

## Pivot tables and data analysis

Business decisions are often based on analysis of correlations among different variables, such as cost, location, and other attributes of an item offered for exchange. For example, the cost of a house may be influenced by the town it is in and the number of rooms in it. The variables that have an effect on others are called "independent," and those that are affected by others are "dependent."

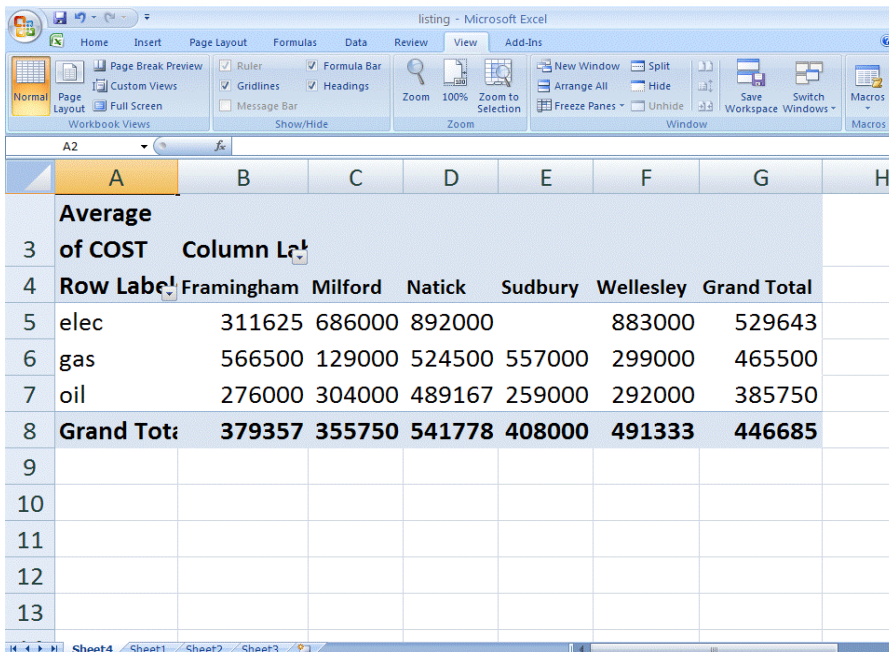
For dependent variables, several choices are possible in summarizing data, including the number of instances, sum, or average. For example, the analyst may wish to know the numbers of houses by town and number of rooms, the average costs of houses by these criteria, or the sums of the costs of all houses by town and number of rooms. (See attached screen shot.)

In Excel 2007, selecting a database table and choosing Insert / Pivot Table allows creation of a pivot table based on the selected data.

In the Field List, choose the three fields by checking their boxes.

Then drag one independent variable field to the "Column Labels" box below the field list, then one independent variable to the Row Labels box, then drag the third to the Sigma Values box. In the example above, the pivot table is displayed as a grid, with costs in the middle, towns on the left, and number of baths across the top.

To display average cost (rather than sum of costs), select the rectangle of cost figures, right click, and choose "Average" in the "Summarize data by" option. The result is the pivot table shown in the attached file.



	Framingham	Milford	Natick	Sudbury	Wellesley	Grand Total
elec	311625	686000	892000		883000	529643
gas	566500	129000	524500	557000	299000	465500
oil	276000	304000	489167	259000	292000	385750
<b>Grand Total:</b>	<b>379357</b>	<b>355750</b>	<b>541778</b>	<b>408000</b>	<b>491333</b>	<b>446685</b>