

Contour Plots for Time Series Analysis

David J. Corliss, University of Toledo / Department of Physics and Astronomy, Toledo, OH

ABSTRACT

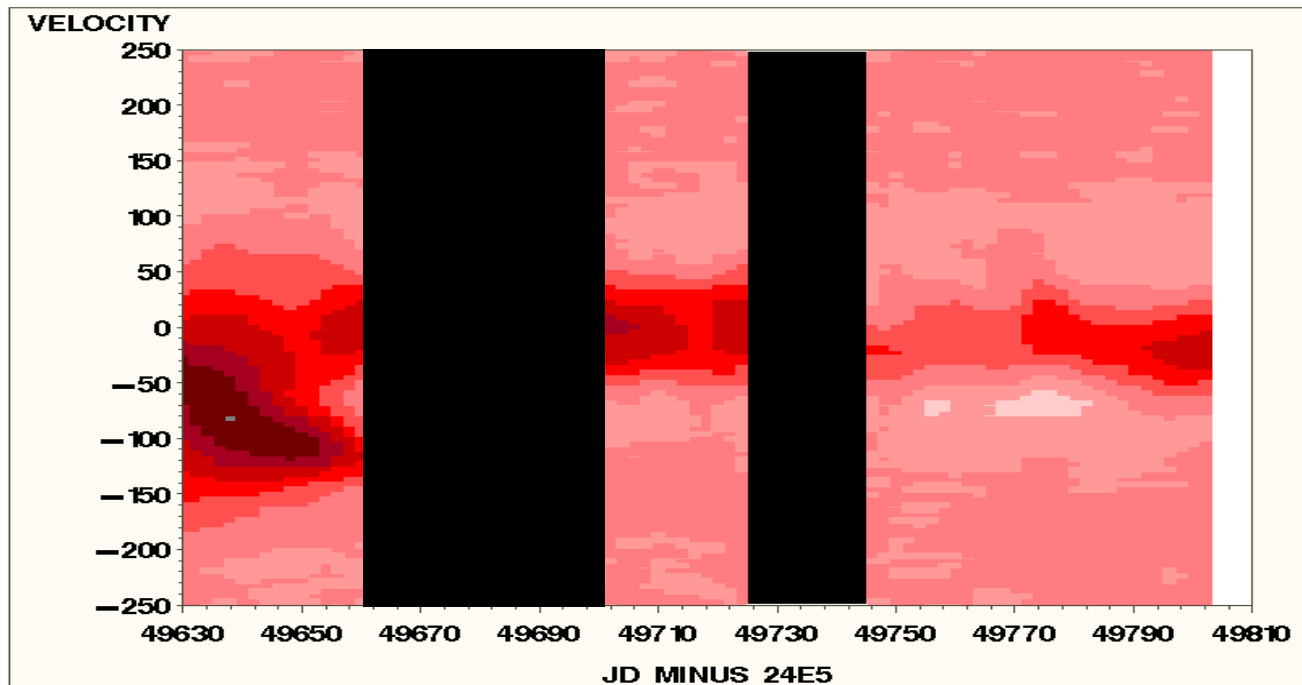
This new SAS tool is two-dimensional color chart for visualizing changes in a population or system over time. Data for one point in time appear as a thin horizontal band of color. Bands for successive periods are stacked up to make a two-dimensional, with the vertical direction showing changes over time. As a system evolves over time, different kinds of events have different characteristic patterns. Creation of Time Contour Plots is explained step by step. Examples are given in astrostatistics, biostatistics, econometrics and demographics.

Keywords: Data Visualization, PROC CONTOUR, Time Series Analysis

INTRODUCTION

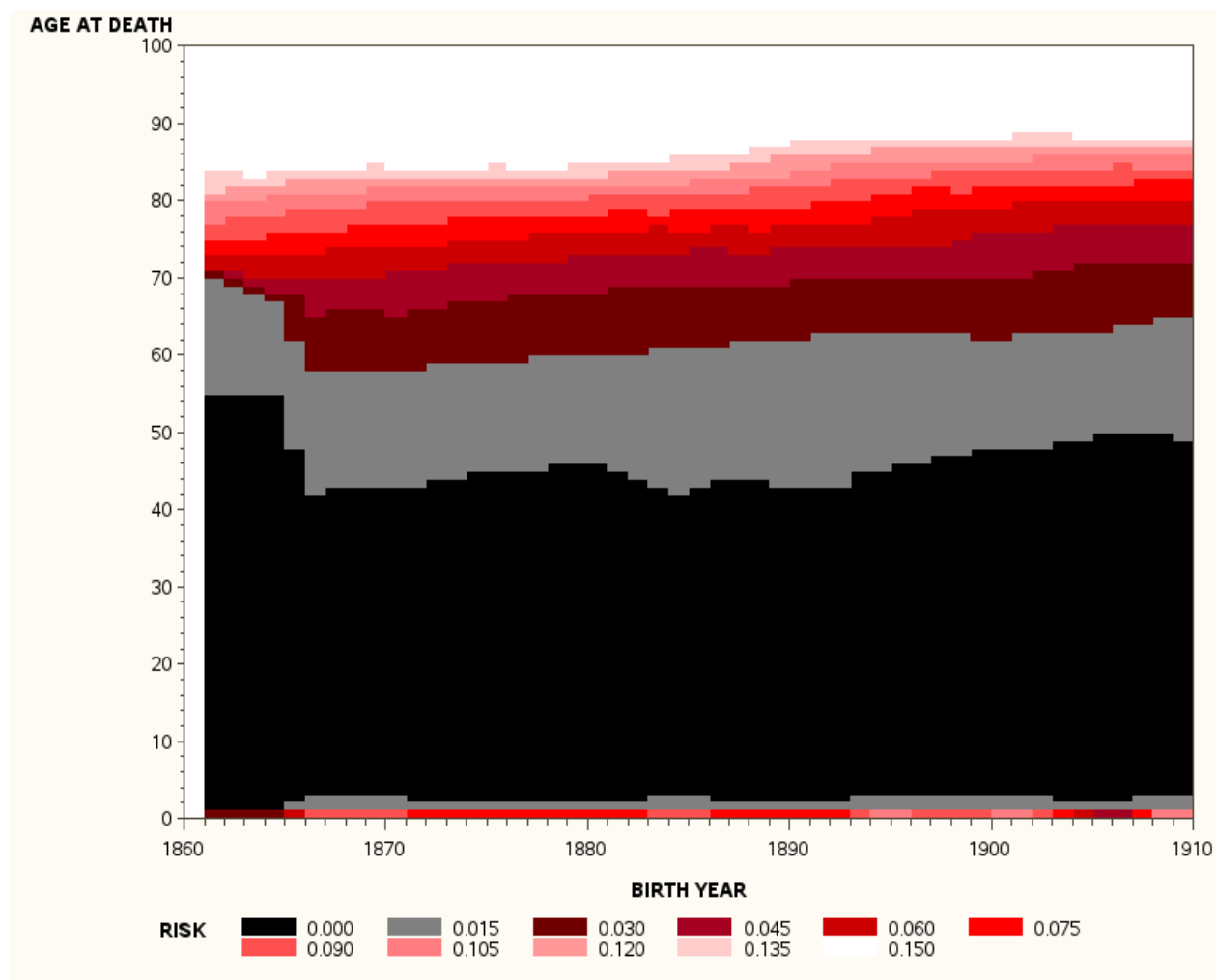
Time Contour Plots, as described in this paper, were first developed in the field of statistical astrophysics (Kaufer et al., 1996). Known as Dynamical Spectra, they are used to graphically display physical characteristics of stars that vary with across different wavelengths of light and by time. In these plots, it is conventional to place the wavelength on the horizontal axis with time of the vertical axis with the values of the physical characteristic to be plotted represented by a color value. Among astrophysicists, the preferred platform for creating these plots is IDL, which is unfamiliar to many researchers. Also, the presentation usual format may seem counter-intuitive to persons more used to seeing time represented on the horizontal axis. In this paper, Time Contour Plots appear in SAS or the first time. A general presentation is given, applicable to many to many sciences, with an option for placing time on the horizontal axis.

AN EXAMPLE OF TIME CONTOUR PLOT

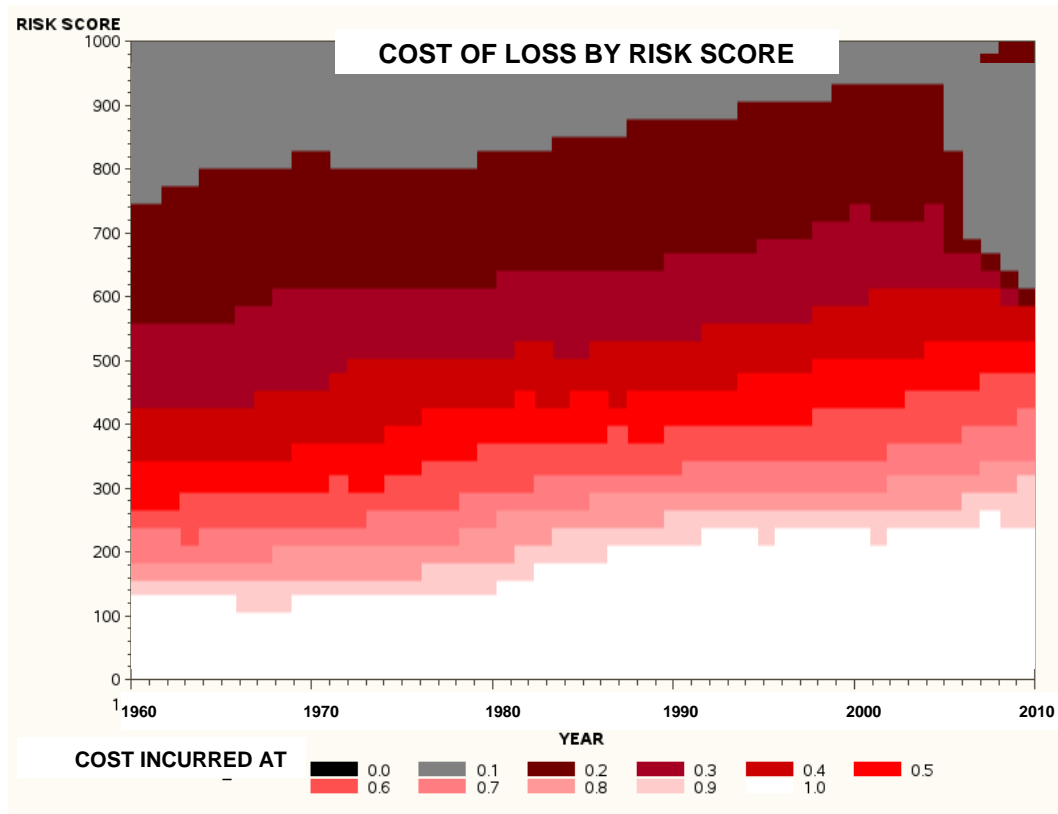


In this example, the date of an observation is shown on the horizontal axis. The velocity of a stream of gas flowing outward from a star is plotted on the vertical axis, with the amount of gas displayed by a color value. (Properly speaking, the color is the fraction of light absorbed by the gas at each wavelength – this is proportional to the mass volume.) A darker color indicates a larger amount of gas flowing. So, the dark area on the left side of the plot indicates a large amount of gas erupting from the surface of the star between dates 2449630 and 2449650 and moving at a very high velocity – more than 100 km/s. To astronomers, this pattern is diagnostic of specific type of stellar eruption known as a High Velocity Absorption event.

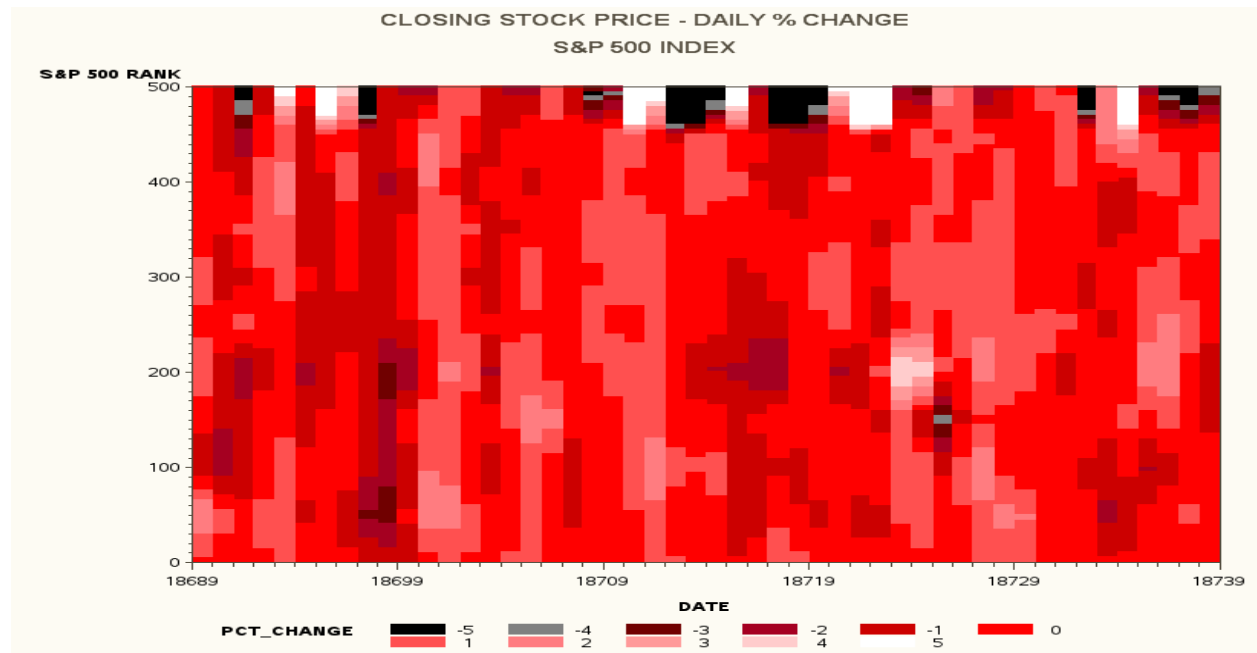
The same plots may be employed in others areas of research to identify events in time by their signature patterns. The following plot presents an example in biometrics, where US census bureau data is used to plot age at time of death, with the color value giving the mortality rate for a given age for each year (horizontal scale) from 1860 through 1910. The high mortality of the aged is plainly visible but careful examination will reveal more information. The infant mortality rate is given by the thin line of color at the bottom of the chart. In general, the color gradually gets lighter over time, reflecting a reduction in the infant mortality rate. However, the dark red points from 1905–1907 indicate an epidemic increasing the mortality of children at the time.



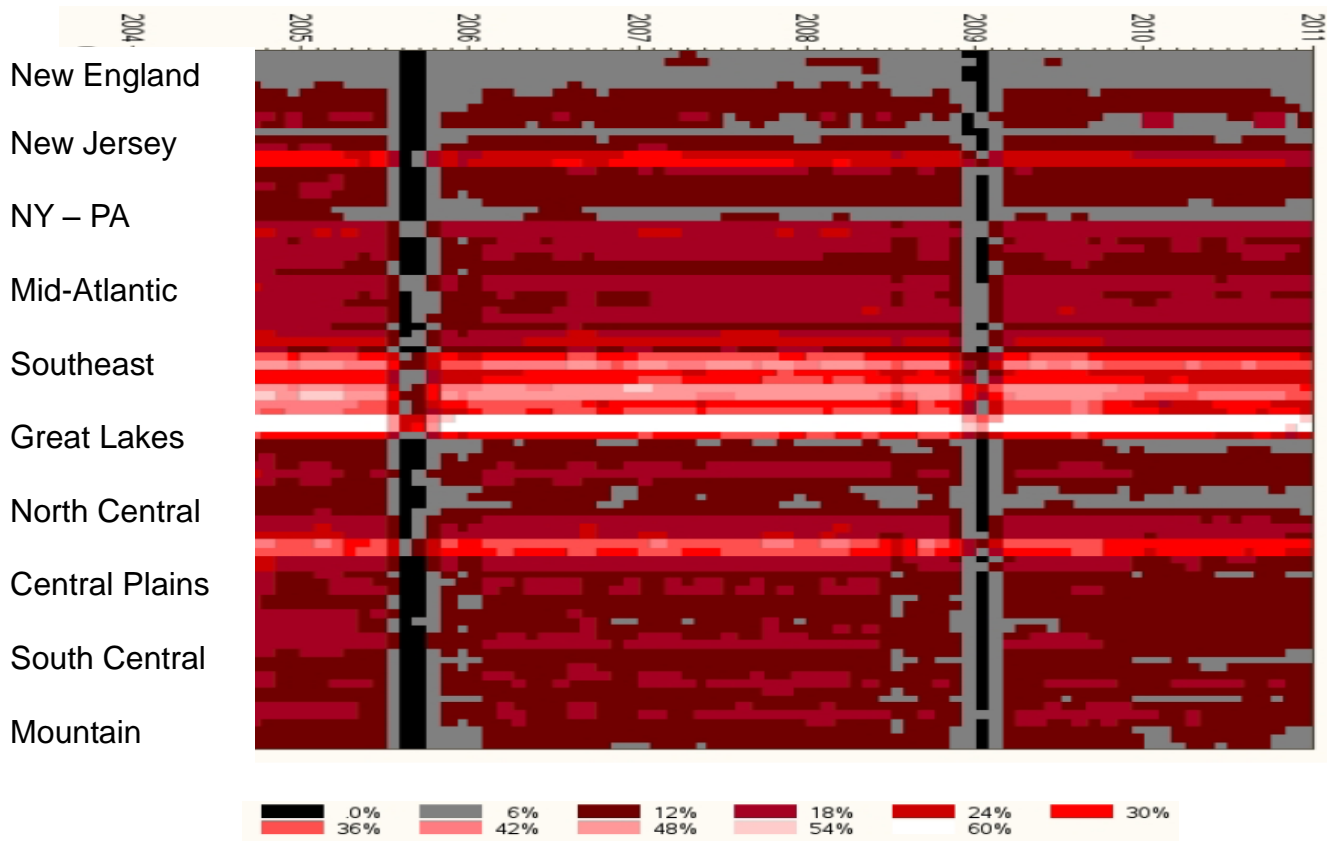
In this example from risk analysis, the amount of loss over time due to credit loss is plotted against risk.



In the next example, from econometrics, trends in stock prices are plotted. The sharp color changes at the bottom reflect the greater volatility near the bottom of the S&P 500 index compared to those stocks at the top.

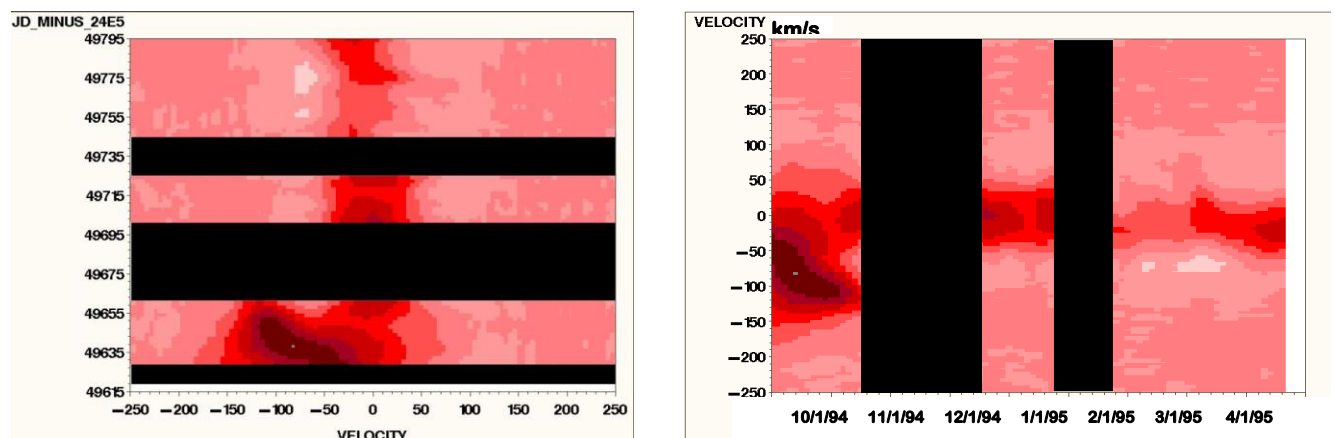


Despite the unconventional appearance, this example is actually a map. The horizontal axis plots the first two digits of the zip code, as recorded by the Census Bureau. This is a dynamical map, showing changes in time of a demographic variable.



PRESENTATION OPTIONS

Time Contour Plots are contour plots produced in SAS® Graph with PROC GCONTOUR. As such, the values of three independent continuous variables are plotted. The user may choose which axis to use for the time value. In the example below, time is vertical on the left with a proportional data scale. On the right, the time is on the horizontal with a natural time scale. The user will need to select which options best serve the situation at hand.



BINNING

Contour plots consist of a mosaic of small rectangles, each of which are assigned a color value. The data used to create the plot, while many-valued, is actually discrete rather than continuous. As a result, it may be necessary to bin continuous variables on the horizontal and vertical axes before creating the plot:

```
/* Binning on the vertical (non-temporal axis) */;
```

```
data work.flux;  
    set hr_1040.flux_h_alpha_cc_flat;  
    wavelength = round(wavelength,0.1);  
run;  
proc sort data=work.flux;  
    by obs_date wavelength;  
run;  
proc univariate data=work.flux noprint;  
    var flux;  
    by obs_date wavelength;  
    output mean=flux out=work.flux;  
run;
```

```
/* Binning on the temporal axis */;
```

```
proc sort data=work.flux;  
    by wavelength period;  
run;  
proc univariate data=work.flux noprint;  
    var flux;  
    by wavelength period;  
    output mean=flux out=work.flux;  
run;  
data work.flux;  
    set work.flux;  
    obs_date = mdy(9,18,1994) + (2 * period) - 1;  
    julian_date = obs_date + 2436935;  
    jd_minus_24e5 = julian_date - 2400000;  
run;
```

CREATION OF A BASIC CONTOUR PLOT

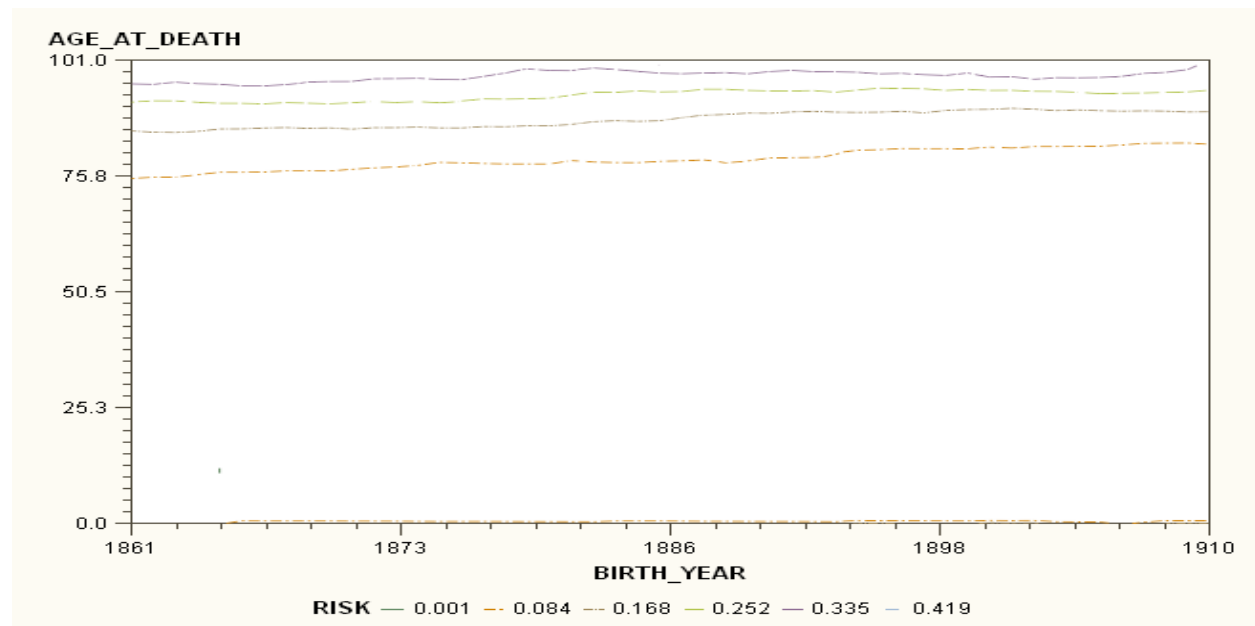
Contour plots are created in SAS Graph using PROC GCONTOUR. Three variables are plotted, with the vertical axis first, the horizontal axis second and the contour value listed last:

```

goptions reset=all border;
proc gcontour data=work.plot_data;
  plot age_at_death * birth_year = risk;
  run;
quit;

```

Here is the result. As this point, the plot looks like an ordinary contour plots, such as are used to represent physical geography, with lines denoting the contour profiles:



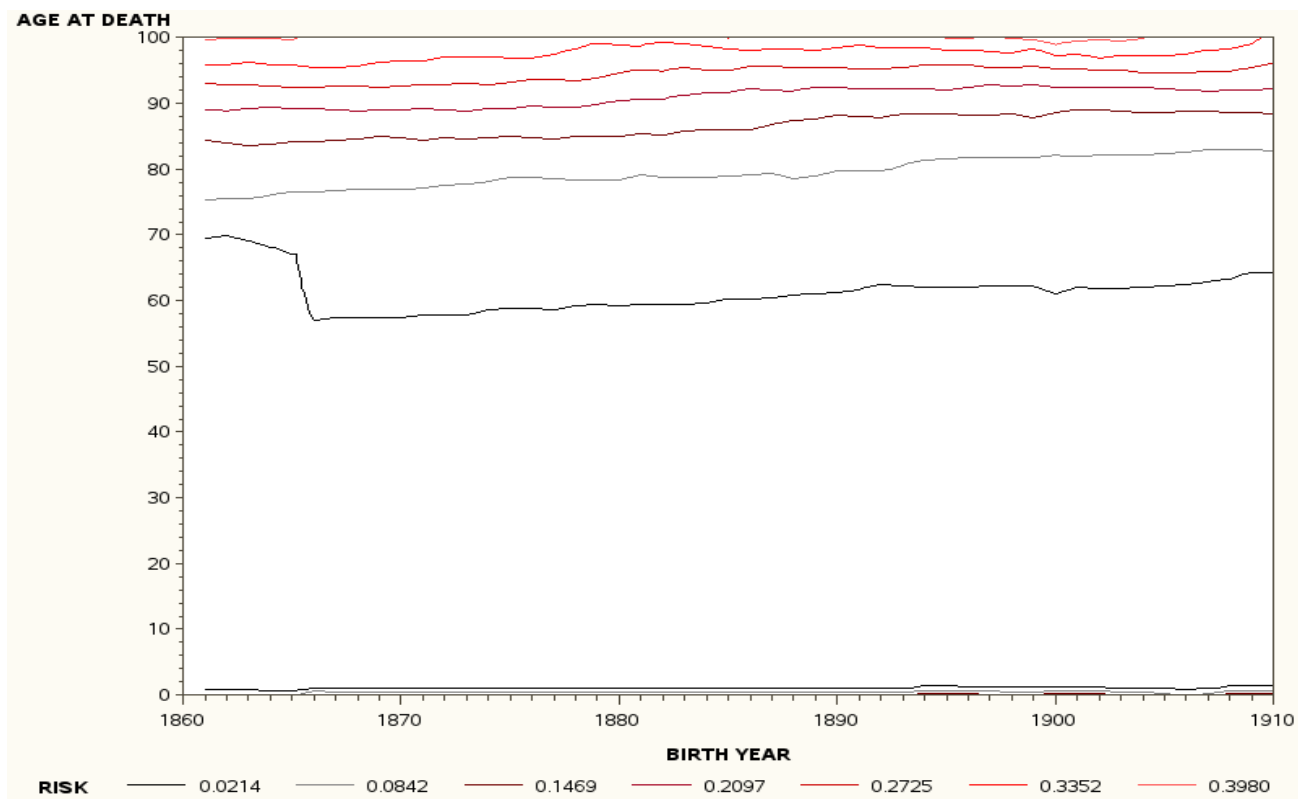
The title, legend, description of the axes and other details are coded in the usual way in SAS graph:

```

goptions reset=all border;
title1 'MORTALITY RATE BY AGE AND YEAR OF BIRTH';
legend1 label=none value=(justify=left 'TEXT 1'
                                'TEXT 2' 'TEXT 3' 'TEXT 4')
footnote j=r height=1.0 font=swiss 'David Corliss 5/17/2011';
goptions colors=(black grey cx700000 cxa50021 cxcc0000 cxff0000
                cxff5050 cxff7c80 cxff9999 cxffc000 cxffffff);
axis1 order=(1860 to 1910 by 10) label=('BIRTH YEAR');
axis2 order=(0 to 100 by 10) label=('AGE AT DEATH');
proc gcontour data=work.plot_data;
  plot age_at_death * birth_year = risk;
  / haxis = axis1
    vaxis = axis2
  ;
  run;
quit;

```

MORTALITY RATE BY AGE AND YEAR OF BIRTH



Specifying the contour levels in PROC GCONTOUR:

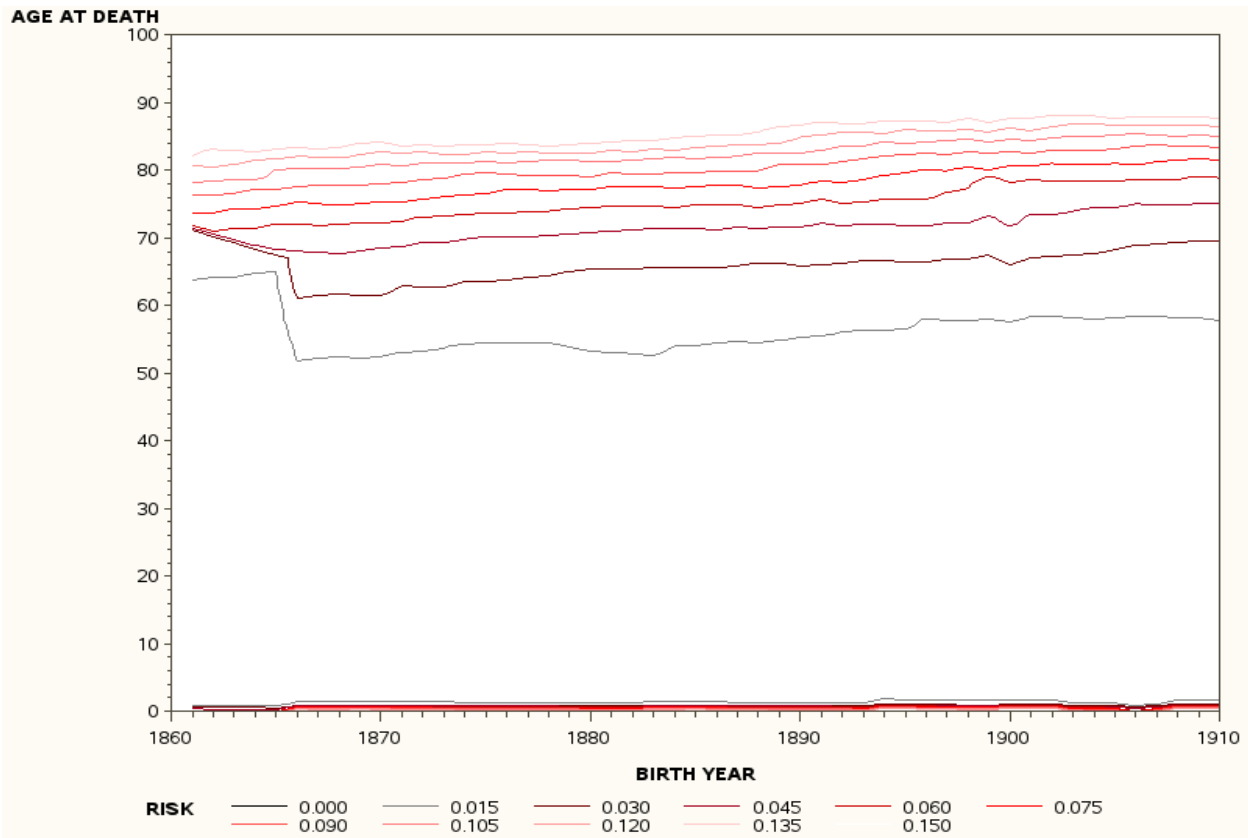
```

goptions reset=all border;
title1 'MORTALITY RATE BY AGE AND YEAR OF BIRTH';
legend1 label=none value=(justify=left 'TEXT 1'
                               'TEXT 2' 'TEXT 3' 'TEXT 4')
footnote j=r height=1.0 font=swiss 'David Corliss 5/17/2011';
goptions device = png
          /* device = png - active-x ignores the levels statement */
          colors=(black grey cx700000 cxa50021 cxcc0000 cxff0000
                  cxff5050 cxff7c80 cxff9999 cxffc000 cxffffff);
axis1 order=(1860 to 1910 by 10) label=('BIRTH YEAR');
axis2 order=(0 to 100 by 10) label=('AGE AT DEATH');
proc gcontour data=work.plot_data;
  plot age_at_death * birth_year = risk;
  / levels = 0.00 to 0.15 by 0.015
    haxis = axis1
    vaxis = axis2
  ;

```

```
run;
quit;
```

MORTALITY RATE BY AGE AND YEAR OF BIRTH



The most important step in creating a time contour plot are the pattern and join options, changing the representation of contours from lines to a mosaic of solidly-colored rectangles:

```
title1 'MORTALITY RATE BY AGE AND YEAR OF BIRTH';
legend1 label=none value=(justify=left 'TEXT 1'
                                'TEXT 2' 'TEXT 3' 'TEXT 4');
footnote j=r height=1.0 font=swiss 'David Corliss 5/17/2011';
goptions device = png
          /* device = png - active-x ignores the levels statement */
          colors=(black grey cx700000 cxa50021 cxcc0000 cxff0000
                  cxff5050 cxff7c80 cxff9999 cxffc000 cxffffff);
axis1 order=(1860 to 1910 by 10) label=('BIRTH YEAR');
axis2 order=(0 to 100 by 10) label=('AGE AT DEATH');
proc gcontour data=work.plot_data;
  plot age_at_death * birth_year = risk;
```

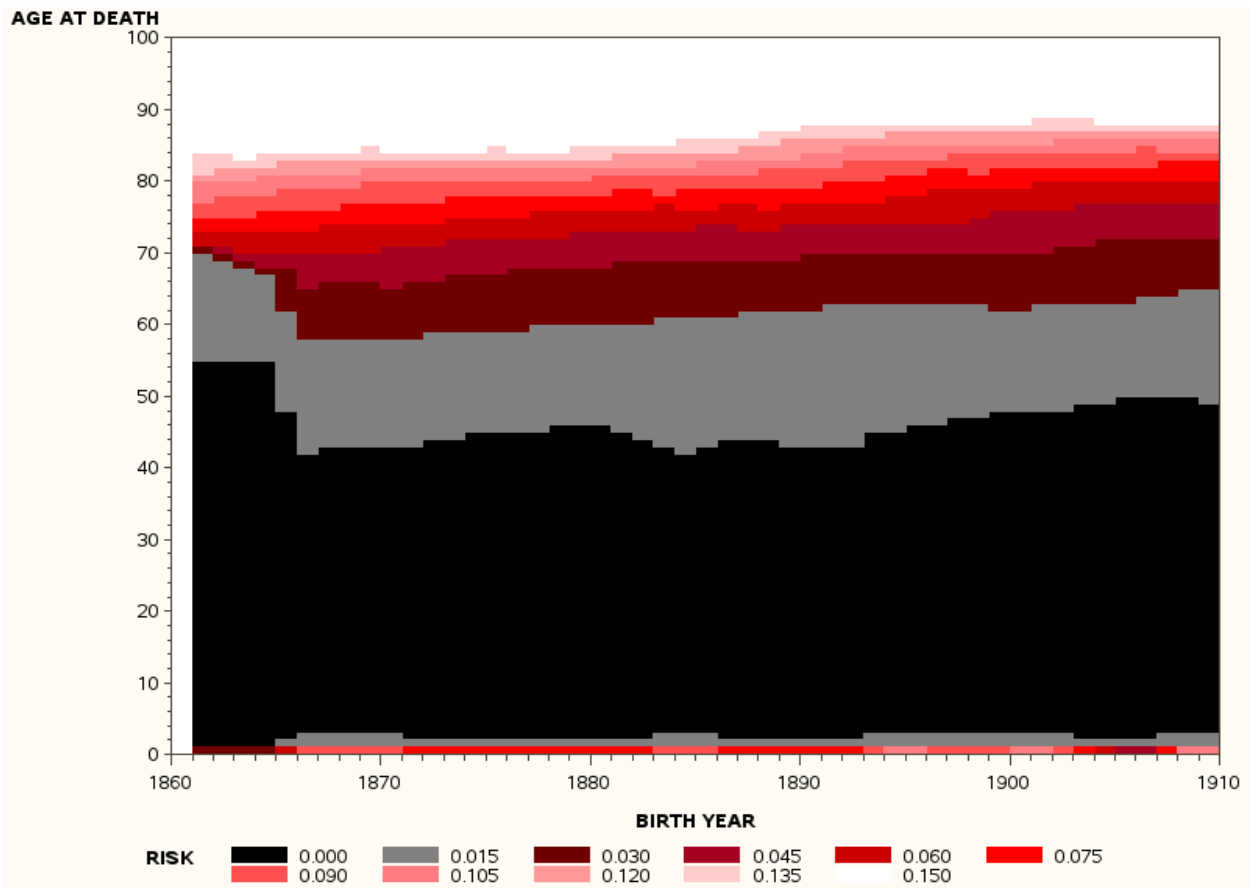


```

    / pattern
      join
      levels = 0.00 to 0.15 by 0.015
      haxis = axis1
      vaxis = axis2
    ;
run;
quit;

```

MORTALITY RATE BY AGE AND YEAR OF BIRTH



A SAS ODS statement is used to create the final plot:

```

title1 'MORTALITY RATE BY AGE AND YEAR OF BIRTH';
legend1 label=none value=(j=left 'TEXT 1' 'TEXT 2' 'TEXT 3' 'TEXT 4')
footnote j=r height=1.0 font=swiss 'David Corliss 5/17/2011';
ods listing close;
ods html path='/home/sas/permanent/DJC/SAS' file='test.html';

```

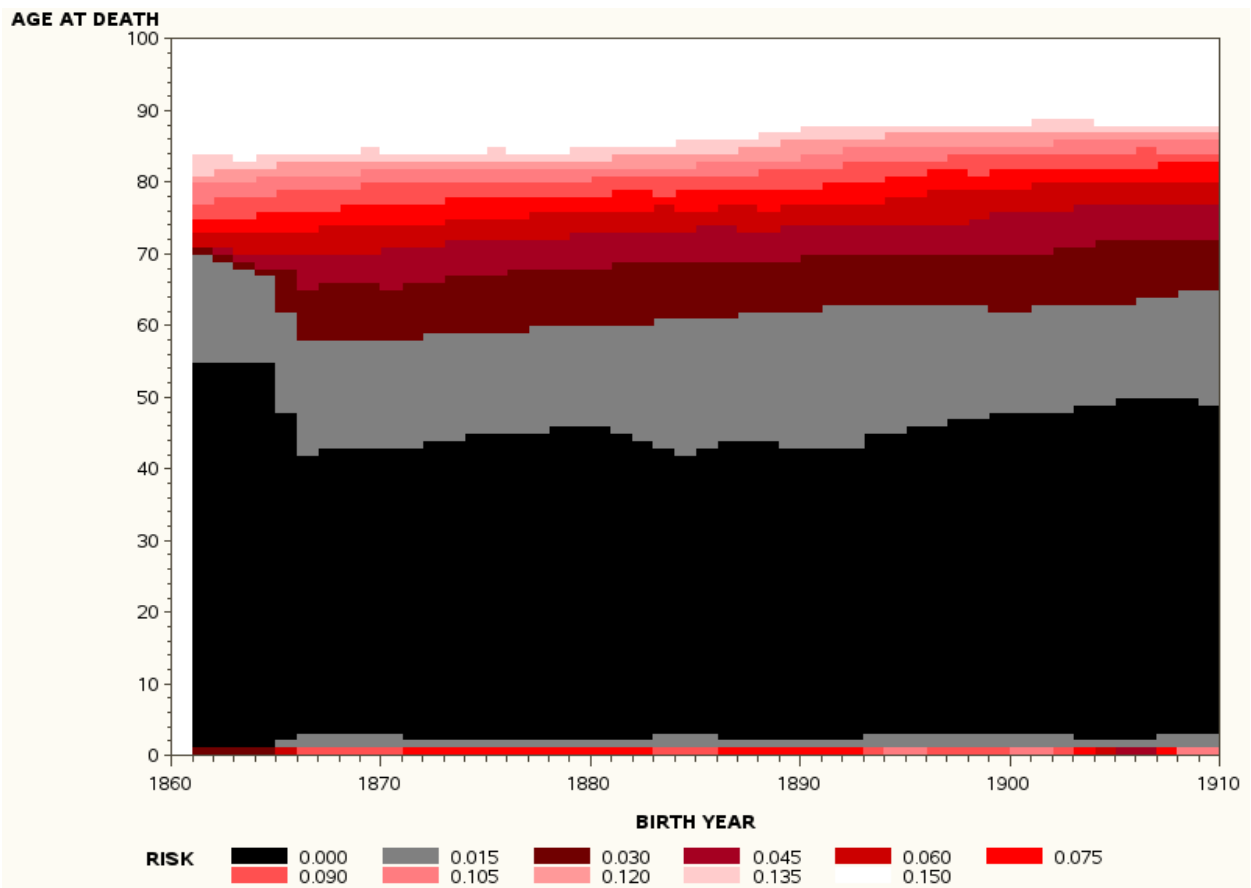
```

goptions device = png /* active-x ignores the levels statement */
      colors=(black grey cx700000 cxa50021 cxcc0000 cxff0000
      cxff5050 cxff7c80 cxff9999 cxffcccc cxffffff);
axis1 order=(1860 to 1910 by 10) label=('BIRTH YEAR');
axis2 order=(0 to 100 by 10) label=('AGE AT DEATH');
proc gcontour data=work.plot_data;
  plot age_at_death * birth_year = risk;
  / pattern join
  levels = 0.00 to 0.15 by 0.015
  haxis = axis1 vaxis = axis2;

run;
quit;
ods html close;
ods listing;

```

MORTALITY RATE BY AGE AND YEAR OF BIRTH



CONCLUSION

The GCONTOUR in SAS Graph can be modified to produce contour plots for data visualization in time series analysis. A time value is plotted with two other continuous variables. The JOIN and PATTERN statements convert the usual contour plot into a mosaic of solidly-colored rectangles. Using these plots, different types of events may identified by the distinctive pattern made in the diagram. In addition to the original applications in astrostatistics, these plots find use in many areas, including biostatistics, econometrics, demographics and other applications.

REFERENCES

Kaufer, A., et al., 1996, Astronomy and Astrophysics, 305, 887K

ACKNOWLEDGMENTS

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

David J Corliss
University of Toledo
Department of Physics and Astronomy
224 McMaster Hall
The University of Toledo, Toledo, OH
Phone: 734.837.9323
Email: davidjcorliss@rockets.utoledo.edu
Web: <http://astro1.panet.utoledo.edu/~dcorliss/>

Marketing Associates, LLC
1 Kennedy Square, Suite 500
Detroit, MI 48224
Phone: 313.202.6323
Email: dcorliss@marketingassociates.com