Star Glyphs and Sunray Plots

Summary

Star Glyphs and Sunray Plots are multivariate visualization techniques that can be very useful in identifying differences and similarities amongst observed cases when the number of dimensions is too large to use a standard scatterplot.

Sample StatFolio: *starglyphs.sgp*

Sample Data:

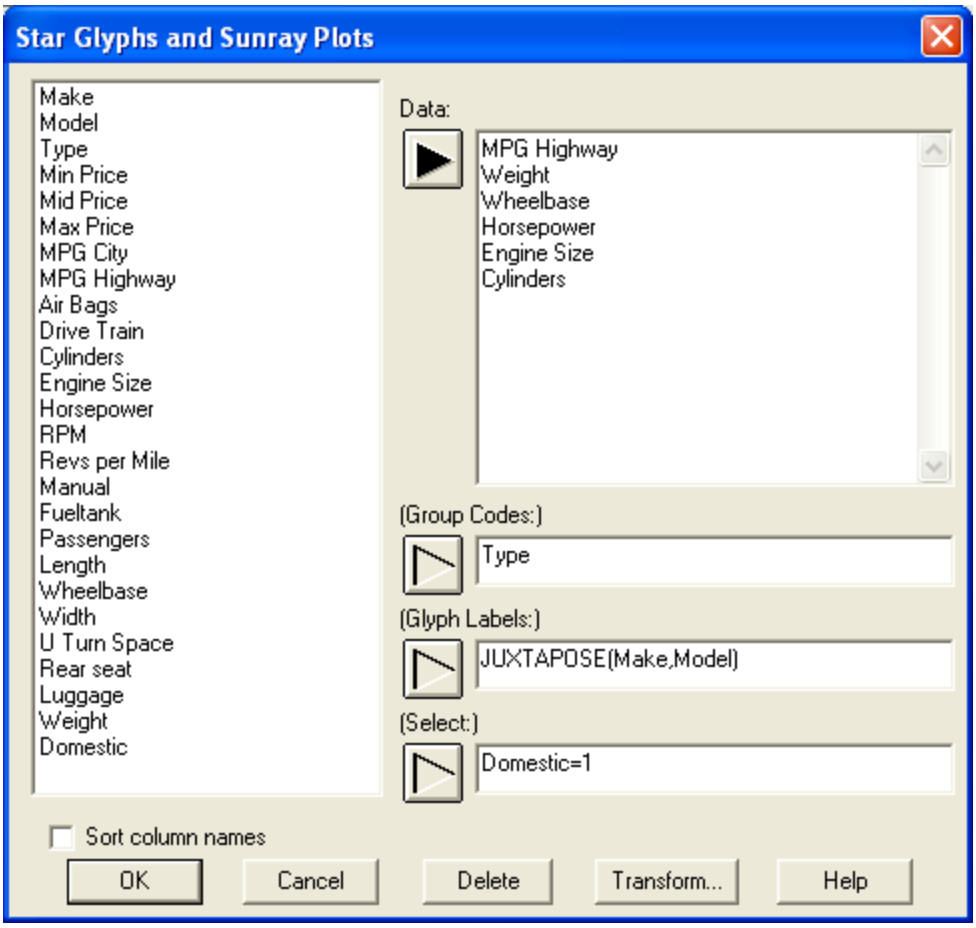
The file *93cars.sgd* contains information on 26 variables for $n = 93$ makes and models of automobiles, taken from Lock (1993). The table below shows a partial list of the data in that file:

Make	*Model*	*MPG Highway*	*Weight*	*Wheelbase*	*Horsepower*	*Engine Size*	*Cylinders*
Acura	Integra	31	2705	102	140	1.8	4
Acura	Legend	25	3560	115	200	3.2	6
Audi	90	26	3375	102	172	2.8	6
Audi	100	26	3405	106	172	2.8	6
BMW	535i	30	3640	109	208	3.5	4
Buick	Century	31	2880	105	110	2.2	4
Buick	LeSabre	28	3470	111	170	3.8	6
Buick	Roadmaster	25	4105	116	180	5.7	6
Buick	Riviera	27	3495	108	170	3.8	6
Cadillac	DeVille	25	3620	114	200	4.9	8
Cadillac	Seville	25	3935	111	295	4.6	8
Chevrolet	Cavalier	36	2490	101	110	2.2	4
Chevrolet	Corsica	34	2785	103	110	2.2	4
Chevrolet	Camaro	28	3240	101	160	3.4	6
Chevrolet	Lumina	29	3195	108	110	2.2	4
Chevrolet	Lumina_APV	23	3715	110	170	3.8	6
Chevrolet	Astro	20	4025	111	165	4.3	6
Chevrolet	Caprice	26	3910	116	170	5.0	8
Chevrolet	Corvette	25	3380	96	300	5.7	8
Chrylser	Concorde	28	3515	113	153	3.3	6

© 2010 by StatPoint Technologies, Inc.

Data Input

The data to be analyzed consist of 2 or more numeric columns and an optional column with group identifiers:



- **Data:** 2 or more numeric columns containing the data to be plotted.
- **Group Codes:** an optional column with levels to be used to identify groups of cases.
- **Glyph Labels:** an optional column with labels corresponding to each row. If not specified, row numbers will be used as labels.
- **Select:** subset selection.

As an example, 6 variables have been selected. The type of vehicle will be used to identify the cases. The JUXTAPOSE operator puts two columns side by side, so that each vehicle may be labeled with both its make and its model. The selection expression “Domestic = 1” specifies that only cars made in the United States should be included.

© 2010 by StatPoint Technologies, Inc.

Analysis Summary

The *Analysis Summary* shows the number of rows with complete data and summary statistics for those rows:

Star Glyphs and Sunray Plots (Domestic=1)
Data variables:
MPG Highway (miles per gallon in highway driving)
Weight (pounds)
Wheelbase (inches)
Horsepower (maximum)
Engine Size (liters)
Cylinders
Selection variable: Domestic=1

Number of complete cases: 48

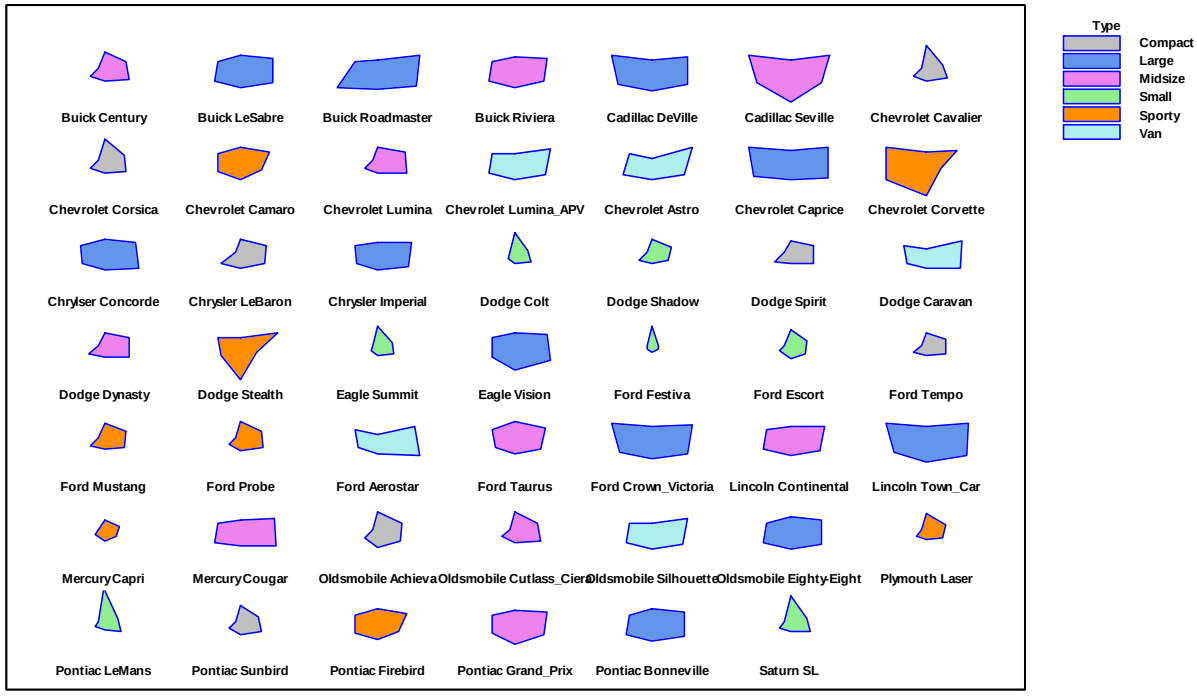
	Sample mean	*Standard deviation*	*Minimum*	*Maximum*
MPG Highway	28.1458	4.15134	20.0	41.0
Weight	3195.31	565.228	1845.0	4105.0
Wheelbase	105.729	6.79092	90.0	119.0
Horsepower	147.521	54.4547	63.0	300.0
Engine Size	3.06667	1.13538	1.3	5.7
Cylinders	5.33333	1.38891	4.0	8.0

There are 48 rows which meet the selection criterion and have data for all of the variables.

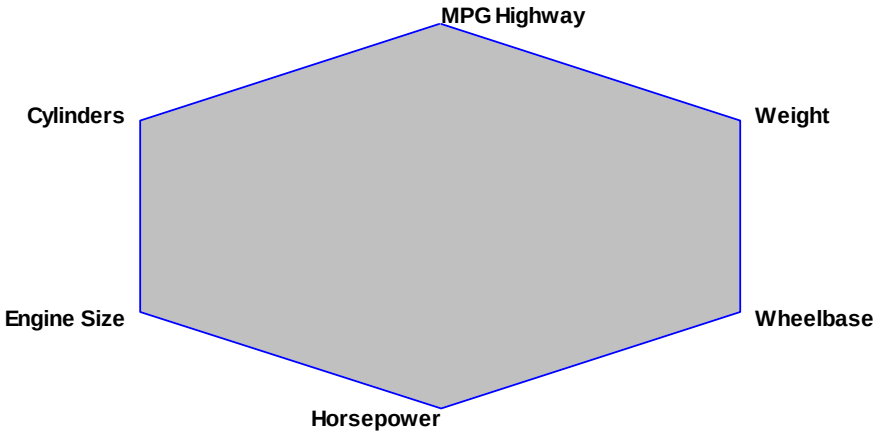
© 2010 by StatPoint Technologies, Inc.

Star Glyphs and Sunray Plots

The plot created displays a glyph for each row:



A glyph is a geometric object that represents the values of each quantitative variable. The default format is a polygon of the general form shown below (select *Key Glyph* from the list of tables and graphs to create this plot):

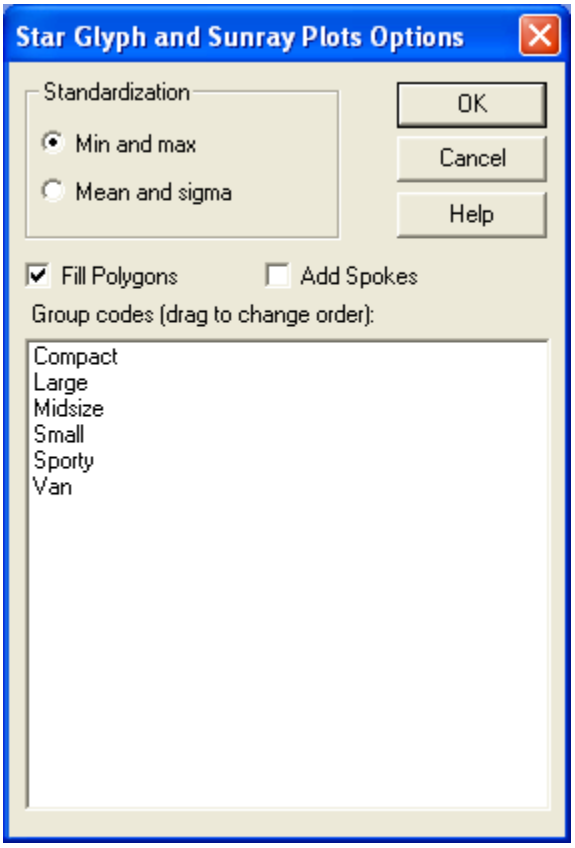


The size of the polygon in each direction is scaled according to the value of each variable for the selected vehicle. Vehicles with similar characteristics will have a similar size and shape.

© 2010 by StatPoint Technologies, Inc.

Analysis Options

The *Analysis Options* dialog box allows you to change various features of the plot:



- **Standardization:** the manner in which to scale the dimensions of the polygon. *Min and max* scales each variable by subtracting its minimum value amongst all of the cases and dividing by the range. *Mean and sigma* subtracts the sample mean and divides by the sample standard deviation.

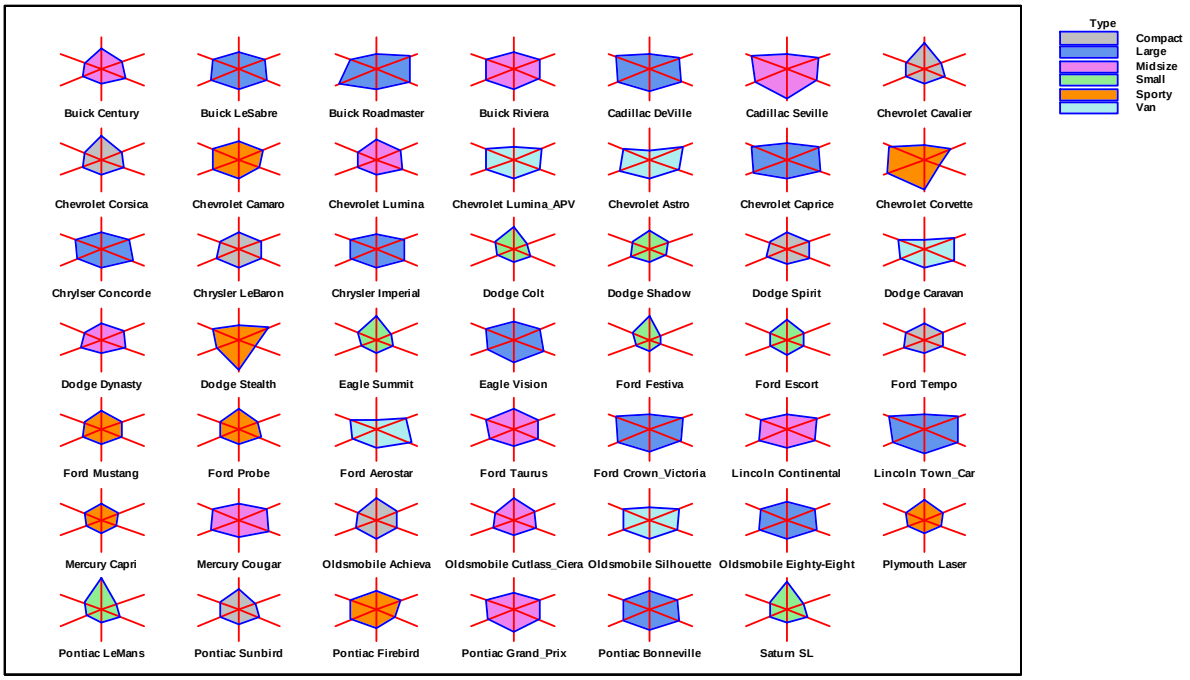
- **Fill Polygons:** fills each polygon with color rather than drawing only the outline.

- **Add Spokes:** add spokes along each dimension.

- **Group Codes:** the order of the group codes in the legend block. You may drag level codes to change their order.

When spokes are added to the plot as shown below, the glyphs are often called *sunray plots*:

© 2010 by StatPoint Technologies, Inc.



The spokes range between the minimum and maximum values of each variable if *Min and Max* scaling is chosen or between the mean plus and minus 3 sigma if *Mean and Sigma* scaling is chosen.

Note that, except for the *Sporty* vehicles, most vehicles of the same type have similar glyphs

© 2010 by StatPoint Technologies, Inc.