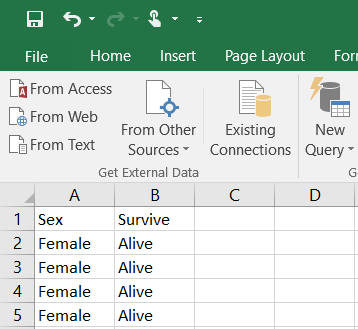
Chapter 2: Displaying and Describing Categorical Data

Part III: Contingency Table

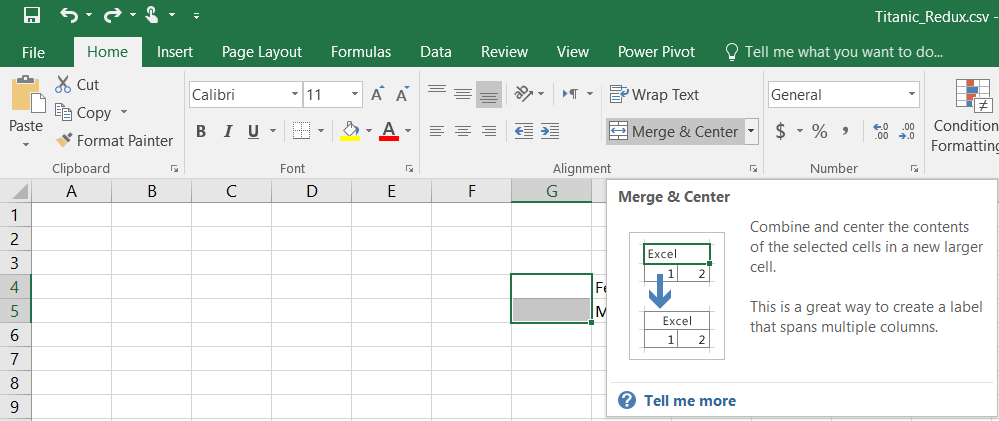
**Titanic Example:**

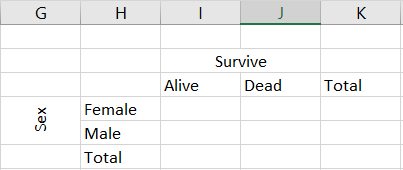
1. Open the data file (Titanic\_Redux.csv). In this file, there are two variables which are sex and survive.



1. Prepare the contingency table; type the categories of one variable as rows and the categories of the second variable as columns. We can add total for the rows and the columns.

To type the variable’s label, we select two cells (or more), then select **Home** tab **Merge and Center**.

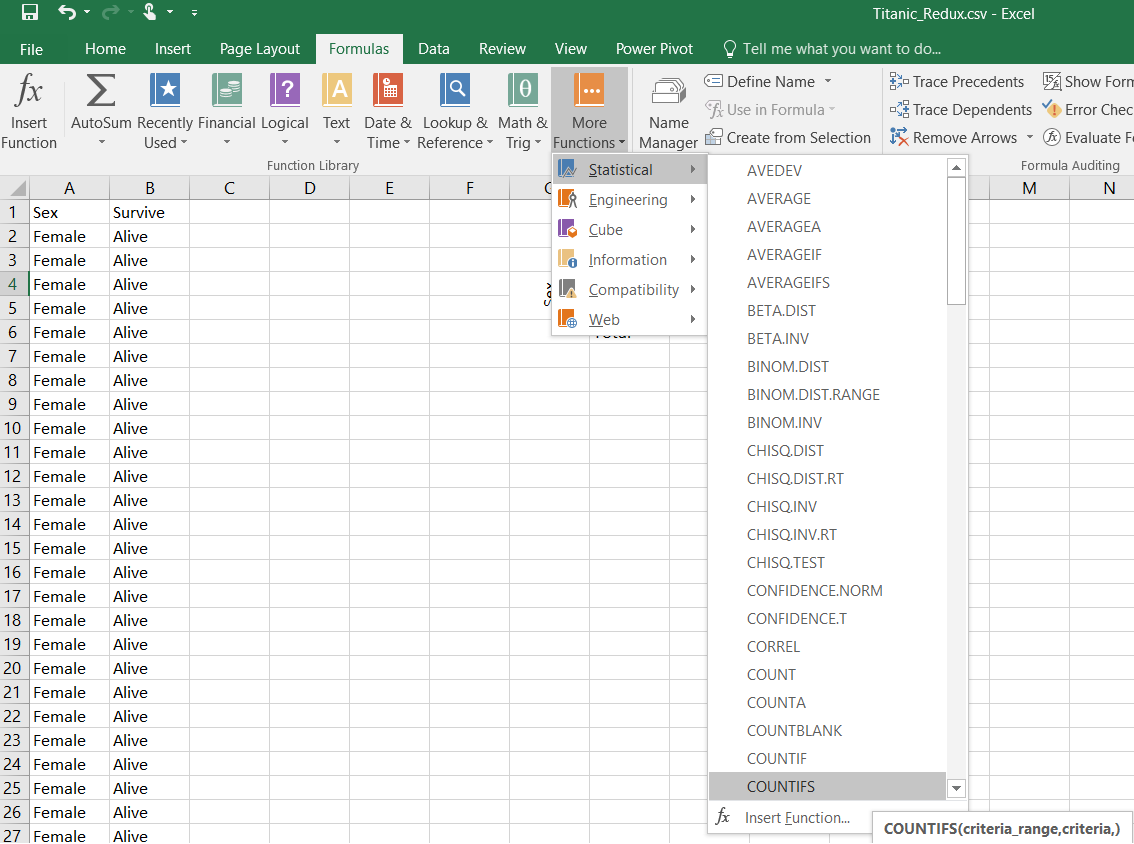


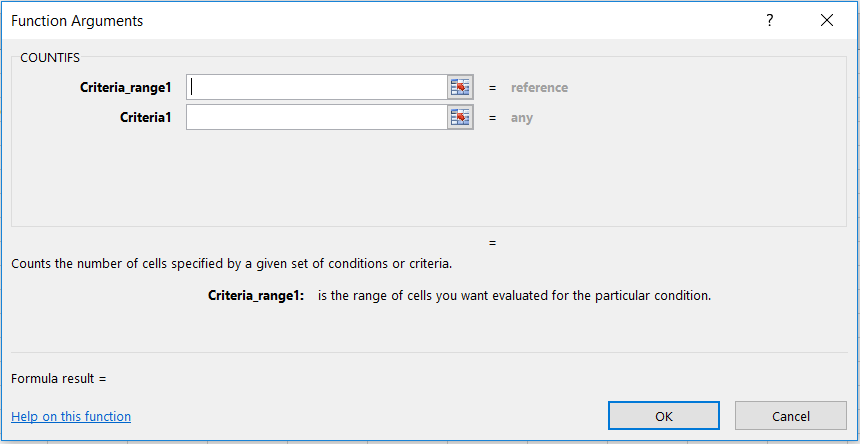


1. Select a cell (I4 – female & alive), and use the function **COUNTIFS**; we can write as “=COUNTIFS”.



or select it from **Formulas** tab **More Functions** **Statistical** **COUNTIFS**.





**Range1** : The first range to evaluate (the range of the first variable - sex).

**Criteria1** : The criteria to use on range1 (female).

**Range2** : The second range to evaluate (the range of the second variable - survive).

**Criteria2** : The criteria to use on range2 (alive).



 : The range of the first variable (sex).

 : The first category of the first variable (female).

 : The range of the second variable (survive).

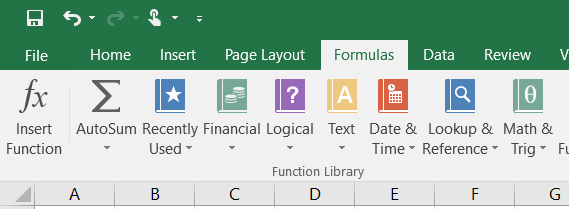
 : The first category of the second variable (alive).

**Note:** use the dollar sign ($) to lock the range, and we can copy and paste the formula to the other cells.

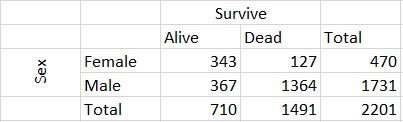
Do so for the other categories. We can just copy and paste the previous cell, and replace A2 (male or female) and B2 (alive or died) as needed.

|  |  |  |
| --- | --- | --- |
|  | Alive | Died |
| Female | A2 and B2 | A2 and B345 |
| Male | A472 and B2 | A472 and B345 |

1. You can add the total to the frequency table by select the total cell (I3) and select **Formulas** tab, click on **AutoSum** (**∑**) and press enter.



1. The contingency table is



1. **Conditional Distributions:** we can add the percentage of row for each cell by divide each cell by the total of its row, the percentage of column for each cell by divide each cell by the total of its column, and the percentage of overall total for each cell by divide each cell by the overall total (the overall total in this example is 2201).