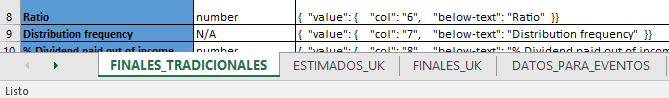
|  |  |
| --- | --- |
| Tipo de fichero | Se refiere al nombre de la pestaña en el fichero de equivalencias |
| ID del fichero | Identificador de la columna dentro de la pestaña especificada (fila 1) |
| Fichero | La dirección completa del fichero que se va a procesar |
| Carpeta Salida | La carpeta donde se pondrán los resultados |
| Fichero Equivalencias | La dirección completa del fichero de equivalencias |

Los argumentos van separados por espacio y rodeados de comillas dobles.

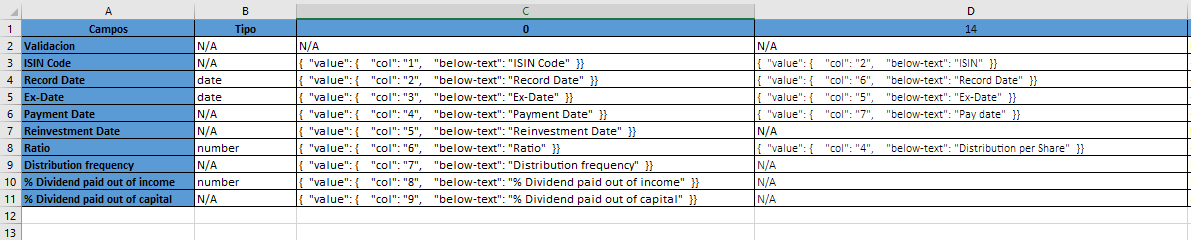
**Fichero de equivalencias**

Es el que contiene toda la información referente a las equivalencias que manejara la herramienta, los datos están representados mediante el formato JSON, su estructura es la siguiente:

Las pestañas indican los tipos de ficheros a procesar, además en base a este nombre se creará el fichero de salida.



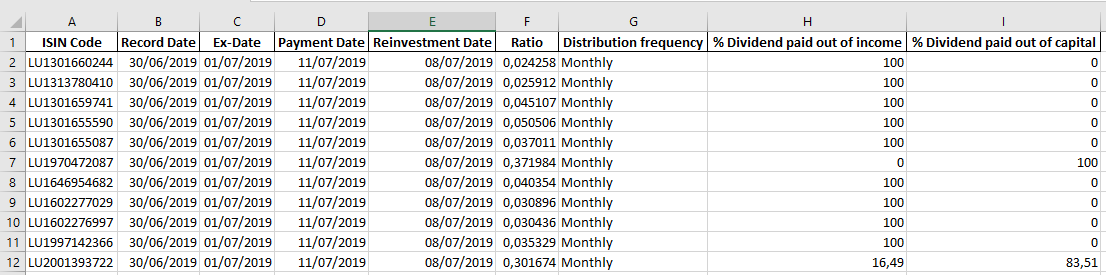
La primera fila de cada pestaña está compuesta por tres cosas:



* **Campos**: Se refiere a las columnas del fichero de salida, los nombres son importantes porque serán los que tenga el fichero.
* **Tipo**: Se refiere al tipo del dato que va a contener la columna, los tipos actuales son:
  + **string**: Una cadena de texto
  + **number**: Un número (flotante o entero)
  + **date**: Una fecha
* **ID**: El identificador del fichero a procesar (es una cadena de texto representativa).

El fichero de salida producido contendrá los **Campos** como columnas y el nombre se deriva de la pestaña.

output\_FINALES\_TRADICIONALES.xlsx



**Funcionalidades Disponibles**

Las funciones disponibles para determinar las columnas del excel resultante.

|  |  |
| --- | --- |
| VALUE | Función básica que devuelve una columna de valores |
| CONSTANT | Función básica que da un valor constante |
| POSITION | Función básica que devuelve el valor de una celda |
| MATH | Función compuesta, que realiza operación matemática |
| DATE | Función compuesta que lee una fecha |
| FILTER | Función para filtrar los datos del excel |

**VALUE**

Sirve para representar una columna COMPLETA indicando una posición inicial, normalmente sería una columna y una fila.

|  |  |  |
| --- | --- | --- |
| **col** | string | Número de la columna |
| **below-text** | string | Busca el texto en la columna indicada, y empieza desde la fila siguiente (+1) |
| **row** | string | Número de la fila de donde empezar. |
| **add-rows** | string | Se le suma a la fila la cantidad especificada. |
| **prepend** | string | Agrega al INICIO del valor obtenido el texto indicado. |
| **append** | string | Agrega al FINAL del valor obtenido el texto indicado. |
| **replace** | Objeto | Reemplaza del valor obtenido el texto indicado por el valor especificado. |
| **Objeto replace** | | |
| **pattern** | string | El patrón a buscar y reemplazar. |
| **text** | string | El texto por el que se va a reemplazar. |

{

"value": {

"col": "1",

"below-text": "ISIN Code"

}

}

* Busca la columna 1
* En la columna 1 busca el texto **“ISIN Code”** y empieza en la siguiente fila

El comando **row** toma precedencia sobre el **below-text**.

{

"value": {

"col":"11",

"below-text":"Currency",

"replace": {

"pattern":"GB",

"text":"G"

}

}

}

* Busca la columna 11
* En la columna 11 busca el texto “**Currency**” y empieza en la siguiente fila
* El valor obtenido en esa celda remplaza el texto **GB** por **G**

**CONSTANT**

Sirve para indicar que la columna será llenada con un valor constante.

{

"constant":"100G"

}

* Llena la columna con el valor 100G

**POSITION**

Sirve para indicar que toda la columna será llenada por el valor de la celda que se encuentra en la posición indicada.

|  |  |  |
| --- | --- | --- |
| **col** | string | Número de la columna |
| **row** | string | Número de la fila de dónde empezar. |
| **prepend** | string | Agrega al INICIO del valor obtenido el texto indicado. |
| **append** | string | Agrega al FINAL del valor obtenido el texto indicado. |
| **replace** | Objeto | Reemplaza del valor obtenido el texto indicado por el valor especificado. |
| **Objeto replace** | | |
| **pattern** | string | El patrón a buscar y reemplazar. |
| **text** | string | El texto por el que se va a reemplazar. |

* Busca el valor que se encuentra en la posición fila 2 columna 3 y llena toda la columna con él.

{

"position": {

"row":"2",

"col":"3"

}

}

**MATH**

Representa una función matemática que se realizara en base a uno o más valores que se encuentran en el excel, se representara por cada operando y una función a ejecutar.

|  |  |  |
| --- | --- | --- |
| **varXXX** | Objeto | Indica el valor del operando, el nombre es el que se utiliza para remplazar en los operando del result, puede ser cualquiera de los tipos básicos ( value , position ) |
| **result** | string | La fórmula matemática a realizar, se evalúa luego de remplazar los operandos. |

{

"math": {

"var1": {

"value": {

"col":"11",

"below-text":"Units in Issue"

}

},

"var2": {

"value": {

"col":"15",

"below-text":"Gross Rate"

}

},

"result":"(var2\*100)/var1"

}

}

* Busca VAR1, evaluando el VALUE que indica su posición.
* Busca VAR2, evaluando el value que indica su posición.
* Sustituye el valor de VAR1 y VAR2 en la función de result y la evalúa.
* El resultado es el valor de la columna final.

**DATE**

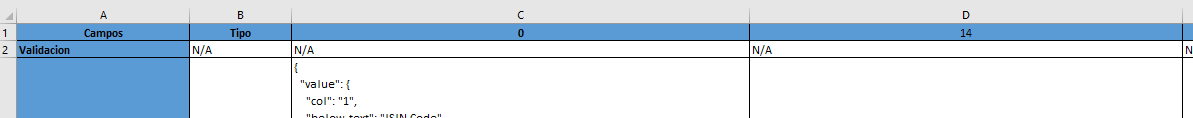
Permite la manipulación de una fecha en la hoja de Excel, si se trata de un fichero CSV indicar el formato.

|  |  |  |
| --- | --- | --- |
| **value** | Objeto | El valor de la fecha puede provenir de un VALUE |
| **position** | Objeto | El valor de la fecha puede provenir de un POSITION |
| **format** | string | El formato de la fecha (formato Python) |
| **transform** | string | La transformación a realizar (**addDays** , **subDays**) |
| **quantity** | string | La cantidad a aplicar a la transformación. |

**addDays** y **subDays**: Agregar o Quitar días laborales. (sábados y domingos, 25/12 y 1/1)

**Filtro Validación**

En la segunda fila se puede indicar un campo llamado Validación, ese campo es ignorado a la hora de crear la salida del fichero, y su función es permitir filtrar los datos del Excel.



**FILTER**

Especifica un filtro a aplicar a los datos del Excel con la finalidad de procesar solo un subgrupo.

|  |  |  |
| --- | --- | --- |
| **column** | Objeto | El valor de la fecha puede provenir de un VALUE o un POSITION |
| **query** | string | Texto que se va a buscar o filtrar |
| **type** | string | El tipo de filtro a aplicar, (**equal** o **diff**) igual o diferente |

{

"filter": {

"column": {

"value": {

"col":"15",

"below-text":"Share Class Type"

}

},

"query":"Accumulative"

"type":"equal"

}

}

* El valor de **column** se busca evaluando el VALUE especificado (col 15 – below-text)
* El **query** y el **type** nos indica que usaremos solo los valores que sean igual a Accumulative

**ISIN y SEDOL**

El campo de identificación será siempre la primera columna, este campo es importante ya que valida que cada línea sea única.

Existe un caso especial que es cuando no se tienen ISIN, en este caso se usara el campo SEDOL como identificación, en el caso de ser de “tipo” SEDOL se genera un fichero especial llamado ouput\_SEDOL

**Incidencias y Errores**

Las validaciones realizadas son las siguientes:

* Verificar si existen líneas completamente iguales, en dicho caso se eliminan todas menos una.
* Verificar si existen líneas parecidas con el mismo identificador (ISIN o SEDOL) y de encontrar alguna se pondrán en el fichero de incidencias.
* Además, cualquier tipo de fallo al manejar un dato se enviará la línea al fichero de incidencias.

**Registros de eventos y errores (LOG)**

La herramienta genera un fichero de eventos en donde se encontrarán todo lo que suceda en su proceso normal de funcionamiento, tanto información de utilidad como errores.

El fichero se generará en la carpeta de salida con el prefijo log\_.

**Salidas de la herramienta (consola)**

Además del log de eventos la herramienta también generara una salida a consola de los resultados o errores que se vaya encontrando.

La primera línea será OK si todo va bien he ira seguido de un listado de los identificadores de línea encontrados, se expresará con la siguiente sintaxis.

Cada línea representará un ítem, el primer valor será el identificador, el segundo será OK o ERROR y el tercero será el nombre del fichero donde se guardó.

***OK***

***ID,OK,Nombre del fichero***

***ID,OK,Nombre del fichero***

***ID,OK,Nombre del fichero***

Si no hay datos para procesar se mostrar un:

***NO DATA: No se encontraron datos para procesar***

Si se encontró algún error se generará un mensaje:

***ERROR : faltan argumentos.. <Tipo Proceso> <ID Proceso> <Fichero> <Directorio Destino>***

***ERROR : No se encuentra la carpeta de destino:***

***ERROR : No se encuentra el fichero a procesar:***

***ERROR : Fichero de equivalencias erroneo o no encontrado***

***ERROR : No se pudo abrir el fichero a procesar***

***ERROR : Fallo en la escritura del fichero de salida.***

***ERROR : Fallo en la escritura del fichero de incidencias.***

|  |  |
| --- | --- |
| **Simbolo** | **Lo que significa** |
| **%a** | Nombre local abreviado de día de semana |
| **%A** | Nombre local completo de día de semana |
| **%b** | Nombre local abreviado de mes |
| **%B** | Nombre local completo de mes |
| **%c** | Representación local de fecha y hora |
| **%d** | Día de mes [01,31] |
| **%H** | Hora (horario 24 horas) [00,23] |
| **%I** | Hora (horario 12 horas) [01,12] |
| **%j** | Número de día del año [001,366] |
| **%m** | Mes [01,12] |
| **%M** | Minuto [00,59] |
| **%p** | Etiqueta AM o PM |
| **%S** | Segundo |
| **%U** | Nº semana del año. Se considera al Domingo como primer día de semana [00,53] |
| **%w** | Establece el primer día de semana [0(Domingo),1(Lunes)... 6]. |
| **%W** | Nº semana del año (Se considera al Lunes como primer día de semana) [00,53] |
| **%x** | Fecha local |
| **%X** | Hora local |
| **%y** | Año en formato corto [00,99] |
| **%Y** | Año en formato largo |
| **%Z** | Nombre de Zona Horaria |

***ERROR : El formato para este excel no es correcto: " + process\_type + " - " + process\_id***

***ERROR : Ha ocurrido un error inesperado.***

***ERROR***

Formatos de fecha Python

**Observaciones**

* Para la función de DATE si se sabe que la columna no es una fecha, pero es una cadena de caracteres se debe indicar el FORMATO de la fecha, este siempre es el caso de los ficheros CSV por lo que siempre es necesario indicar el formato.
* Los números se les debería indicar el Tipo como number para evitar problemas de formato.
* Para los ficheros CSV es mejor indicar el tipo de campo porque siempre llegaran como una cadena de caracteres (string)
* Si la primera fila completa no se encuentra ningún valor entonces se asume que no es el fichero correcto.
* En cada ejecución se cargan los datos de los ficheros de salida que se encuentren en el directorio de salida (para hacer validaciones y detectar incidencias).
* La primera columna siempre debería ser la del ISIN y la segunda de SEDOL de existir.
* Si se consigue un error en la primera fila se asume que el fichero es erróneo, esto solo en la primera fila, si se encuentran errores en las subsecuentes líneas se ignoran y se pasan al de incidencias.
* Todos los valores obtenidos desde VALUE y POSITION si son string se les quita los espacios en blanco al principio y al final.