



EVERY
ALGORITHM
HAS **ART** IN IT

Treemap Art Project

By Ben Shneiderman

Visit Exhibitions @ www.cpnas.org



This exhibit is a project of the
Cultural Programs of the National Academy of Sciences
www.cpnas.org

Printing was paid for by the
University of Maryland Institute for Advanced Computer Studies (Amitabh Varshney, Director)
www.umiacs.umd.edu

Booklet design by Minhaz Rahman Kazi

Exhibit on view at the
National Academies Keck Center, 500 Fifth Street, NW, Washington, DC 20001
October 16, 2014 through April 15, 2015



Introduction

JD Talasek

JTalasek@nas.edu, @jtalasek

Director, Cultural Programs
of the National Academy of Sciences

In his book, *“Visual Complexity: Mapping Patterns of Information”*, Manuel Lima coins the term *networkism* which he defines as “a small but growing artistic trend, characterized by the portrayal of figurative graph structures- illustrations of network topologies revealing convoluted patterns of nodes and links.” Explaining networkism further, Lima reminds us that the domains of art and science are highly intertwined and that complexity science is a new source of inspiration for artists and designers as well as scientists and engineers. He states that this movement is equally motivated by the unveiling of new knowledge domains as it is by the desire for the representation of complex systems.

The name Ben Shneiderman has become intrinsically linked with current advancements in the representation of complex systems. Dr. Shneiderman is a Distinguished University Professor in the Department of Computer Science, Founding Director (1983-2000) of the Human-Computer Interaction Laboratory, and a member of the Institute for Advanced Computer Studies at the University of Maryland, College Park. In addition to pioneering the highlighted textual link in 1983, his work in information visualization led to the development of treemaps (a method of visualizing hierarchical,

tree-structured data as a set of nested rectangles) which has had a rippling impact on systems of data visualization since they were first conceived in the 1990s.

True innovation, by definition, never rests on accepted practices but continues to investigate by finding new perspectives. In this spirit, Shneiderman has created a series of prints that turn our perception of treemaps on its head – an effort that resonates with Lima’s idea of networkism. In the exhibition, *Every AlgoRiThm has ART in it: Treemap Art Project*, Shneiderman strips his treemaps of the text labels to allow the viewer to consider their aesthetic properties thus laying bare the fundamental property that makes data visualization effective. That is to say that the human mind processes information differently when it is organized visually. In so doing Shneiderman seems to daringly cross disciplinary boundaries to wear the hat of the artist – something that has garnered him some criticism from colleagues. Yet at the heart of the creation of treemaps is the need to investigate and promote the human interface with technology and information. In this sense, exploring how aesthetic meaning is created and conveyed– a yet relatively unexplored terrain–is a logical extension of Shneiderman’s work.

The office of Cultural Programs of the National Academy of Sciences’ mission is to explore the intersection between science, technology, art and culture. The exhibition of Shneiderman’s collection of prints creates a platform for considering this area between epistemologies and the

potential for cultural impact. How is meaning conveyed through the display of art? How can the human interaction with data visualization be better understood through the engagement of art processes? Rather than providing answers to these questions, the inventor of treemaps blurs the boundaries between disciplines and thus opens the door for a dialogue that considers a different perspective.

The year 2014 is significant because it marks the 50th anniversary of the establishment of the National Academy of Engineering. CPNAS is pleased to celebrate this landmark occasion by hosting an exhibit at the Keck Center in Washington, DC of thought-provoking work by Shneiderman, a member of the NAE. We are grateful to Shneiderman and the University of Maryland for donating a copy of this print series for exhibit and thus opening up a discourse on the relationship between technology, creativity, and human interaction.



From Algorithms to Art: Treemap Art Project History

Ben Shneiderman

ben@cs.umd.edu, @benbendc

Distinguished University Professor,
University of Maryland

Although I conceived treemaps for purely functional purposes (understanding the allocation of space on a hard drive), I was always aware that there were aesthetic choices in making appealing treemaps, such as:

1. layout design (slice-and-dice, squarified, ordered, strip, etc.),
2. color palette (muted, bold, sequential, divergent, rainbow, etc.),
3. aspect ratio of the entire image (square, golden ratio, wide, tall, etc.) and
4. prominence of borders for each region, each hierarchy level, and the surrounding box.

In addition, certain treemaps are inherently interesting because of the data displayed or patterns revealed. The aesthetic choices appeal to different viewers and the data displayed also triggers interest by different communities. So even though treemaps are constrained by algorithmic rules, the aesthetic and data choices leave much room for creative exploration.

The absolute size as hung on a wall or seen on a screen also changes the impact on viewers, as does the texture or reflectivity of the medium. Other choices that influence the impact are the number of regions and number of levels in the

hierarchy. More complex visual presentations suggest the richness and complexity of the data, while simpler examples convey clarity and simplicity. My awareness of these issues was enriched by my interest in the 1960s OP-ART movement through exhibits I saw at the Museum of Modern Art while growing up in New York.

I believe that some topics yield data that have greater vitality and interest for viewers: maybe sports data, political elections, Hollywood films, popular music, nature, pets, health, science, etc. For example, would the carbon emissions of countries around the world produce a treemap that could be laid out by the slice-and-dice or squarified algorithms and then colored in a way that would engage the eye and mind? Would the batting averages and runs scored for all professional baseball players organized by teams interest sports fans? Some data might be personal such as weight loss-gain or personal income-expenditures, while others might be corporate such as number of employees per division or profit/loss over 10 years.

Colored rectangular regions have been a popular theme in 20th century art, most notably in the work of Piet Mondrian, whose work was often suggested to have close affinity with treemaps. Not all his designs are treemaps, but many are. His choice of colors, aspect ratios, and layout are distinctive, so simulating them with a treemap is not as trivial as you might think. Gene Davis's large horizontal paintings with vertical stripes of many colors were more easily generated with slide-and-dice treemap algorithms. The rectangles in Josef Albers "Homage to the Square" or Mark Rothko's imposing paintings are not treemaps, but these works influenced our artistic explorations. Other modern artists such as Paul Klee, Kenneth Noland,

Barnett Newman, and Hans Hofmann gave further provocations for the images in this collection.

The treemaps were developed at the University of Maryland Human-Computer Interaction Lab starting in 1990 and have a significant history:
<http://www.cs.umd.edu/hcil/treemap-history/index.shtml>
 Our software tool to produce treemaps is free to download and use: <http://www.cs.umd.edu/hcil/treemap> .

Several interesting treemaps are available to explore at the Hive Group website (www.hivegroup.com), which has licensed our software. Their examples include iTunes, nutrition, earthquakes, politics, etc. The treemaps generated with other commercial tools such as Macrofocus TreeMap, Panopticon (acquired by Datawatch), Spotfire, and Tableau show a variety of styles as do free tools such as SequoiaView, WinDirStat, ManyEyes, and d3. Dozens of interesting variations have emerged such as the Voronoi treemaps and circular treemaps. Wikipedia offers further perspectives on the topic as well as links to treemapping software.

My appreciation to Minhaz Rahman Kazi for helping me realize this idea and contributing much to the aesthetics of the results. Thanks also for the financial support for printing from the University of Maryland Institute for Advanced Computer Studies (Amitabh Varshney, Director) and for permission to hang the exhibit on campus from the Department of Computer Science (Samir Khuller, Chair).

<http://treemapart.wordpress.com>



The Evocative Charm of Treemaps

Manuel Lima

m@mslima.com, @mslima

Designer, author, researcher, and lecturer based in New York. Founder of VisualComplexity.com. Design Lead @ Codecademy.

History is filled with symbiotic exchanges between art and science, and more specifically between art and cartography. David Woodward, in his *“Art and Cartography”*, published in 1987, describes in detail how numerous artists were influenced by cartography, and how maps themselves, as witnessed in the work of numerous Dutch painters, were hung on walls as pieces of art.

Ben Shneiderman’s eminent work in creating a tiling mechanism to visualize a tree structure by means of nested rectangles was a pivotal contribution to the infant domain of data visualization in the early 1990s. The resourceful model devised by Shneiderman, and later identified as a treemap, quickly became one of the most widespread methods for visualizing hierarchy, primarily due to its efficient and adaptable use of space. Perhaps most importantly, the pioneering nesting algorithm opened the door to a huge array of new visualization models and approaches.

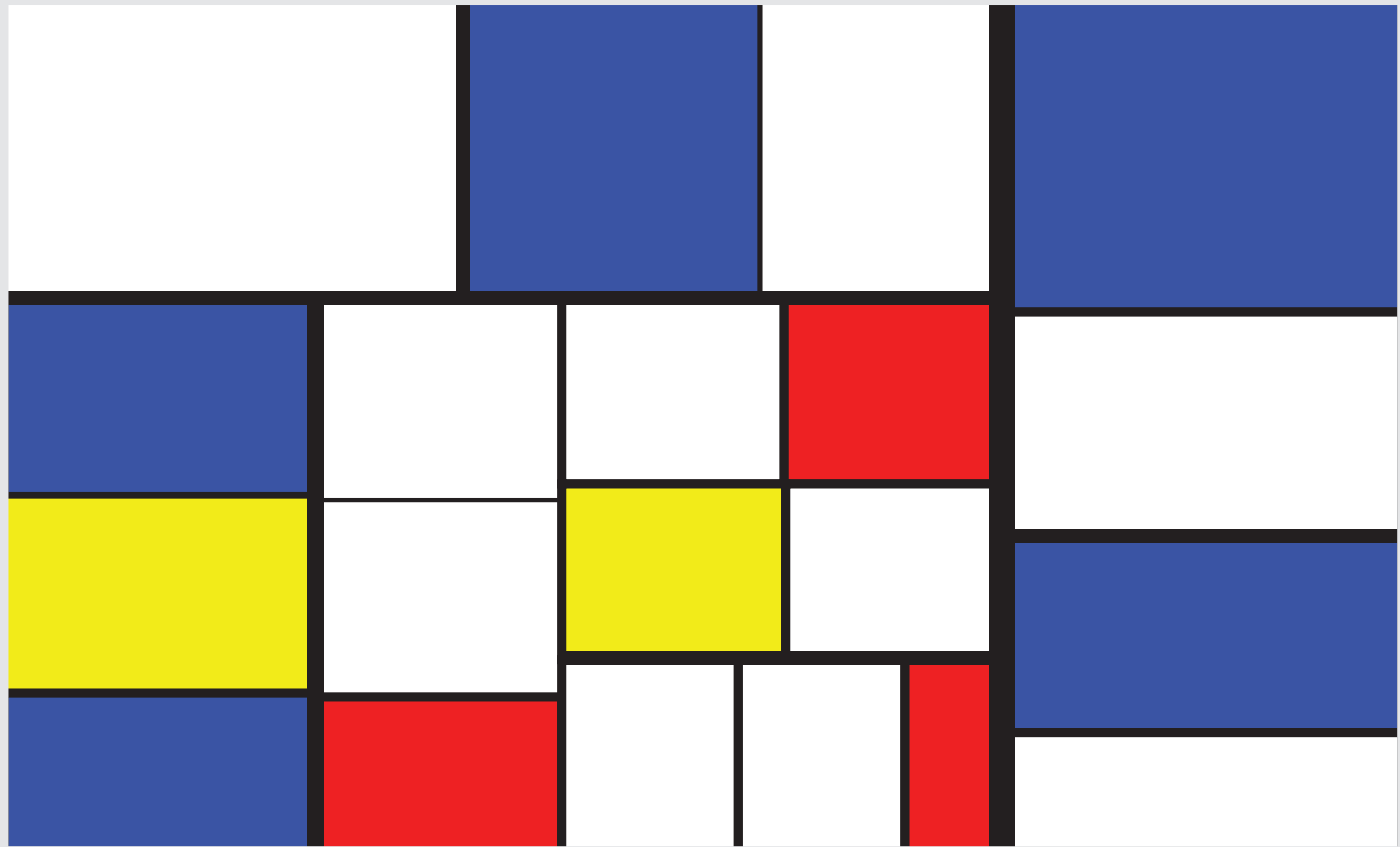
The original rectangular treemap, like many of its subsequent variations, such as the Voronoi treemap and the circular treemap, is not just effective at mapping tree structures. All models without exception share a remarkably expressive

quality. Their elaborate cell layout and colorful display, which frequently looks like intricate stained glass, is as visually captivating as it is pragmatically functional. Treemaps not only epitomize the recent growth of data visualization but also its occasional, enthralling aesthetic charm.

Shneiderman’s exhibit “Every AlgoRiThm has ART in it: Treemap Art Project” gathers twelve compelling pieces of evidence for the treemap’s pictorial force. Some of the pieces intentionally pursue the lure of other artists’ style, such as *The Singing Mondrian*, while others like *Dazzling Talks* have their own unique, magnetic charisma. The combination of colors, the scaling and subdivision of cells, the orthogonality and confluence of its multiple lines, create an enticing mesh that resembles at times the essentials of form and color found in neoplasticism, orphism, minimalism, and abstract expressionism. All pieces can immediately draw you in, as you marvel at the details of its intricate matrix. Once you realize each piece is depicting real data attributes from a specific subject, like international flight routes or the individual scoring of basketball players, then its allure becomes stronger, more powerful, and surprisingly more meaningful.

With this influential exhibit, Ben Shneiderman completes the full circle of the treemap’s adaptability, first as a revolutionary utilitarian tool to map large hierarchies, and now as a powerful visual element able to entice and evoke emotions. Most importantly, the exhibit successfully bridges the gap between art and science, by showing they do not represent two sides of the same coin, but one single side, or should we say algorithm, seen through different lenses.

The Singing Mondrian

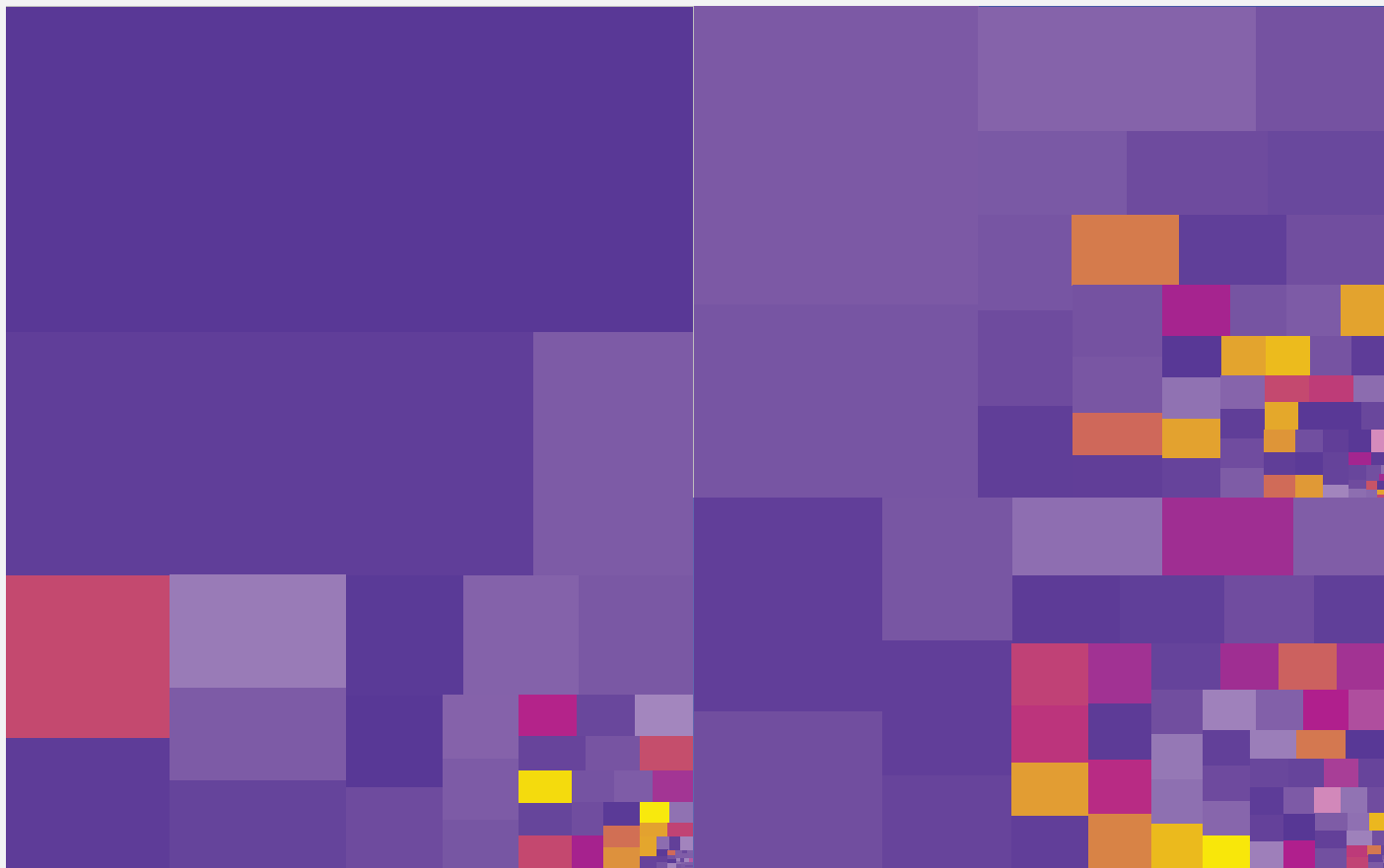


This visualization was inspired by Piet Mondrian's compositions with red, yellow, and blue colors. The left one is titled "Composition C (no.III), with Red, Yellow and Blue."

This data set contains artist data from Last.fm. On their 10 year anniversary, Last.fm published a list of top 100 artists based on their popularity as per user data. From that list, we have taken the top 20 artists, the total number of times their

tracks were 'scrobbled' or played, and the number of unique listeners for each artists. We also identified broad genres of the artists based on their most popular tags.

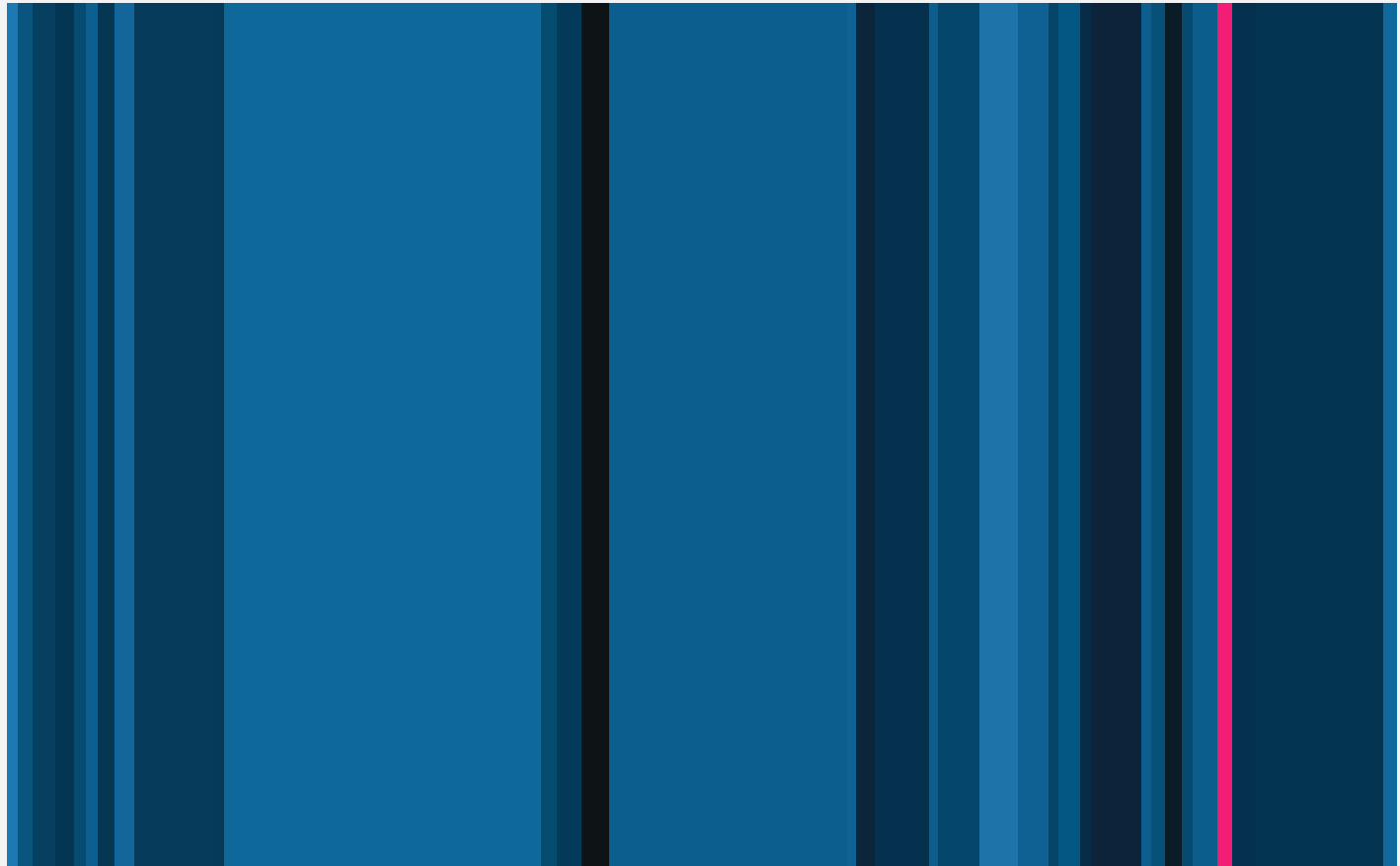
The boxes represent individual artists where the size of the box is the number of times their tracks were played while the color represents the genre of the artist – rock is white, alternative is blue, pop is yellow and hip-hop is red.



We Are the World

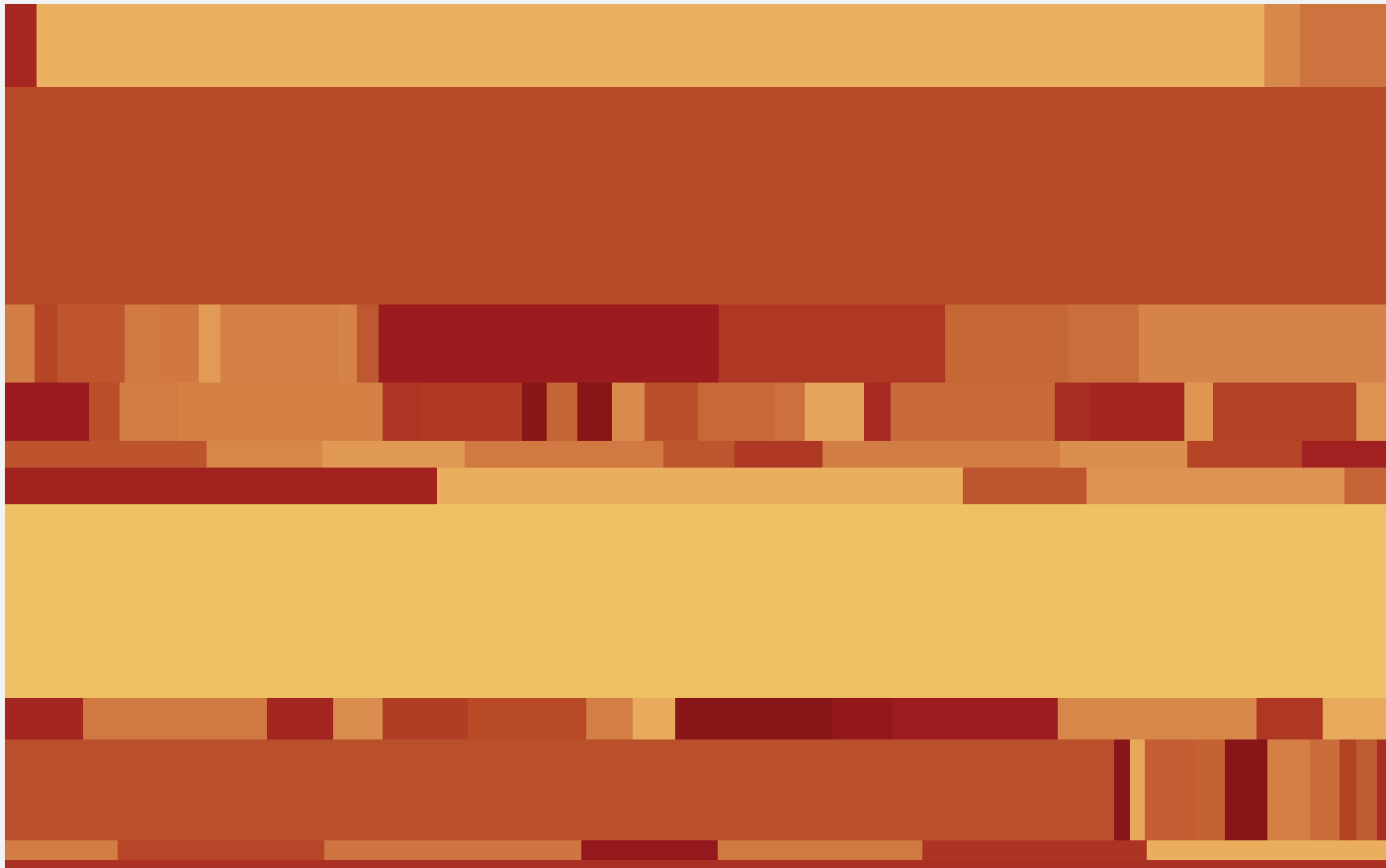
The dataset was collected from the World Bank's website. The size of the boxes represent the urban population counts of countries. They are colored according to the urban population growth percentage - negative values are colored magenta. Positive values start at light purple for lower values and then gradually shift to dark purple and then to yellow for higher values.

Urban Blues



The dataset was collected from the World Bank's website. This contains the urban population counts and annual urban population growth percentages for all countries for the year 2010. We filtered the data to include countries with urban populations of 20 million or more. This resulted in 36 records. The box sizes represent urban population count for an individual country.

The wide horizontal layout and vertical stripes are inspired by Gene Davis's style. Countries with negative urban population growth are colored in pink – here we find only Ukraine under this criteria. Other countries are colored in a black to blue scale where black represents zero urban population growth and blue represents the highest among these countries (6.25%).



The dataset was collected from the World Bank's website. In this strip tree map, the box sizes represent population densities of people per square kilometer of land area. The colors represent annual urban population growth percentages. Negative values are colored in lighter brown. Positive values are colored from yellowish brown (0%) to dark brown (6.25%).

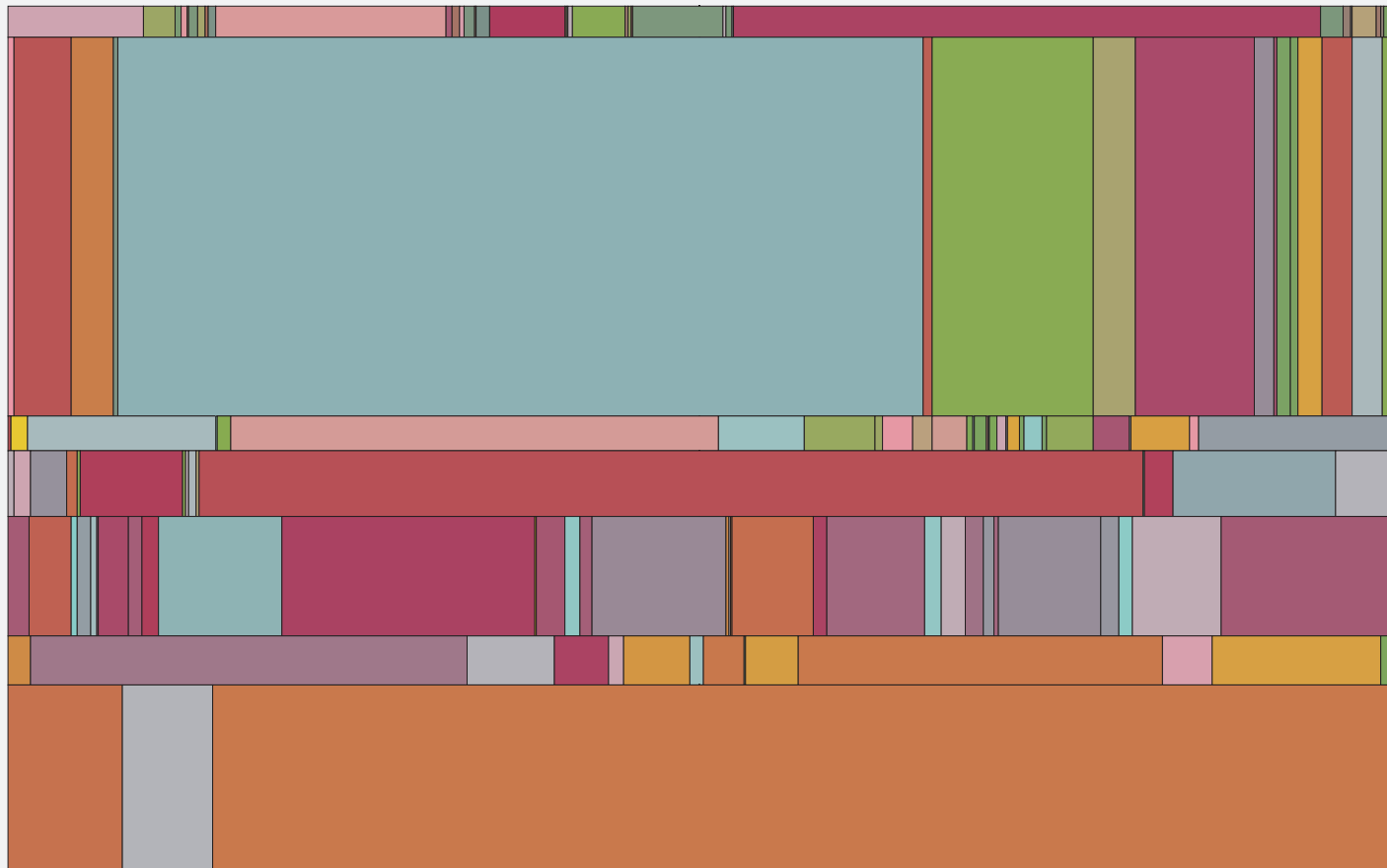
The data is filtered to show countries with population densities of 100 or more people per square kilometer of land area (91 records).

The Big Urbans



This urban population dataset for the year 2010 was collected from the World Bank's website. The visualization shows countries with urban populations of 23.3 million or more. The box sizes are urban population counts. The colors represent population density (people per square kilometer of land area) with the highest being greenish yellow (Bangladesh: 1159) and the lowest being purple (Canada: 4).

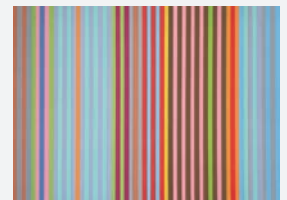
The coloring was done in 4 equally dense bins and using linear scale. The color palette was inspired by Josef Albers' works while the composition and arrangement was inspired by Gene Davis' works. The diversity of city life is conveyed by the lively colors.



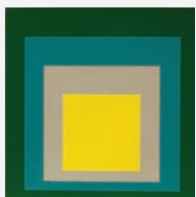
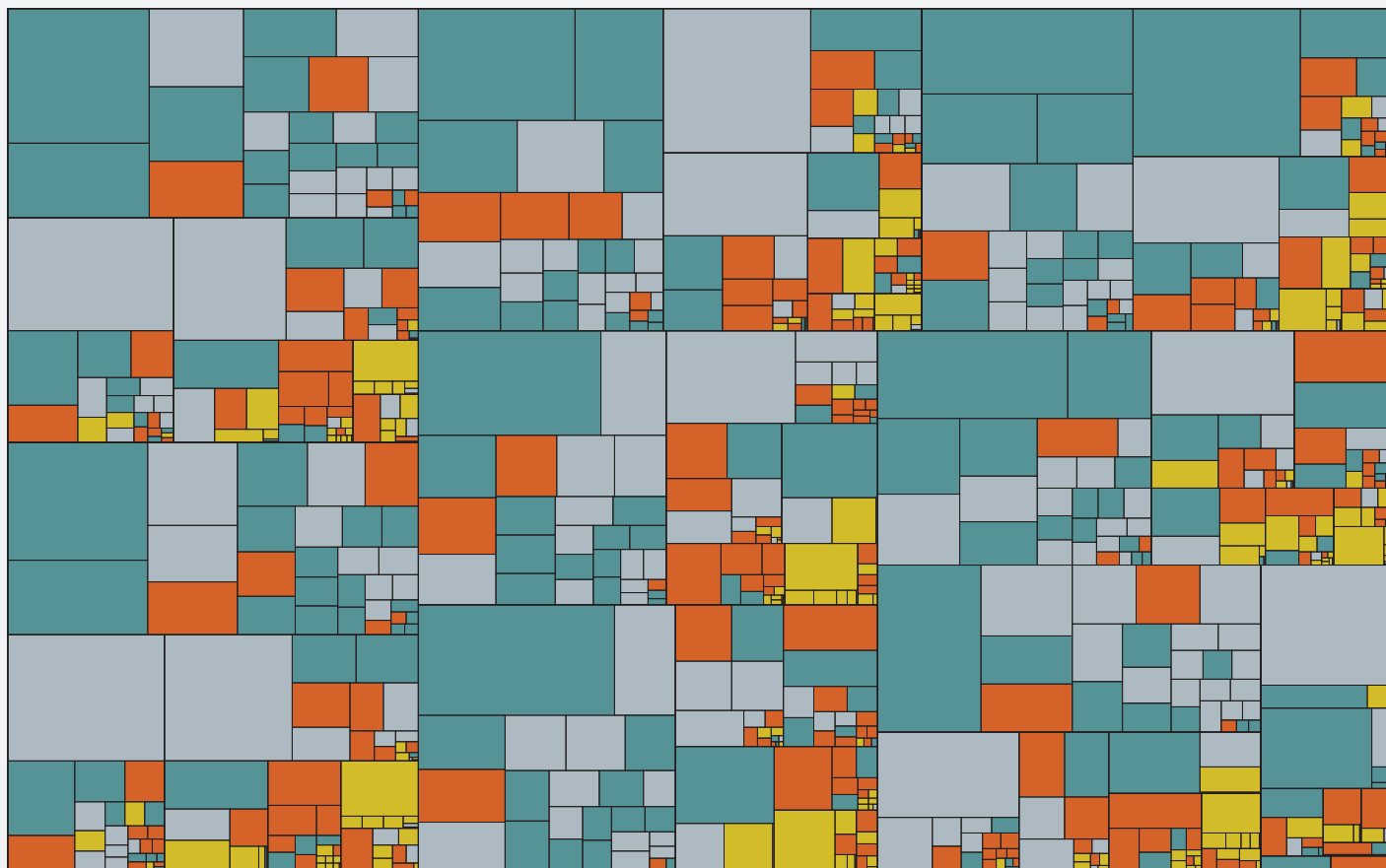
The muted color palette was inspired by Gene Davis' Apricot Ripple (right). The dataset was collected from the US Energy Information Administration. It contains 2010 data on total CO2 emissions (million metric tons) as well as per capita CO2 emissions (metric tons of CO2 per person) for all countries categorized into 7 continents. There are 224 records in total.

The size of the boxes represents CO2 emissions for countries while the color varies by per capita CO2 emissions. The coloring was done via categorizing countries into 6 equally dense bins for per capita CO2 emissions.

Out There in the Air



Blooming Businesses



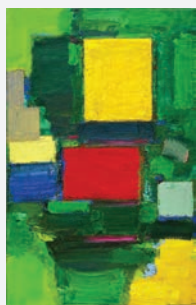
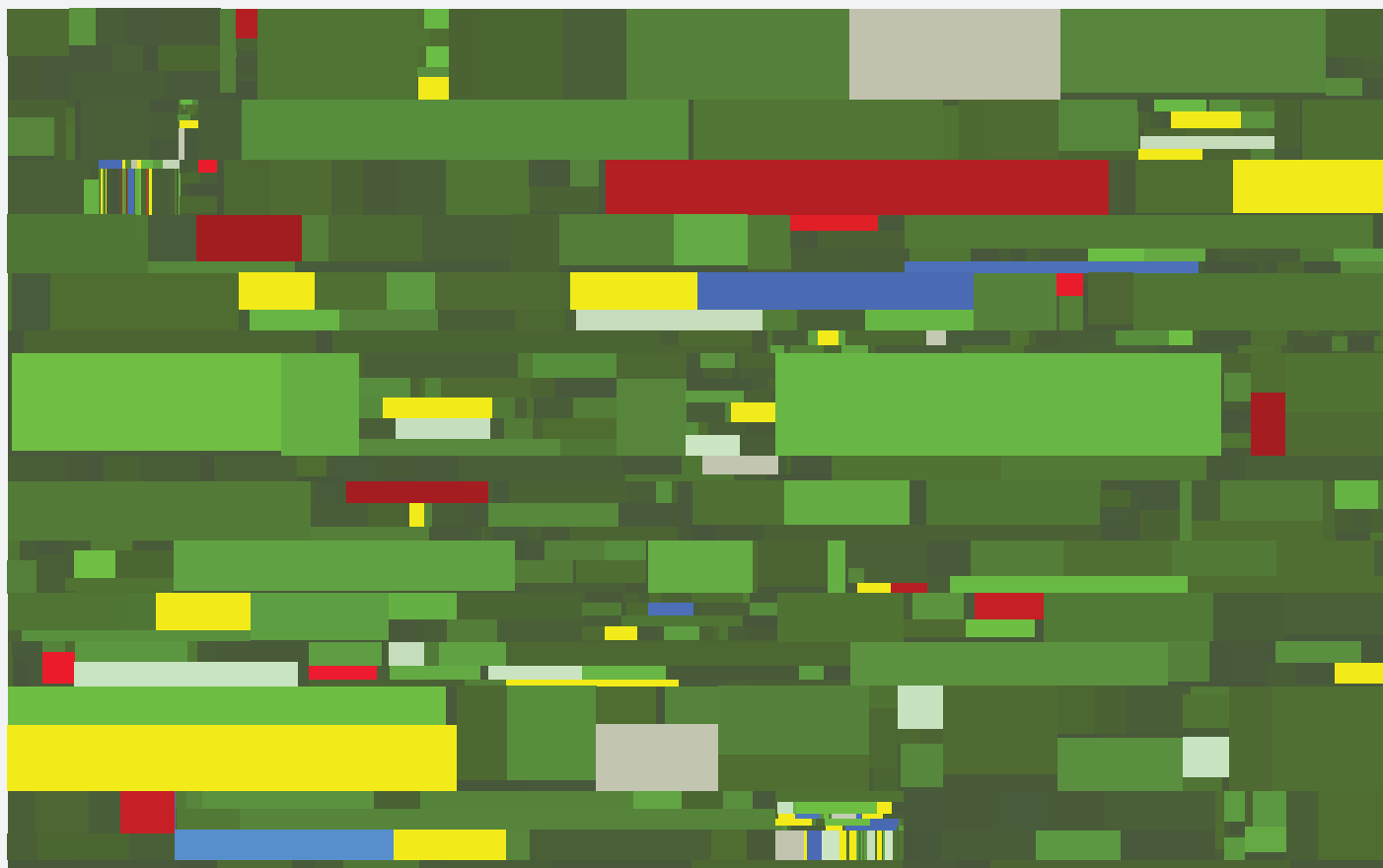
The color theme was inspired by works from Josef Albers' Homage to the Square series, including the one on the left (Homage to the Square: Park). This dataset was collected from the World Bank's Doing Business website. The dataset contains new firm count (the number of newly registered corporations during the calendar year) and new firm density (the number of newly registered limited liability companies per 1,000 working-age people) for all countries for individual years from 2004 to 2011.

In the visualization, the box size represents new firm count and the colors represent various degrees of firm density - grey is lowest, then yellow, orange and teal is highest. The countries are grouped into major economic groups. The countries are again subdivided into different years. There are 864 records in this dataset.



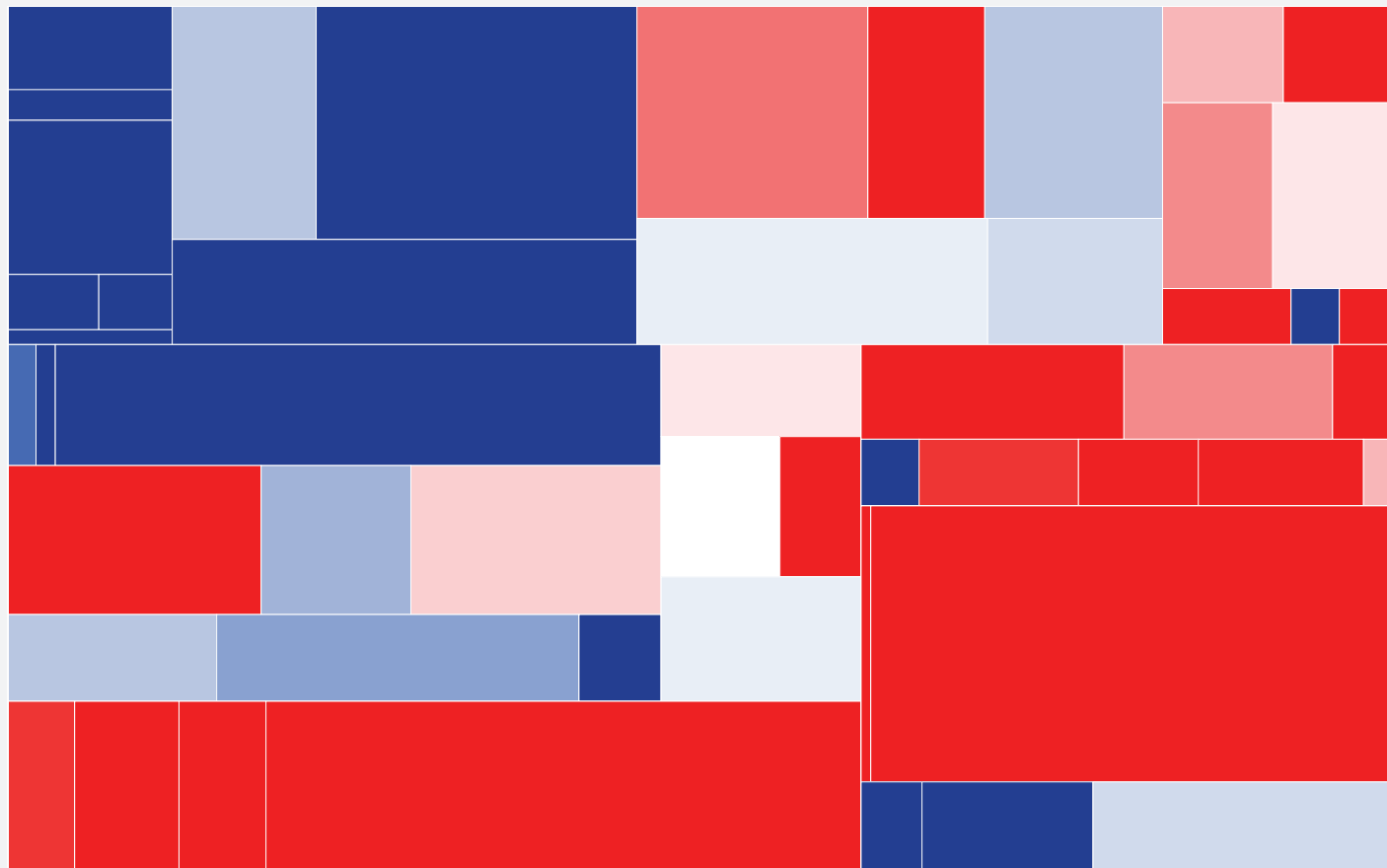
This visualization shows statistics about certain TED talks. The dataset was compiled by Sebastian Wernicke for his TED talk, “Lies, Damned Lies and Statistics.” Each of the boxes represents the engagement score for a certain TED talk. The colors depend on the total number of del.icio.us bookmarks for that TED talk. The coloring was done in 8 equally dense bins with pink being highest and red being lowest. The colors here try to capture the variety and charismatic excellence of the TED talks.

Frequent Flyers



This dataset was collected from openflights.org via visualizing.org. The boxes represent the ratio of international to domestic flights for individual airports. Larger sizes indicate more international flights. The different colors represent the variation in total number of routes served by that airport. Greens represent less busy airports while the bright hues show ones with a large number of routes.

The color theme was inspired by Hans Hofmann's *The Gate* (left).

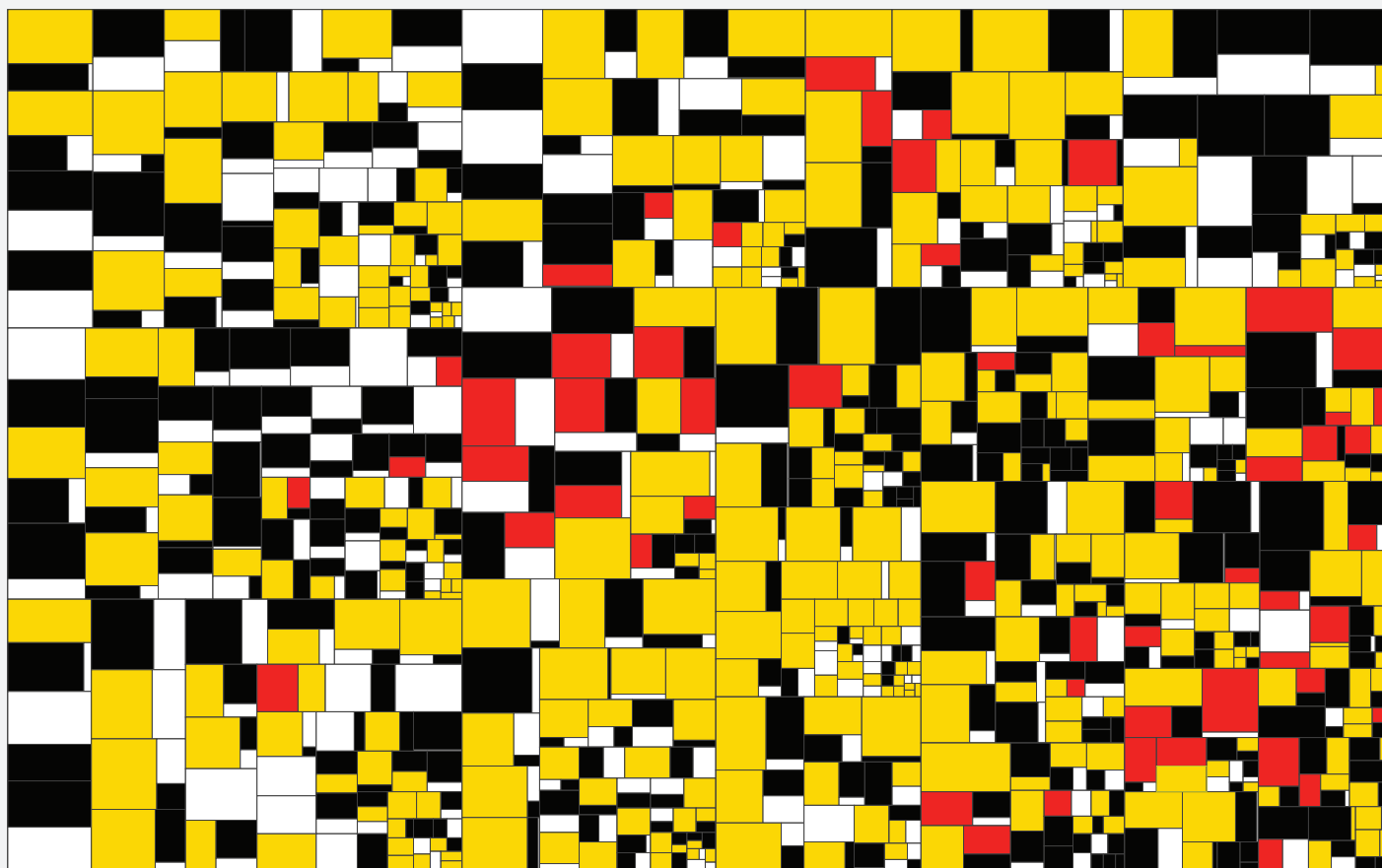


The Men and Women of Red, White and Blue

For this visualization, US population data for 2012 was collected from Census.gov. This dataset contains statewide total population and the percentage of 18+ population for the year 2012. The box size represents the total population of a state and the color varies as per the percentage of 18+ population.

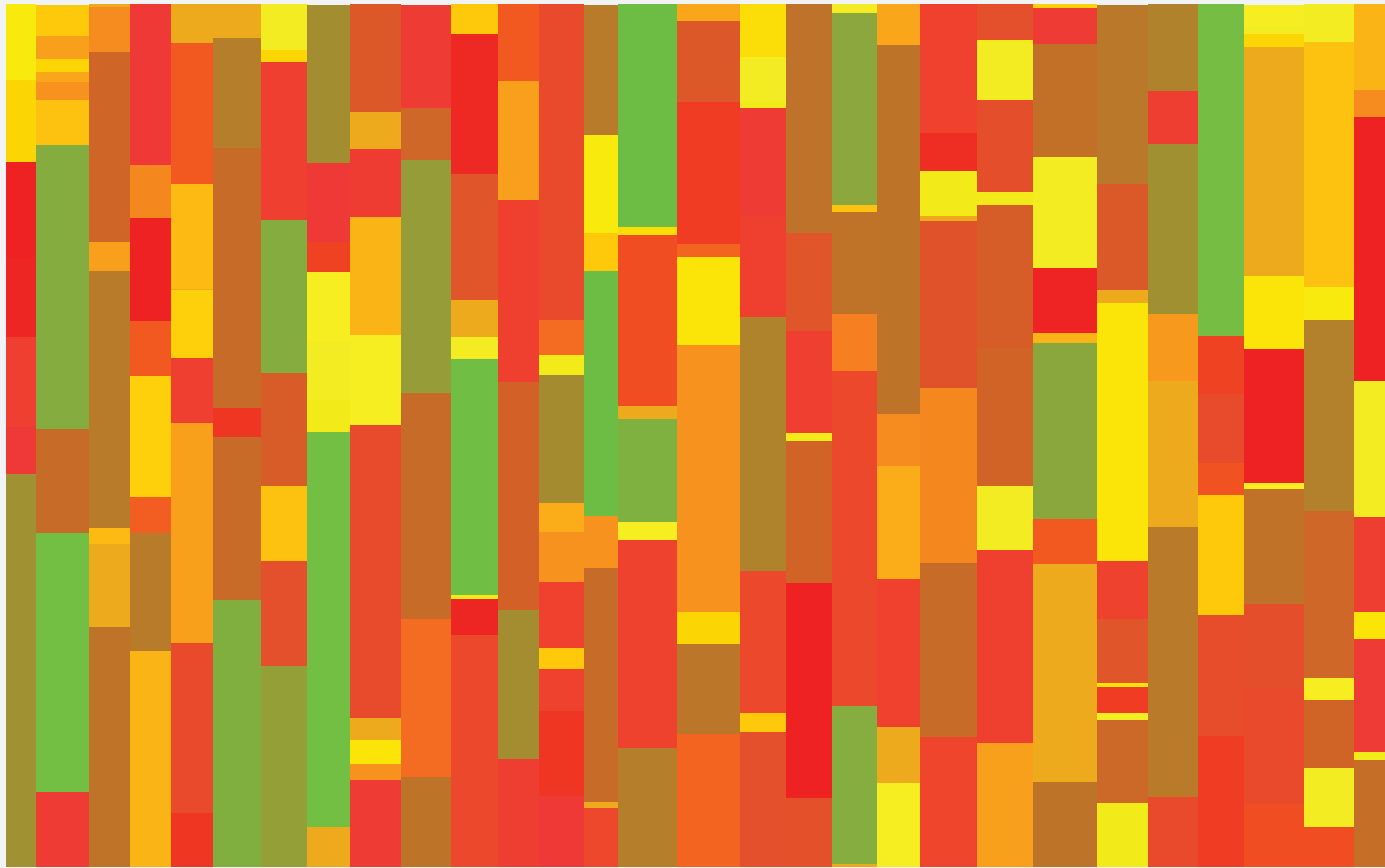
The red, white and blue colors were chosen to represent the US, with gradients to allow more accurate representations of the data. The color binning was adjusted to bring darker blue regions to the upper left.

Green Terps



This dataset was collected from the Maryland Open Data website. The dataset contains historical data on total grant amount awarded to different clean energy projects. The projects are grouped first by county and then by zip code. They are aggregated into four different technology types and were colored accordingly: solar hot water (white), solar photovoltaic (yellow), geothermal (black) and wind (red).

The colors were chosen to represent the Maryland flag. Each box represents the total amount grant awarded for a certain technology in a certain zip code. The items were filtered to show only values below \$40,000 for grant award amounts.



The Better Players

This dataset of 441 players in the National Basketball Association is organized into 29 teams. For each player, area indicates number of points scored and color indicates number of personal fouls. Darker, redder colors indicate more personal fouls. There is clearly a wide range of scoring and personal fouling, but the large green rectangles highlight those who manage to score many points while avoiding personal fouling.



BEN SHNEIDERMAN (<http://www.cs.umd.edu/~ben>) is a Distinguished University Professor in the Department of Computer Science, Founding Director (1983–2000) of the Human-Computer Interaction Laboratory (<http://www.cs.umd.edu/hcil/>), and a Member of the Institute for Advanced Computer Studies (UMIACS) at the University of Maryland. He is a Fellow of the AAAS, ACM, and IEEE, and a Member of the National Academy of Engineering, in recognition of his pioneering contributions to human-computer interaction and information visualization. His contributions include the direct manipulation concept, clickable highlighted web-links, touchscreen keyboards, dynamic query sliders for Spotfire, development of treemaps, novel network visualizations for NodeXL, and temporal event sequence analysis for electronic health records.

Ben is the co-author with Catherine Plaisant of *“Designing the User Interface: Strategies for Effective Human-Computer Interaction”* (5th ed., 2010) <http://www.awl.com/DTUI/> . With Stu Card and Jock Mackinlay, he co-authored *“Readings in Information Visualization: Using Vision to Think”* (1999). His book *“Leonardo’s Laptop”* appeared in October 2002 (MIT Press) and won the IEEE book award for Distinguished Literary Contribution. With Derek Hansen and Marc Smith, he co-authored *“Analyzing Social Media Networks with NodeXL”* (www.codeplex.com/nodexl, 2010).

EVERY
ALGORITHM
HAS **ART** IN IT

<http://treemapart.wordpress.com>

CPNAS
CULTURAL PROGRAMS
OF THE
NATIONAL ACADEMY
OF SCIENCES

The National Academy of Sciences (NAS) is a private, nonprofit, self-perpetuating society to which distinguished scholars are elected for their achievements in research, and is dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the Congress in 1863, the NAS has a mandate to advise the federal government on scientific and technical matters. The mission of the office of Cultural Programs of the National Academy of Sciences (CPNAS) is to explore the intersections of art, science, and culture through the presentation of public exhibitions, lectures, and other cultural programs.