Sub FindAndExtractDuplicates()

Dim mainFolder As String, duplicatesFolder As String

Dim yearFolder As String, filePath As String

Dim year As Integer, month As Integer

Dim monthNames As Variant

Dim wbOriginal As Workbook, wbDuplicates As Workbook

Dim wsOriginal As Worksheet, wsDuplicates As Worksheet

Dim ssnitCol As Range, cell As Range

Dim dict As Object, duplicateRows As Collection

Dim outputFileName As String

' Initialize

mainFolder = "C:\Users\USER\Desktop\STALLION TIGER COMPANY LTD\"

duplicatesFolder = "C:\Users\USER\Desktop\Duplicates\"

monthNames = Array("jan", "feb", "mar", "apr", "may", "jun", "jul", "aug", "sep", "oct", "nov", "dec")

' Create duplicates folder if missing

On Error Resume Next

MkDir duplicatesFolder

On Error GoTo 0

' Loop through years (2014 to 2024)

For year = 2014 To 2024

yearFolder = mainFolder & year & "\"

' Loop through months (jan to dec)

For month = LBound(monthNames) To UBound(monthNames)

filePath = yearFolder & monthNames(month) & year & ".xlsx"

' Check if file exists

If Dir(filePath) <> "" Then

Set dict = CreateObject("Scripting.Dictionary")

Set duplicateRows = New Collection

' Open the workbook

Set wbOriginal = Workbooks.Open(filePath)

Set wsOriginal = wbOriginal.Sheets(1) ' Assuming data is in Sheet1

' Check Column A for duplicates (skip header)

For Each cell In wsOriginal.Range("A2:A" & wsOriginal.Cells(Rows.Count, 1).End(xlUp).Row)

If dict.exists(cell.Value) Then

' Add duplicate row to collection

duplicateRows.Add cell.EntireRow

Else

dict.Add cell.Value, 1 ' Track SSNIT

End If

Next cell

' If duplicates found, create a new workbook

If duplicateRows.Count > 0 Then

Set wbDuplicates = Workbooks.Add

Set wsDuplicates = wbDuplicates.Sheets(1)

' Copy headers

wsOriginal.Rows(1).Copy wsDuplicates.Rows(1)

' Copy duplicate rows

For Each row In duplicateRows

row.Copy wsDuplicates.Rows(wsDuplicates.Cells(Rows.Count, 1).End(xlUp).Row + 1)

Next row

' Save duplicates file

outputFileName = duplicatesFolder & monthNames(month) & year & "\_Duplicates.xlsx"

wbDuplicates.SaveAs outputFileName

wbDuplicates.Close

End If

' Close original workbook

wbOriginal.Close False

End If

Next month

Next year

MsgBox "Duplicate extraction complete!", vbInformation

End Sub