# INTERACTIVE TABLES AND MAPS — A GLANCE AT EPA'S CUMULATIVE EXPOSURE PROJECT WEB PAGE

**Jürgen Symanzik[*], Utah State University,**
**Daniel B. Carr, George Mason University,**
**Daniel A. Axelrad, U.S. Environmental Protectation Agency,**
**Jingfang Wang, George Mason University,**
**David Wong, George Mason University,**
**Tracey J. Woodruff, U.S. Environmental Protectation Agency.**
[*]Department of Mathematics and Statistics, Logan, UT 84322–3900,
e–mail: `symanzik@sunfs.math.usu.edu`

**Key Words:** Micromaps, Hazardous Air Pollutants (HAPs), Visualization, Map Generalization.

## Abstract

The distribution of statistical data through printed tables has never been an effective and fast approach for understanding large data sets. Multiple rows and columns, spread across multiple pages, make it difficult to locate interesting information in a printed table at a glance. However, the usefulness of tables has been increased considerably when used on the World Wide Web. Interactive tables that can be rearranged by the click of a mouse are now a very helpful tool to quickly find structure in data.

Such interactive tables, together with hierarchical clickable micromaps paired with graphical representations of data through the Graphics Production Library (GPL), were constructed as the core of EPA's Cumulative Exposure Project Web page. The main intention of this Web page is to provide easy access to EPA's hazardous air pollutant (HAP) data for 1990. Concentrations of 148 air pollutants were estimated for each of the 60,803 census tracts in the 48 contiguous states. The Web page design allows the user to easily move through the data set to find information on different pollutants at different locations and at different levels of geographic resolution.

## 1. Introduction

Under EPA's Cumulative Exposure Project (CEP), modeled concentration estimates of 148 hazardous air pollutants (HAPs) for each of the 60,803 census tracts in the continental United States have been developed. These modeled concentrations are annual averages for the year 1990. Additional uncertainty bounds have been developed for each estimate. The development of the modeled concentrations is described in Rosenbaum, Axelrad, Woodruff, Wei, Ligocki & Cohen (1999) and the analysis of the modeled concentrations is presented in Woodruff, Axelrad, Caldwell, Morello–Frosch & Rosenbaum (1998) and Caldwell, Woodruff, Morello–Frosch & Axelrad (1998).

One of the original main purposes of the CEP Web page, accessible through the URL `http://www.epa.gov/CumulativeExposure`, was to provide easy, fast, and understandable access to the modeled HAP concentrations. Since there are 148 × 60,803 × 3 numeric values that resulted from the modeling process, we developed two hierarchical spatial approaches for the display of this data — hierarchical clickable micromaps, described in Section 2, and interactive tables that can be rearranged by the click of a mouse button, described in Section 3.

Additional details regarding the access and visualization of HAPs and the CEP Web page can be found in Symanzik, Axelrad, Carr, Wang, Wong & Woodruff (1999) and Symanzik, Wong, Wang, Carr, Woodruff & Axelrad (1999). These articles also contain additional examples of linked micromap plots at the state level for Michigan and Pennsylvania, respectively.

## 2. Micromaps

Linked micromap plots, often simply called micromaps, have been fully described in Carr & Pierson (1996), Carr, Olsen, Courbois, Pierson & Carr (1998), Carr, Olsen, Pierson & Courbois (1998), and Carr, Olsen, Pierson & Courbois (1999). In short, the main idea behind micromaps is to focus the viewer's attention on the statistical information presented in such a plot. Multiple, small generalized maps are used to provide the appropriate geographic reference for the data.

Figure 1 shows a micromap of the median (with respect to the underlying census tract estimates) modeled 1990 lead concentrations for Iowa. Lead is a HAP that occurs in highly industrialized areas and, in 1990, still was widely used as an addendum to gas. Not surprisingly, four of the five highest lead concentrations were modeled for highly industrialized counties in Iowa, i.e., for Scott County (with Davenport), Dubuque County (with Dubuque), Linn County (with Cedar Rapids), and Black Hawk County (with Waterloo). Even though Story County is not highly industrialized, it has the fourth highest modeled median lead concentration.

Polk County (with Des Moines, Iowa's largest City) is a region of service businesses and a large number of everyday commuters but little heavy industry and therefore only appears as #7 with respect to the modeled median lead concentration. Air pollution does not stop at state boundaries, which becomes obvious for #8 in the list, Pottawattamie County, which directly faces the highly industrialized metropolitan area of Omaha, Nebraska. Lowest modeled median lead concentrations were found in the mostly rural and sparsely populated counties in western and central–eastern Iowa.

Figure 2 shows a micromap at the census tract level of modeled 1990 lead concentrations for Polk County. As already mentioned earlier, Polk County only has little industry but a large, commuting population. No wonder that the highest modeled lead concentrations were found in census tracts of Des Moines and its suburbs with a high population density, while lowest modeled lead concentrations were found mostly in rural census tracts surrounding the city.

## 3. Interactive Tables

Interactive tables have been developed for the CEP Web page as a hierarchical navigational tool and for the rearrangement of data. When starting at the top US level, a user can select a state by mouse–clicking on one of the state names in the table and immediately get a new tabular display at the state level. When selecting a county inside a table, a new tabular display at the census tract level will be displayed. Additional menu buttons also allow to move up and down in this geographical hierarchy.

Figure 3 demonstrates the effects of an interactive rearrangement of the same data, once again the modeled 1990 lead concentrations for Iowa. In the top of this figure, counties are arranged with respect to their minimum lead concentration. Polk County is listed as #2. When rearranging the data with respect to median lead concentration (middle), Polk County is listed as # 7. Finally, when rearranging the data with respect to the 75th percentile (bottom), Polk County is not listed among the counties with the 10 highest concentrations. This is not surprising since the lead distribution through mobile sources such as cars is more evenly spread through an entire county while stationary sources such as industry cause highest concentrations closest to the source but are much less spread through an entire county.

Figure 4 shows a screenshot from the CEP Web page that displays the modeled 1990 lead concentrations for Polk County, sorted decreasingly by concentration. In addition, the control buttons that allow to move up and down in the geographical hierarchy are visible above the tabular display. The frame left of the tabular display provides links to the textual part of the CEP Web page.

## 4. Discussion

EPA originally planned to provide access to the HAP data using interactive tables in December 1998. Micromap displays, accessed through the Graphics Production Library (GPL), were scheduled for release in early 1999. The GPL was originally developed within the Bureau of Labor Statistics (BLS) in 1996 (Carr, Valliant & Rope 1996) and provides tools for the creation of statistical graphics on the Web. However, due to concerns that the estimates for 1990 may not be representative of current conditions, EPA ultimately decided not to make the modeled air toxics concentrations available through its public Web site.

A preview Web page that contains the fully implemented interactive tables is accessible for federal agencies. Please contact Daniel Axelrad (`axelrad.daniel@epa.gov`) how to access the preview site. The combination of micromaps with the (old) GPL has been halted. At this point, it seems to be more advisable to combine micromaps with a new commercial version of the GPL scheduled for release in the Fall of 1999. Even if the CEP Web page will never provide public access to the modeled 1990 HAP concentrations, it seems to be likely that interactive maps and hierarchical clickable micromaps in the new GPL environment might be used for other federal projects in the near future.

## Acknowledgments

EPA funded the majority of the work behind this paper under contracts No. 8W–0662–NAEX and No. 8W–1712–NTEX and cooperative agreement No. CR825564–01–0. Additional federal agencies, BLS and NCHS, supported some facets of this work. The article has not been subject to review by any of these agencies and so does not necessarily reflect the view of the agencies. No official endorsement should be inferred. The conclusions and opinions are solely those of the authors.

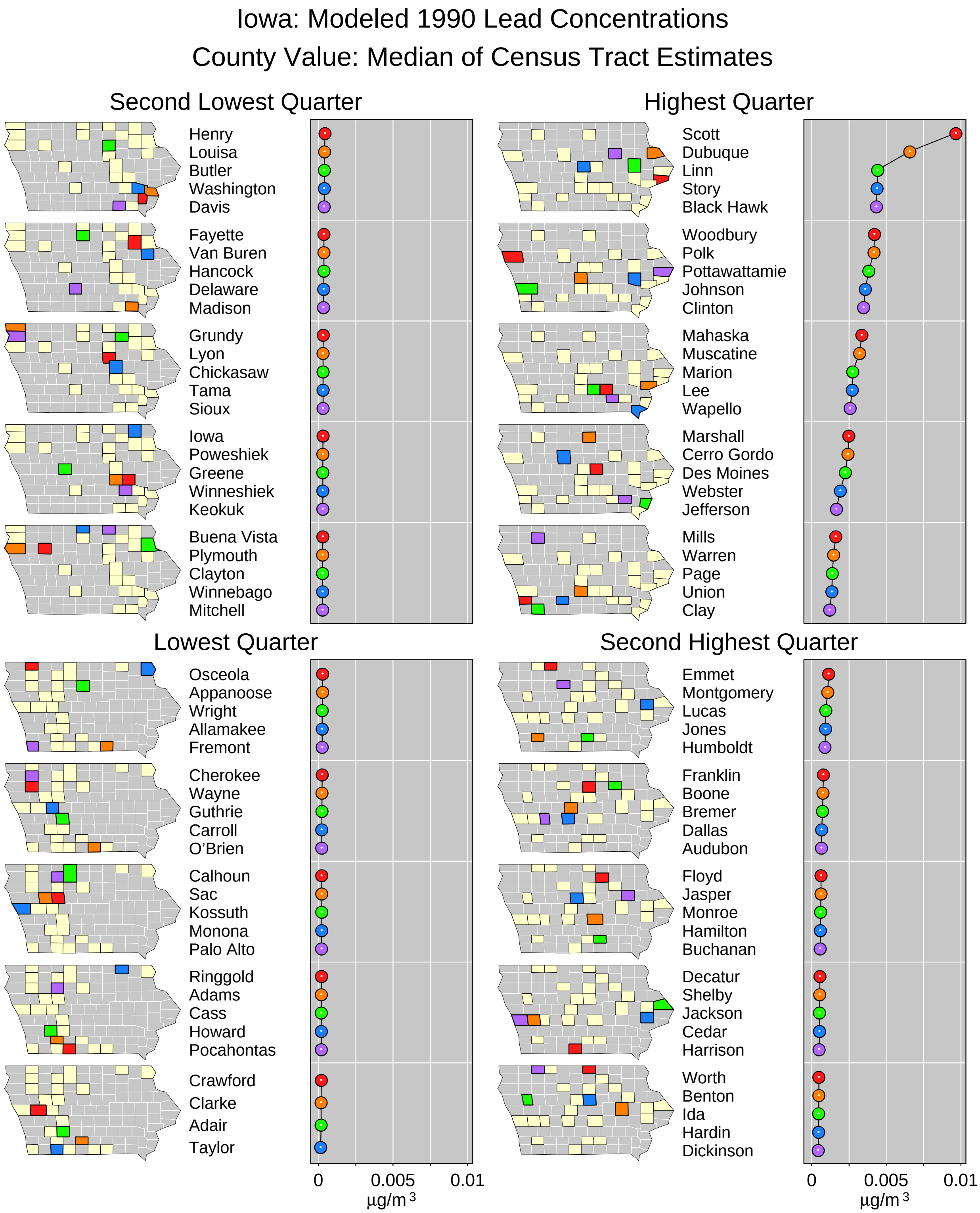



Figure 1: Micromap display at the state level for Iowa for all of its 99 counties. Displayed is the median (with respect to the census tract estimates within each county) of the modeled 1990 lead concentrations in $\mu g/m^3$. All counties that have concentrations in the same quarter of the data are highlighted in the 5 related maps to provide a better geographical relationship. Micromaps are best viewed in full color. You can find a color version of this paper on the Web at `http://www.math.usu.edu/~symanzik/`. A collection of micromap plots (from this and other papers) is available at `http://www.galaxy.gmu.edu/~dcarr/graphics/graphics.html`.

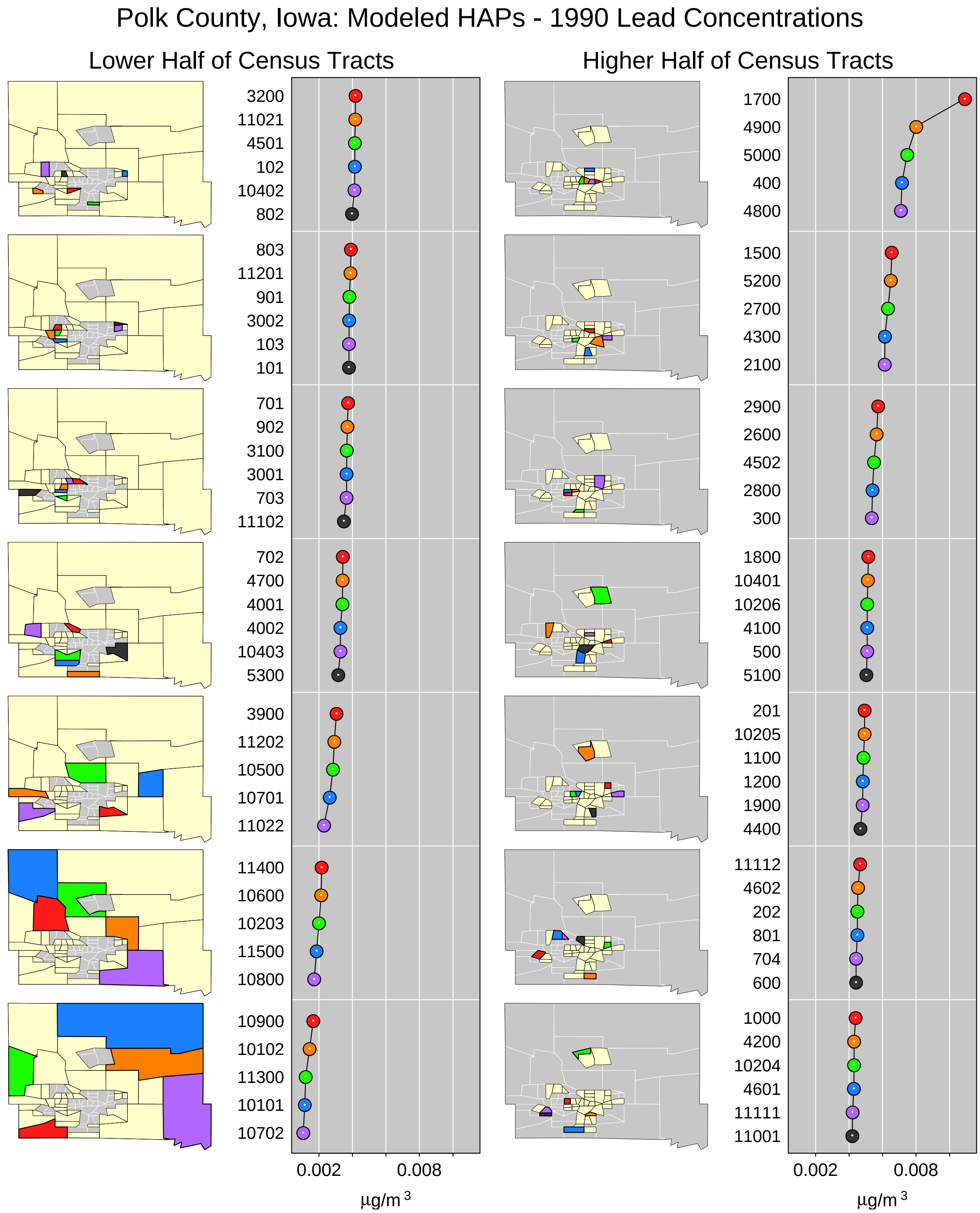



Figure 2: Micromap display at the county level for Polk County, Iowa, for all of its 78 census tracts. Displayed is the modeled 1990 lead concentration in $\mu g/m^3$. All census tracts that have concentrations below (above) the median concentration are highlighted in the left (right) 7 maps to provide a better geographical relationship.

## Iowa – Summary

**1990 Modeled Concentration of HAP lead in µg/m³**

| County | Tracts | Mean | Median | Min | P25 | P75 | Max |
|---|---|---|---|---|---|---|---|
| Scott County | 39 | **0.00936** | **0.00967** | 0.00111 | 0.00741 | 0.01124 | 0.01889 |
| Polk County | 78 | **0.00426** | **0.00419** | 0.00106 | 0.00341 | 0.00508 | 0.01092 |
| Linn County | 40 | **0.00449** | **0.00443** | 0.00063 | 0.00198 | 0.00616 | 0.01535 |
| Muscatine County | 10 | **0.00396** | **0.00321** | 0.00057 | 0.00115 | 0.00702 | 0.00945 |
| Black Hawk County | 36 | **0.00412** | **0.00435** | 0.00054 | 0.00238 | 0.00602 | 0.00876 |
| Des Moines County | 12 | **0.00215** | **0.00226** | 0.00049 | 0.00111 | 0.00290 | 0.00406 |
| Johnson County | 25 | **0.00412** | **0.00360** | 0.00047 | 0.00151 | 0.00612 | 0.01236 |
| Cerro Gordo County | 10 | **0.00319** | **0.00243** | 0.00045 | 0.00132 | 0.00582 | 0.00798 |
| Dubuque County | 26 | **0.00603** | **0.00658** | 0.00044 | 0.00189 | 0.00885 | 0.01421 |
| Wapello County | 11 | **0.00295** | **0.00259** | 0.00043 | 0.00112 | 0.00400 | 0.00854 |

## Iowa – Summary

**1990 Modeled Concentration of HAP lead in µg/m³**

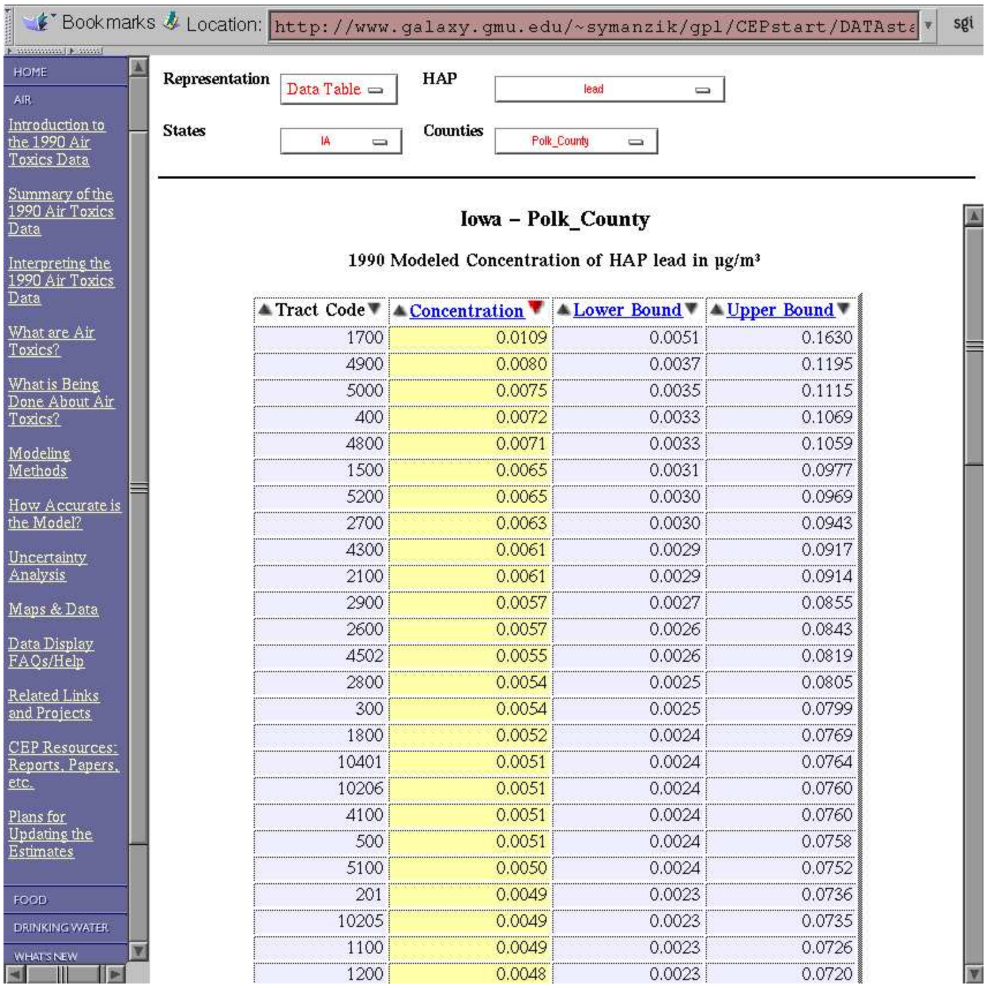
| County | Tracts | Mean | Median | Min | P25 | P75 | Max |
|---|---|---|---|---|---|---|---|
| Scott County | 39 | **0.00936** | **0.00967** | 0.00111 | 0.00741 | 0.01124 | 0.01889 |
| Dubuque County | 26 | **0.00603** | **0.00658** | 0.00044 | 0.00189 | 0.00885 | 0.01421 |
| Linn County | 40 | **0.00449** | **0.00443** | 0.00063 | 0.00198 | 0.00616 | 0.01535 |
| Story County | 20 | **0.00486** | **0.00439** | 0.00041 | 0.00158 | 0.00730 | 0.01182 |
| Black Hawk County | 36 | **0.00412** | **0.00435** | 0.00054 | 0.00238 | 0.00602 | 0.00876 |
| Woodbury County | 26 | **0.00397** | **0.00421** | 0.00022 | 0.00199 | 0.00518 | 0.01058 |
| Polk County | 78 | **0.00426** | **0.00419** | 0.00106 | 0.00341 | 0.00508 | 0.01092 |
| Pottawattamie County | 26 | **0.00369** | **0.00383** | 0.00023 | 0.00195 | 0.00532 | 0.00967 |
| Johnson County | 25 | **0.00412** | **0.00360** | 0.00047 | 0.00151 | 0.00612 | 0.01236 |
| Clinton County | 12 | **0.00397** | **0.00349** | 0.00039 | 0.00109 | 0.00572 | 0.00987 |

## Iowa – Summary

**1990 Modeled Concentration of HAP lead in µg/m³**

| County | Tracts | Mean | Median | Min | P25 | P75 | Max |
|---|---|---|---|---|---|---|---|
| Scott County | 39 | **0.00936** | **0.00967** | 0.00111 | 0.00741 | 0.01124 | 0.01889 |
| Dubuque County | 26 | **0.00603** | **0.00658** | 0.00044 | 0.00189 | 0.00885 | 0.01421 |
| Story County | 20 | **0.00486** | **0.00439** | 0.00041 | 0.00158 | 0.00730 | 0.01182 |
| Muscatine County | 10 | **0.00396** | **0.00321** | 0.00057 | 0.00115 | 0.00702 | 0.00945 |
| Linn County | 40 | **0.00449** | **0.00443** | 0.00063 | 0.00198 | 0.00616 | 0.01535 |
| Johnson County | 25 | **0.00412** | **0.00360** | 0.00047 | 0.00151 | 0.00612 | 0.01236 |
| Black Hawk County | 36 | **0.00412** | **0.00435** | 0.00054 | 0.00238 | 0.00602 | 0.00876 |
| Cerro Gordo County | 10 | **0.00319** | **0.00243** | 0.00045 | 0.00132 | 0.00582 | 0.00798 |
| Clinton County | 12 | **0.00397** | **0.00349** | 0.00039 | 0.00109 | 0.00572 | 0.00987 |
| Pottawattamie County | 26 | **0.00369** | **0.00383** | 0.00023 | 0.00195 | 0.00532 | 0.00967 |

Figure 3: Tabular display at the state level for Iowa showing the modeled 1990 lead concentrations in $\mu g/m^3$. Displayed are the top 10 counties when the data is ordered from highest to lowest with respect to the Minimum (top), Median (middle), and 75th Percentile (bottom), based on the underlying census tract estimates within each county.




| Tract Code | Concentration | Lower Bound | Upper Bound |
|---|---|---|---|
| 1700 | 0.0109 | 0.0051 | 0.1630 |
| 4900 | 0.0080 | 0.0037 | 0.1195 |
| 5000 | 0.0075 | 0.0035 | 0.1115 |
| 400 | 0.0072 | 0.0033 | 0.1069 |
| 4800 | 0.0071 | 0.0033 | 0.1059 |
| 1500 | 0.0065 | 0.0031 | 0.0977 |
| 5200 | 0.0065 | 0.0030 | 0.0969 |
| 2700 | 0.0063 | 0.0030 | 0.0943 |
| 4300 | 0.0061 | 0.0029 | 0.0917 |
| 2100 | 0.0061 | 0.0029 | 0.0914 |
| 2900 | 0.0057 | 0.0027 | 0.0855 |
| 2600 | 0.0057 | 0.0026 | 0.0843 |
| 4502 | 0.0055 | 0.0026 | 0.0819 |
| 2800 | 0.0054 | 0.0025 | 0.0805 |
| 300 | 0.0054 | 0.0025 | 0.0799 |
| 1800 | 0.0052 | 0.0024 | 0.0769 |
| 10401 | 0.0051 | 0.0024 | 0.0764 |
| 10206 | 0.0051 | 0.0024 | 0.0760 |
| 4100 | 0.0051 | 0.0024 | 0.0760 |
| 500 | 0.0051 | 0.0024 | 0.0758 |
| 5100 | 0.0050 | 0.0024 | 0.0752 |
| 201 | 0.0049 | 0.0023 | 0.0736 |
| 10205 | 0.0049 | 0.0023 | 0.0735 |
| 1100 | 0.0049 | 0.0023 | 0.0726 |
| 1200 | 0.0048 | 0.0023 | 0.0720 |

Figure 4: Tabular display at the county level for Polk County, Iowa, showing the modeled 1990 lead concentrations in $\mu g/m^3$. The data is ordered from highest to lowest. The interval (*Lower Bound, Upper Bound*) represents the 90% confidence interval for these estimates. Control buttons for navigation (top) and links to textual pages (left) are visible as well.

# References


Caldwell, J. C., Woodruff, T. J., Morello–Frosch, R. & Axelrad, D. A. (1998), 'Application of Health Information to Hazardous Air Pollutants Modeled in EPA's Cumulative Exposure Project', *Toxicology and Industrial Health* **14**(3), 429–454.

Carr, D. B., Olsen, A. R., Courbois, J. P., Pierson, S. M. & Carr, D. A. (1998), 'Linked Micromap Plots: Named and Described', *Statistical Computing and Statistical Graphics Newsletter* **9**(1), 24–32.

Carr, D. B., Olsen, A. R., Pierson, S. M. & Courbois, J. P. (1998), 'Boxplot Variations in a Spatial Context: An Omernik Ecoregion and Weather Example', *Statistical Computing and Statistical Graphics Newsletter* **9**(2), 4–13.

Carr, D. B., Olsen, A. R., Pierson, S. M. & Courbois, J. P. (1999), Using Linked Micromap Plots to Characterize Omernik Ecoregions, *in* 'Data Mining and Knowledge Discovery', Kluwer. In Press.

Carr, D. B. & Pierson, S. M. (1996), 'Emphasizing Statistical Summaries and Showing Spatial Context with Micromaps', *Statistical Computing and Statistical Graphics Newsletter* **7**(3), 16–23.

Carr, D. B., Valliant, R. & Rope, D. (1996), 'Plot Interpretation and Information Webs: A Time–Series Example from the Bureau of Labor Statistics', *Statistical Computing and Statistical Graphics Newsletter* **7**(2), 19–26.

Rosenbaum, A. S., Axelrad, D. A., Woodruff, T. J., Wei, Y.-H., Ligocki, M. P. & Cohen, J. P. (1999), 'National Estimates of Outdoor Air Toxics Concentrations', *Journal of the Air and Waste Management Association*. In Press.

Symanzik, J., Axelrad, D. A., Carr, D. B., Wang, J., Wong, D. & Woodruff, T. J. (1999), HAPs, Micromaps and GPL — Visualization of Geographically Referenced Statistical Summaries on the World Wide Web, *in* 'Annual Proceedings (ACSM–WFPS–PLSO–LSAW 1999 Conference CD)', American Congress on Surveying and Mapping.

Symanzik, J., Wong, D., Wang, J., Carr, D. B., Woodruff, T. J. & Axelrad, D. A. (1999), Web–based Access and Visualization of Hazardous Air Pollutants, *in* 'Proceedings GIS in Public Health 1998'. Forthcoming.

Woodruff, T. J., Axelrad, D. A., Caldwell, J. C., Morello–Frosch, R. & Rosenbaum, A. S. (1998), 'Public Health Implications of 1990 Air Toxics Concentrations Across the United States', *Environmental Health Perspectives* **106**(5), 245–251.