

## Household Sample Data Explanation

This document will explain some of the thinking and steps to prepare your data for mail merge with a few householding scenarios we discussed during the Office Hour and the excel spreadsheet example called, **sampledata\_household.xls**.

### Preparing your Data

There are a few basic tips to consider in maintaining clean data, which will save you from headaches whenever you need to prepare your data for doing searches and sorting, creating lists for mailings, walk or phone lists, or before manipulating the data using calculations and formulas, as we did with excel. (also see the document called, “clean\_data.pdf”, at <http://network.progressivetech.org/trainings/back-office/techcamp-online-mail-merge>).

- Standardize how you enter addresses (the US Postal office has a document describing postal addressing standards called, Publication 28, at <http://pe.usps.gov/cpim/ftp/pubs/Pub28/pub28.pdf>)
- If you need to update and standardize addresses, phone numbers, etc., for a list of contacts, Melissa Data, <http://www.melissadata.com>, can provide this service.
- Periodically check and verify data being entered is correct
- Periodically check for duplicate information

### The Scenarios

Before we started applying formulas to create households, we did the following first:

- Apply a filter for each column, which makes things easier to sort and search
- Sorted by zip code (the column “Zip”) and then address (the column “StreetAddress”)

Then we had to take some time to study our data and look at the different scenarios how names were being entered in each record (each line) and in multiple records (multiple lines) by looking those records that shared the same address. So here are the scenarios that exist in our example:

1. One record will contain 2 different first names (in First Name 1 & 2) and 2 different last names (in Last Name 1 & 2) (see line 17)
2. One record contains 2 different first names (in First Name 1 & 2) and the same last names in (Last Name 1 & 2) (see line 2)
3. One record contains multiple first names (in First Name 1) and the same last name (in Last Name 1) (see lines 5 and 24)
4. Separate records that have the same address (see lines 11 & 12, and lines 29-31)

We then created four new columns to build the household names:

- **New household** (column A) – indicates if we’re creating a new household

- **Last in household** (column B) – indicates if the record (or line, in this case) contains the last name in the household
- **names** (column C) – builds the names from First Name 1 & 2 and Last Name 1 & 2 fields
- **new\_household names** (column D) – builds the household name to be used

Since we want to print a label or mail to ONE household only, the solution we used to handle the different scenarios is to search **where “Last in household” = “YES”**, then just use the **“new\_household names” field to print the label or mailing.**

## Using Excel formulas to create households for mailing

So now, let’s break down the thinking behind the logic and formulas we used to create the values in our new columns: “New household,” “Last in household,” “names,” and “new\_household names”. The following formulas will begin building the household names for each record, then you’ll be able to filter by the “Last in Household” to get the full household name so you can use it in greetings and mailing labels.

We will be using the IF formula for much of the work which looks like:  
**=IF(logical test, value if true, [value if false])**

Also, to make referencing fields within formulas easier, we named the columns with values using the same names as the column headings, such as **FirstName1** and **StreetAddress**. You’ll see examples of this in the formulas below. To create these names you can do the following:

In MS Excel 2003:

- click on the column you want to name first
- INSET-> NAME -> DEFINE
- In the dialog box, the value of the name will default to the column name you entered. You can change this if you needed.
- Then click OK

In MS Excel 2007:

- click on the column you want to name first
- click on the FORMULAS tab
- In the NAME MANAGER area, click on DEFINE NAME and the value of the name will default to the column name you entered. You can change this if you needed.
- Then click OK

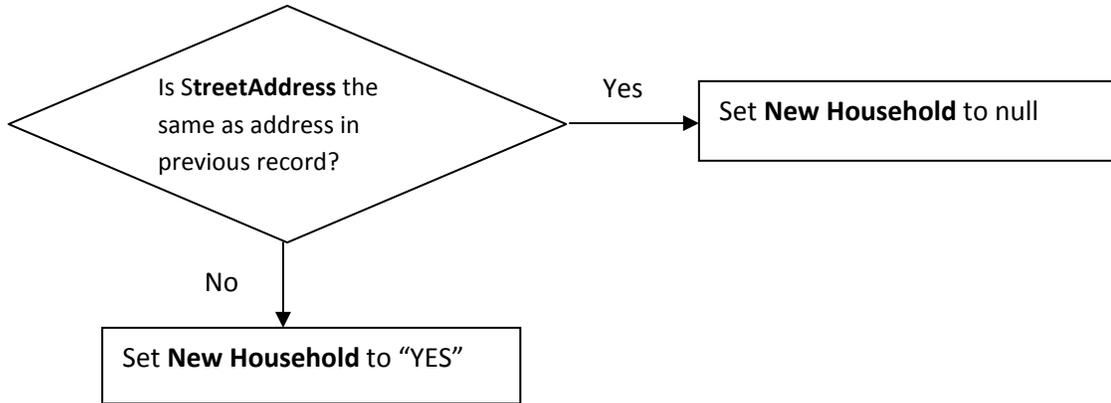
Here’s the householding process:

1. **Sort your data by zip and then by address first.**
2. Add a new column called **New household** - First we’re going to check if we need to create a new household by comparing the address of the current record, to the address of the record

ABOVE (or previous record). If the addresses are not the same (which means they are in the same household), then set the value to “yes”.

Here’s the logic (see **New Household**, column A, line 11):

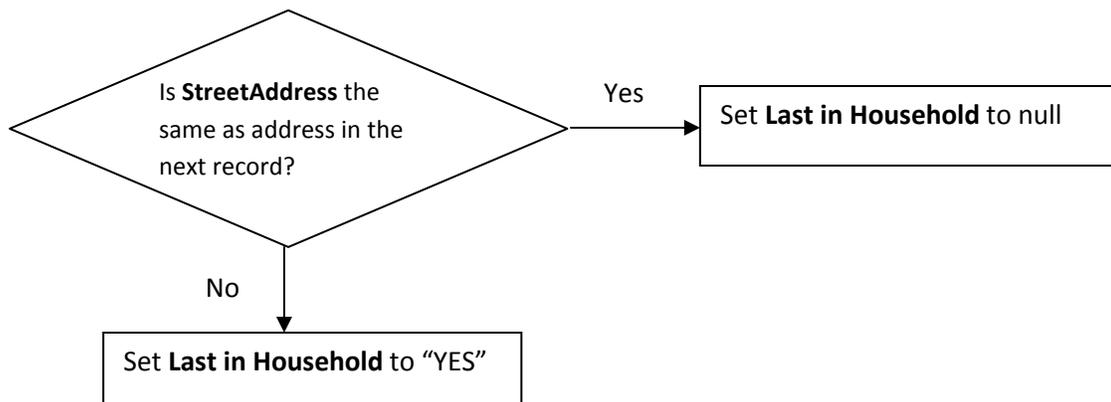
**=IF(StreetAddress<>I10,"yes", "")**



3. Add a new column called **Last in household** – since we’ve sorted the list, we can no look for the last record associated with a household, by comparing the address of the record BELOW (or next record). If the addresses are not the same, then the value is set the “yes” indicating that it’s the last record in the household.

Here’s the logic (see **Last In Household**, column B, line 12):

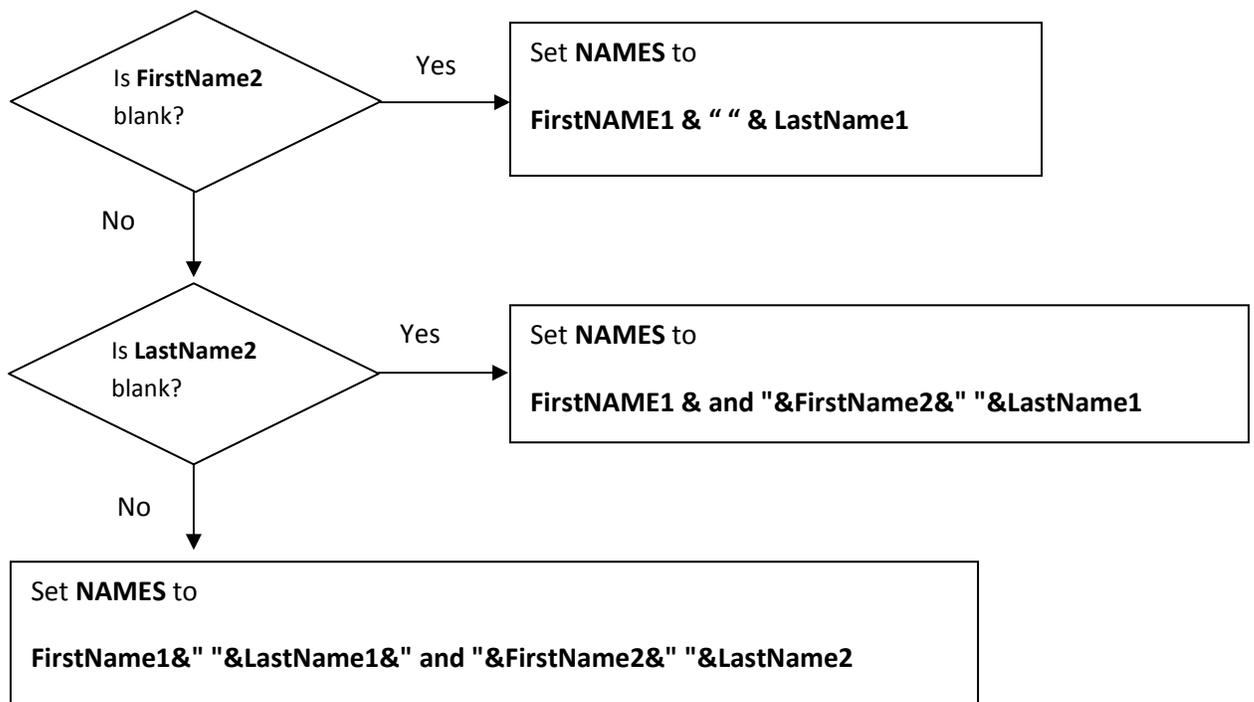
**=IF(I13<>StreetAddress,"yes", "")**



4. Add a new column called **Names** – next we will build the names for each record by checking if there’s information in FirstName1, FirstName2, LastName1, and LastName2 fields.

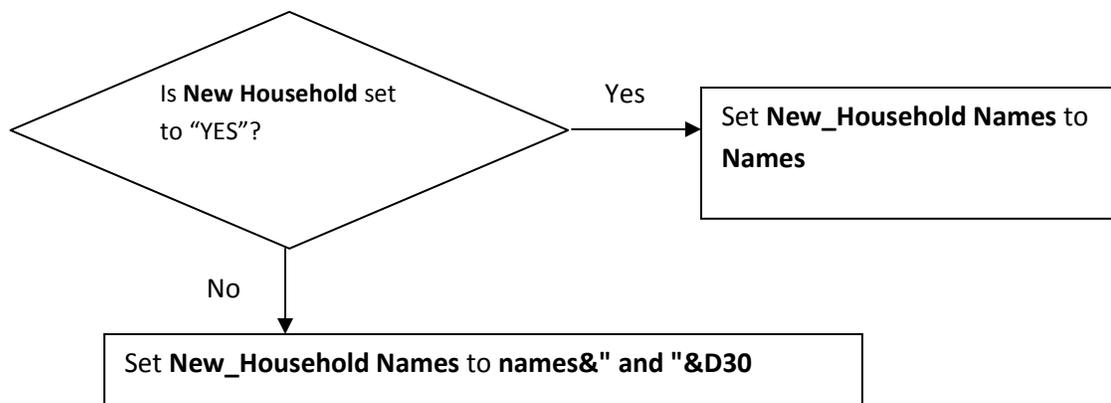
Here’s the logic (see **NAMES**, column C, line 20):

**=IF(FirstName2<>"",IF(LastName2<>"",FirstName1&" "&LastName1&" and "&FirstName2&" "&LastName2, FirstName1&" and "&FirstName2&" "&LastName1), FirstName1&" "&LastName1)**



5. Add a new column called **New\_Household Names** – next we will build the household names used for each record by checking if the status of the current record’s the **New\_Household** field and building the **New\_Household Names** field as follows:

Here’s the logic (see **New\_Household Names**, column D, line 31):  
 =IF(New household="yes",names,names&" and "&D30)



6. When you’re ready to send mailings to a household, all you need to do is filter for “**YES**” in the **Last In Household** field (column B) and use **New\_Household Names** field for your greeting line.