

# VBA Code Documentation

## Module: ThisWorkbook.cls

### Full Code:

```
Attribute VB_Name = "ThisWorkbook"
Attribute VB_Base = "0{00020819-0000-0000-C000-000000000046}"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = True
Attribute VB_TemplateDerived = False
Attribute VB_Customizable = True
```

## Module: Sheet1.cls

### Full Code:

```
Attribute VB_Name = "Sheet1"
Attribute VB_Base = "0{00020820-0000-0000-C000-000000000046}"
Attribute VB_GlobalNameSpace = False
Attribute VB_Creatable = False
Attribute VB_PredeclaredId = True
Attribute VB_Exposed = True
Attribute VB_TemplateDerived = False
Attribute VB_Customizable = True
```

## Module: Module1.bas

### Functions and Subroutines:

Name	Parameters
PreprocessData	

### Variables:

Name	Type
ws	Worksheet
lastRow	Long
i	Long
rowDict	Object
tempRow	String
cellValue	Variant

### Comments:

Set the worksheet  
Change to your sheet name

Find the last used row and column  
Initialize the dictionary  
Loop through each row from bottom to top to avoid skipping rows after deletion  
Loop through each cell in the row  
Check if the cell is empty or not an integer  
Build the row string  
Check for duplicate rows

### Logic:

```
For i = lastRow To 1 Step -1
For j = 1 To lastCol
If IsEmpty(cellValue) Or Not IsNumeric(cellValue) Or cellValue <> Int(cellValue) Then
ws.Rows(i).Delete
If rowDict.exists(tempRow) Then ws.Rows(i).Delete
```

### Full Code:

```
Attribute VB_Name = "Module1"
Sub PreprocessData()
Dim ws As Worksheet
Dim lastRow As Long, lastCol As Long
Dim i As Long, j As Long
Dim rowDict As Object
Dim tempRow As String
Dim cellValue As Variant
' Set the worksheet
Set ws = ThisWorkbook.Sheets("Sheet1") ' Change to your sheet name
' Find the last used row and column
lastRow = ws.Cells(ws.Rows.Count, 1).End(xlUp).row
lastCol = ws.Cells(1, ws.Columns.Count).End(xlToLeft).Column
' Initialize the dictionary
Set rowDict = CreateObject("Scripting.Dictionary")
' Loop through each row from bottom to top to avoid skipping rows after deletion
For i = lastRow To 1 Step -1
tempRow = ""
' Loop through each cell in the row
For j = 1 To lastCol
cellValue = ws.Cells(i, j).value
' Check if the cell is empty or not an integer
If IsEmpty(cellValue) Or Not IsNumeric(cellValue) Or cellValue <> Int(cellValue) Then
ws.Rows(i).Delete
GoTo NextRow
End If
' Build the row string
tempRow = tempRow & cellValue & Chr(0)
Next j
' Check for duplicate rows
If rowDict.exists(tempRow) Then
ws.Rows(i).Delete
Else
rowDict.Add tempRow, True
End If
NextRow:
Next i
MsgBox "Data preprocessing completed successfully."
End Sub
```