

## Example 4 - Spreadsheet functions: Ranking, weighted ranking, order of formula processing.

Other features illustrated: Summing calculation, hiding columns, creating column spanners, viewing hidden columns, IF conditions

Background: This example shows how you can use ranking and apply weights to those ranks in order to answer questions like "What is the best county for ...?"

Goals: 1) Create a table that will rank the best county in the US in which to live, based on our three criteria - commute time, median household income, and median housing value. 2) Limit our rankings to counties that meet our population size criteria - between 50,000 and 150,000 people.

### Steps

**Goal 1** Create a table that will rank the best county in the US in which to live, based on our three criteria - commute time, median household income, and median housing value.

1. Open the Ferrett Session File (or saved data basket) named [sf3\\_bestcounty\\_tolive.fsf](#).
2. Go to the Step 2 tab to see the variables in the data basket - number of workers not working at home, aggregate travel time to work, median household income, median value of owner-occupied housing units, and county (all counties in US).
3. "Make a Table"
4. Turn off the "Automatic Totals" (in the "Options" menu) and drag the county variable to column 1 to define the rows.
5. Drag the aggregate travel time variable, P033001, to column 2 - C2, R1.
6. Drag the workers variable, P031002, to column 3 - C3, R1.
7. Skip two columns and drag the median household income variable, P053001, to column 6 - C6, R1.
8. Skip 2 columns and drag the median housing unit value variable, P085001, to column 9 - C9, R1.
9. "Go Get Data"

*Create calculations and rankings.*

1. Calculate the average commute time in column 4.
  - o Click in the C4 column header to select the column in which to create the calculation.
  - o In the formula bar, enter **=comp(c2/c3)** == aggregate travel time divided by workers not working at home equals average travel time.
2. Create the county ranking for commute time to work. The lower the commute time, the higher the rank number.
  - o Click the C5 column header to select the column in which to create the ranking.
  - o In the formula bar, enter **=rank(c4)**
3. Create the county ranking for median household income. In this case, we want the higher value to have the higher rank number. Therefore, we will need to create an inverse income column on which to create the ranking.
  - o Click the C7 column header to select the column in which to create the calculation for inverting the income.
  - o In the formula bar, enter **=comp(0-c6)**
  - o Click the C8 column header to select the column in which to create the ranking.
  - o In the formula bar, enter **=rank(c7)**

4. Create the county ranking for median housing unit value. The lower the median value, the higher the rank number. If for some reason you were to want to rank the counties higher based on HIGH housing values, then you would need to create an inverted housing value column like we did for income.
  - o Click the C10 column header to select the column in which to create the ranking.
  - o In the formula bar, enter **=rank(c9)**

### Hide Columns

Since several of our columns were used to create our desired columns, we can hide them. In our example, we do not need to see the columns used to calculate the commute time or the inverted income column used for ranking income.

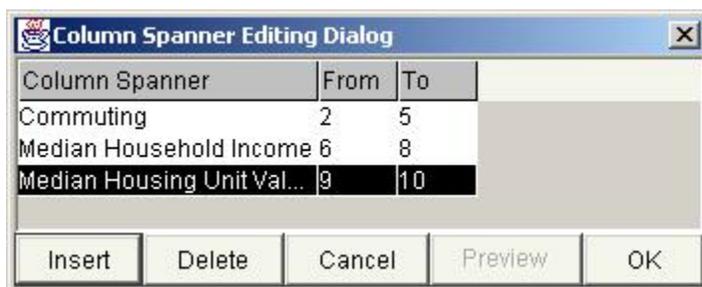
1. Click the C2 column header to select the column you wish to hide, then select the "Edit" menu, then select "Hide".
2. Repeat for columns 3 and 7. Currently you can only highlight one column at a time.
3. If you ever want to see the hidden columns, select the "View" menu, then select "Hidden Rows/Columns".
4. If you ever want to change a hidden column back to not being hidden (unhide), first view the hidden columns as in the previous step, then highlight the column, then select the "Edit" menu and select "Unhide". The "Unhide" selection allows you to unhide only the selected column, or unhide all hidden columns.

### Create Custom Column Spanners

You can create your own custom column spanners to group columns together. In our example, we can create spanners to group the three different measures and their rankings together.

1. Select the "Edit" menu, then select "Column Spanners". This opens the "Column Spanner Editing Dialog" box. There are three fields for you to supply - 1) the text for the spanner, 2) the beginning column, and 3) the ending column. See the image.

Screenshot of column spanner creation:



2. Press the "Insert" button at the bottom to enter the first spanner.
3. Click in the empty "Column Spanner" field and type in your first spanner - **Commuting** - then press the "Enter" key.
4. Click in the empty "From" column and enter the column number the spanner should start with - **2** - then press the "Enter" key.
5. Click in the empty "To" column and enter the column number where the spanner should end - **5** - then press the "Enter" key.
6. Press the "Insert" button at the bottom to enter another spanner.

7. Enter the information as show in the image above.
8. Once the three spanners have been entered, press the "OK" button.

### Create the Combined Ranking

Now we want to sum all three rankings together to find which counties rank the best based on the three criteria. The summed rankings will show us which counties have the shortest commute times, the highest median income, and the lowest median housing values.

1. Click in the C11 column header to select the column in which to create the summed ranking.
2. In the formula bar, enter **=sum(c5,c8,c10)** == this sums the rankings together.
3. A small dialog box will appear showing the computations and the last one shows a conflict. See the image.

Screenshot of calculation sequence conflict dialog:

Sequence	Conflict		Number	Definition	Label
0		Col	4	=COMP(C2/...	commute ti...
0		Col	5	=RANK(C4)	rank
0		Col	7	=COMP(0-C6)	
0		Col	8	=RANK(C7)	rank
0		Col	10	=RANK(C9)	rank
0	Conflict	Col	11	=SUM(C5,C...	

The computations must happen in order, or they cannot be completed properly. In this case, all previous ranking functions must be completed prior to our summing the ranks together, therefore we want to change the conflicting computation's sequence.

4. Click in the "Sequence" cell of the "Conflict" row and change the 0 to a 2 and press the "Enter" key. This will make this computation occur last.
5. Press the "OK" button.  
To see which county ranks the highest or lowest, highlight column 11 and use the sort buttons in the toolbar.

### Adding Weights to the Ranking Criteria

We can customize our final ranking by weighting the criteria that are of more importance to us. For example, if the commute time is the most important criterion to us, we can weight it higher than the others. The weights must add up to 1. We will weight commute time as 0.5, and the other two as 0.25 each.

1. Unsort the table by pressing the "Return to original order" button next to the sort buttons on the toolbar.
2. Click in the C12 column header to select the column in which to create the weighted ranking.
3. In the formula bar, enter `=comp((c5*0.5)+(c8*0.25)+(c10*0.25))` == this weights the ranks and sums them together.
4. We notice that we get all 0s in our table. We have a sequence conflict again, but we were not notified this time. We can change the sequence manually.
5. Go to the "Options" menu and select the "Specify Sequence" option. This will bring up the dialog we saw before. Now, we want to change the last two calculations to have a sequence of 2, making sure to press the "Enter" key after each change. See the image.

Screenshot of calculation sequence specification dialog:

Sequence	Conflict	Number	Definition	Label
0		4	=COMP(C2/...	commute ti...
0	Col	5	=RANK(C4)	rank
0	Col	7	=COMP(0-C6)	
0	Col	8	=RANK(C7)	rank
0	Col	10	=RANK(C9)	rank
2	Col	11	=SUM(C5,C...	Ranking
2	Col	12	=COMP((C5...	

Once the sequence is changed, the weighted rankings appear. You can easily change these weighted rankings by changing the weights in the formula for column 12. To edit a formula, simply highlight the column and the formula should appear in the formula bar. Edit the formula and press the "Enter" key and the changes will be applied.

You can sort the table to find the best and worst counties based on your weighted ranks.

**Goal 2** Limit our rankings to counties that meet our population size criteria - between 50,000 and 150,000 people.

If we only want to see the best counties of a certain size, we can include the population variable and then use an IF condition in a formula to remove counties from the top of the list.

1. DO NOT CLOSE your spreadsheet window.
2. Go back to the Step 1 tab and find the P1. Total Population table in Summary File 3.
3. Get the variable, and add the P001001 variable to your data basket.
4. Now go back to the spreadsheet window and you will see the P001001 variable in your variable list.
5. Drag P001001 into column 13. Press the "Go Get Data" button.
6. Click in the C14 column header to select the column in which to create the weighted ranking for counties of a certain size.

7. In the formula bar, enter `=IF(((C13>50000)AND(C13<250001)),C12,3500)` == this condition will use the weighted rank (C12) if the county population is between 50,001 and 250,000 and if the population is smaller or greater than that, it will set the rank to 3500 (an arbitrary number higher than the highest county rank).
8. We notice that we get all 0s and 3500s in our table. We have a sequence conflict again, and again we were not notified. We must change the sequence manually.
9. Go to the "Options" menu and select the "Specify Sequence" option. This will bring up the dialog we saw before. Now, we want to change the last three calculations - two to have a sequence of 2, and the last one to a sequence of 3, making sure to press the "Enter" key after each change. See the image.

Screenshot of calculation sequence specification dialog:

Sequence	Conflict	Number	Definition	Label
0	Col	4	=COMP(C2/...	commute ti...
0	Col	5	=RANK(C4)	rank
0	Col	7	=COMP(0-C6)	
0	Col	8	=RANK(C7)	rank
0	Col	10	=RANK(C9)	rank
2	Col	11	=SUM(C5,C...	Ranking
2	Col	12	=COMP((C5...	Weighted
3	Col	14	=IF(((C13>5...	

Once the sequence is changed, we now can sort the list and see the best counties based on the weighted criteria we used and limiting the counties to the size we specified.