**🧠 Training Worksheet: Build a Dynamic Fake Data Generator in Excel VBA**

**📘 Objective**

Create a macro that generates realistic fake employee data based on column headers, using a UserForm to specify the number of rows. This exercise teaches modular macro design, type inference, and user interaction.

**🧩 Step 1: Prepare Your Workbook**

1. Open Excel and save your file as **Macro-Enabled Workbook (.xlsm)**
2. In **Row 1**, enter sample headers like:

Employee ID | Name | Email | Gender | Ethnicity | Location | Grade | Record Date

**🧩 Step 2: Insert a VBA Module**

1. Press Alt + F11 to open the **VBA Editor**
2. Go to Insert > Module
3. Rename the module to modFakeDataGenerator
4. Paste the following code:

Function InferType(header As String) As String

If header Like "\*Date\*" Then InferType = "Date"

ElseIf header Like "\*ID\*" Then InferType = "ID"

ElseIf header Like "\*Name\*" Then InferType = "Name"

ElseIf header Like "\*Email\*" Then InferType = "Email"

ElseIf header Like "\*Gender\*" Then InferType = "Gender"

ElseIf header Like "\*Ethnicity\*" Then InferType = "Ethnicity"

ElseIf header Like "\*Location\*" Then InferType = "Location"

ElseIf header Like "\*Grade\*" Then InferType = "Grade"

Else InferType = "Text"

End Function

Function GenerateValue(dataType As String) As Variant

Select Case dataType

Case "Date": GenerateValue = DateSerial(2020, Int(Rnd() \* 12 + 1), Int(Rnd() \* 28 + 1))

Case "ID": GenerateValue = "EMP" & Format(Int(Rnd() \* 1000 + 100), "000")

Case "Name": GenerateValue = Choose(Int(Rnd() \* 5 + 1), "Alicia", "Marcus", "Priya", "Daniel", "Fatima")

Case "Email": GenerateValue = "user" & Int(Rnd() \* 1000) & "@example.com"

Case "Gender": GenerateValue = Choose(Int(Rnd() \* 2 + 1), "Male", "Female")

Case "Ethnicity": GenerateValue = Choose(Int(Rnd() \* 4 + 1), "Asian", "White", "Black or African American", "Hispanic or Latino")

Case "Location": GenerateValue = Choose(Int(Rnd() \* 5 + 1), "Chicago", "Seattle", "Austin", "Denver", "Boston")

Case "Grade": GenerateValue = Choose(Int(Rnd() \* 3 + 1), "Grade: 7 and Below", "Grade: 8 and Above", "Grade 08-11")

Case Else: GenerateValue = "Sample " & Int(Rnd() \* 100)

End Select

End Function

Sub GenerateFakeData(rowCount As Integer)

Dim headers As Collection

Set headers = New Collection

Dim cell As Range

For Each cell In Range("A1", Cells(1, Columns.Count).End(xlToLeft))

headers.Add cell.Value

Next cell

Dim r As Integer, c As Integer

For r = 2 To rowCount + 1

For c = 1 To headers.Count

Dim header As String: header = headers(c)

Dim dataType As String: dataType = InferType(header)

Cells(r, c).Value = GenerateValue(dataType)

Next c

Next r

MsgBox rowCount & " rows of fake data generated successfully!", vbInformation

End Sub

Sub LaunchFakeDataForm()

frmRowCount.Show

End Sub

**🧩 Step 3: Create the UserForm**

1. In the VBA Editor, go to Insert > UserForm
2. Rename it to frmRowCount using the **Properties window** (F4)
3. Add the following controls:

|  |  |  |
| --- | --- | --- |
| **Control Type** | **Name** | **Caption / Purpose** |
| Label | lblPrompt | "Enter number of rows to generate:" |
| TextBox | txtRowCount | (blank) — user types a number |
| CommandButton | btnGenerate | "Generate" — runs the macro |

**🧩 Step 4: Add Code to the UserForm**

Double-click the **Generate** button and paste this code:

Private Sub btnGenerate\_Click()

Dim rowCount As Integer

If IsNumeric(txtRowCount.Value) Then

rowCount = CInt(txtRowCount.Value)

If rowCount > 0 Then

GenerateFakeData rowCount

Unload Me

Else

MsgBox "Please enter a number greater than 0.", vbExclamation

End If

Else

MsgBox "Please enter a valid number.", vbExclamation

End If

End Sub

**🧪 Step 5: Run and Test**

* In the VBA Editor, run LaunchFakeDataForm
* Enter a number (e.g., 30) and click **Generate**
* Review the generated data starting from row 2

**🧠 Reflection & Extension**

* ✅ How could you add edge cases like nulls or duplicates?
* ✅ Could you use a config sheet to override type inference?
* ✅ What would change if you exported this to CSV or integrated with Power BI?