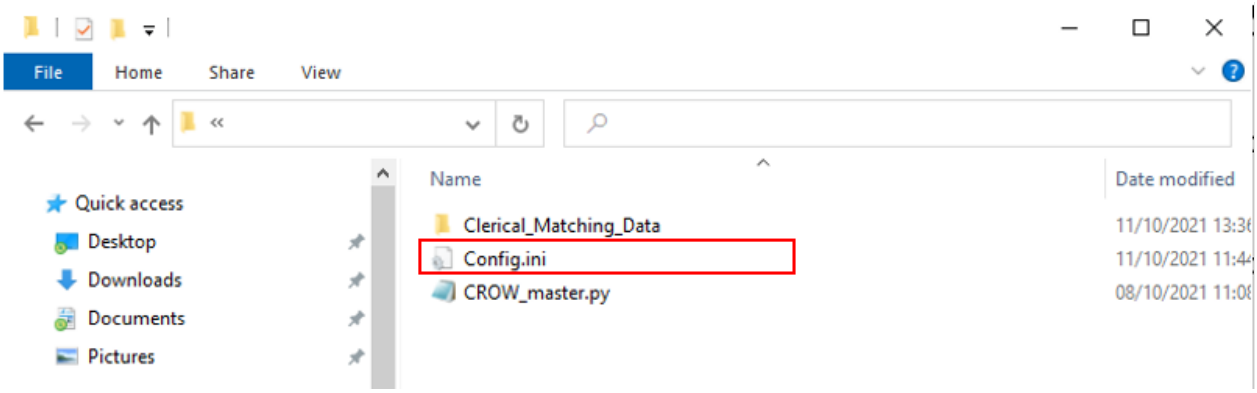
Using the CROW – Information for Clerical Coordinators

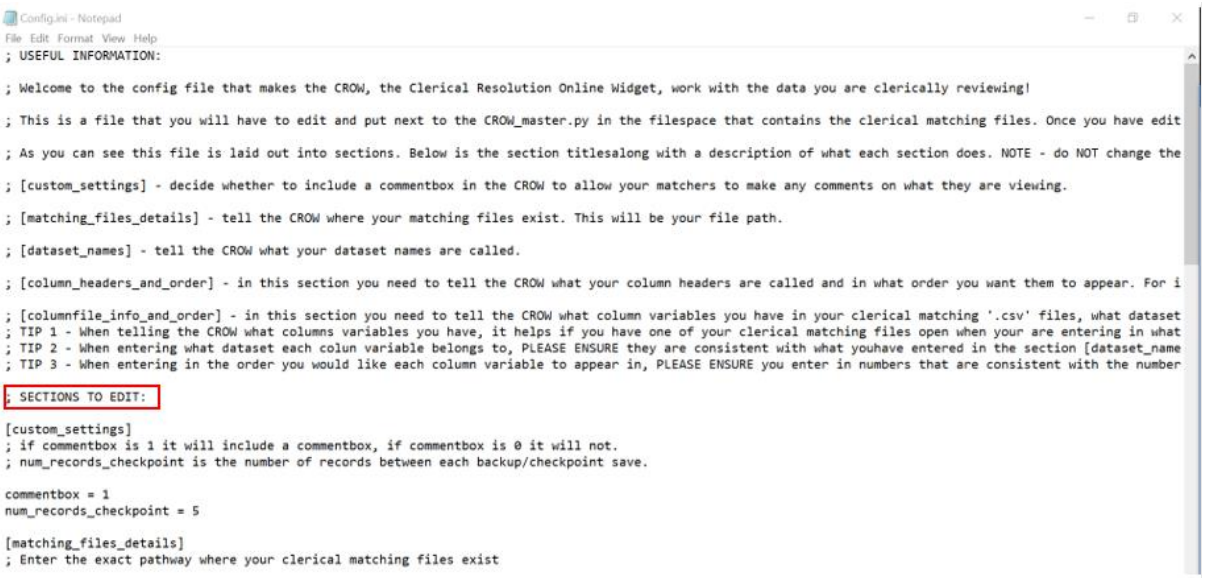
Setting up the CROW for your projects

Instructions for Clerical Coordinators to set up the CROW for the clerical matchers are as followed:

1) Once you have pulled the CROW from GitLab, put the Config.ini file and CROW_master.py file in your workspace. Open the Config.ini file in a text editor of your choice (e.g. NotePad, Word, Spyder). Note: The rest of these instructions will be shown using Notepad



2) Navigate to the 'SECTIONS TO EDIT' part.
NOTES:
   - More information can be found in the 'USEFUL INFORMATION' part should you wish to find out more.
   - In the 'SECTIONS TO EDIT' part, the sections are text contained within the [ ] (e.g. the first section is called '[custom_settings]'). **The text within each squared brackets should not be edited.**



3) Navigate to the '[custom_settings]' section. In this section you can change features that you may want to add or change about the CROW. See comments for more information on the variables. **NOTE:** This section may have more variables added in the future.

```
; SECTIONS TO EDIT:

[custom_settings]
; if commentbox is 1 it will include a commentbox, if commentbox is 0 it will not.
; num_records_checkpoint is the number of records between each backup/checkpoint save.

commentbox = 1
num_records_checkpoint = 5
```

4) Navigate to the '[matching_files_details]' section. In this section you need to add in the file pathway where you have stored your clerical matching files. To do this add in the pathway after the 'file_pathway =' part. **NOTE:** The CROW will still work if you leave this blank but it wont direct your clerical matchers to where the files are.

```
[matching_files_details]
; Enter the exact pathway where your clerical matching files exist

file_pathway = _____Your Pathway Directory_____
```

5) Navigate to the '[dataset_names]' section. In this section you can tell the CROW what your datasets are called. You can do this changing the bit after the equals signs.

```
[dataset_names]
; tell the CROW what your linked datasets are called.
; Simply edit A or B to what ever your datasets are called.
; if you wanted to add a dataset C then you can do this by
; creating a new variable and entering in its name.
; E.G. dataset3_name = C

dataset1_name = A
dataset2_name = B
```

6) Navigate to the '[column_headers_and_order]' section. In this section you need to think about what variables you have available in your data, that you want to the CROW to display. For instance, if I have 2 datasets that both have a first name variable I may want to include that as a column under the heading 'First_Name' so my clerical matchers can view that variable.

```
[column_headers_and_order]
; tell the CROW what your column headers are and in what order
; you want to order them in. This should be entered after the '=' sign
; in the following order:
;                 Example = <Column_Title> , <Column_Order>

fname = First_Name , 1
mname = Middle_Name , 2
sname = Surname , 3
DoB = DoB , 4
Sex = Sex , 5
pcode = Postcode , 6
Address = Address , 7
```

7) Navigate to the '[columnfile_info_and_order]' section. In this section you need to look at the headers of your clerical matching files and tell the CROW what they are, what dataset they belong to and in what order they should be placed. For instance, if I have 2 datasets both with a first name variable I would create 2 variables and after the '=' sign to those variables I would then add in the what the dataset label is called in the CSV, tell the crow which dataset it belongs to and then say which column header it belongs to using the same numbers I assigned under the relevant column header.

```
[columnfile_info_and_order]
; tell the CROW what column variables you want to display
; and be clerically reviewed. Tell The CROW what dataset this
; belongs to and in what order.
; This should be entered after the '=' sign in the following order:

; Example = <Column_Variable_From_CSV> , <Dataset_Column_Variable_Belongs_To> , <Column_Order>

fname_a = forename_a , A ,  1
fname_b = forename_b , B , 1
mname_a = middlename_a , A , 2
mname_b = middlename_b , B , 2
sname_a = surname_a , A , 3
sname_b = surname_b , B , 3
dob_a = dob_a , A , 4
dob_b = dob_b , B , 4
sex_a = sex_a , A , 5
sex_b = sex_b , B , 5
pcode_a = postcode_a , A , 6
pcode_b = postcode_b , B , 6
address_a = address_a , A , 7
address_b = address_b , B , 7
```

Note: the Column_Variable_From_CSV must be exactly the same as the column header in the CSV file or the app will not work.

Adding or removing a column and data source attributes

Adding or removing a column and its corresponding data source attribute is simple.

To remove a column and data source attributes simply delete it. For instance, if you do not have an middle name column in your CSV files simply remove the 'mname', 'mname_a' and 'mname_b' variables in the '[column_headers_and_order]' and the '[columnfile_info_and_order]' sections.

To add a column and its corresponding data source attributes you will have to follow these two simple steps below. For instance, if you have an ethnicity attribute that you want to include and display using the CROW, you can:

1. Navigate to the '[column_headers_and_order]' section and create a variable for your column header. Do this by creating a variable (see example below – 'ethnic' is variable) and then tell the CROW what your column header title is and what order you want it displayed in. **Don't forget to change the orders of the other columns so that they will all display accurately.**

```
[column_headers_and_order]
; tell the CROW what your column headers are and in what order
; you want to order them in. This should be entered after the '=' sign
; in the following order:
;                  Example = <Column_Title> , <Column_Order>

fname = First_Name , 1
mname = Middle_Name , 2
sname = Surname , 3
ethnic = Ethnicity , 4
DoB = DoB , 6
Sex = Sex , 5
pcode = Postcode , 7
Address = Address , 8
```

2. Navigate to the '[columnfile_info_and_order]' section and add in what your dataset attributes are called. Do this by creating a variable first (see example below - 'ethnic_a' and 'ethnic_b' are variables) then telling the CROW what the variable is called in your csv file (see example below - 'ethnic_Cen' and 'ethnic_Hog' are headers in my csv file), what dataset it belongs to (see example below – 'Census' and 'Hogwarts' are my datasets these are the same dataset titles outlined in the [dataset_names] section) and what order it should be put in (see example below – '4' is entered in both which corresponds to the Ethnicity column title position in the '[column_headers_and_order]' section). **Don't forget to change the orders of the other column variables you have to reflect this change.**

```
[columnfile_info_and_order]
; tell the CROW what column variables you want to display
; and be clerically reviewed. Tell The CROW what dataset this
; belongs to and in what order.
; This should be entered after the '=' sign in the following order:

; Example = <Column_Variable_From_CSV> , <Dataset_Column_Variable_Belongs_To> , <Column_Order>

fname_a = fname_Cen , Census ,  1
fname_b = fname_Hog , Hogwarts , 1
mname_a = midname_Cen , Census , 2
mname_b = midname_Hog , Hogwarts , 2
sname_a = sname_Cen , Census , 3
sname_b = sname_Hog , Hogwarts , 3
ethnic_a = ethnic_Cen , Census , 4
ethnic_b = ethnic_Hog , Hogwarts , 4
dob_a = dob_Cen , Census , 6
dob_b = dob_Hog , Hogwarts , 6
sex_a = sex_Cen , Census , 5
sex_b = sex_Hog , Hogwarts , 5
pcode_a = pcode_Cen , Census , 7
pcode_b = pcode_Hog , Hogwarts , 7
address_a = address_Cen , Census , 8
address_b = address_Hog , Hogwarts , 8
```