

SOCR MotionCharts CAOzoneData

Contents

- [1 SOCR MotionCharts Activities - California Ozone Pollution: Geographic and Time Effects on Health Activity](#)
- [2 Summary](#)
- [3 Goals](#)
- [4 Activity video](#)
- [5 Background](#)
- [6 Case Study](#)
 - [6.1 Temporal changes in California Ozone](#)
 - [6.2 Geographic distribution of California Ozone pollution](#)
- [7 Initial setting](#)
- [8 Hands-on activity](#)
- [9 Other types of exploratory and statistical analyses of the Ozone data](#)
 - [9.1 Spider Chart](#)
 - [9.2 Box-and-whisker plot](#)
 - [9.3 Simple Linear Regression analysis](#)
- [10 Data type and format](#)
- [11 Applications](#)
- [12 See also](#)

SOCR MotionCharts Activities - California Ozone Pollution: Geographic and Time Effects on Health Activity

Summary

This activity demonstrates a complete, interactive, self-contained and technology-enhanced pedagogical study of geographic and time effects of Ozone pollution on human health. In addition, this activity illustrates the usage and functionality of [SOCR MotionCharts](#) for exploratory multivariate data analysis using the [SOCR California Ozone dataset](#).

Goals

The aims of this activity is to:

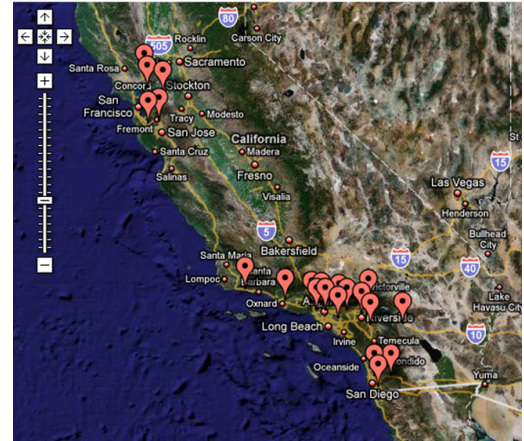
- use MotionChart to address 2 specific health-related case-studies
- demonstrate data import, MotionChart data manipulations and graphical data interpretation
- explore the interactive graphical visualization of real-life multidimensional datasets
- data navigation from different directions (using data mappings).

Activity video

There is a 6-minute [YouTube video of this California Ozone Pollution: Geographic and Time Effects on Health Activity](#).

Background

Suppose we are asked to analyze a complex dataset that included observational multivariate ozone depletion data. The data included [California Ozone measurements from 20 locations between 1980 and 2006](#). The figure on the right illustrates a dynamic interactive map of the geographic locations of the data measurements. This dataset consists of 540 rows and 22 variables. The goals of the study were to identify relationships and associations between the variables and map geographically the significant ozone layer effects. Any such quantitative study requires a preliminary exploratory data analysis. The complexity of the dataset and the intrinsic measurement characteristics of the ozone data demand a new approach to visualization and exploration of these heterogeneous measurements.



Case Study

This Ozone pollution case study addresses the following specific driving environmental challenges:

- *Are there temporal changes in California Ozone?*
- *What is the geographic distribution of the California Ozone pollution and is it changing with time?*

The following chart illustrates the health-related interpretation of the Ozone data in terms of the particulate (particles per million, **ppm**) recordings, according to the National Oceanic and Atmospheric Administration's (NOAA) Air Quality Index (AQI). [EPA](#) guidelines on air quality are as follows:

Air Quality Index for Ozone		
Concentration Range (ppm)	Air Quality Description	Cautionary Statements for Ozone
0.00 – 0.060 ppm	Good	No health impacts are expected
0.061 – 0.075 ppm	Moderate	Unusually sensitive people should consider limited prolonged outdoor exertion
0.076 – 0.104 ppm	Unhealthy for Sensitive Groups	Active children and adults, and people with respiratory conditions (e.g., asthma) should limit prolonged outdoor exertion
0.105 – 0.115 ppm	Unhealthy	Active children and adults, and people with respiratory conditions (e.g., asthma) should avoid prolonged outdoor exertion . Everyone else, especially children and elderly, should limit prolonged outdoor exertion
0.116 – 0.374 ppm	Very Unhealthy	Active children and adults, and people with respiratory conditions (e.g., asthma) should avoid all outdoor exertion . Everyone else, especially children and elderly, should limit outdoor exertion

- (2010) The Obama administration’s proposal sets a primary standard for ground-level ozone of no more than 0.060 to 0.070 ppm, to be phased in over two decades.
- (2005) The Bush administration imposed a limit of 0.075 ppm.
- (1997) The Clinton administration introduced a standard of 0.084 ppm.

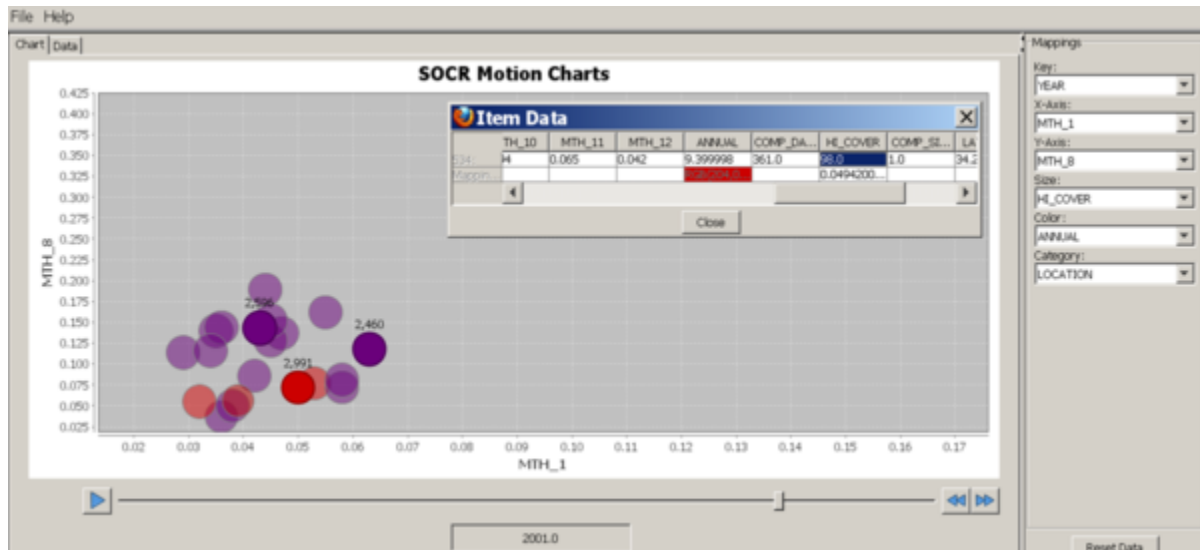
Temporal changes in California Ozone

- Observation: Notice the annualized increase of the ozone pollution with time (increase of the proportion of *hot-colored bubbles* with time).
- Motion Chart: Use the following *variable-mapping* to demonstrate the significant time effect on the increase of the ozone pollution as measured by ppm recordings:

	Variables					
<u>SOCR MotionChart Property</u>	Key	X-Axis	Y-Axis	Size	Color	Category
<u>Data Column Name</u>	Year	MTH_1	MTH_8	HI_COVER	ANNUAL	Location

- Note that the variable mapping we used above allows us to track seasonal changes of ozone pollution (Winter, month1, on the X-axis; and Summer, month8, on the Y-axis). If we play this motion chart, the following interpretation of the motion chart is appropriate: For a given bubble (representing one measurement location):
 - up-down movement* indicate annual Winter (January) increase or decrease of pollution levels, respectively;
 - left-right movement* indicate annual Summer (August) increase or decrease of pollution levels, respectively;
 - bubble-size increase or decrease* indicate corresponding changes in the annual percent coverage during typical periods of high concentration;
 - color change from cool-to-hot* indicates an *annual increase* of the ozone measurement for that specific location from one year to the next.

You should see an image like this one shown below. Play this motion charts by clicking the **Play** button and observe the increase of hot-colored bubbles in the chart as time goes from 1980 to 2006.



Geographic distribution of California Ozone pollution

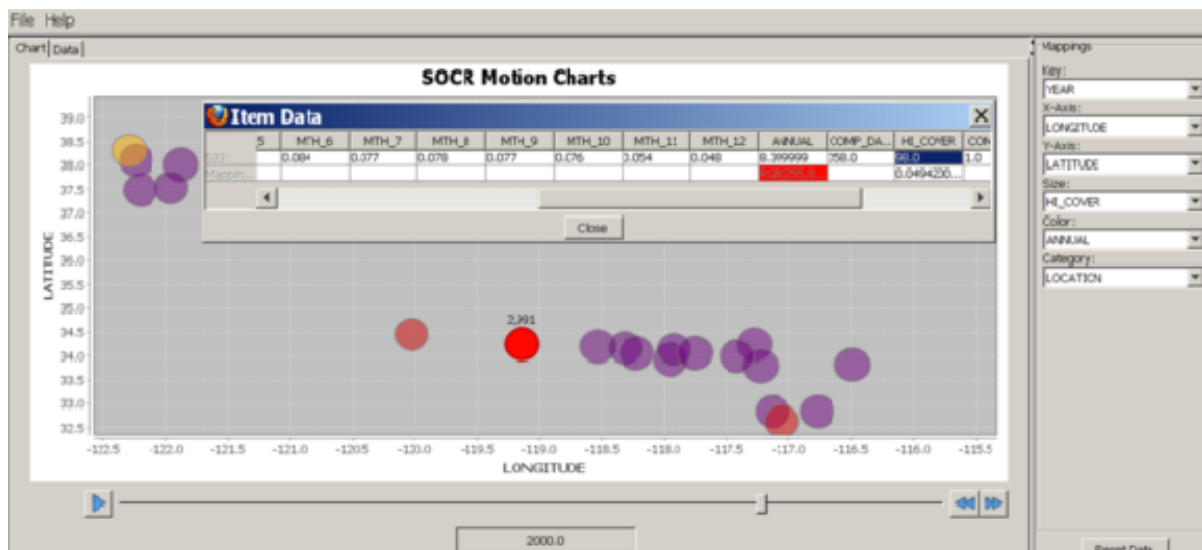
- Observation: The ozone pollution appears to be a more geographically spread out phenomenon in the 2000's, compared to the 1980's -- most of the bubbles cluster together in later years, whereas there were wider geographic-driven fluctuations in the ozone particles in the earlier years. The size of the bubbles reflects the maximum annual pollution and the bubble color indicates the average annual ozone pollution -- *hot-colors* represent high and *cool-colors* represent low ozone pollution levels, respectively.
- Motion Chart: Use the following variable-mapping to demonstrate the significant geographic temporal re-distribution of the ozone pollution as measured by ppm recordings:

Variables

SOCR MotionChart Property	Key	X-Axis	Y-Axis	Size	Color	Category
Data Column Name	Year	LONGITUDE	LATITUDE	HI_COVER	ANNUAL	Location

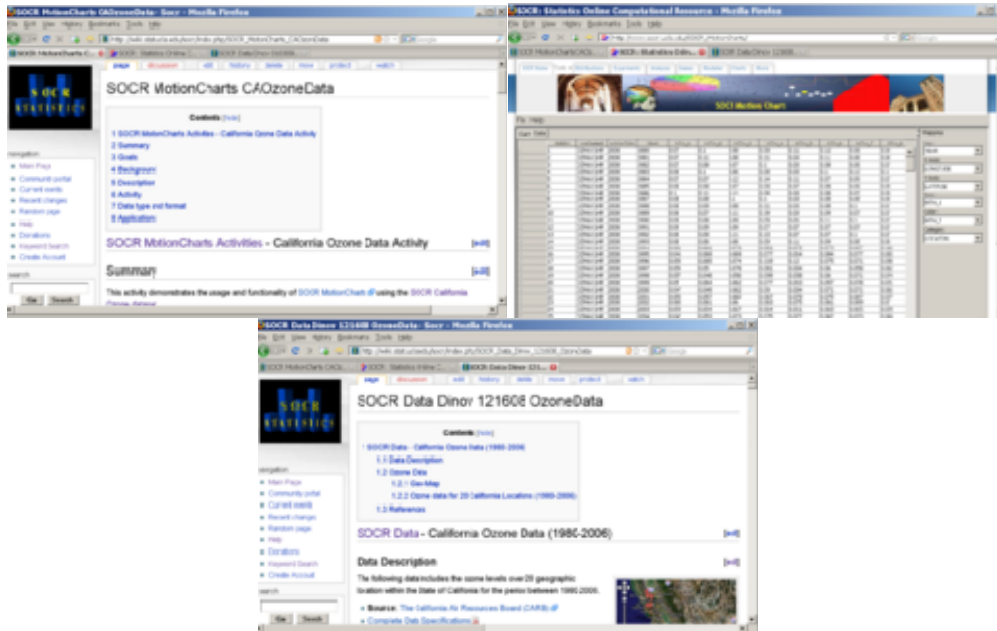
- This second variable mapping allows us to track seasonal changes of ozone pollution for each geographic location. Here the X and Y locations of bubbles are locked at the [GIS longitude and latitude coordinates](#). By playing this motion chart, we have the following interpretation: For a given bubble (representing one GIS location):
 - bubble-size increase* or *decrease* indicate corresponding changes in the annual percent coverage during typical periods of high concentration;
 - color change* from *cool-to-hot* or *hot-to-cool* indicate *annual increase* or *annual decrease*, respectively, of the ozone measurement for that specific location from one year to the next.

You should see an image like this one shown below. In this mapping, each bubble corresponds spatially to a geographic location, just like in the [geographic-map above](#). Play this motion charts by clicking the **Play** button and observe the increase of hot-colored bubbles in later years at geographic locations which did not show unhealthy ozone pollution levels in the early years.



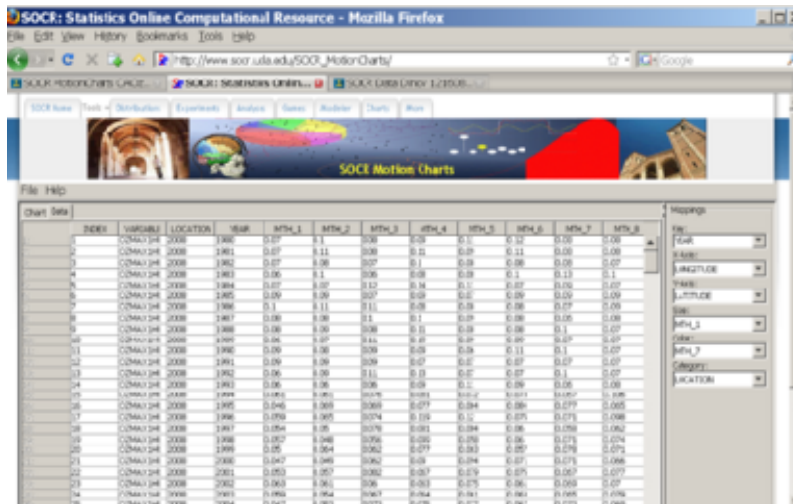
Initial setting

In addition to this activity, open 2 more browser tabs - one is pointing to the [SOCR MotionCharts applet](#) and the other is displaying the [SOCR California Ozone dataset](#). The image below shows this setting.



Hands-on activity

- Using the mouse, copy the [SOCR California Ozone dataset](#), click on the first cell (top-left) in the DATA tab of the [SOCR Motion Charts applet](#), and paste the data in the spreadsheet.



- Next, you need to map the column-variables to different properties in the SOCR MotionChart. For example, you can use the following mapping:

Mappings

Key:

X-Axis:

Y-Axis:

Size:

Color:

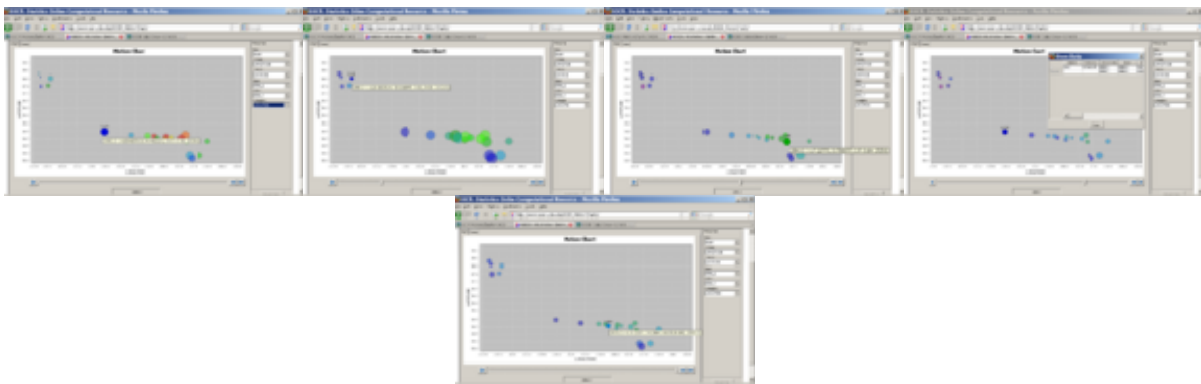
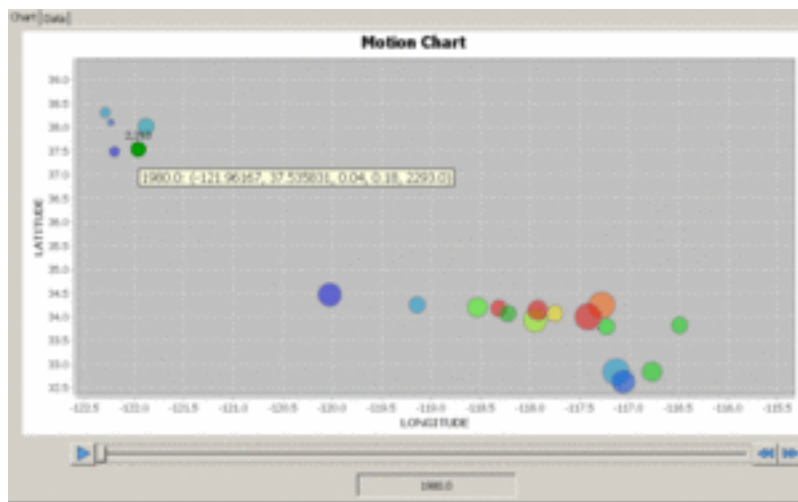
Category:

	Variables					
SOCR MotionChart Property	Key	X-Axis	Y-Axis	Size	Color	Category
Data Column Name	Year	Longitude	Latitude	MTH_1	MTH_7	Location

The figures below represent snapshots of the generated dynamic SOCR motion chart. In the real applet, you can *play* (animate) or *scroll* (1-year steps) through the years (1980, ..., 2006). Notice the position change between different snapshots of the time slider on the bottom of these figures. Also, mouse-over a blob triggers a dynamic graphical pop-up providing additional information about the data for the specified blob in the chart.

You can also change what variables are mapped to the following SOCR MotionCharts properties:

- *Key, X-Axis, Y-Axis, Size, Color and Category.*



We can also overlay the Motion-chart above on the [Geo-map of the 20 locations](#) of the ozone recording stations. This of course, stretches vertically the motion chart, as longitude and latitude coordinates are not isotropic.



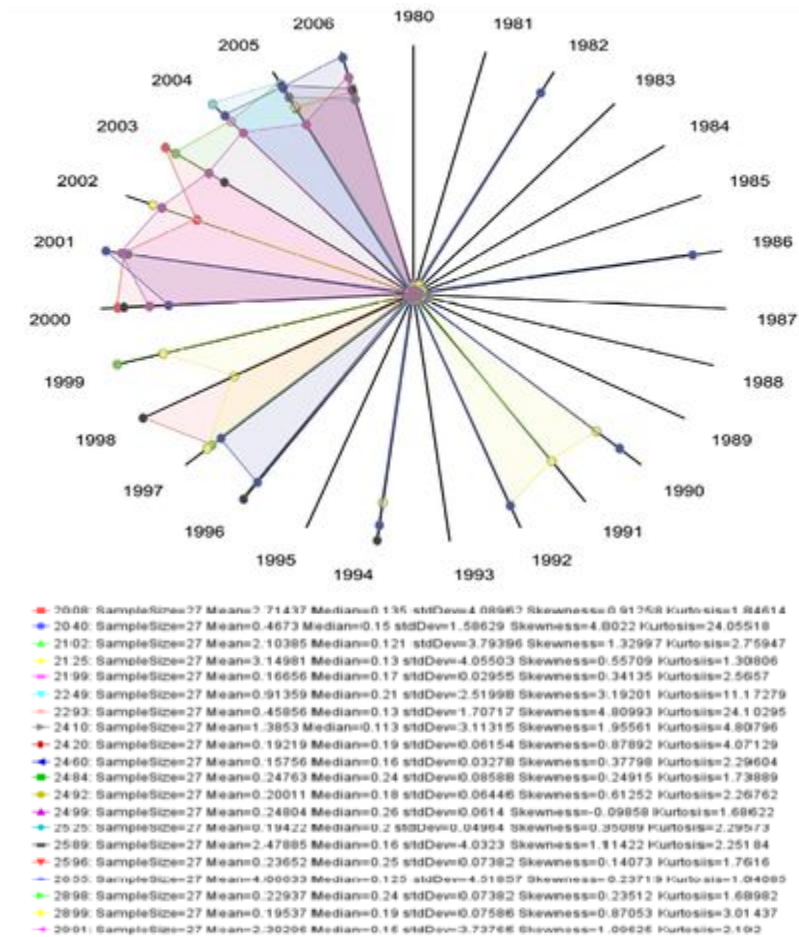
Other types of exploratory and statistical analyses of the Ozone data

Various [SOCR online tools](#) can also be used to analyze (visually or quantitatively) the Ozone pollution data. The table below contains a summary of the annual pollution rates (ppm) for each of the 20 locations across the span of 27 years:

LOCATION	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
2008	0.12	0.11	0.15	0.14	0.14	0.13	0.11	0.17	0.11	0.19	0.11	0.11	0.13	0.11	0.106	0.135	0.12	8.7	9.8	0.088	9.3	9.2	7.4	9.7	0.109	8.2	8.2
2040	0.18	0.2	0.23	0.28	0.28	0.22	0.15	0.17	0.22	0.23	0.2	0.18	0.16	0.146	0.102	0.12	0.12	0.115	0.125	0.104	0.118	0.135	0.112	0.107	0.105	8.3	0.108
2102	0.13	0.11	0.1	0.14	0.16	0.14	0.1	0.15	0.12	0.11	0.11	7.9	0.11	0.13	0.111	0.124	0.121	8.7	0.097	9.7	0.107	0.118	0.111	9.3	8.9	9.3	0.105
2125	0.15	0.13	0.1	0.17	0.11	0.13	0.1	0.12	0.1	0.1	7.9	7.9	8.9	0.1	8.3	0.14	0.097	8.9	6.5	8.2	0.083	0.105	8.9	0.113	0.097	8.3	8.4
2199	0.21	0.19	0.19	0.19	0.2	0.24	0.18	0.17	0.2	0.19	0.17	0.18	0.15	0.17	0.165	0.16	0.16	0.155	0.173	0.126	0.124	0.137	0.136	0.141	0.125	0.139	0.126
2249	0.31	0.27	0.32	0.27	0.32	0.34	0.25	0.24	0.29	0.26	0.21	0.21	0.21	0.19	0.252	0.16	0.15	0.134	0.182	0.116	0.137	0.114	0.121	0.165	9.8	9.3	0.146
2293	0.19	0.16	0.14	0.16	0.15	0.15	0.14	0.16	0.13	0.12	0.13	0.12	0.12	0.13	0.12	0.153	0.1	0.109	0.115	0.133	0.102	0.109	0.11	0.123	8.9	0.105	0.102
2410	0.14	0.12	0.1	0.13	0.14	0.12	8.9	0.11	0.12	0.12	0.11	0.11	0.1	0.11	0.1	0.133	0.112	0.103	0.119	0.113	0.079	9.1	0.109	0.101	0.104	8.7	7.9
2420	0.38	0.25	0.22	0.26	0.26	0.25	0.22	0.22	0.25	0.23	0.19	0.22	0.17	0.19	0.14	0.145	0.205	0.121	0.161	0.1	0.109	0.14	0.152	0.179	0.131	0.138	0.158
2460	0.23	0.19	0.18	0.18	0.16	0.22	0.16	0.17	0.19	0.2	0.17	0.15	0.17	0.187	0.147	0.146	0.138	0.136	0.164	0.124	0.121	0.135	0.121	0.125	0.106	0.113	0.121
2484	0.41	0.35	0.36	0.39	0.31	0.36	0.31	0.3	0.3	0.33	0.23	0.28	0.27	0.24	0.251	0.212	0.196	0.162	0.195	0.137	0.174	0.189	0.136	0.15	0.134	0.145	0.165
2492	0.35	0.27	0.25	0.31	0.26	0.3	0.28	0.23	0.24	0.2	0.2	0.22	0.22	0.18	0.167	0.165	0.142	0.134	0.177	0.12	0.152	0.129	0.128	0.134	0.137	0.142	0.166
2499	0.32	0.35	0.32	0.28	0.34	0.3	0.26	0.29	0.29	0.27	0.33	0.27	0.28	0.24	0.265	0.256	0.234	0.205	0.244	0.174	0.176	0.17	0.161	0.163	0.163	0.182	0.164
2525	0.29	0.24	0.28	0.26	0.22	0.29	0.22	0.2	0.23	0.21	0.19	0.2	0.21	0.2	0.184	0.202	0.18	0.136	0.149	0.112	0.164	0.152	0.147	0.155	0.128	0.126	0.169
2589	0.16	0.17	0.2	0.21	0.15	0.2	0.14	0.16	0.22	0.16	0.15	0.15	0.15	0.133	9.8	0.14	9.7	0.117	9.8	0.105	9.1	0.102	0.115	7.4	0.097	9.2	8.3
2596	0.37	0.3	0.31	0.36	0.32	0.35	0.25	0.29	0.28	0.27	0.29	0.24	0.26	0.26	0.253	0.213	0.203	0.187	0.195	0.142	0.14	0.143	0.155	0.169	0.141	0.144	0.151
2655	0.12	0.11	8.9	0.11	0.11	0.11	8.9	0.11	0.1	0.1	8.9	0.11	8.9	0.12	9.2	0.13	8.9	8.3	0.125	0.115	7.7	9.8	0.116	0.105	9.2	9.1	9.6
2898	0.37	0.33	0.31	0.34	0.31	0.33	0.27	0.29	0.29	0.25	0.24	0.24	0.26	0.21	0.241	0.215	0.187	0.156	0.184	0.141	0.152	0.144	0.15	0.161	0.131	0.14	0.151
2899	0.29	0.32	0.4	0.26	0.29	0.3	0.22	0.22	0.21	0.25	0.2	0.19	0.2	0.16	0.193	0.167	0.144	0.12	0.148	0.128	0.136	0.116	0.122	0.152	0.11	0.121	0.108
2991	0.13	0.16	0.15	0.15	0.14	0.15	0.18	0.18	0.16	0.19	0.12	0.12	0.141	0.138	0.115	0.124	0.121	0.102	0.106	0.103	8.3	9.3	8.6	8.0	8.3	7.5	8.8

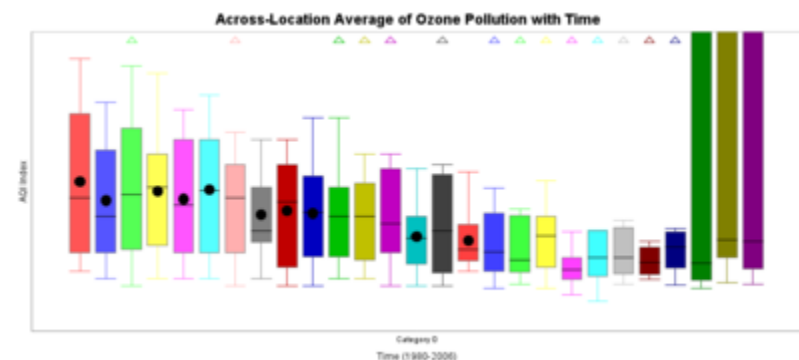
Spider Chart

This [spider chart](#) visually demonstrates the rapid increase of the pollution levels across the years (radial spokes). For each site the 27 annual measurements are connected with lines of the same color as shown on the 20-locations color-mapping below.



Box-and-whisker plot

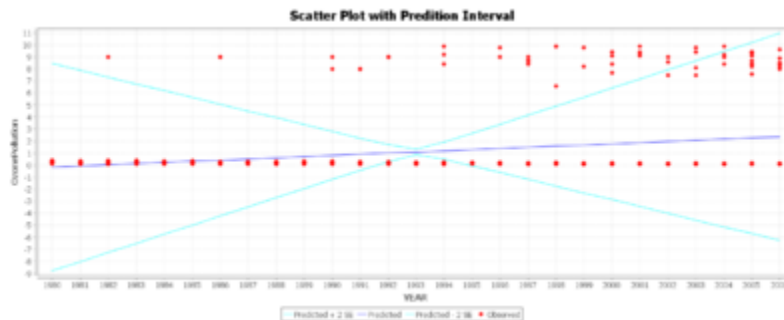
This [Box and Whiskers Plot](#) illustrates the same annual (across locations) averages of the ozone pollution for the 27 years on record. Notice the increase in (high-level) outliers, denoted by colored triangles on the top, and the atypically average high pollution levels in the last few years.



Simple Linear Regression analysis

We can employ the [SOCR simple linear regression analysis](#) to determine the relation between the (average) **Annual** Ozone pollution (response variable) relative to **Year** (time predictor variable). The quantitative results and graphs of this analysis, using only the [Year and Annual variables in the Ozone dataset](#), are included below:

- Quantitative analysis:
 - Regression Line:
$$\text{OzonePollution} = -191.6095127656 + 0.09667 \times \text{YEAR}$$
 - Intercept:
 - Parameter Estimate: -191.6095127656
 - Standard Error: 28.4291851368
 - T-Statistics: -6.7398876135
 - P-Value: 0.0000000000
 - Slope:
 - Parameter Estimate: 0.0966959040
 - Standard Error: 0.0142644095
 - T-Statistics: 6.7788228014
 - P-Value: 0.0000000000
 - Correlation(YEAR, OzonePollution) = 0.2805211066
- The prediction interval and the regression line have a slight (but statistically significant) upward trend, which is mostly due to the sporadic extremely high pollution levels in the last few years.



Data type and format

SOCR Motion Charts currently accepts three types of data: numbers, dates/time, and strings. With these data types, we feel that the application is able to handle the majority of data out here. We use the natural ordering of these types as defined by Java however. While many types of data can be interpreted as a string, it may not make sense to use lexicological ordering on all the different types. When designing SOCR Motion Charts, we took this into consideration and designed the application so that it can easily be extended to provide a greater variety of interpreted types. Thus, a developer should be able to easily provide better type interpretation for particular types of data.

Applications

The SOCR MotionCharts can be used in a variety of applications to visualize dynamic relationships in multidimensional data in up to four dimensions and a fifth temporal component. Its design and implementation allow for extensions allowing and supporting higher dimensions plug-ins. The overall

purpose of SOCR MotionCharts is to provide users with a way to visualize the relationships between multiple variables over a period of time in a simple, intuitive and animated fashion.

See also

- [E.P.A. Seeks Stricter Rules to Curb Smog \(New York Times article\)](#)
-

Translate this page:

(default) 	Deutsch 	Español 	Français 	Italiano 	Português 	日本語 	България 	تارامالا ڈیبرعرا قدحتول 	Suomi 	☐☐☐☐ ☐☐☐☐ ☐☐☐☐ 	Norge 
한국어 	中文 	繁体中文 	Русский 	Nederlands 	Ελληνικά 	Hrvatska 	Česká republika 	Danmark 	Polska 	România 	Sverige 