In [1]:

```
import pandas as pd
import openpyxl
import os
import numpy as np
import warnings
warnings.filterwarnings('ignore')
os.listdir()
```

Out[1]:

```
['.ipynb_checkpoints', 'Consolidator.ipynb', 'Datafiles', 'Outputs']
```

In [2]:

```
datadir="Datafiles"
os.chdir(datadir)
```

In [3]:

```
os.listdir()
```

Out[3]:

```
['Account Report - Gweru Mon 29 08 2022.xlsx',
 'Expense Report -  Gweru Mon 29 08 2022.xlsx',
 'Income Report -  Gweru Mon 29 08 2022.xlsx',
 'Loan Given Report -  Gweru Mon 29 08 2022.xlsx',
 'Loan Taken Report -  Gweru Mon 29 08 2022.xlsx']
```

Reading the income reports

In [4]:

```
income_reports=[] ### To create an empty dataframe where we can append our income reports d
data =pd.read_excel('Income Report -  Gweru Mon 29 08 2022.xlsx', sheet_name =2) ### Readin

## Converting the uppercases to lower cases and replacing spaces with underscores(_) on eac

data.columns=data.columns.str.replace(' ','_')

data.columns=data.columns.str.lower()

data['file']='income'# Adding a file variable so as to distinquish various types for report

## Selecting the variables of interest in the dataset data
data1 = data[['respondent_id',  'city_of_residence',
       'citizenship', 'income_report_data_id',
       'income_report_income_source_name', 'income_report_amount',
       'income_report_currency',
       'income_report_date_received', 'income_report_account',
       'income_report_account_report_id', 'income_report_surveyor',
       'income_report_customer', 'income_report_timeline',
       'income_report_transaction_type', 'income_report_payment_type',
       'income_report_date_created', 'income_report_last_updated',
       'income_report_edit_count', 'income_report_description', 'file']].copy()


    # Renaming the variables names
data1=data1.rename(columns={
                        'income_report_data_id':'data_id',
                        'income_report_amount':'amount',
                        'income_report_currency':'currency',
                        'income_report_income_source_name':'reason_name',
                        'income_report_account':'account',
                        'income_report_date_received':'transaction_date',
                         'income_report_account_report_id':'transaction_id',
                        'income_report_transaction_type':'transaction_type',
                        'income_report_timeline':'timeline',
                        'income_report_payment_type':'payment_type',
                        'income_report_date_created':'date_created',
                        'income_report_last_updated':'date_last_updated',
                        'income_report_edit_count':'edit_count',
                        'income_report_surveyor':'surveyor',
                        'income_report_customer':'tool',
                        'income_report_description':'description'})

income_reports.append(data1)
income_reports=pd.concat(income_reports)## joining the data into the initialized empty data
```

Reading the expense reports

In [7]:

```
expense_reports=[]
data =pd.read_excel('Expense Report -  Gweru Mon 29 08 2022.xlsx', sheet_name =2)
data.columns=data.columns.str.replace(' ','_')

data.columns=data.columns.str.lower()
data['file']='expense'
data1=data[['respondent_id', 'city_of_residence',
       'citizenship', 'expense_report_data_id', 'expense_report_reason',
        'expense_report_amount','expense_report_currency',
        'expense_report_date',
       'expense_report_account', 'expense_report_account_report_id',
       'expense_report_surveyor', 'expense_report_vendor',
       'expense_report_timeline', 'expense_report_transaction_type',
       'expense_report_payment_type', 'expense_report_date_created',
       'expense_report_last_updated', 'expense_report_edit_count',
       'expense_report_description', 'file']].copy()
data1=data1.rename(columns={'expense_report_account_report_id':'transaction_id',
                            'expense_report_data_id':'data_id',
                         'expense_report_amount': 'amount',
                         'expense_report_currency':'currency',
                         'expense_report_reason':'reason_name',
                         'expense_report_account':'account',
                         'expense_report_date':'transaction_date',
                         'expense_report_transaction_type':'transaction_type',
                         'expense_report_timeline':'timeline',
                         'expense_report_payment_type':'payment_type',
                         'expense_report_date_created':'date_created',
                            'expense_report_last_updated':'date_last_updated',
                         'expense_report_edit_count':'edit_count',

                         'expense_report_surveyor':'surveyor',
                         'expense_report_vendor':'tool',
                         'expense_report_description':'description'})
expense_reports.append(data1)
expense_reports=pd.concat(expense_reports)
```

Reading the account reports

In [9]:

```
accounts_reports=[]
data =pd.read_excel('Account Report - Gweru Mon 29 08 2022.xlsx', sheet_name =2)
data.columns=data.columns.str.replace(' ','_')

data.columns=data.columns.str.lower()

data['file']='Accounts'

data1 = data[['respondent_id','city_of_residence',
       'citizenship', 'account_report_data_id', 'account_report_account_name',
       'account_report_deposit/withdrawal', 'account_report_amount',
       'account_report_currency', 'account_report_date',
       'account_report_data_type', 'account_report_surveyor',
       'account_report_date_created', 'account_report_last_updated',
       'account_report_edit_count', 'account_report_description','file']].copy()

# # ###.....................................................................................
data1=data1.rename(columns={
                        'account_report_data_id':'data_id',
                        'account_report_amount':'amount',
                        'account_report_currency':'currency',
                        'account_report_account_name':'account',
                        'account_report_deposit/withdrawal':'reason_name',
                        'account_report_date':'transaction_date',
                        'account_report_data_type':'transaction_type',
                        'account_report_date_created':'date_created',
                        'account_report_last_updated':'date_last_updated',
                        'account_report_edit_count':'edit_count',
                        'account_report_surveyor':'surveyor',

                        'account_report_description':'description'})
accounts_reports.append(data1)
accounts_reports=pd.concat(accounts_reports)
```

Loans taken

In [12]:

```
loans_taken_data=[]


repayments = pd.read_excel('Loan Taken Report -  Gweru Mon 29 08 2022.xlsx',
                     sheet_name = 'Loan Repayment Reports')
repayments.columns=repayments.columns.str.replace(' ','_')

repayments.columns=repayments.columns.str.lower()

repayments['file']='loans_taken_repayment'
repayments1 = repayments[['respondent_id','city_of_residence',
       'citizenship', 'loan_taken_repayment_report_data_id',
       'loan_taken_repayment_report_source_name',
       'loan_taken_repayment_report_repaid_amount',
       'loan_taken_repayment_report_currency',
       'loan_taken_repayment_report_date_paid',
       'loan_taken_repayment_report_surveyor',
       'loan_taken_repayment_report_account',
       'loan_taken_repayment_report_account_report_id',
       'loan_taken_repayment_report_date_created',
       'loan_taken_repayment_report_last_updated',
       'loan_taken_repayment_report_edit_count',

       'loan_taken_repayment_report_description','file']].copy()



repayments1=repayments1.rename(columns={
                                'loan_taken_repayment_report_data_id':'data_id',
                                'loan_taken_repayment_report_repaid_amount':'amount',
                                'loan_taken_repayment_report_currency':'currency',
                                'loan_taken_repayment_report_source_name':'reason_name',
                                "loan_taken_repayment_report_account":'account',
                                'loan_taken_repayment_report_account_report_id':'transactio
                                'loan_taken_repayment_report_date_paid':'transaction_date',
                                'loan_taken_repayment_report_date_created':'date_created',
                                'loan_taken_repayment_report_last_updated':'date_last_updat
                                'loan_taken_repayment_report_edit_count':'edit_count',
                                'loan_taken_repayment_report_description':'description',
                                'loan_taken_repayment_report_surveyor':'surveyor'})


loans_taken_data.append(repayments1)
loans_taken_data = pd.concat(loans_taken_data)
```

Loans given report

In [13]:

```
loans_given_data=[]


repayments = pd.read_excel('Loan Given Report -  Gweru Mon 29 08 2022.xlsx',
                     sheet_name = 'Loan given repayments')
repayments.columns=repayments.columns.str.replace(' ','_')

repayments.columns=repayments.columns.str.lower()



repayments['file']='loans_given_repayment'
repayments1 = repayments[['respondent_id', 'country_of_residence',
       'citizenship', 'loan_given_repayment_report_data_id',
       'loan_given_repayment_report_source_name',
       'loan_given_repayment_report_repaid_amount',
       'loan_given_repayment_report_currency',
       'loan_given_repayment_report_date',
       'loan_given_repayment_report_surveyor',
       'loan_given_repayment_report_account',
       'loan_given_repayment_report_account_report_id',
       'loan_given_repayment_report_date_created',
       'loan_given_repayment_report_last_updated',
       'loan_given_repayment_report_edit_count',
       'loan_given_repayment_report_description',
       'file']].copy()


repayments1=repayments1.rename(columns={'loan_given_repayment_report_description':'descript
                                'loan_given_repayment_report_data_id':'data_id',
                                'loan_given_repayment_report_account_report_id':'transactio
                                'loan_given_repayment_report_repaid_amount':'amount',
                                'loan_given_repayment_report_currency':'currency',
                                'loan_given_repayment_report_source_name':'reason_name',
                                "loan_given_repayment_report_account":'account',
                                'loan_given_repayment_report_date':'transaction_date',
                                'loan_given_repayment_report_date_created':'date_created',
                                'loan_given_repayment_report_last_updated':'date_last_updat
                                'loan_given_repayment_report_edit_count':'edit_count',
                                'loan_given_repayment_report_surveyor':'surveyor'})


loans_given_data.append(repayments1)
loans_given_data = pd.concat(loans_given_data)
```

Data consolidation for the Mali corner shop

In [14]:

```
consolidated=income_reports.append([expense_reports,
                                    accounts_reports,
                                    loans_taken_data,
                                    loans_given_data],sort=False).copy()

# consolidated.transaction_type = np.where((consolidated.file=='Accounts'),'cash',
# #                               consolidated.transaction_type)
consolidated.file = np.where((consolidated.file=='accounts') &
(consolidated.reason_name=='Deposited'), 'savings_deposit', consolidated.file)
consolidated.file = np.where((consolidated.file=='accounts') &
(consolidated.reason_name=='Withdrawn'), 'savings_withdrawal', consolidated.file)
```

In [15]:

```
consolidated.columns
```

Out[15]:

```
Index(['respondent_id', 'city_of_residence', 'citizenship', 'data_id',
       'reason_name', 'amount', 'currency', 'transaction_date', 'account',
       'transaction_id', 'surveyor', 'tool', 'timeline', 'transaction_type',
       'payment_type', 'date_created', 'date_last_updated', 'edit_count',
       'description', 'file', 'country_of_residence'],
      dtype='object')
```

In [16]:

```
consolidated.to_csv("../Outputs/Consolidated_Gweru.csv")
```

In [ ]:

```

```