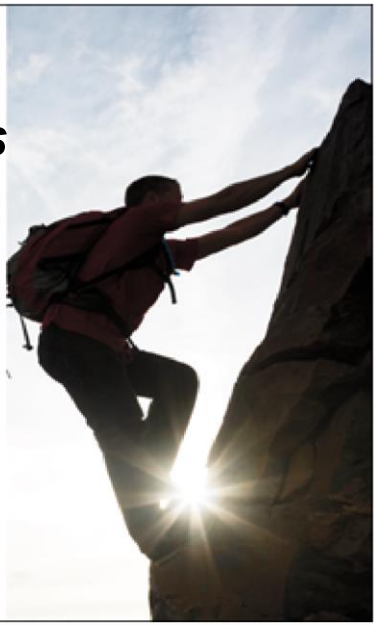
Three Dimensional Waterfall Plots

statgraphics® centurion



Revised: 8/11/2019

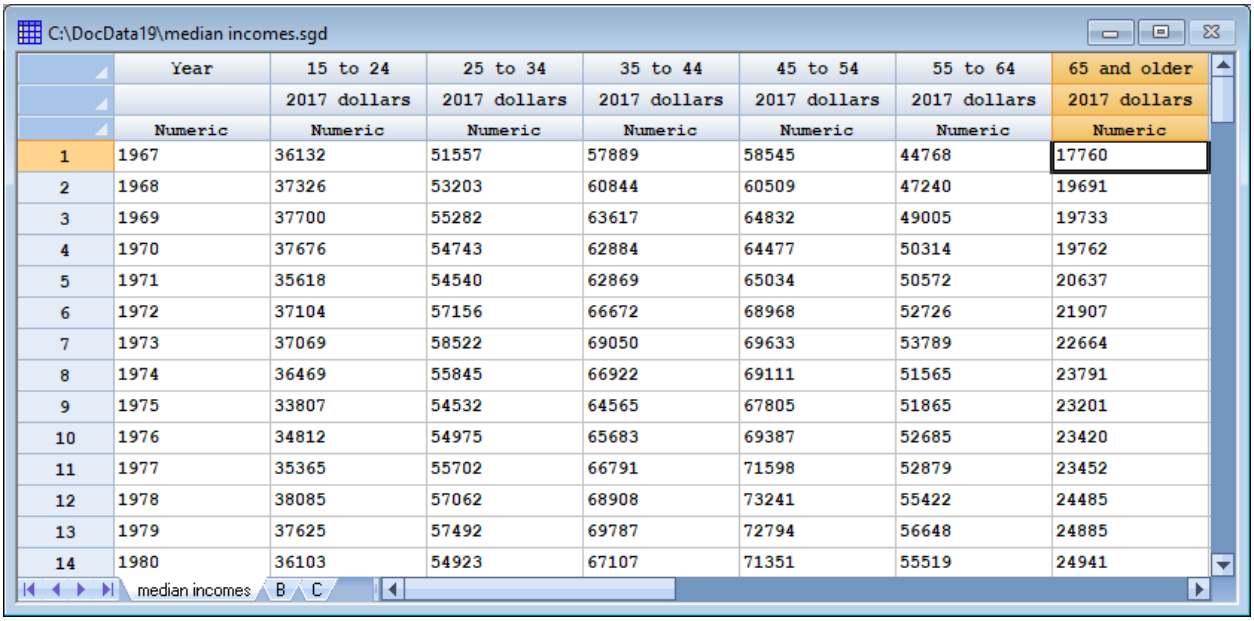
Summary

Three Dimensional Waterfall Plots are widely used to display multiple columns of data versus a common variable. One commonly encountered example is a Cumulative Spectral Decay plot, in which spectra are plotted at multiple times to illustrate changes in amplitude as a function of both frequency and time. In general, the plots are used to show changes in a quantitative variable versus both time and some other factor.

Sample StatFolio: *Waterfall3.sgp*

Sample Data

The U.S. Census Bureau collects data describing many aspects of the U.S. population and economy. One interesting data set, contained in the data file named *median incomes.sgd*, tabulates the median income of U.S. households between 1967 and 2017 based on the age of the primary householder:



C:\DocData19\median incomes.sgd

	Year	15 to 24	25 to 34	35 to 44	45 to 54	55 to 64	65 and older
		2017 dollars	2017 dollars	2017 dollars	2017 dollars	2017 dollars	2017 dollars
	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric	Numeric
1	1967	36132	51557	57889	58545	44768	17760
2	1968	37326	53203	60844	60509	47240	19691
3	1969	37700	55282	63617	64832	49005	19733
4	1970	37676	54743	62884	64477	50314	19762
5	1971	35618	54540	62869	65034	50572	20637
6	1972	37104	57156	66672	68968	52726	21907
7	1973	37069	58522	69050	69633	53789	22664
8	1974	36469	55845	66922	69111	51565	23791
9	1975	33807	54532	64565	67805	51865	23201
10	1976	34812	54975	65683	69387	52685	23420
11	1977	35365	55702	66791	71598	52879	23452
12	1978	38085	57062	68908	73241	55422	24485
13	1979	37625	57492	69787	72794	56648	24885
14	1980	36103	54923	67107	71351	55519	24941

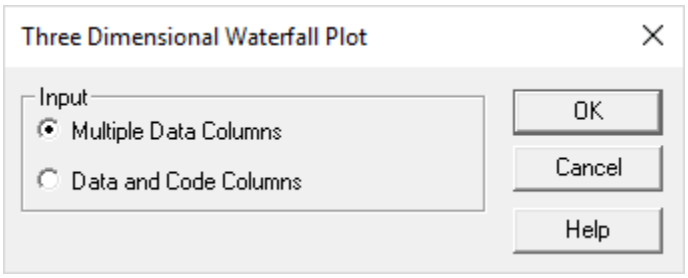
median incomes | B | C

Households are divided into 6 age categories. All data are given in 2017 dollars.

Data Input

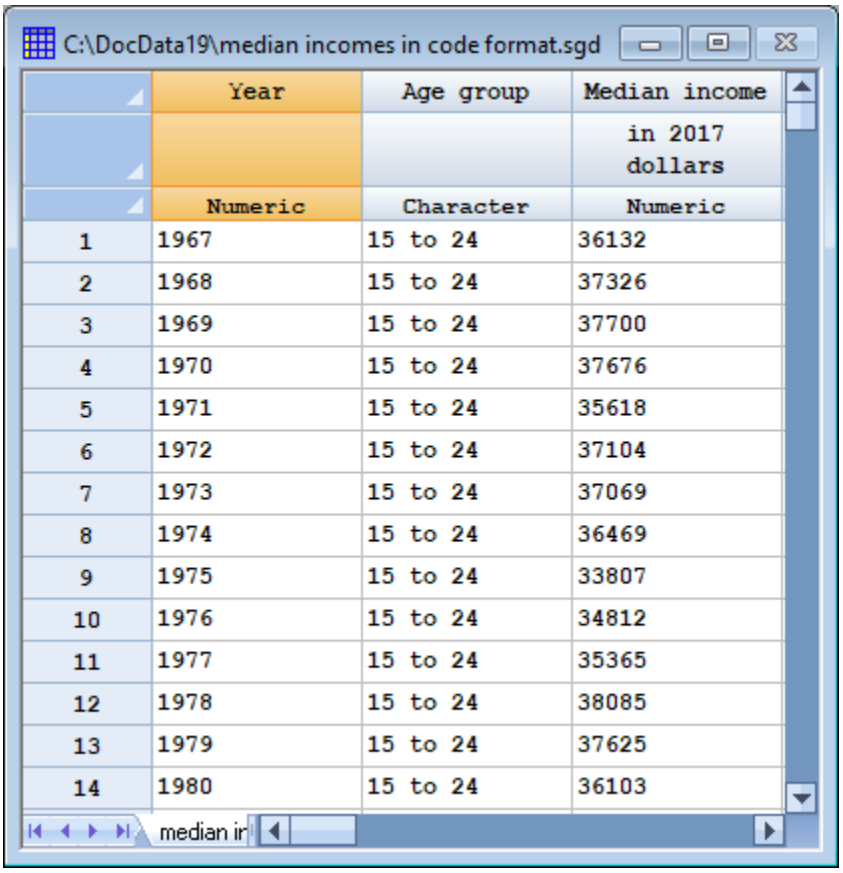
Data Input Dialog Box

Data to be displayed by the *Three Dimensional Waterfall Plot* may be arranged in either of 2 ways. When selected from the main menu, the procedure first displays a dialog box which specifies how the data are structured:



1. **Multiple Data Columns:** The data for each plane are placed in a separate column, as in the sample data file shown above.

2. **Data and Code Columns:** All data to be plotted are placed in a single data column. Columns are then constructed to identify the 2 categorical factors. This format allows for more than one variable to be stored in the same data file. The file below has the median incomes arranged in this fashion:



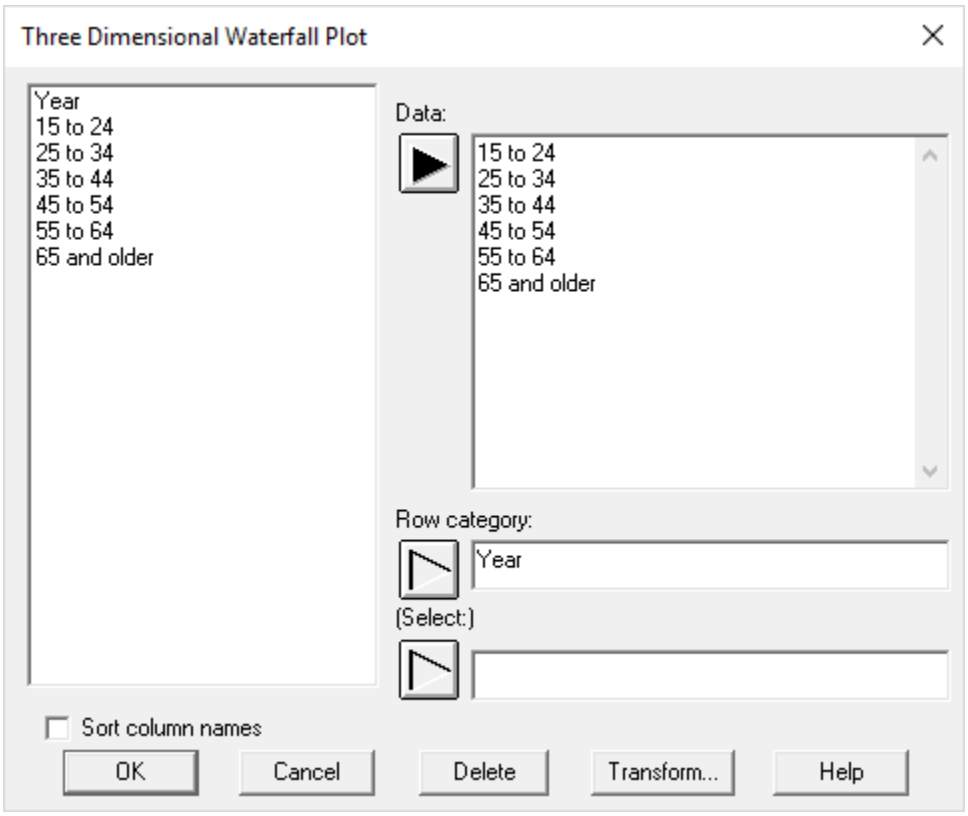
C:\DocData19\median incomes in code format.sgd

	Year	Age group	Median income in 2017 dollars
	Numeric	Character	Numeric
1	1967	15 to 24	36132
2	1968	15 to 24	37326
3	1969	15 to 24	37700
4	1970	15 to 24	37676
5	1971	15 to 24	35618
6	1972	15 to 24	37104
7	1973	15 to 24	37069
8	1974	15 to 24	36469
9	1975	15 to 24	33807
10	1976	15 to 24	34812
11	1977	15 to 24	35365
12	1978	15 to 24	38085
13	1979	15 to 24	37625
14	1980	15 to 24	36103

median in

Multiple Data Columns

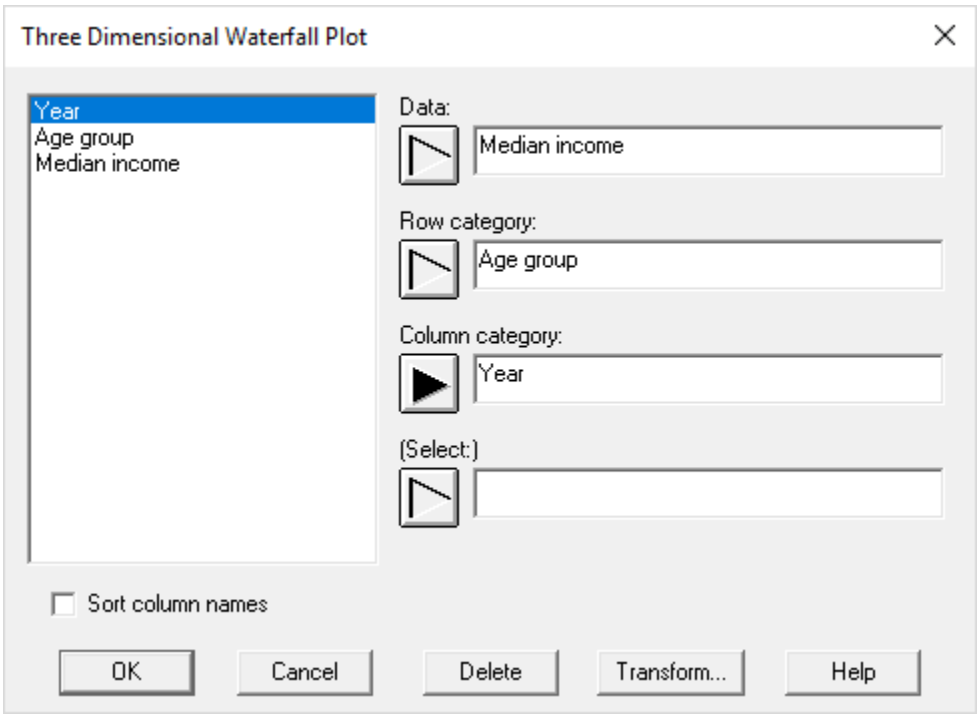
After specifying the structure of the data, a second data input dialog box requests the names of the columns containing the data values to be analyzed. For structure #1, the dialog box requests the names of multiple data columns:



- **Data:** names of 2 or more numeric columns containing the observations to be plotted.
- **Row category:** name of the column used to define the location of the data along the X-axis. The variable may be either numeric or non-numeric.
- **Select:** optional subset selection.

Data and Code Columns

For structure #2, the data input dialog box has the following form:

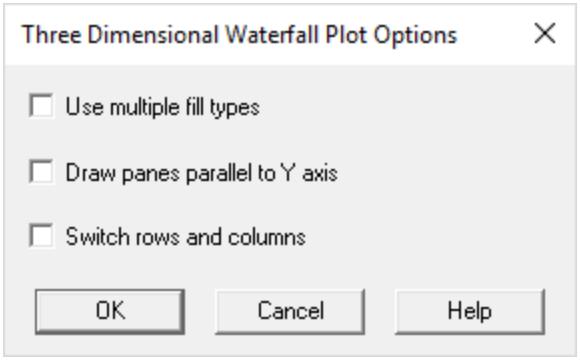


- **Data:** name of a single numeric column containing the observations to be plotted.
- **Row category:** name of the data column used to define the rows or planes of the waterfall plot. The variable may be either numeric or non-numeric.
- **Column category:** name of the data column used to define the X-axis locations of the waterfall plot. The variable may be either numeric or non-numeric.
- **Select:** optional subset selection.

Each row of the data file represents one combination of the row and column categories.

Analysis Options

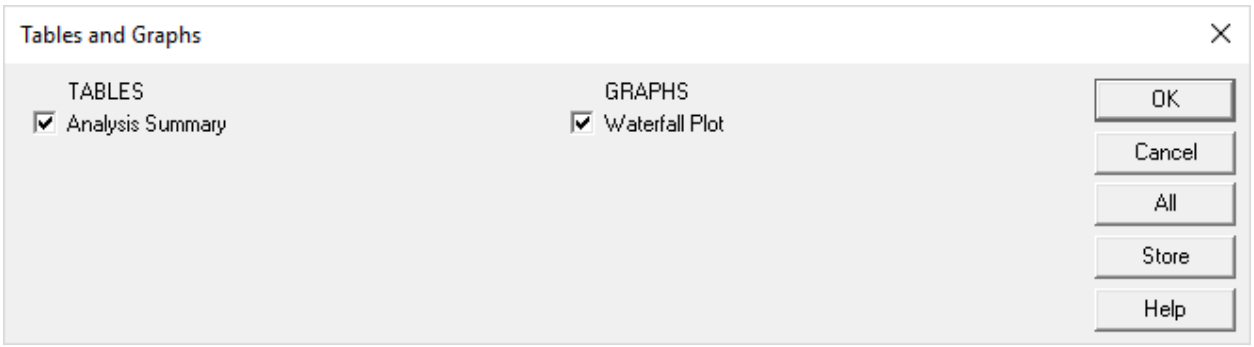
The *Analysis Options* dialog box sets various options for the plot:



- **Use multiple fill types**: if selected, each plane will be plotted using a different fill color and type.
- **Draw panes parallel to Y axis:** if selected, the panes will be drawn parallel to the Y-axis. Otherwise, they will be drawn parallel to the X-axis.
- **Switch rows and columns**: if selected, the variables defining the X-axis and Y-axis locations will be reversed.

Tables and Graphs

The following tables and graphs may be created:



Analysis Summary

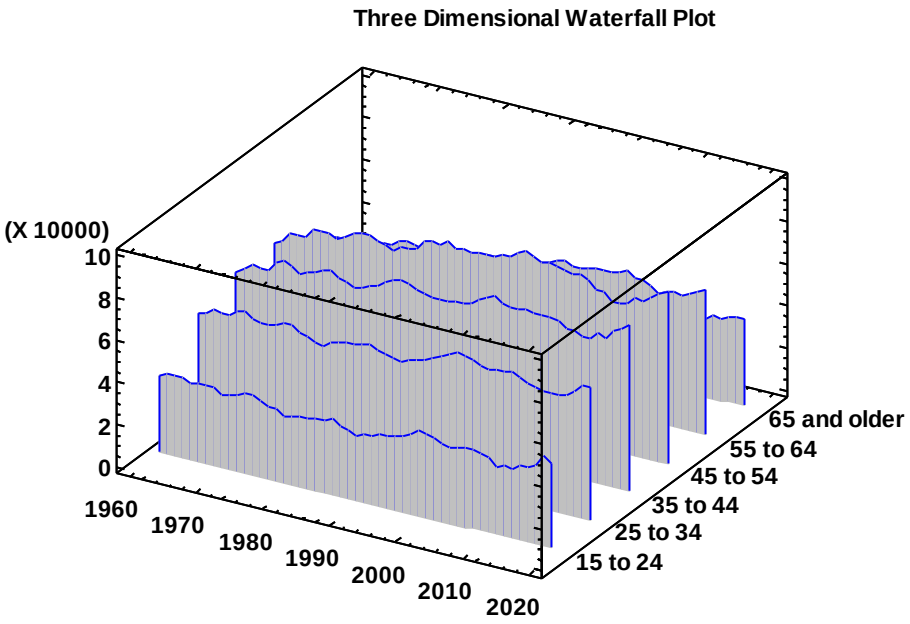
The *Analysis Summary* displays the input values. A portion of the output is shown below:

Three Dimensional Waterfall Plot
Row variable: Year

Year	15 to 24	25 to 34	35 to 44	45 to 54	55 to 64	65 and older
1967	36132	51557	57889	58545	44768	17760
1968	37326	53203	60844	60509	47240	19691
1969	37700	55282	63617	64832	49005	19733
1970	37676	54743	62884	64477	50314	19762
1971	35618	54540	62869	65034	50572	20637
1972	37104	57156	66672	68968	52726	21907
1973	37069	58522	69050	69633	53789	22664
1974	36469	55845	66922	69111	51565	23791
1975	33807	54532	64565	67805	51865	23201
1976	34812	54975	65683	69387	52685	23420
1977	35365	55702	66791	71598	52879	23452
1978	38085	57062	68908	73241	55422	24485
1979	37625	57492	69787	72794	56648	24885
1980	36103	54923	67107	71351	55519	24941
…	…	…	…	…	…	…

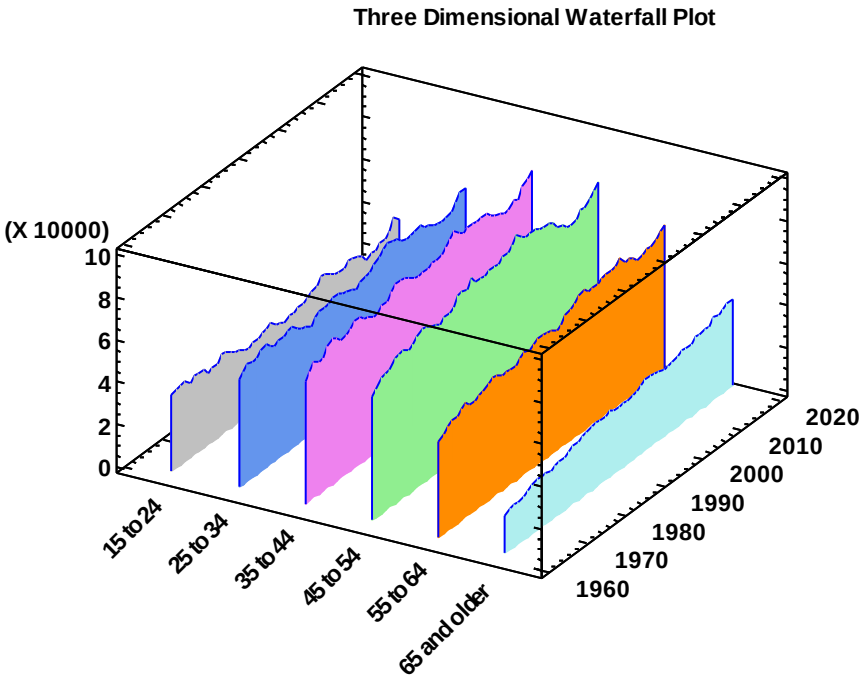
Waterfall Plot

This pane shows the waterfall plot:



In the graph above, planes are drawn connecting the data values for each age category.

A second example is shown below:



In this graph:

1. Multiple fill colors are used.
2. Panes are drawn parallel to the Y axis.
3. The variables defining the X and Y axes have been switched.