

Documentación de Enriquecimiento Certero

Autor: Jorge Tillería

Para realizar el proceso de enriquecimiento hay que seguir los pasos especificados en este archivo y ejecutar los scripts señalados.

1.1. Descarga de datos.

Primeramente debemos de extraer los datos de la base de datos para hacer el pre-procesamiento para luego mandarlo a Certero para hacer el enriquecimiento.

Para esto debemos proceder a conectarnos a la máquina de Marcos a través de SSH con el siguiente comando.

```
ssh marcos-rago@192.168.130.18
```

Luego debemos ubicarnos en la carpeta
/home/marcos-rago/Documents/IA/pruebas_desde_19_8_2021/

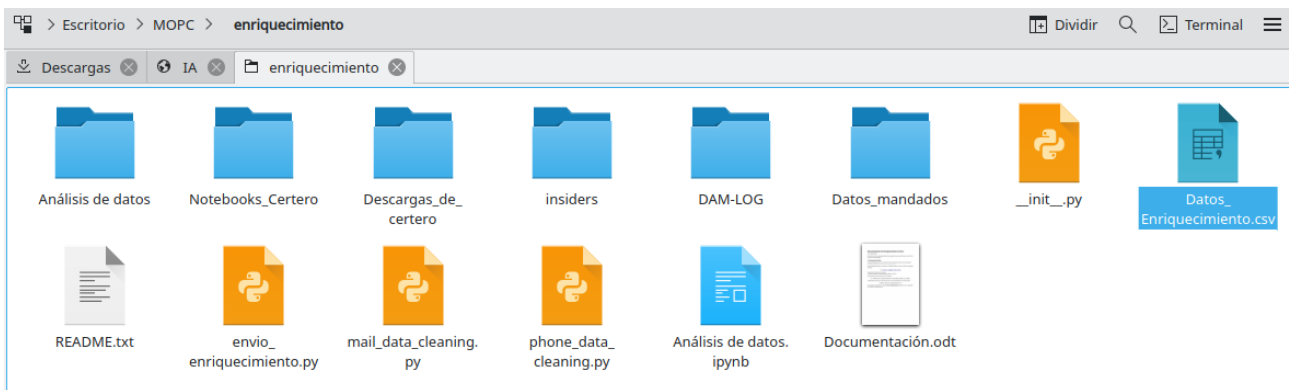
Esto podemos hacerlo directamente con el comando

```
cd /home/marcos-rago/Documents/IA/pruebas_desde_19_8_2021/
```

y luego debemos ejecutar el script para que envíe al correo especificado en el script los datos.

```
python3 datos_enriquecimiento.py
```

Luego podemos descargar el csv llamado **Datos_Enriquecimiento.csv** desde el correo y guardarlo en la carpeta de enriquecimiento /.



1.2. Envío de datos

También existe la posibilidad que se mande por correo automáticamente para enriquecimiento semanal, también se puede utilizar este archivo pero hay que colocarlo en la carpeta Datos_Enriquecimiento y colocarle el nombre genérico *Datos_Enriquecimiento_AAAA-MM-DD* para luego ejecutar el script de **0.Preprocesamiento de datos.py**

2. Envío para enriquecimiento

Para poder enviar el archivo a enriquecimiento primero debemos de preprocesarlo para sacar los datos de interés para CERTERO. Para ello debemos ejecutar el script `envio_enriquecimiento.py`

```
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$ python envio_enriquecimiento.py
Los archivos fueron creado con éxito
Se guardaron en:
Datos_mandados/enriquecimiento_fis_2021-08-31.csv
Datos_mandados/enriquecimiento_jur_2021-08-31.csv
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$
```

Nos aparecerá este mensaje y podremos ver los archivos guardados.

Dichos archivos debemos de subirlos a la página de certero entrando en <https://www.certero.precorp.com.ar/IWC/Procesados> con el usuario y contraseña de Facundo.

Usuario : facundo.ferrari@mopc.cc

Contraseña: 123456

Y luego haremos click en nuevo enriquecimiento y subiremos uno de los archivos (preferentemente el fis).

Una vez que carguemos el archivo, debemos de seleccionar las siguientes opciones.

Archivo enriquecimie...21-08-31.csv

Tipo Debe especificar si son personas fisicas o juridicas

Columnas

COLUMNA	VALOR
IDMOROSO	<input type="text" value="Id"/>
DOCUMENTO	<input type="text" value="Documento"/>
NOMBRE	<input type="text" value="Nombre"/>
TIPO	<input type="text" value="Ignorar"/>

Notificación Email para notificar cuando finalice el proceso

Tipo de Cruce ☒ Cruce estándar ☐ Cruce segmentado

Y para personas jurídicas de la siguiente forma:

Archivo enriquecimie...21-08-31.csv

Tipo

Debe especificar si son personas físicas o jurídicas

Columnas

COLUMNA	VALOR
IDMOROSO	<input type="text" value="Id"/>
DOCUMENTO	<input type="text" value="CUIL/CUIT"/>
NOMBRE	<input type="text" value="Razon Social"/>
TIPO	<input type="text" value="Ignorar"/>

Notificación

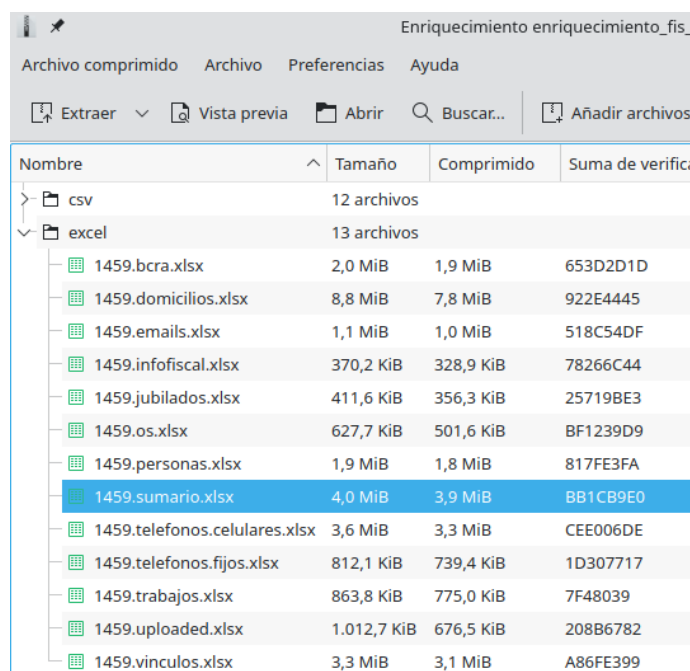
Email para notificar cuando finalice el proceso

Tipo de Cruce ☒ Cruce estándar ☐ Cruce segmentado

Y luego simplemente a esperar a recibir la notificación por parte de CERTERO.

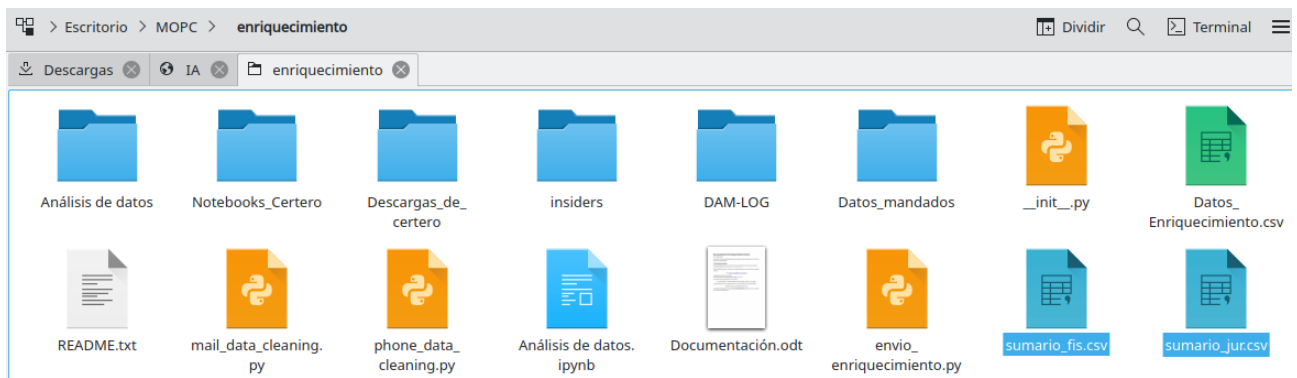
3. Procesamiento de datos

Ahora iniciamos el procesamiento de datos, debemos descargar de CERTERO el archivo zip y buscar entre los archivos de tipo excel el que diga Sumario. Y el archivo zip debemos guardarlo en la carpeta **Descargas_de_certero**.



Nombre	Tamaño	Comprimido	Suma de verificación
csv	12 archivos		
excel	13 archivos		
1459.bcrs.xlsx	2,0 MiB	1,9 MiB	653D2D1D
1459.domicilios.xlsx	8,8 MiB	7,8 MiB	922E4445
1459.emails.xlsx	1,1 MiB	1,0 MiB	518C54DF
1459.infofiscal.xlsx	370,2 KiB	328,9 KiB	78266C44
1459.jubilados.xlsx	411,6 KiB	356,3 KiB	25719BE3
1459.os.xlsx	627,7 KiB	501,6 KiB	BF1239D9
1459.personas.xlsx	1,9 MiB	1,8 MiB	817FE3FA
1459.sumario.xlsx	4,0 MiB	3,9 MiB	BB1CB9E0
1459.telefonos.celulares.xlsx	3,6 MiB	3,3 MiB	CEE006DE
1459.telefonos.fijos.xlsx	812,1 KiB	739,4 KiB	1D307717
1459.trabajos.xlsx	863,8 KiB	775,0 KiB	7F48039
1459.uploaded.xlsx	1.012,7 KiB	676,5 KiB	208B6782
1459.vinculos.xlsx	3,3 MiB	3,1 MiB	A86FE399

Dicho archivo debemos de abrirlo y guardarlo en formato csv en la carpeta de enriquecimiento y renombrarla con el siguiente nombre **sumario_fis.csv** para personas físicas y **sumario_jur.csv** para personas jurídicas.



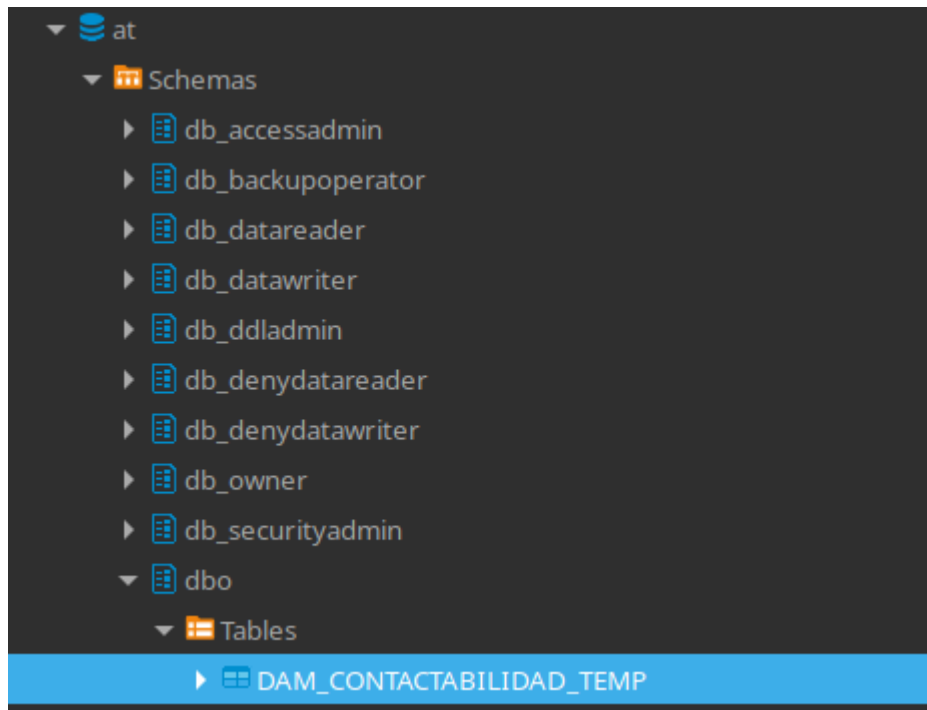
Luego debemos correr los scripts **mail_data_cleaning.py** y **phone_data_cleaning.py**

```
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$ python mail_data_cleaning.py
Resultado exitoso.
Guardado en: insiders/mail_2021-08-31.csv
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$ python phone_data_cleaning.py
Resultado exitoso.
Guardado en: insiders/phone_2021-08-31.csv
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$
```

Luego tendremos los archivos listos para cargarlos en la base de datos en la carpeta insiders.

4. Subida de datos.

En el Dbeaver nos ubicamos en la base de datos at y buscamos la tabla DAM_CONTACTABILIDAD_TEMP y luego le damos click derecho e Importar datos.



Una vez que cargamos los datos y no hay errores, procedemos a ejecutar el siguiente script.

```
insert into DAM_CONTACTABILIDAD
SELECT
    [IDMOROSO]
    , [TIPO_DOC]
    , [NRO_DOC]
    , [CUIT_CUIL]
    , [TIPO]
    , [SUBTIPO]
    , [DATO]
    , [PROPIETARIO_DATO]
    , [PROVEEDOR_DATO]
    , getdate () as [FECHA_INGRESO_DATO]
    , [VERIFICADO]
    , convert ( datetime, ([FECHA_VERIFICACION]),102) as [FECHA_VERIFICACION]
    , [OPERADOR_VERIFICACION]
    , [SCORE_DATO]
    , null as rechazado
    , null as fecha_rechazo
    , null as operador_rechazo
    , null as razon_rechazo
    , null as SCORE_DATO_100
FROM at.dbo.DAM_CONTACTABILIDAD_TEMP with (nolock)
where at.dbo.DAM_CONTACTABILIDAD_TEMP.idmoroso not in
```

```
(select DAM_CONTACTABILIDAD.IDMOROSO from DAM_CONTACTABILIDAD where
DAM_CONTACTABILIDAD.IDMOROSO =at.dbo.DAM_CONTACTABILIDAD_TEMP.idmoroso and
DAM_CONTACTABILIDAD.dato = at.dbo.DAM_CONTACTABILIDAD_TEMP.dato and
DAM_CONTACTABILIDAD.[PROVEEDOR_DATO] =
at.dbo.DAM_CONTACTABILIDAD_TEMP.[PROVEEDOR_DATO])
```

Luego procedemos a borrar los datos de DAM_CONTACTABILIDAD_TEMP y repetir el proceso para el otro archivo.

```
DELETE FROM DAM_CONTACTABILIDAD_TEMP
```

6. Inicio del proceso DAM_LOG

Una vez que actualizamos la tabla debemos proceder a correr el archivo **script_DAM_LOG.py**.

```
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$ python script_DAM_LOG.py
Ingrese la fecha que se mandó a enriquecer. Ej:2021-08-26
2021-08-26
Ingrese la fecha en que se subió a la BD. Ej:2021-08-27
2021-08-27
El programa fue exitoso.
Se guardó en dam_log_2021-08-31.csv
(mopc) tetotille@tetotille-SVF15215CLB:~/Escritorio/MOPC/enriquecimiento$
```

Debemos controlar que el número de registros de salida coincida con el número de registros mandados.

7. Subir el DAM_LOG a la BD

Nuevamente debemos subir en el Dbeaver

Para eso simplemente debemos cargar haciendo los mismos pasos que para DAM_CONTACTABILIDAD_TEMP, es decir, esta vez directamente se suben a la base de datos de DAM_LOG.

8. Mandar el correo

Debemos de mandar el correo electrónico a las siguientes personas comunicando el enriquecimiento.

Facundo Ferrari <facundo.ferrari@mopc.cc>, Vanessa Fondado <vanessa.fondado@mopc.cc>, Luis Castillo <luis.castillo@mopc.cc>, Nicolás Caruso <Nicolas.Caruso@mopc.cc>

Con cc a:

Patricio Viñuela <patricio.vinuela@mopc.cc>, FRANCISCO LONARDI FRANCISCO <flonardi@itba.edu.ar>, Francisco Cañas <francisco.canas@mopc.cc>, Marcos Rago <marcos.rago@mopc.cc>, Tecnologia <it@mopc.cc>, Matias Alvarez <matias.alvarez@mopc.cc>, Santiago Usach <santiago.usach@mopc.cc>, Julieta Galarza <julieta.galarza@mopc.cc>, Nicolas Rodriguez <nicolas.rodriguez@mopc.cc>, Marisol Blanco <marisol.blanco@mopc.cc>, Ale Muller <amuller@afmabogados.com>, Natalia Gomez <ngomez@afmabogados.com>, Andrea Pereiro <andrea.pereiro@mopc.cc>, Paola Cardena <paola.cardena@mopc.cc>

El enriquecimiento debe seguir este formato:

Estimados,

Se generó una nueva inserción de **datos de contactabilidad** a partir de la última devolución de ZERT. **Esta actualización corresponde al enriquecimiento *masivo* por incorporación de nuevo proveedor.**

En total, tenemos 17.646 IDMorosos de los que se actualizaron los datos de contactabilidad, conformados por **teléfonos y mails**. Se agregaron 42229 teléfonos y 16272 mails para esos IDMorosos.

Les dejo adjunta la query que pueden utilizar para tomar estos datos, donde tienen IDMoroso, IDCliente, Estado, Filtro A, Scoring, DeudaTotal y DiasSinLLamar.

Por último, les dejo adjunta otra query para que puedan visualizar los IDMorosos que volvieron con **datos que no teníamos en nuestras bases**. Van a poder observar IDMoroso, IDCliente, Estado y FiltroA según cada caso. Además, en este caso les agregué la columna Dato, para que vean efectivamente cuál fue el dato que se agregó en cada caso.

Saludos,

Jorge

PD: Estos datos se verán reflejados en el front a partir de mañana.

Query 1: para visualizar la **totalidad** de casos insertados, por IDCliente.

```
WITH SUBQUERY1 AS (  
SELECT DISTINCT DAM_CONTACTABILIDAD.IDMOROSO, IDCLIENTE, ESTADO, FILTROA  
FROM Active.dbo.DAM_CONTACTABILIDAD  
LEFT JOIN Active.dbo.CARTERA  
ON Active.dbo.CARTERA.IDMOROSO = Active.dbo.DAM_CONTACTABILIDAD.IDMOROSO  
WHERE PROVEEDOR_DATO in ('ZERT')  
AND FECHA_INGRESO_DATO BETWEEN '27/08/2021' AND '28/08/2021'),  
SUBQUERY2 AS(  
SELECT dp.IDMOROSO, p.IDCLIENTE, dp.SCORING_MOPC, SUM(TOTALDEUDA) AS TOTALDEUDA  
FROM ACTIVE.dbo.DAM_DATOS_PERSONALES dp  
INNER JOIN PRODUCTO p  
ON p.IDMOROSO = dp.IDMOROSO  
GROUP BY dp.IDMOROSO, p.IDCLIENTE, dp.SCORING_MOPC)  
SELECT SUBQUERY1.IDMOROSO, SUBQUERY2.IDCLIENTE, SUBQUERY1.ESTADO, SUBQUERY1.FILTROA,  
SUBQUERY2.SCORING_MOPC, SUBQUERY2.TOTALDEUDA, [dbo].[DiasSinLlamar]  
(SUBQUERY1.IDMOROSO, SUBQUERY2.IDCLIENTE) AS [DIAS SIN LLAMAR]  
FROM SUBQUERY1  
LEFT JOIN SUBQUERY2  
ON SUBQUERY1.IDMOROSO = SUBQUERY2.IDMOROSO  
AND SUBQUERY1.IDCLIENTE = SUBQUERY2.IDCLIENTE
```

```
WHERE SUBQUERY2.IDCLIENTE = 'FRANCES_2012'  
ORDER BY TOTALDEUDA DESC
```

Nota: para cambiar de IDCliente, pueden reemplazar en la línea WHERE SUBQUERY2.IDCLIENTE = 'FRANCES_2012' por sus IDClientes correspondientes.

Query 2: trae los casos que tienen datos que **no teníamos** en nuestras bases:

```
WITH SUBQUERY1 AS (  
SELECT ACTIVE.dbo.DAM_CONTACTABILIDAD.IDMOROSO, DATO,  
ACTIVE.dbo.CARTERA.IDCLIENTE, ACTIVE.dbo.CARTERA.ESTADO, ACTIVE.dbo.CARTERA.FILTROA  
FROM ACTIVE.dbo.DAM_CONTACTABILIDAD  
LEFT JOIN ACTIVE.dbo.CARTERA ON ACTIVE.dbo.CARTERA.IDMOROSO =  
ACTIVE.dbo.DAM_CONTACTABILIDAD.IDMOROSO  
WHERE ACTIVE.dbo.DAM_CONTACTABILIDAD.PROVEEDOR_DATO = 'ZERT'  
AND FECHA_INGRESO_DATO BETWEEN '27/08/2021' AND '28/08/2021'  
AND ESTADO <> 'DEVUELTO'  
AND IDCLIENTE = 'FRANCES_2012'  
)  
, SUBQUERY2 AS (  
SELECT IDMOROSO, DATO, COUNT (*) AS CANTIDAD_APARICIONES FROM  
ACTIVE.dbo.DAM_CONTACTABILIDAD  
GROUP BY IDMOROSO, DATO)  
SELECT SUBQUERY1.IDMOROSO, SUBQUERY1.DATO, SUBQUERY1.IDCLIENTE, SUBQUERY1.ESTADO,  
SUBQUERY1.FILTROA FROM SUBQUERY1  
INNER JOIN SUBQUERY2  
ON SUBQUERY1.IDMOROSO = SUBQUERY2.IDMOROSO  
AND SUBQUERY1.DATO = SUBQUERY2.DATO  
AND CANTIDAD_APARICIONES = 1  
ORDER BY IDCLIENTE, FILTROA, IDMOROSO
```

Nota: para consultar los IDMorosos correspondientes a sus carteras, simplemente cambiar el IDCliente que aparece resultado.

Debemos de corregir los datos que enviamos y listo.