PROCESS OF SIMULATING TREE RINGS FOR IMMIGRATION IN THE U.S.

Pedro M. Cruz
Art+Design
Northeastern University
p.cruz@northeastern.edu

Steve Costa
Art+Design
Northeastern University
costa.st@husky.neu.edu

John Wihbey
School of Journalism
Northeastern University
j.wihbey@northeastern.edu

Ruan Chao
School of Art & Design
Zhejiang Sci-Tech Univ.
ruanchao6688@126.com

Avni Ghael
College of Computer and
Information Science
Northeastern University
ghael.a@husky.neu.edu

Felipe Shibuya
Independent Artist
felipe.shibuya@gmail.com

Abstract
This article presents the iterative design process of representing the growth of populations in the United States as tree rings. It explains why this metaphor was chosen, and how the iterative process went through several ideas and implementations in order to make the metaphor more visible. For that, several algorithmic approaches are discussed as their graphical results are presented.

Authors Keywords
Information visualization, data visualization, information design, data art, graphic design, figurative metaphors

Introduction
Immigration has been present in the United States since its inception and thus is central to the identity of the country. Waves of immigrants from around the globe have shaped the country's population. The objective of this work is to make visible the fact that immigration had a pronounced impact in the population growth of the U.S. For this, instead of showing immigrants as a percentage of the whole population, it shows how much immigration has increased in contrast with the number of national newborns. In other words, it shows the number of people added to the population in a given time period (i.e. the growth of population). These people can either be born in the U.S. (i.e. natural-born), or they can be immigrants who came from outside.

In order to give form to these numbers, the metaphor of tree rings was specifically chosen for this theme. Tree rings are the circular marks seen in a tree's trunk when it is cut. They are made of cells that multiplied and died. The growth rate of these cells varies throughout the seasons, resulting in the visible marks that are the tree rings. They generally mark one-year periods of growth. The thickness of the ring can tell how much the tree has grown in a year.

Trees can be hundreds, even thousands of years old. The cells grow slowly, and their pattern of growth influences the shape of the tree's trunk. They are the result of a slow process that occurred a long time ago. This idea lends itself to the representation of history itself, as it shows a sequence of events that have left a mark and shaped the present. If cells leave a mark in the tree, so can incoming immigrants and natural-borns be seen as natural contributors to the growth of a trunk that is the U.S. It carries the idea that these marks in the past are immutable and that cannot be erased regardless how you read them. Additionally, it embodies the concept that all cells contributed to the organism's growth and that they are all part of it. The cells and rings in a tree are nature's own way of organizing information, which can serve as a reference for how we can spatially organize immigration data for presentation.

The U.S. is currently formed by fifty states. Each state has grown at different rates and with varying immigration profiles. Some will be larger, some will be smaller, some will have complex shapes that represent waves of immigrants, and others will be perfectly circular due to the absence of immigration. Each state has its own signature and can be characterized individually. The U.S. can then be envisioned as a forest of trees, opening the possibility of presenting more information through portrayals. The visualization of such trees are cross sections of their trunks that reveal the tree rings inside. In fact, when one looks at a set of tree rings, what is presented is a sample of all of the tree's cells. This happens in this context as well, in the sense that the visualized data is a sample of the universe of study.

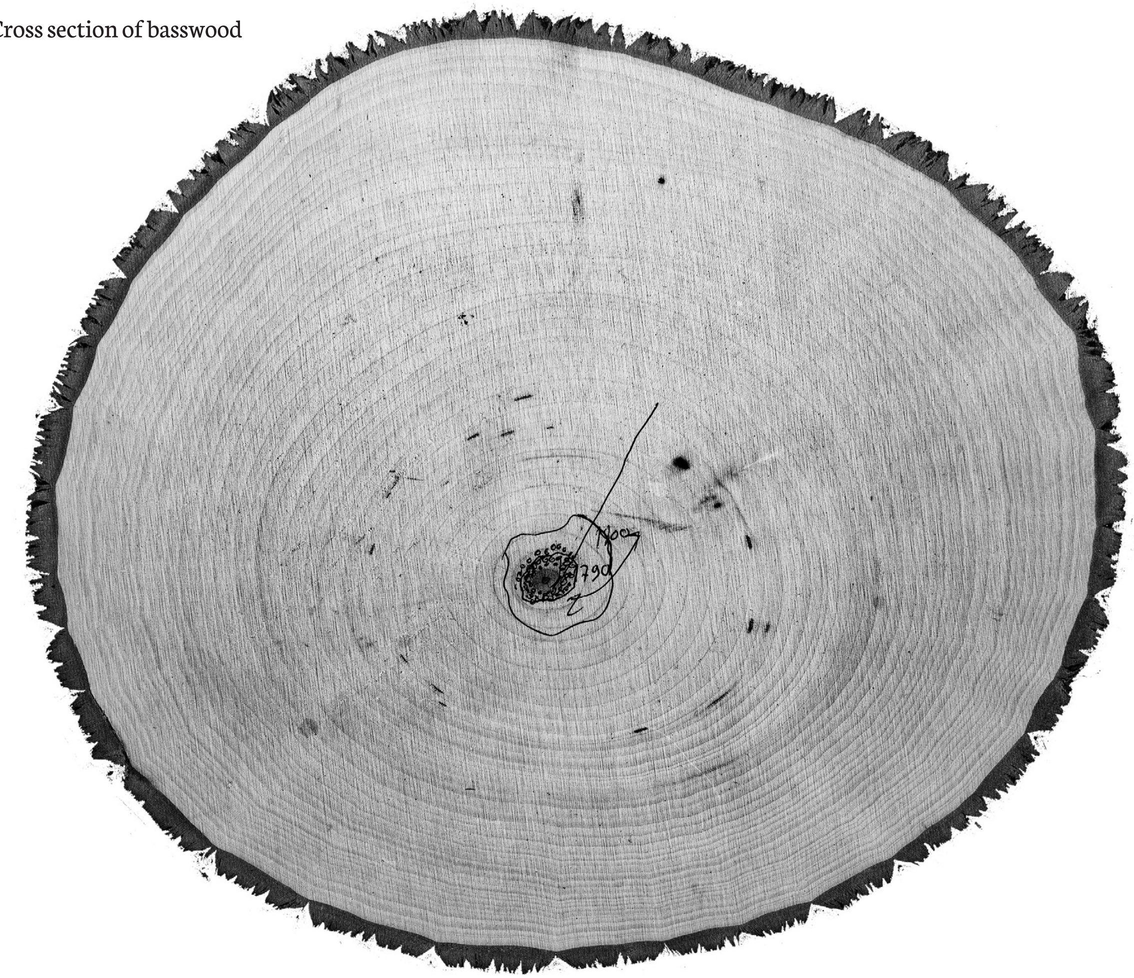
Data and first design decisions

The dataset consists of millions of samples of questionnaires from the U.S. Census made available through IPUMS. We queried the U.S. state of residence, age, and place of origin of each person since 1790. The places of origin have originally 571 denominations. Using this data, we calculated estimates for the number of native-born persons and the number of immigrants who arrived in each decade. This article focuses on the visual and algorithmic design of the visualization and hence does not discuss how these numbers were extracted. Nonetheless, the fact that the dataset describes each person individually inspired the idea to make a cell in a tree represent a specific number of people. Furthermore, given that the temporal resolution is ten years, a tree ring will represent how population grew in a decade. Naturally, for some states, there is no data before a certain year. In these cases the first tree ring seen is the first decade for which data is available.

Before prototyping

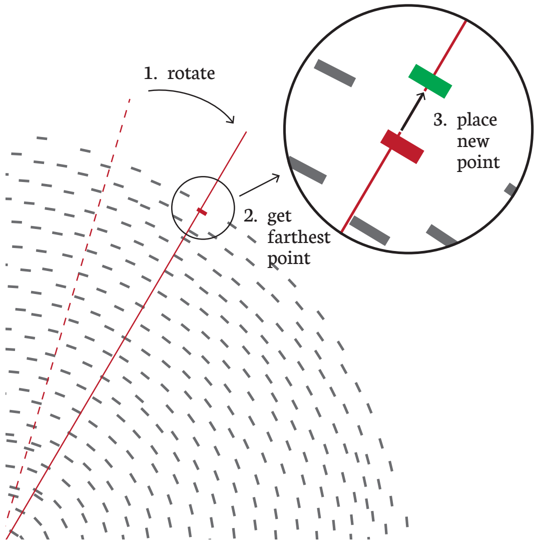
Tree rings come in many shapes and configurations, from the more concentric and circular to the more wiggly. Their shape is a result of environmental conditions during the tree's growth. A cross section of basswood

Cross section of basswood



(Tilia americana) is shown here that served as a reference throughout the process. It can be observed that tree rings can be tenuous, that they vary in thickness, and that they are not perfect circles. In fact most of the rings appear to be concave, but some can be convex as well as clearly seen by the bark's outline.

This article shows the main design iterations of building a set of visualizations of immigration and natural-borns that not only resemble tree rings, but that also follow several principles inherent to the formation of tree rings. The focus is to expose and explain all the iterations and design decisions that were taken until the results embodied a visual and functional resemblance with tree rings. The emphasis is on the iterative process that started delivering promising results, showing that the process is not a black box, and exposing the ideas that eventually met a dead end.



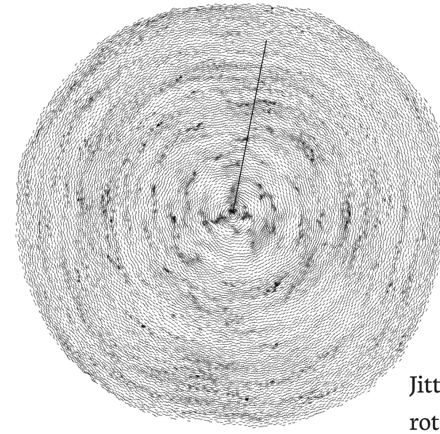
Prototyping an algorithm

The main idea behind the algorithm is, to a certain extent, to simulate how a tree grows. For this, instead of starting with a set of rings or circles and distorting them in order to resemble tree rings, a set of dots are placed in space as if they were the cells of a tree. The rules that dictate such placement should result in a graphical arrangement that as an aggregate resembles tree rings. No two cells can occupy the same physical space. Therefore, the dots also obey this constraint as they can push among each other in order to make space in case they overlap. One property of the algorithm is the minimum spacing between points, which can be fixed or can vary. If two points are closer than the minimum space, they are considered to be overlapping and thus a cascade of pushes is enacted until the minimum space constraint is resolved.

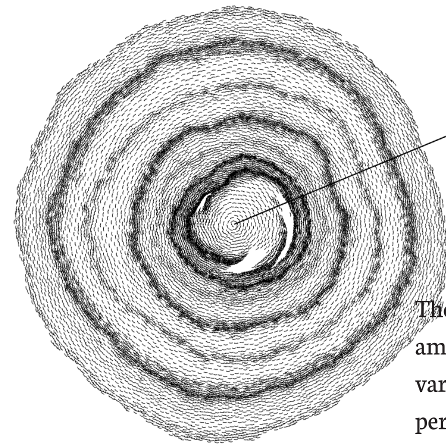
Dots start being placed in the center of the canvas. The dots are not solely represented as a pixel or a small circle. Instead, they are rendered as very small lines. The orientation of these lines follows the perimeter formed by the dots. This gives the impression of continuity, as if the image was made out of curved paths.

The first attempt to implement the abovementioned is based on the rotation of a ray that starts at the center of the canvas. The ray is rotated to a certain angle, and a new dot is placed just beyond the farthest dot it intersects. If noise, or jitter, is added to this rotation angle it can result in an arrangement that is not perfectly circular, but that has an smooth outline. As the ray rotates, if the minimum spacing varies in a periodical fashion, different densities in the concentration of cells make rings more evident.

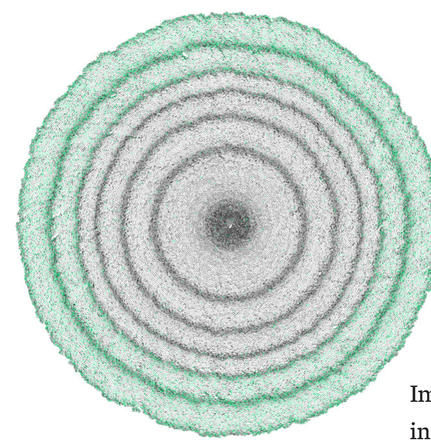
The suitability of this first algorithm was tested first with data from Massachusetts. Each newly placed dot represents a certain number of people *n* that were added to the population. The algorithm queries *n* people from the dataset starting in 1790. A dot is placed and the ray is rotated. It repeats until there are no more people to place from that decade, in which case it advances to the following decade. When immigrants are represented as green dots, diffuse waves of immigration start appearing. Because the order by which data is queried is not guaranteed, the representation does not allow for comparision of the number of immigrants with the number of natural-borns.



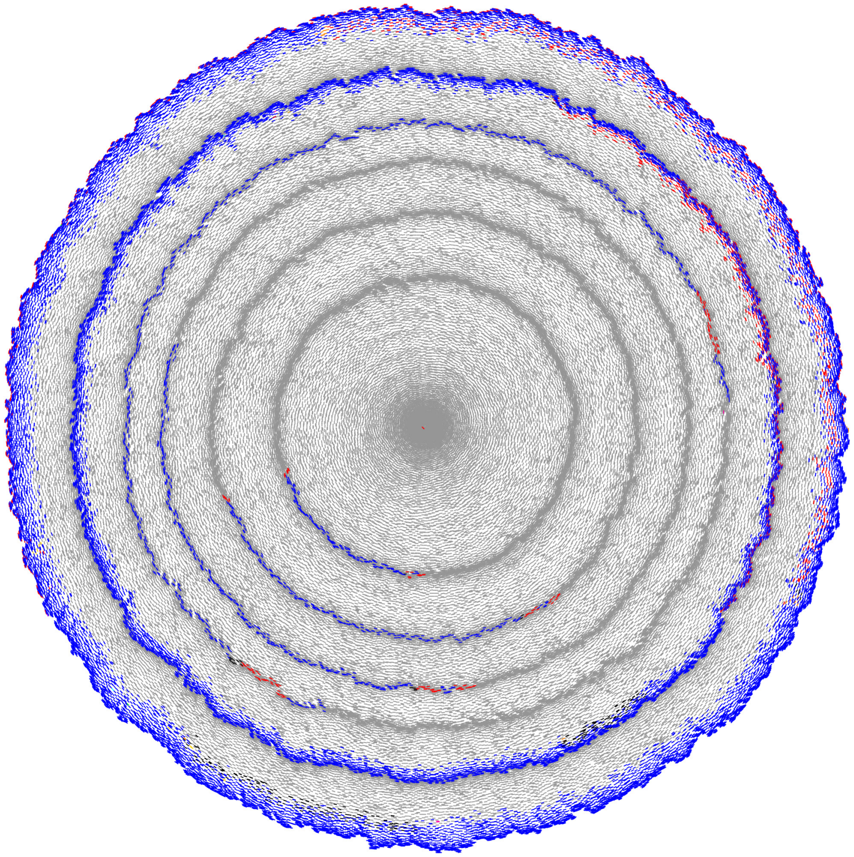
Jittering the rotation angle

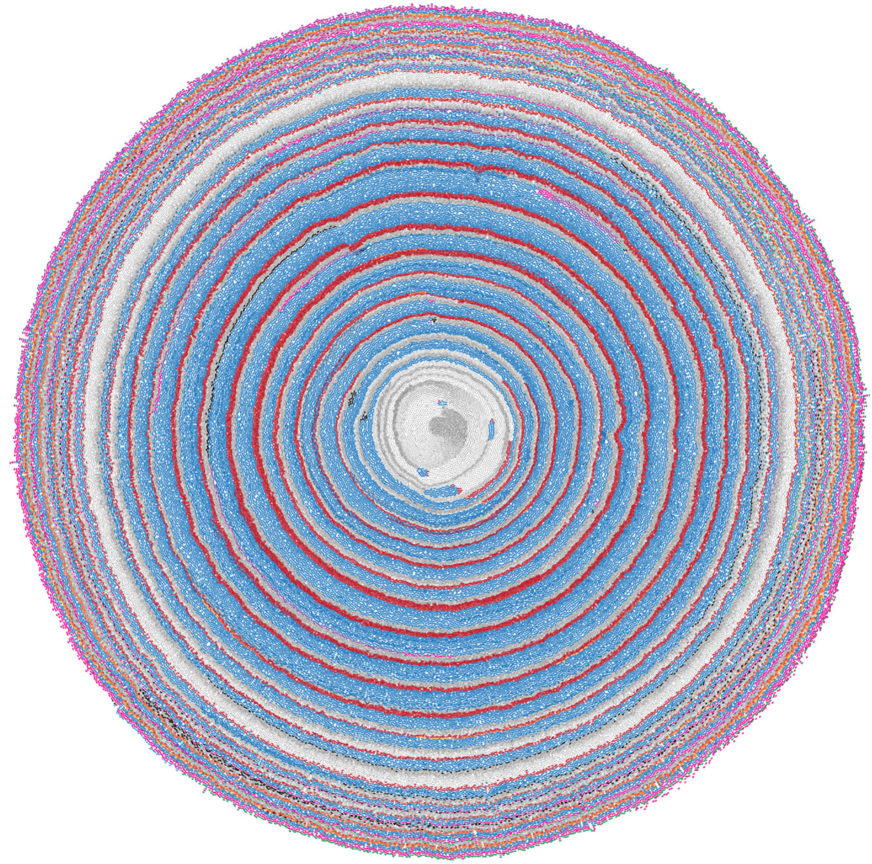


The spacing among points varies and is periodical



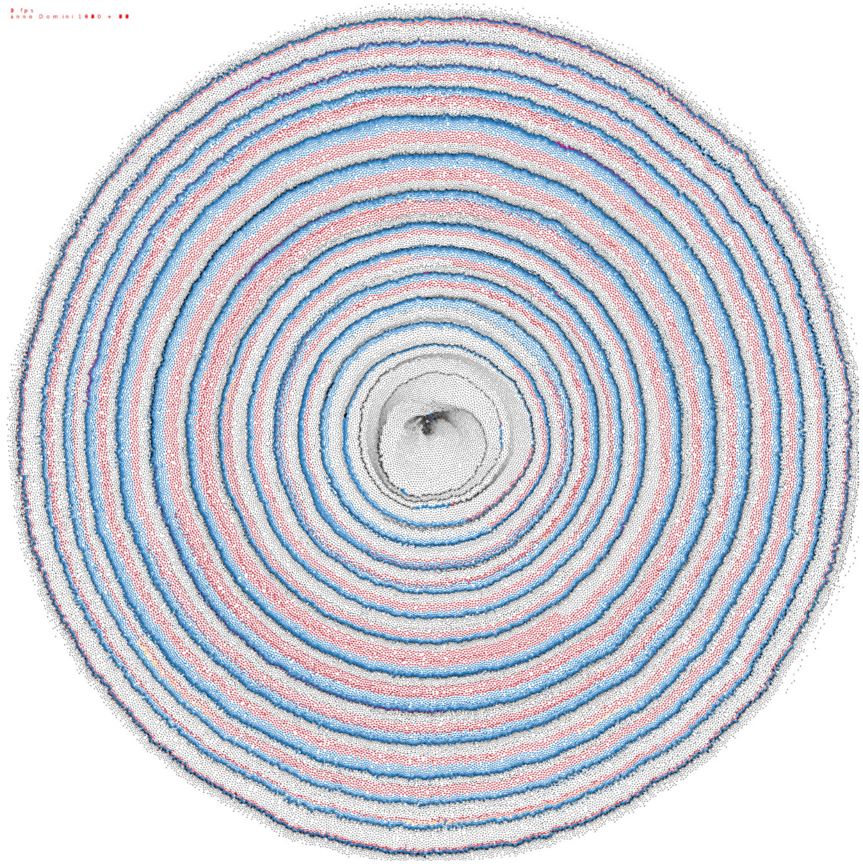
Immigrants in one color

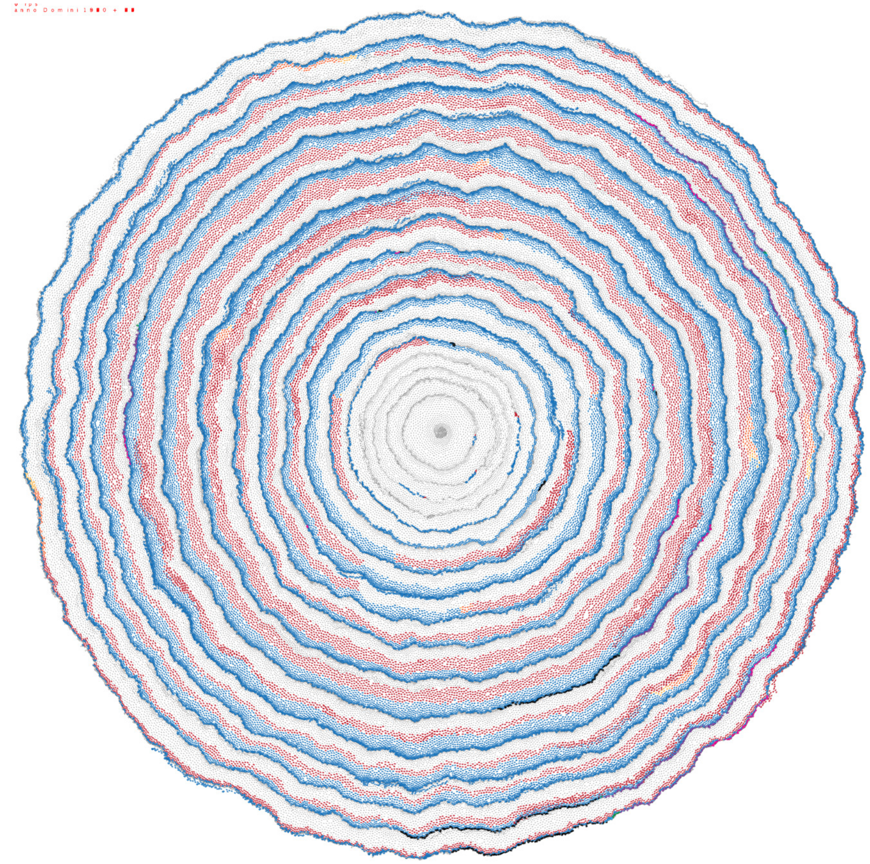




The algorithm was modified in order to first query natural-borns, and them to query immigrants. Having 571 different denominations by place of origin makes it impossible to distinguish among them by using color. By observing the data, these places of origin were grouped into seven cultural-geographical groups: Canada, Europe, Latin America, Asia, Oceania, Africa, and Middle East. The algorithm ran for Massachusetts for the first six decades where Canadians were red and Europeans were blue. Natural-borns were rendered gray.

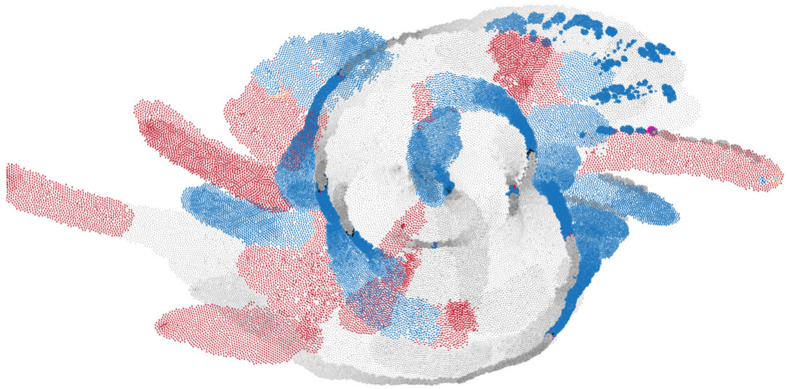
It should be noticed that as data is queried, and as a certain decade is running out of data, the minimum distance among dots diminishes. This causes immigrants to be under-represented as the space they take is more compressed than the space of natural-borns. As more decades and groups of immigrants are represented, recent decades become harder to decipher.



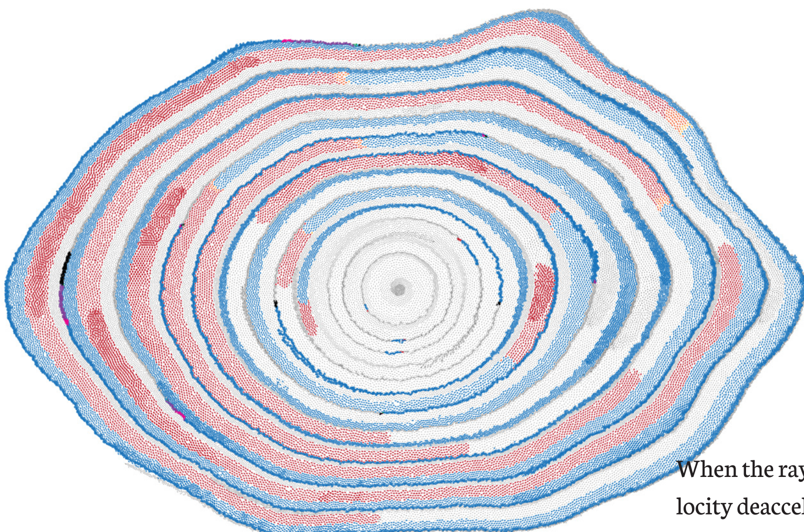


The way in which the minimum distance diminishes can be made subtler. At the same time, natural-borns can be represented with slightly smaller lines. This increases contrast among population groups and enables the visualization to better distinguish them.

As the attempt of creating a graphical arrangement that strongly resembles tree rings resulted in rings that are too circular and concentric, the jitter of the ray's rotation was modified. The ray now jitters too often during a full rotation, which did not provide better results.



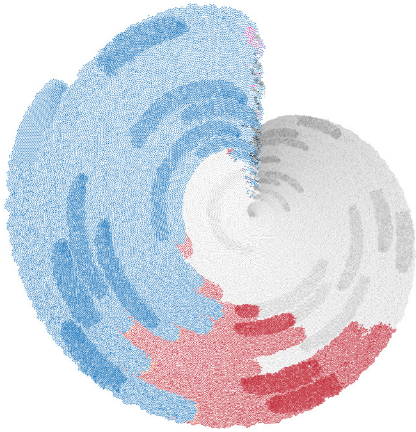
When the ray rotates too slowly



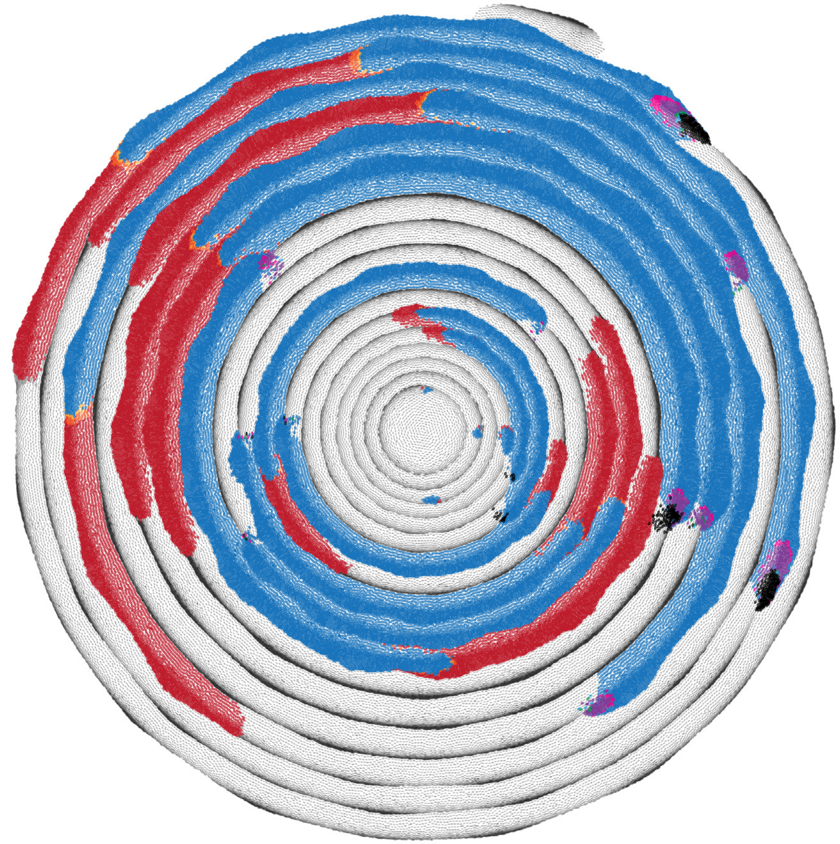
When the ray's velocity deaccelerates in certain angles

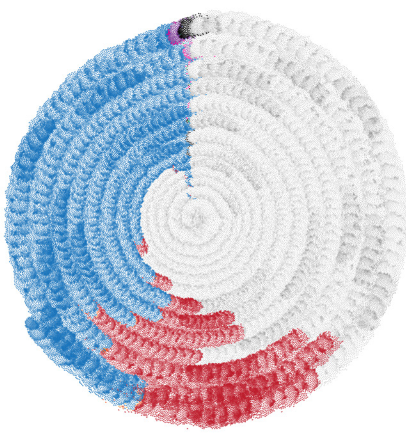
On the other hand, if the ray rotates so slowly that a decade's data is not guaranteed to be dispersed in a full rotation, the results produced become too abstract and too far from our idea of tree rings. The ray can deaccelerate only for certain angles, always with a gentle jitter. This results in a bark shape too skewed and with several protuberances.

It should be noticed that this early attribution of red to represent Canadians and blue to represent Europeans is problematic. In several states, just like in Massachusetts, these are the earliest and some of the most pronounced waves of immigration. Given the cultural significance of blue and red in the American context, the color attribution was changed as not to imply that these two immigration groups are more American than others.



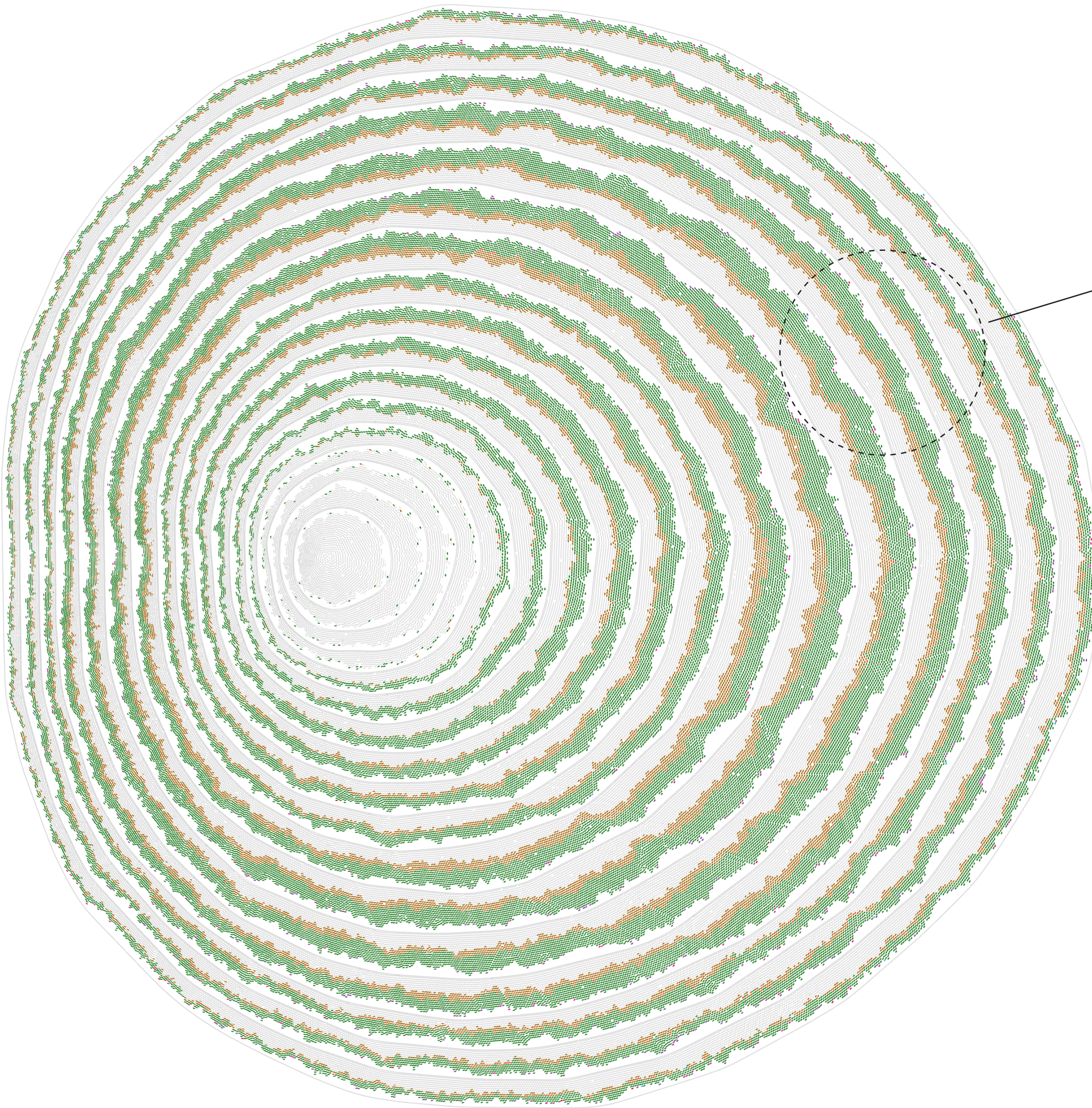
The minimum spacing can
vary as a ring is formed

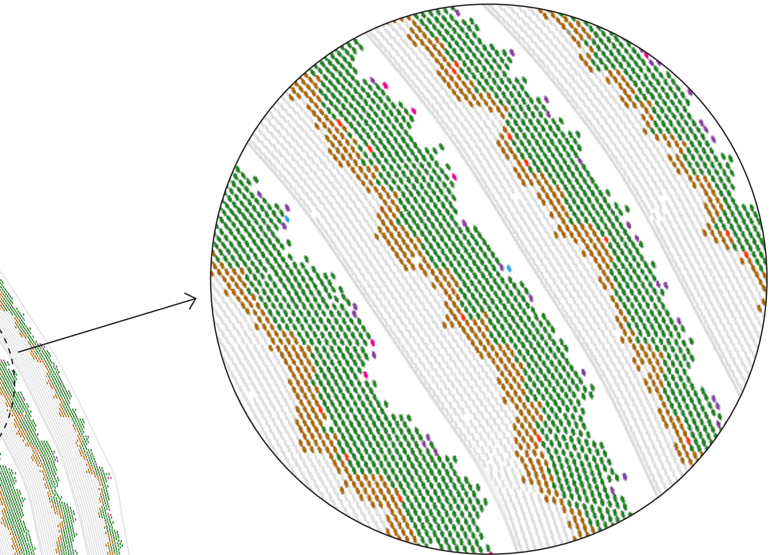




The algorithm can ensure that all the data for a decade is displayed during a single full rotation, in contrast with the previous approach where a decade results from several full rotations of the ray. With this, the minimum spacing among dots can vary in several ways as the ray rotates: it can increase linearly with the rotation; or it can increase and decrease periodically with it.

If the variations of the minimum rotation are made subtle, the visual result functions as a traditional set of donut charts. Since sometimes the space gets compressed, the result appears to have been made with ink that spills and mixes. It confers an analog appearance to the visual, but attaining this type of expressiveness was never the objective of this work, for which much better results could be attained in drawing a donut chart with a paint brush.





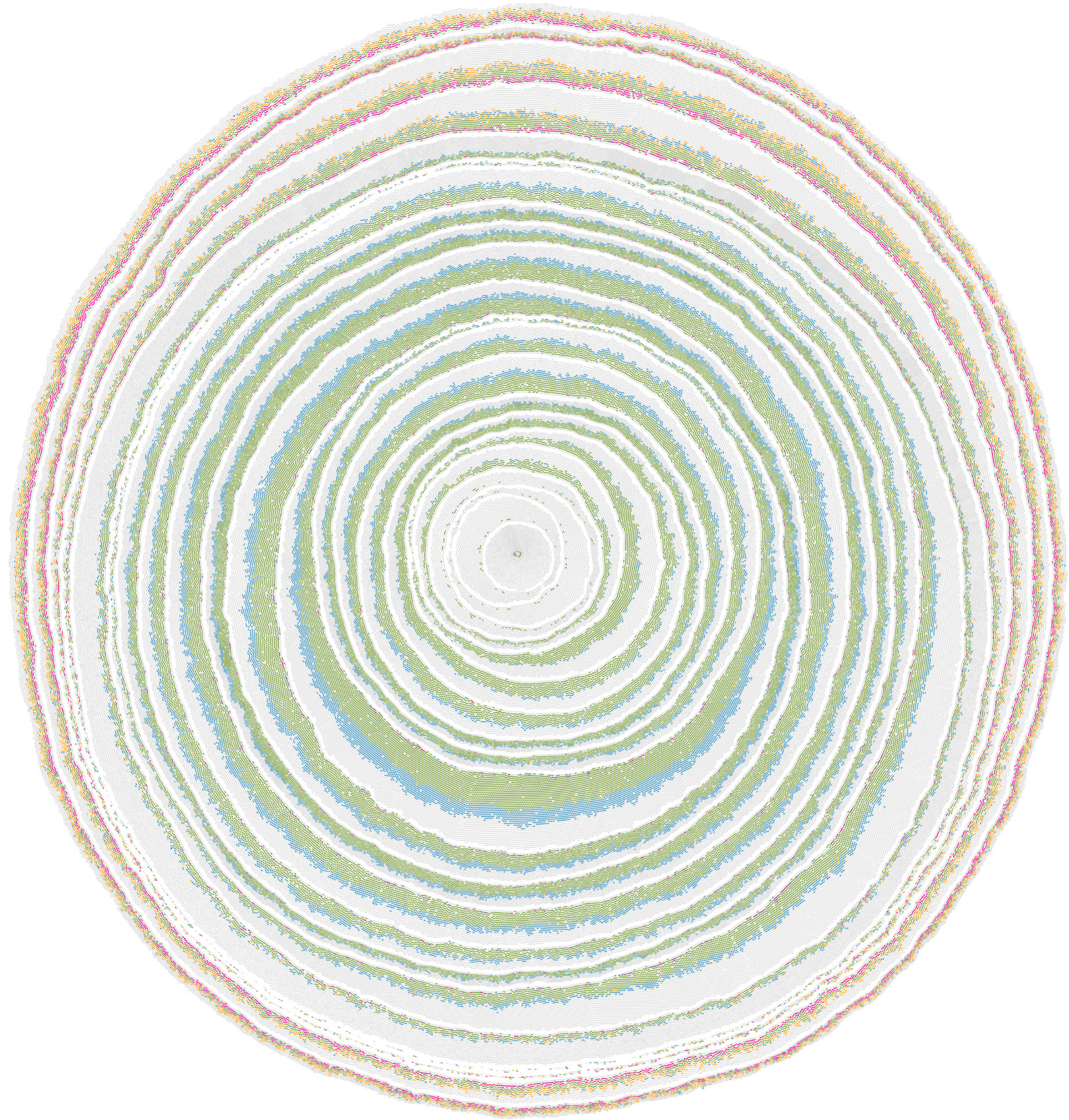
Spacing between rings is introduced. The cells are placed in a ring, and them by smoothing the outline of such ring a new path is created which is the baseline on which the next ring grows.

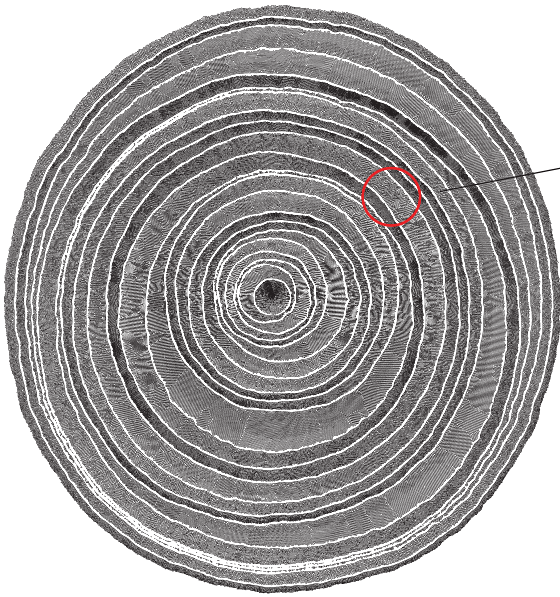
Refinement

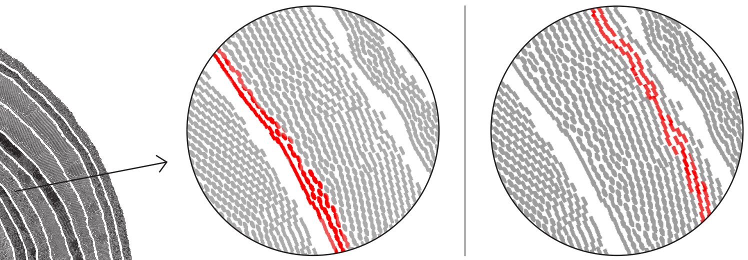
Wanting to part with the appearance of a donut chart, so that immigrants can be visualized as waves and so that they also follow the spatial arrangement of tree rings, the dots are laid out in layers inside each tree ring. Looking at the basswood again, it was observed that if its tree rings were unfolded over a straight line, their thickness profiles resembled a gaussian distribution. Given this observation, instead of using the ray sequentially, the angles that the ray hits are now random, but following a probability model. This is more on par with how cells multiply, as they certainly do not do it in an ordered, clockwise manner. The first implementation of this approach can be seen here, where there is a higher probability of placing a dot on the right of the center. The placing follows a gaussian distribution centered around the direction *right*, and has a specific standard deviation.

Although the resulting image has a subtle shape that can resemble a bark, the monotony in which the thickness of the rings varies appears alien to the idea of tree rings. This result is also a consequence of the data as the number of people added to population increases throughout the decades. Furthermore, it should be noticed that the result has very rugged outlines for every ring, which would be acceptable for a bark, but not for the rings themselves.

In order to tackle this, each ring was created following its own unique gaussian distribution: some distributions skew rings more in one direction, and other distributions to another. In order to fix the rugged outlines, new dots are not placed anymore in the outer periphery of the ring, but instead are inserted in the baseline of the ring. This makes the placed cell push all others outwards, and the ring expands naturally in a way that results in smooth outlines. This also causes immigrants to appear in the image before natural-borns: since immigrants are still the last to be queried, if they are positioned on the rings' baseline, they push the natural-borns outwards. This approach results in a set of non-concentric rings, but that overall are still very circular, and that lack the sense of order seen in basswood. The minimum spacing among dots also varies with the gaussian distribution: in this case the dots are more concentrated in areas where they have a lower probability of being placed. This variation of contrast adds visual granularity to the image, being from a visual perspective perhaps the most organic generated so far.

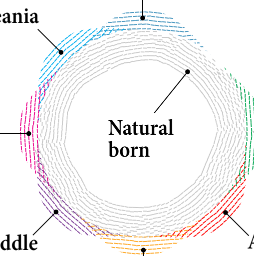






New dots inserted at the baseline of a ring

New dots inserted below the outer layers of ring

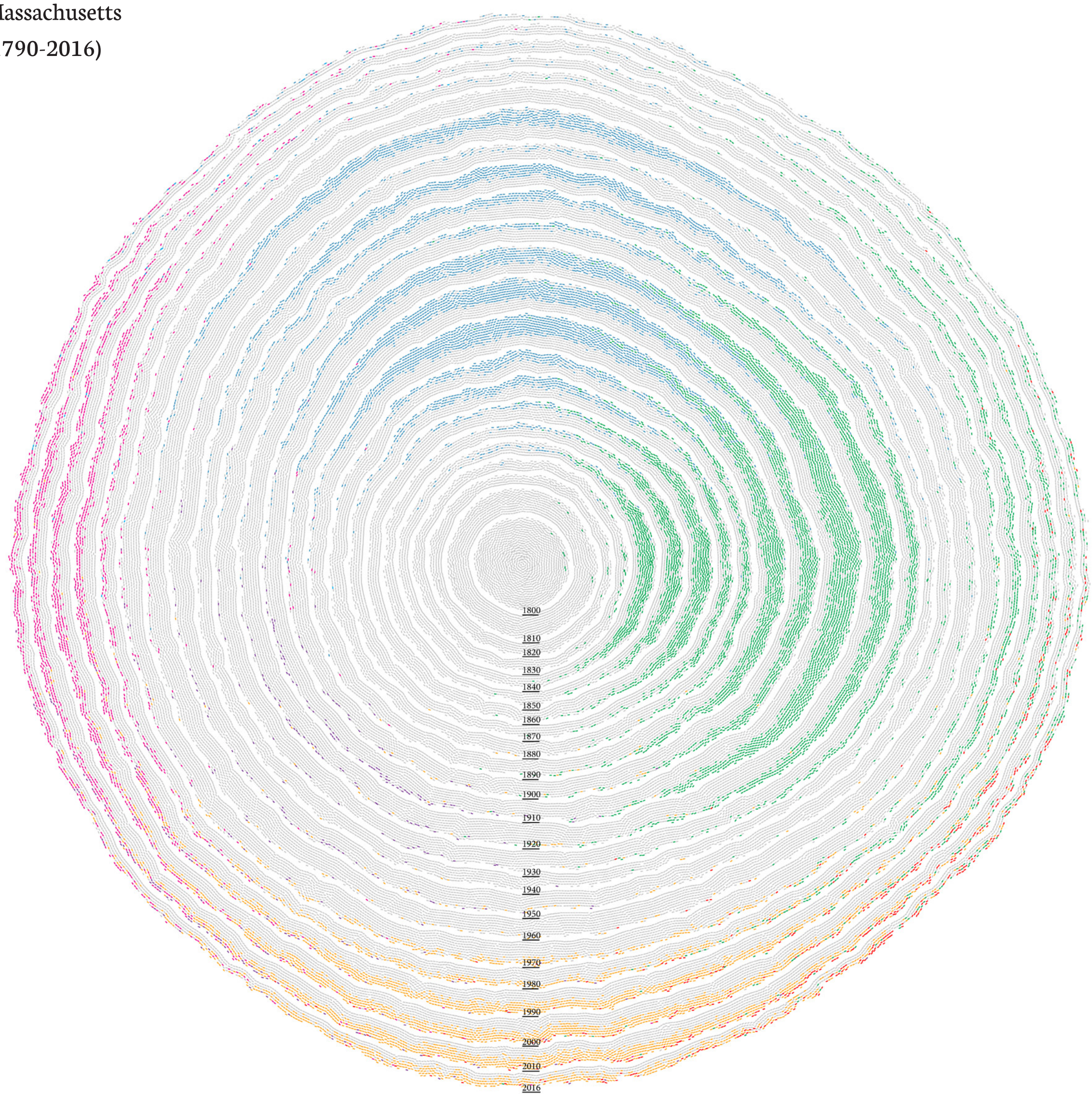


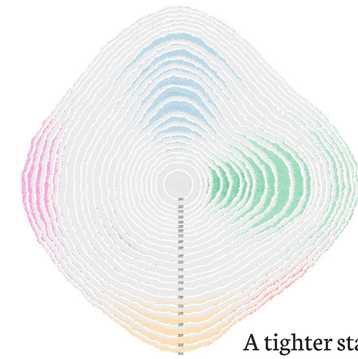
The gaussian distributions utilized to position the dots were picked randomly. Doing this makes certain properties of the visualization neither data-related nor related to the process in how tree rings form. Furthermore, inserting dots at the baseline of a ring, in spite of creating smoother rings, is not how new cells appear in a tree. In fact, new cells originate in the *vascular cambium*, a layer of cells just below the bark. In order to mimic this while preserving the smoothness of the rings, new dots are now placed just a couple of layers below the outer layers of a ring.

Knowing that trees can grow more in one direction than others, this idea was translated to the visualization. When considering the U.S., immigrants come from several geographical directions, so it makes sense that a tree can grow more in the direction where immigration is coming from. In order to do this, the seven cultural-geographical groups were attributed to specific directions (e.g. Canada → North, Europe → East, Latin America → South, and so on). With these directions, a gaussian distribution can be created for each immigration group, with the average centered on the corresponding direction. This results in each state having its own form derived from data. Rings that are more skewed toward East, for example, show more immigration from Europe, while rings skewed South show more immigration from Latin America. Fifty sets of tree rings were simulated to show different profiles of growth and immigration for each U.S. state.

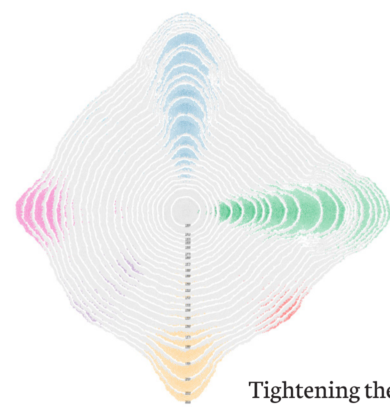
The next visualizations show this approach implemented for the fifty U.S. states which are then laid out geographically as in a cartogram. For each ring, first natural-borns are placed according with an uniform distribution, and secondly the immigration groups are placed, dot by dot, according with their own gaussian distribution. The standard deviation of these distributions can be modified in a way to disperse the lumps created by immigration. Each dot, or cell, corresponds to 150 people. The visualizations can work in small sizes as shown in the geographical layout of the states, but they were devised to be presented in a large format. This way each cell and each dot can be connected with the overall patterns that they form.

Massachusetts
(1790-2016)

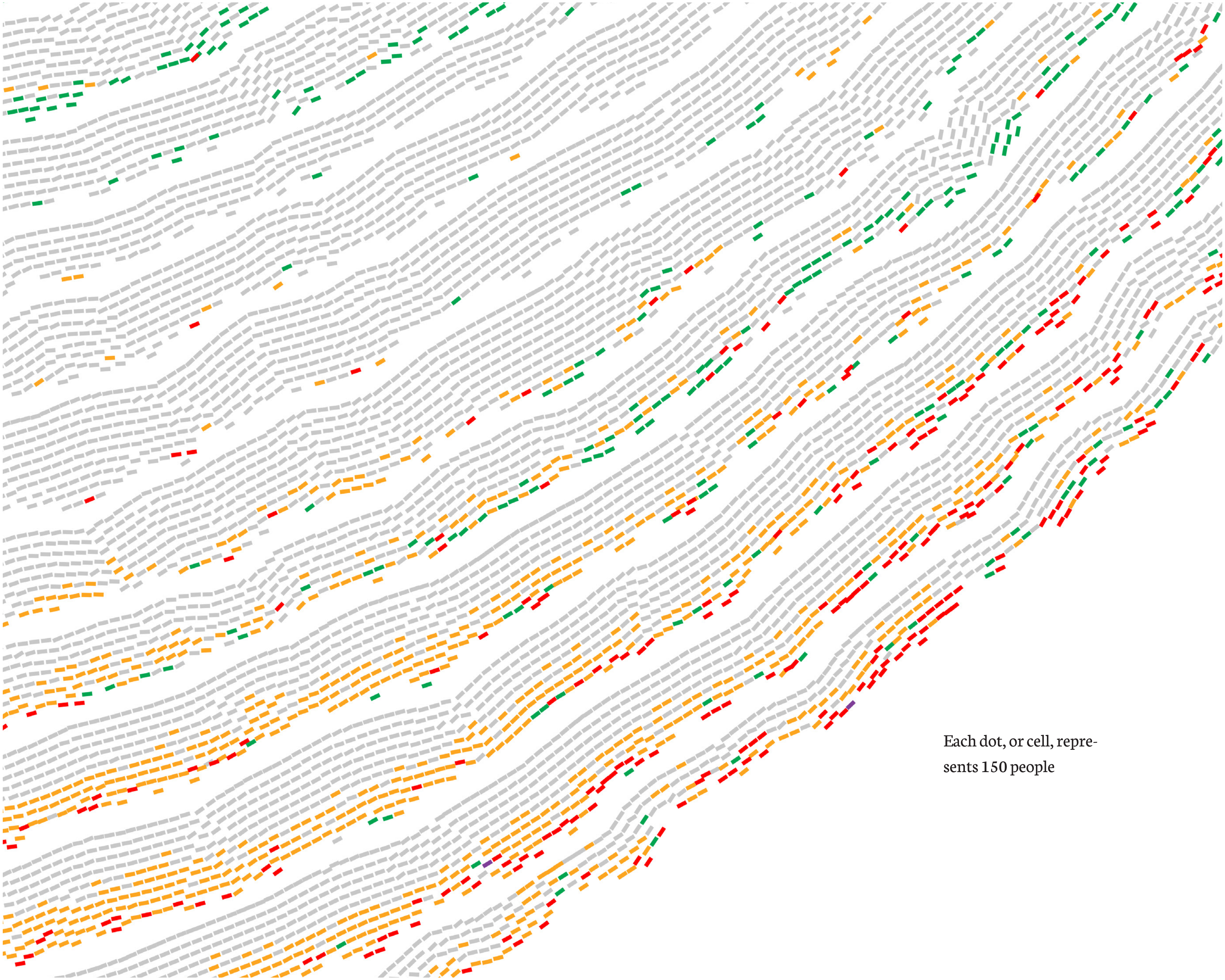




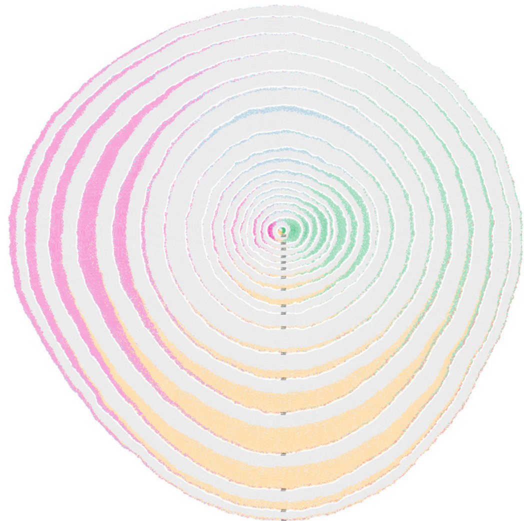
A tighter standard deviation
for the gaussian distribution



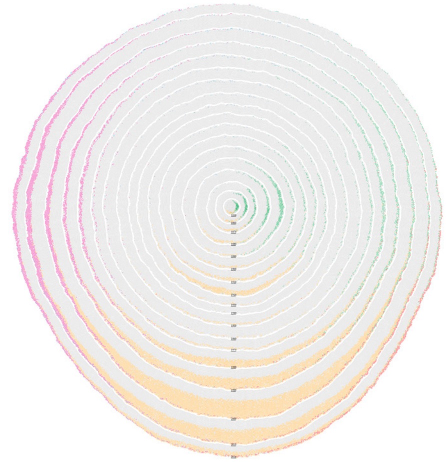
Tightening the standard deviation creates lumps in the rings



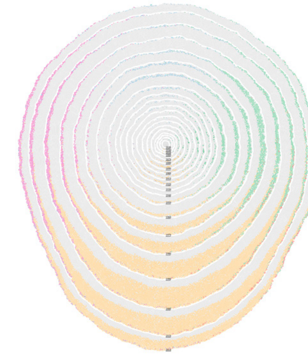
Each dot, or cell, represents 150 people



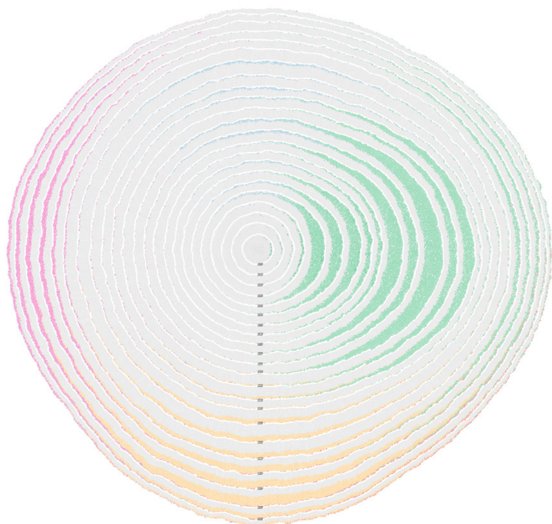
California (1840-2016)



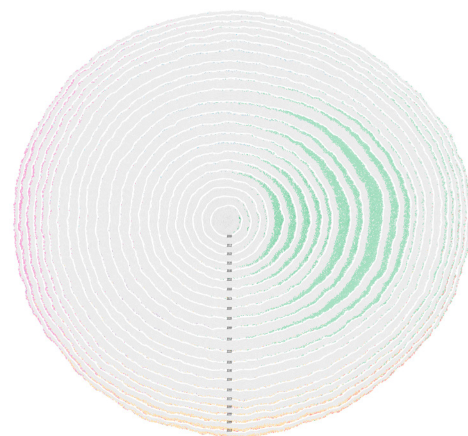
Texas (1840-2016)



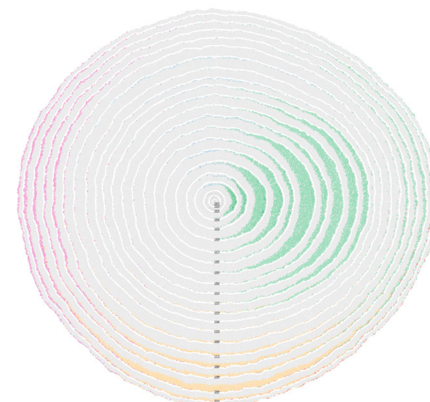
Florida (1820-2016)



New York (1790-2016)

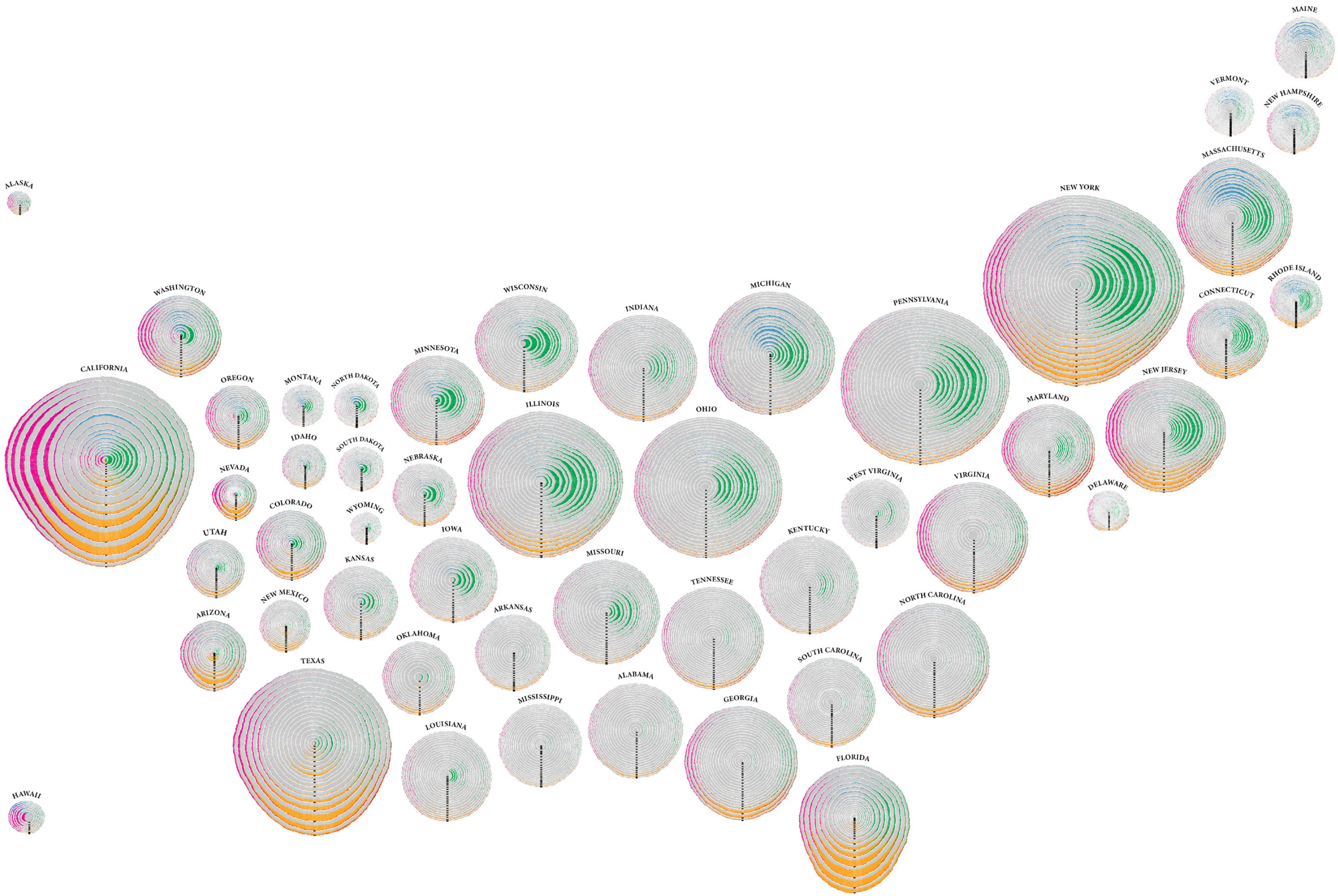


Pennsylvania (1790-2016)

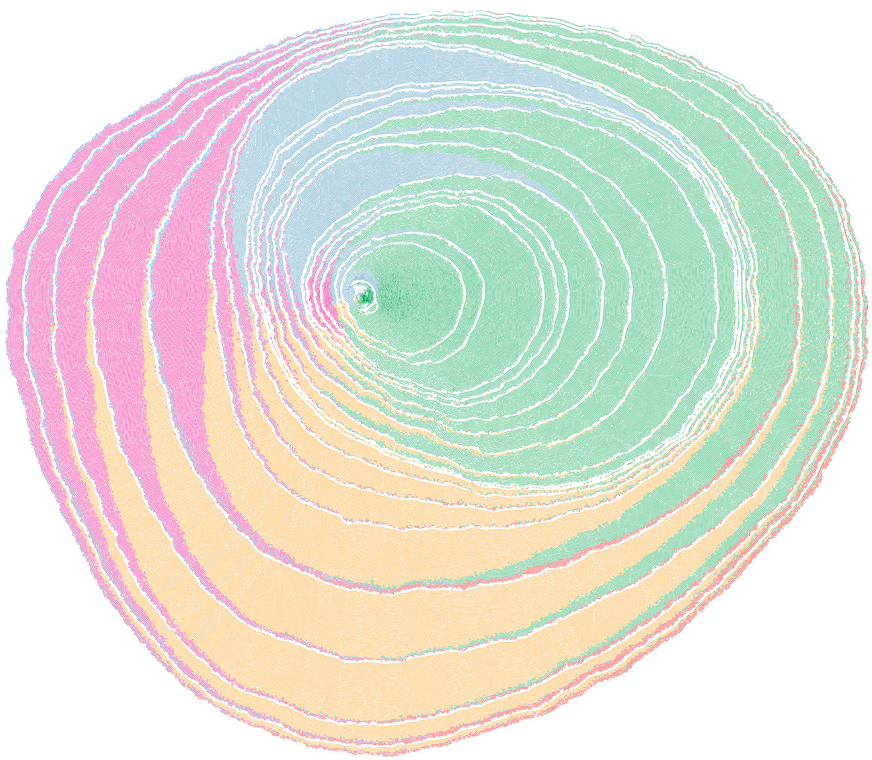


Illinois (1800-2016)

All fifty states as a cartogram

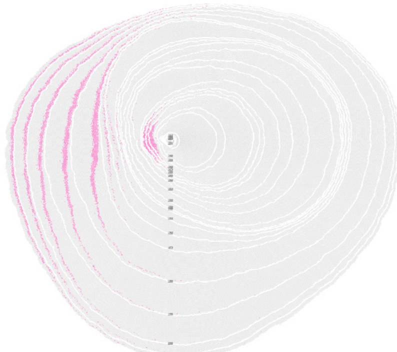


A complementary take
The dataset contains information on the countries where immigration came from. In spite of having all countries represented it is not possible to distinguish among them except for knowing that they are part of a different cultural-geographical group. If one is to show some of the main countries of origin, a tree with only immigrants can be generated. This loops back to the idea of having the U.S. as a single tree, made out of immigration. Additionally, it provides more spatial resolution to color some of the cells and visualize the history of immigration as coming from a certain country.

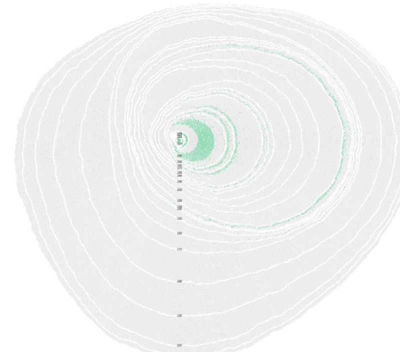


One last comment
This ongoing work started producing results in which the profiles of population growth are portrayed as rings in a tree. The visualizations are the result of seeking a visual resemblance with the appearance of tree rings, while at the same time visually organizing data in a way that is inspired by how cells grow in a tree. This approach can be utilized to produce visualizations that show immigration with a poetic take, but that also show specific information in an organized space. There are several presentation possibilities for these visualizations, be it in an animated or interactive form, or by statically annotating historical events that explain the ring's shape.

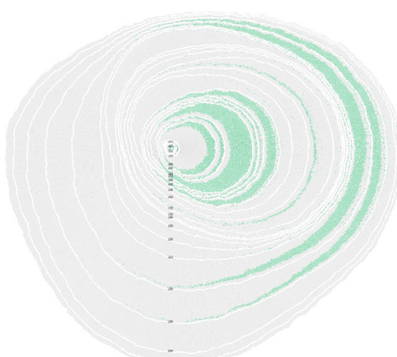
In order to make the visualization stand out by its form, while conferring an unique identity to each state, we applied the metaphor of tree rings. The application of a metaphor enables the visualization to carry its own messages about the theme. In fact, if the metaphor is made evident, the visualization can act as a set of devices to build high-level rationales on the data using the metaphor's own network of concepts: a ring skewed in a certain direction tells where immigration came from; a thicker ring indicates that the growth was higher that decade; counting the number of rings is counting the number of decades; the outer rings correspond to more recent portrayals, while the inner rings correspond to older portrayals; the bigger the tree's cross section, the greater the number of people who passed through that state. Finally, all people are part of the tree, and all cells contributed to population growth and have left their mark in the tree's history.

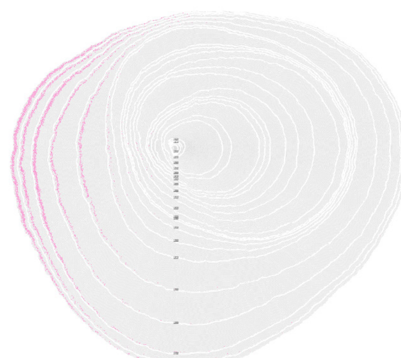

China

Cuba

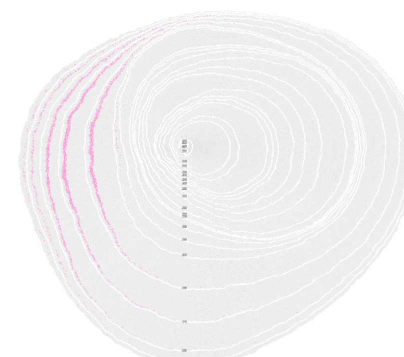

Ireland

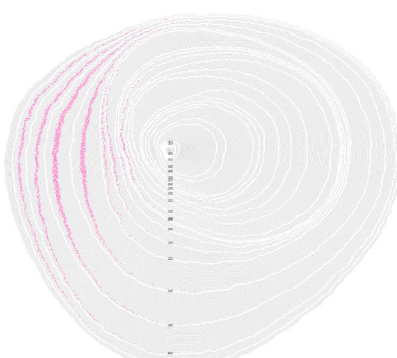
Mexico

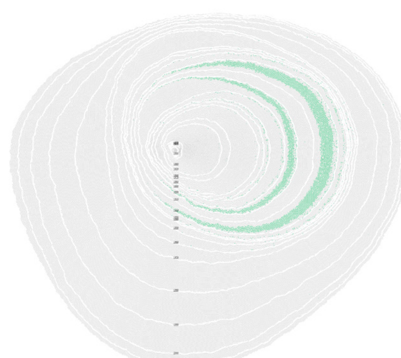

Germany


India

Puerto Rico


Vietnam


Philippines


Poland

References

1. A. Fahn, J. Burley, K. A. Longman and A. Mariaux. 1981. Possible contributions of wood anatomy to the determination of the age of tropical trees. In: F. H. Bormann and G. Berlyn (eds.). Age and growth rate of tropical trees: new directions for research: New directions for research. pp. 31–54 (Bull. 94). Yale University, New Haven.
2. Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. Integrated Public Use Microdata Series: Version 7.0 [dataset]. Minneapolis: University of Minnesota, 2017. https://doi.org/10.18128/D010.V7.0.
3. Peter H. Raven, Ray Franklin Evert and Susan E. Eichhorn. 2005. Biology of Plants. W. H. Freeman and Company, New York.