Rough Procedures

Please remember that there are almost always two or more ways to do things in Excel. The procedures offered in this handout are suggestions.

The data that accompanies the scenario includes two tables – one with transactional data (determinations, which is on the benefits tab) and one with characteristic data (dates of birth and gender, which is on the lookup tab). Having two tables like this is very common in relational databases. In order to answer the questions above, it is best to combine information from the two tables into a single table (array).

Unfortunately, the data on the lookup tab doesn't have ages. We could use filters and tediously determine ages by year – or by age group such as those requested above; or we could use the functions within Excel to quickly convert the text entries to serial dates and calculate ages. The following are general procedures to convert text fields to dates and to and calculate ages:

1. Change the text entries to years, months, and dates using LEFT, MID, and RIGHT functions.
2. Use the DATE function to convert those entries into serial dates.
3. Use the DATEDIF function to calculate age making reference to the date of birth and the current date (or date given). This will be contrasted with using other mathematical approaches.

Once the age and gender are available to match with the transactions table, we want to sort and name the table array. Then we're ready to merge the tables using the VLOOKUP function:

1. In a cell adjacent to the transactional table (usually in the cell to the right of the amount, we enter the VLOOKUP function making reference to the table on the lookup tab (remember, we gave it a name).
2. The function uses the format: =VLOOKUP(lookup_value, table_array, col_index_num, [range_lookup]). These are explained:
    a. Lookup_value – This is the cell on the benefits table from which the matching element is located. This is in column A.
    b. Table_array – This is the associated table. In this case it's the table on the lookup tab. We gave a name to this table array.
    c. Col_index_num – This refers to the cell in the matching table whose data you want in the benefits table. The col_index_num is a number – meaning the *nth* column from the left (if the data is in the ninth column from the left, we enter 9 for the col_index_num.
    d. [Range_lookup] – Refers to the matching value in the lookup table. This is the value that matchesa the lookup_value explained above. In this case, it's the first column so we enter 1.
3. We'll do this twice – with different col_index_num values because we want the age and the gender – this will be demonstrated.

We notice in the questions from the LFA that we need to determine the unduplicated counts of individuals by program. We'll discuss at least two ways this can be achieved – but one of those ways is to create one additional field (add a column) to the transactional table using SORTS and a NESTED IF statement to identify duplicates. We can use the same approach to find individuals participating in both programs A and B. This will be demonstrated.

Now we're ready to begin making calculations to answer the questions above. We'll use advanced IF statements described in the pre-training exercises to answer some of the basic questions – then we'll use a pivot table. Pivot tables provide superior speed and management of the data over single functions. Pivot tables will be demonstrated using this data.

PIVOT Tables

The key to using pivot tables is to understand what they do … Among other functions, a pivot table can automatically sort, count, sum, or give the average of the data stored in a single table, displaying the results in a second table showing the summarized data. Users create or change the summary's structure by dragging and dropping fields into columns, rows, report filters, or the summation values (body) of the pivot table. Moving data from columns to rows or to the report filter allows quick “rotation” or “pivoting” of the information. Much more information about pivot tables is available online.

Select the range – or give the table a name. Please note that you MUST have a contiguous array (no blank rows or spaces between whole columns) for pivot tables to view all data in a table. Individual cells may be null; however, an entire record may not be blank. Then create the pivot table (this will be demonstrated). Fields may be dragged and dropped into any of four areas as follows (info courtesy of Wikipedia with my own comments added):

Report Filter is used to apply an Excel filter to an entire table. For example, if you drag the program field into to this area, then the table constructed will have a report filter inserted above the table. This report filter will have drop-down options (A or B since there are only two options). When you choose an option from this drop-down list ("B" for example), then the table that would be visible will contain only the data from those rows that have the "B" program.

Column Labels are used to apply an Excel filter to one or more columns that have to be shown in the pivot table. For instance if we drag the field "Gender" into this area then data in the table constructed will have values corresponding to gender (M and F). There will also be one added column of Total. In the example above, this instruction will create 3 columns in the table - M, F, and Grand Total. There will be a filter above the data - Column Labels, from which you can select or deselect a particular value (M or F).

Row Labels are used to apply an Excel filter to one or more rows that have to be shown in the pivot table. For instance if we drag the field "Age" into this area then the table constructed will have values corresponding to age (by year). There will also be one added row of "Grand Total". There will be a filter above the data - Row Labels, from which you can select or deselect a particular age or multiple ages for the Pivot table.

Summation Values - This usually receives a field that has numerical values that can be used for different types of calculations. However, using text values will provide counts only. If we drag "Amount" to this area then this will sum (or average, etc) the values corresponding to the report filter, rows, and columns selected.