Tabulation

The **Tabulation** procedure is designed to summarize a single column of attribute data. It tabulates the frequency of occurrence of each unique value within that column. The frequencies are displayed both in tabular form and graphically as a barchart or piechart.

Sample StatFolio: *tabulation.sgp*

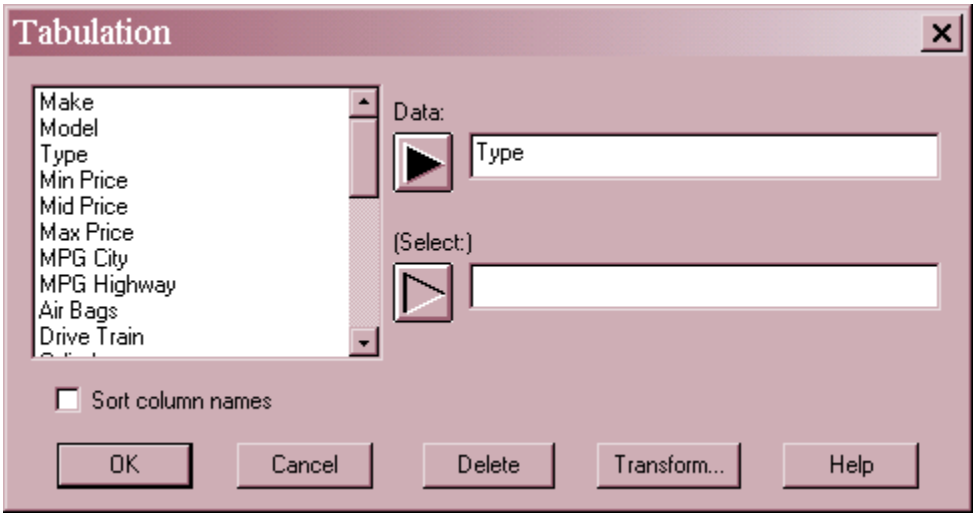
Sample Data:

The file *93cars.sgd* contains information on 26 variables for $n = 93$ makes and models of automobiles, taken from Lock (1993). The table below shows a partial list of 3 columns from that file:

Make	*Model*	*Type*
Acura	Integra	Small
Acura	Legend	Midsize
Audi	90	Compact
Audi	100	Midsize
BMW	535i	Midsize
Buick	Century	Midsize
Buick	LeSabre	Large
Buick	Roadmaster	Large
Buick	Riviera	Midsize
Cadillac	DeVille	Large
Cadillac	Seville	Midsize
Chevrolet	Cavalier	Compact

Data Input

The data input dialog box specifies the column containing the data to be tabulated.



- **Data:** numeric or non-numeric column with the data to be tabulated.

- **Select:** subset selection.

Analysis Summary

The *Analysis Summary* shows the number of rows with non-missing data and the number of unique values in the column.

Tabulation - Type
Data variable: Type

Number of observations: 93
Number of unique values: 6

Frequency Table

This pane presents the results in tabular form.

Frequency Table for Type

			Relative	*Cumulative*	*Cum. Rel.*
Class	*Value*	*Frequency*	*Frequency*	*Frequency*	*Frequency*
1	Compact	16	0.1720	16	0.1720
2	Large	11	0.1183	27	0.2903
3	Midsize	22	0.2366	49	0.5269
4	Small	21	0.2258	70	0.7527
5	Sporty	14	0.1505	84	0.9032
6	Van	9	0.0968	93	1.0000

This table displays the following information for each unique value or "class":

- **Value** - the value of the data variable.
- **Frequency** - the frequency of occurrence f_j of each value in the data column.
- **Relative Frequency** - the proportion of times each value occurred, given by f_j/n.
- **Cumulative Frequency** - the number of observations lying in the current or previous classes:

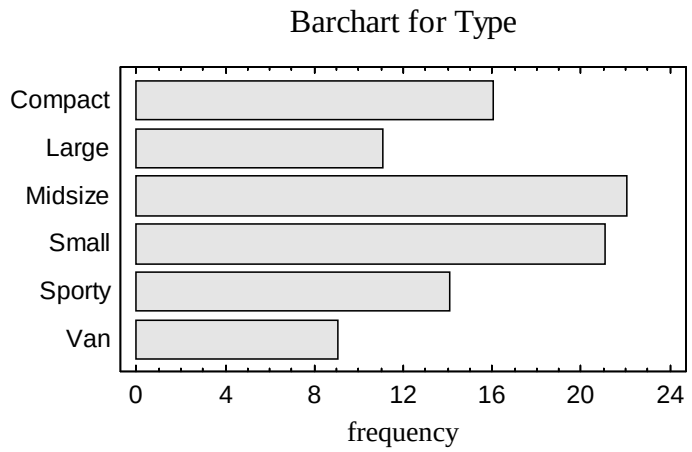
$$\sum_{i=1}^{j} f_i \qquad (1)$$

- **Cumulative Relative Frequency** - the proportion of observations lying in the current or previous classes:

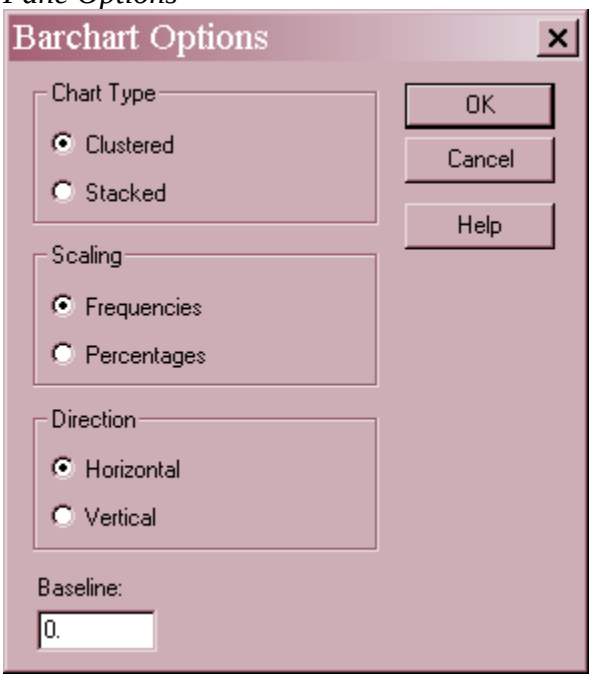
$$\frac{\sum_{i=1}^{j} f_i}{n} \qquad (2)$$

Barchart

The *Barchart* plots the frequency associated with each unique value as a vertical or horizontal bar.



Pane Options



- **Chart Type:** The bars may be clustered side by side as shown in the example or stacked one upon the other.

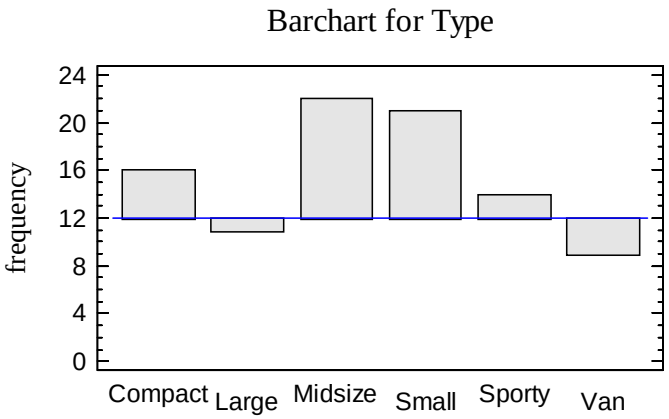
- **Scaling**: whether the axis scale shows the frequencies f_j or the percentages given by

$$p_j = 100\frac{f_j}{n}\% \qquad (3)$$

- **Direction**: whether the bars extend horizontally or vertically.

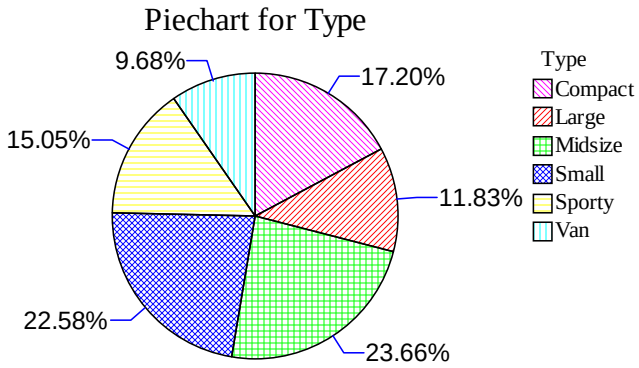
- **Baseline**: the value from which the bars extend.

Example: Vertical Barchart with Non-Zero Baseline
The following chart plots vertical bars extending from a baseline of 12.

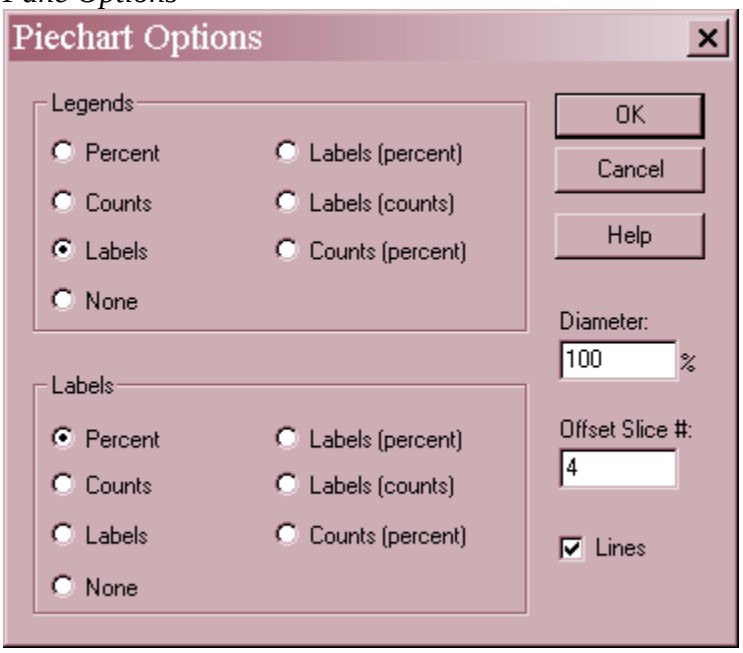


Piechart

The *Piechart* plots each unique value using slices of a pie to represent their relative magnitude.

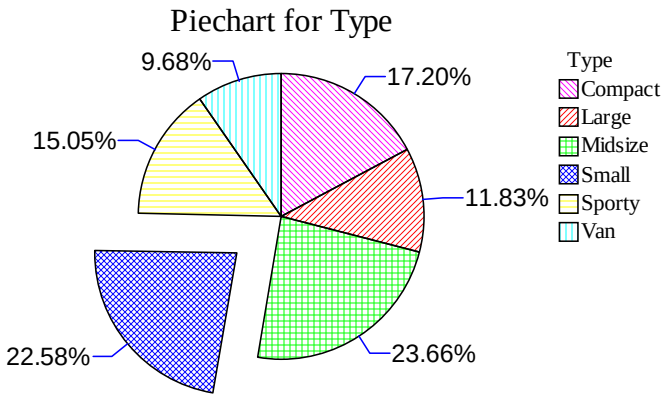


Pane Options



- **Legends:** the desired information to be placed in the legend block to the right of the pie.
- **Labels**: the desired information to be placed next to each slice of the pie.
- **Diameter**: the size of the pie relative to the size of the plotting area. Reducing this value makes the pie smaller.
- **Offset #**: number of a slice between 1 and *k* to offset from the rest of the pie.
- **Lines**: whether to extend lines from each slice to its label.

Example: Piechart with Offset Slice



Save Results

The following results can be saved to the datasheet:

1. *Class Frequencies* – the frequencies of occurrence of each unique value.

2. *Class Labels* – the labels associated with each unique value.

