Interactive Visualization of Business Births and Deaths in the U.S. Economy using a Novel Visualization Technique Called HiFi Pie

Leonidas Deligiannidis  
Professor of Computer Science  
Wentworth Institute of Technology  
Boston, MA, USA  
deligiannidis1@wit.edu

Erik Noyes  
Associate Professor of Entrepreneurship  
Babson College  
Babson Park, MA, USA  
enoyes@babson.edu

Abstract
HiFi Pie is a novel technique for interactive information visualization. To illustrate its strength, we explore historic data on new business births and deaths in the U.S. economy. As U.S. Economic Census data continually improves to track the birth and death of new businesses, one can visualize patterns of creative destruction in the U.S. economy, particularly broad historic patterns of whole-economy expansion, all aggregated from industry-level views of new business births and deaths. A novel, interactive viewer built on the visual language the common pie-chart, HiFi Pie allows economic development organizations focused on entrepreneurship, as well as economists and other policy makers, to visualize new business creation trends, including dynamics of entrepreneurial job creation and industry innovation.

Keywords: Information Visualization, Creative Destruction, Entrepreneurship, Economic Development, Job Creation, Visualization Methods

1. Introduction
As data improves to track new business births and deaths in the U.S. economy, new methods are needed to visualize historic patterns of new business creation and industry growth. Economic development organizations focused on entrepreneurship, as well as economists and policy makers, understand that new entrepreneurial ventures are the single greatest driver of new job creation and economic growth, but fast-growth industries are also characterized by high rates of new business failure.

Creative destruction, an idea first advanced by Joseph Schumpeter [1], is the concept that the constant process of new business creation, even driving other businesses to their death, accelerates innovation and economic development. The birth and death of new ventures, while a source of disruption and unemployment, is on the whole an economic win for societies because surviving, competitive ventures create new innovations, economic value and jobs.

These two perspectives, one of whole-economy growth and constituent new venture birth and death are rarely combined to visually provide a coherent sense of their interrelationship.

HiFi Pie represents an effort to explore the potential usability of uniting these two perspectives in a browsable, controllable visual interface. The user-friendly visualization leverages the visual language of the common pie chart, while including added dimensional information about new business births and deaths and industry expansion. Diverging from traditional pie chart conventions, total circle area has precise numerical meaning (i.e., beyond 100%), varies depending on data depicted, and here represents concentric rings of economic and business activity similar to the growth rings of a tree, conveying both percentages and totals.

While the paper details a first, functioning prototype of HiFi Pie, examining nine major industry groups, or sectors, the tool is being developed to visualize multiple levels of industry classification data (e.g., according to the 6-digit, hierarchical, North American Industrial Classification System)

2. Background
The concept of using a pie chart as a visualization mechanism was conceived two centuries ago [2]. William Playfair used the pie chart concept in [2] to visualize the areas, population, and revenues of European countries. In [3] Ian Spence provides a great picture of how the original pie chart evolved over the years and how a pie chart could be used as a useful visualization metaphor.

The authors in [4] provide a wonderful historical review of radial visualization tracking its roots in centuries in the past. One of the concepts that is mentioned in this paper and is related to HiFi Pie is that longer radii of concentric circles have bigger area. The radii, thus, must be computed accurately to reflect the areas added onto the concentric circles. This introduces confusion for the reader and that's the reason the pie-charts do not have widespread acceptance. However, we address this issue by our second stage of our algorithm in which we align the segments each pie slice is composed of.

Krona [5] combines a variant of radial displays with parametric coloring and interactive polar-coordinate zooming. The display resembles pie-charts with inter-segmented pie slices producing an embedded hierarchical visualization. The inter-segmented slices are arranged from the top level of the hierarchy and progress outwards.
In [6], the authors use tree maps with a pie transformation which results in twisted rectangles around a center point to create pie-chart-like visualizations. In [7] a new framework is proposed for visualizing tables, proportions, and probabilities. They produce pie-charts with different radii for each concentric circle. PieTrees [8] are area based tree visualization that can be used to represent hierarchical numerical data. They map size directly onto area into a circular layout. Users can expand and collapse any nodes or the entire graph. Circle View [9] combines pie charts with a novel arrangement of time events on circle segments where each segment is further divided into sub-segments to visualize the distribution and changes over time. This enables it to present a visualization with both local detail and global context in a single view. The resulting visualization graphs resemble pie-charts with different concentric ring widths.

3. The Data

We obtained the data from the US Department of Commerce, United States Census Bureau [10]. Description of the fields in the dataset is located at [11]. The dataset consists of nine sectors:

1. AGR (Agriculture)
2. MIN (Mining)
3. CON (Construction)
4. MAN (Manufacturing)
5. TCU (Transportation)
6. WHO (Wholesale Trade)
7. RET (Retail Trade)
8. FIRE (Finance)
9. SRV (Services)

The dataset contains data from 1977 to 2011. There are many fields in the dataset such as, firms (number of firms), estabs (number of establishments), job_creation (number of jobs created over the last 12 months), etc. The complete list of fields is located at [11].

4. The Algorithm

The algorithm to generate the HiFi Pie consists of two stages. Stage’s one goal is to calculate the correct radii of all concentric circles, and thus, the area of each pie slice is represented with high accuracy. The second stage consists of the alignment process where each layer of the pie slice is aligned with all the other layers to produce the final HiFi Pie slices.

![Image](https://via.placeholder.com/150)

**Figure 1.** Layers added to each pie-slice, which is data added each year.

Each pie slice consists of layers, data that is added each year and with different rate as shown in figure 1. These layers are centered along their center axis.

4.1. Stage One – Calculating Radii

The algorithm of the first stage consists of several short steps. We present the algorithm in detail (as pseudo-code) and for each step of the algorithm we show how we fill in our data-structures; with the ultimate goal to calculate the radii of each concentric circle to build the pie-slices. We present the data in our data-structures as a table (in the appendix) with the results filled in by the execution of each step of the algorithm. To explain the algorithm with numerical values, we use a small dataset consisting of three years and three sectors (AGR, MIN, and CON). This algorithm needs to be run for each field; in this example we run it for just one field. Initially, the data loaded from our database looks like Table 1.

We first calculate the overall sum of all years of all sectors as shown below; in this example this total sum is equal to 300.

```plaintext
for each year Y do
    for each sector S do
        TOTAL += getVal(Y, S)
    done
done
```

We then calculate the percentage of each sector compared to the overall sum calculated in the previous step (also see Table 2 in the appendix).

```plaintext
for each year Y do
    for each sector S do
        ValuePercentage = getVal(Y, S) * 100 / TOTAL
        Store(Y, S, ValuePercentage)
    done
done
```

We then compute the sum of the percentages for each year as well as the cumulative sum, as a percentage, starting at the first year in our dataset. In other words, for each year we calculate the percentage sums of all sectors (in the column TotalForRow) as well as the cumulative percentages sum of all years (in the column Cum) (also see Table 3 in the appendix).

```plaintext
Cum = 0
for each year Y do
    for each sector S do
        V = V + getVal(Y, S)
    done
    storeTotal(Y, S, V)
    Cum += V
    storeCum(Y, S, Cum)
done
```

Based on the percentages of each sector, we calculate the degrees/angle of the pie slice that corresponds to each sector for each year (also see Table 4 in the appendix).
for each year Y do
for each sector S do
  Val = get_ValuePerc (Y,S)*360/getRowTotal(Y)
  storeDegrees(Y,S,Val)
done
done

In the final step of stage one of the algorithm, we calculate the corresponding radii for each year in relation to the unit circle. The radius of the unit circle is calculated as $R = \sqrt{\frac{\text{Area}}{\pi}} = \sqrt{\frac{\text{getCum(year=1979)}}{\pi}}$ where the year is equal to the last – most resent – year in the dataset; all data will be within the circle defined by this radius. The total area of all our data is equal to the cumulative sum which is stored in the Cum column for the last year. We use this radius as the radius of the unit circle and we calculate all other radii for all other circles as shown below (also see Table 5 in the appendix):

$$R = \sqrt{\frac{\text{Area}}{\pi}} = \sqrt{\frac{\text{getCum(year=1979)}}{\pi}}$$

for each year Y do
  Val = $\sqrt{\frac{\text{getCum(Y)}}{\pi}} / R$
  storeRadius(Y,Val)
done

The user, using a slider in the application, can adjust the radii of the circles to better view the pie-charts. We call this user-specified value sliderVal. To draw each concentric circle, we simply need to multiply each radii of a corresponding circle with this value:

$$\text{radius} = \text{getRadius}(Y) * \text{sliderVal}$$

where Y is the corresponding year.

Now that we have all the data we need, we can contrast the difference in the pie charts we generate using HiFi Pie and using Excel’s built-in 2D-piechart Doughnut representation. Figure 2 shows the two generated pie charts. The left one is the data representation using Microsoft Excel’s doughnut 2D pie chart. All circles are of the same width. The right one is the HiFi Pie chart before stage two, which is the alignment process of the algorithm. The concentric circles here are shown with different widths. Each radii is correctly calculated to show the correct area each sector occupies, keeping the overall area consistent with the data and the overall area.

Figure 2. Data representation using Microsoft Excel’s Doughnut 2D pie chart graph (left), and using HiFiPie (right) before the alignment process. Notice the width of the concentric circles that are different in size in HiFi Pie, which accurately visualizes the area of each year.

To present the data as an overall conventional pie-chart using Microsoft Excel, we need to calculate the sums of each sector and their corresponding angle of each pie-slice as shown below (also see Table 6 in the appendix):

for each sector S do
  Val=0
  for each year Y do
    Val = Val + getVal(Y,S)
  done
  StoreVtotal(S,Val)
  Degrees = (Val * 100 / TOTAL) * 360/100
  storeTotalDegrees(S, Degrees)
done

This produces the chart shown in figure 3. Even though the overall data is represented just fine here, what is missing is how each pie slice evolved over the years. HiFi Pie is able to visualize how the data from each year shapes each pie-slice. Stage two of the algorithm will align each layer of each pie slice to produce the final product.

![Figure 3](image)

**Figure 3.** Using a conventional pie-chart to visualize overall data.

### 4.2. Stage Two – Alignment Process

The results of stage one is fed to stage two of the algorithm. Each year adds a layer to the previous years. When considering all sectors’ data for a particular year, this additional data that is added appears as a ring of data. For this example, the three rings of data for the three year are shown in figure 4. We construct the layers by drawing layer N and then subtracting layers N-1 and beyond, recursively, where N in the most current layer representing the data for the most recent year. These layers (or rings) contain the data for a particular year for all sectors. We then extract the segments of the rings of each sector and center/align them producing the HiFi Pie slices as shown in figure 5.
5. The Graphical Interface

The graphical interface presents to the user the tools to select the different fields to be visualized, modify the range of years the HiFi Pies represent, to zoom in and out on a HiFi Pie chart, and to expand the pie slices. Expanding the slices means that we move all slices away from their common center to enable the user to rotate them later and fit them next to each other. It also enables the user to direct-manipulate the entire chart or individual HiFi pie slices: the user can mouse click on a slice to move it and rotate it; left mouse button moves the entire HiFi Pie, shift-left mouse button moves individual HiFi Pie slices, and right mouse button rotates a HiFi Pie slice around the HiFi Pie’s common center. From the menu bar the user can choose the quality in which the HiFi Pie slices are drawn; select/deselect antialiasing, and level of detail to render the HiFi Pie slices faster. The user can also select the Universal Radius option so that each HiFi Pie chart is drawn with respect to the largest HiFi Pie. We first calculate the total values of each field and we get the maximum. From then on, each HiFi Pie is drawn with respect to this maximum radius. This enables us to compare different HiFi Pie chart side by side. Figure 6 shows the HiFi Pie before we run the stage two of the algorithm. The radii are shown on top of the pie chart and by themselves alone to show that the radii are different for each concentric circle. The area of the outside rings grows faster than the rings’ area closer to the center.

After stage two, where we align and center each sector data from Figure 6, we produce an in-memory image (for each sector) where we render on it the aligned pie slices to produce the HiFi Pie slices as shown in figure 7. Because of the shapes of the HiFi Pies, it is very common, that the individual HiFi Pie slices overlap. The user, however, can expand the chart, moving all sectors away from their common center, then rotate and move the individual slices to produce HiFi Pie charts where there is no slice overlaps as shown in figure 8.

The High Fidelity of the HiFi Pie slices can be seen in figure 9. This is a close up view of figure’s 8 bottom left area. Here we can see with high precision how the data from individual years affected the formation of each pie slice.
6. Conclusion
HiFi Pie demonstrates the potential viability and usability of a multidimensional, interrelated representation of creative destruction leveraging the pie chart structure. The visualization captures economic growth, industry growth, and new business births and deaths. Applications of HiFi Pie may include use by policy makers or economic development organizations focused on entrepreneurship to contextualize policy, as well as use in entrepreneurship educators to depict and explore processes of creative destruction. Further research and testing is needed to explore industry views for multiple levels of analysis (e.g., according to the 6-digit, hierarchical, North American Industry Classification System) to provide finer-grained views of new business creation and job creation beyond broad industry groups such as agriculture, mining and services. One challenge with representing multiple, smaller slices of HiFi Pie is the breakdown of the classic pie chart structure and the potential that relative percentages and areas for different slices become more difficult to compare. However, it is quite possible that the basic pie chart structure is flexible and adaptable to integrate added dimensions of information while retaining core visualization properties of establishing relative areas. One important area for future research is the opportunity to establish meaning in the circular layout of industry categories. For example, the Internet and web industry has most shaped the services industry, suggesting a logic to put most-related industries side by side (potentially based on related products or related customers). While some arrays of basic pie charts have an underlying logic for category proximity (e.g., percentage of respondents who are 18-24 years old adjoining percentage of respondents who are 25-32), there are likely means to automatically array industry segments based on relationships.

Because the study of composition and change is central to many types of analysis and presentation, of course HiFi Pie may be extended to represent and browse other types of hierarchical classification data. Broadly, this research
aims to bring together research on entrepreneurship and economic development with research on interactive information visualization to suggest new opportunities and directions at the intersection of the two specialized fields.

References
Appendix

Table 1. Initial load of data in our data-structure.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGR Value</th>
<th>% degrees</th>
<th>MIN Value</th>
<th>% degrees</th>
<th>CON Value</th>
<th>% degrees</th>
<th>TotalFor Row %</th>
<th>Cum %</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>20</td>
<td>20/3</td>
<td>50</td>
<td>50/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>100/3</td>
<td>0.577</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>10/3</td>
<td>60</td>
<td>60/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>200/3</td>
<td>0.815</td>
</tr>
<tr>
<td>1979</td>
<td>50</td>
<td>50/3</td>
<td>40</td>
<td>40/3</td>
<td>10</td>
<td>10/3</td>
<td>100/3</td>
<td>300/3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2. Calculating percentages of each sector for each year.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGR Value</th>
<th>% degrees</th>
<th>MIN Value</th>
<th>% degrees</th>
<th>CON Value</th>
<th>% degrees</th>
<th>TotalFor Row %</th>
<th>Cum %</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>20</td>
<td>20/3</td>
<td>50</td>
<td>50/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>100/3</td>
<td>0.577</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>10/3</td>
<td>60</td>
<td>60/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>200/3</td>
<td>0.815</td>
</tr>
<tr>
<td>1979</td>
<td>50</td>
<td>50/3</td>
<td>40</td>
<td>40/3</td>
<td>10</td>
<td>10/3</td>
<td>100/3</td>
<td>300/3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3. Calculate total percentages for each year as well as their cumulative sum.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGR Value</th>
<th>% degrees</th>
<th>MIN Value</th>
<th>% degrees</th>
<th>CON Value</th>
<th>% degrees</th>
<th>TotalFor Row %</th>
<th>Cum %</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>20</td>
<td>20/3</td>
<td>50</td>
<td>50/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>100/3</td>
<td>0.577</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>10/3</td>
<td>60</td>
<td>60/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>200/3</td>
<td>0.815</td>
</tr>
<tr>
<td>1979</td>
<td>50</td>
<td>50/3</td>
<td>40</td>
<td>40/3</td>
<td>10</td>
<td>10/3</td>
<td>100/3</td>
<td>300/3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 4. Calculate the angle of the pie slice for each year for each sector.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGR Value</th>
<th>% degrees</th>
<th>MIN Value</th>
<th>% degrees</th>
<th>CON Value</th>
<th>% degrees</th>
<th>TotalFor Row %</th>
<th>Cum %</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>20</td>
<td>20/3</td>
<td>50</td>
<td>50/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>100/3</td>
<td>0.577</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>10/3</td>
<td>60</td>
<td>60/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>200/3</td>
<td>0.815</td>
</tr>
<tr>
<td>1979</td>
<td>50</td>
<td>50/3</td>
<td>40</td>
<td>40/3</td>
<td>10</td>
<td>10/3</td>
<td>100/3</td>
<td>300/3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 5. Calculate the radii of each sector for each year.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGR Value</th>
<th>% degrees</th>
<th>MIN Value</th>
<th>% degrees</th>
<th>CON Value</th>
<th>% degrees</th>
<th>TotalFor Row %</th>
<th>Cum %</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>20</td>
<td>20/3</td>
<td>50</td>
<td>50/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>100/3</td>
<td>0.577</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>10/3</td>
<td>60</td>
<td>60/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>200/3</td>
<td>0.815</td>
</tr>
<tr>
<td>1979</td>
<td>50</td>
<td>50/3</td>
<td>40</td>
<td>40/3</td>
<td>10</td>
<td>10/3</td>
<td>100/3</td>
<td>300/3</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6. Calculating the overall area to display the data using a conventional Pie Chart.

<table>
<thead>
<tr>
<th>Year</th>
<th>AGR Value</th>
<th>% degrees</th>
<th>MIN Value</th>
<th>% degrees</th>
<th>CON Value</th>
<th>% degrees</th>
<th>TotalFor Row %</th>
<th>Cum %</th>
<th>Radius</th>
</tr>
</thead>
<tbody>
<tr>
<td>1977</td>
<td>20</td>
<td>20/3</td>
<td>50</td>
<td>50/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>100/3</td>
<td>0.577</td>
</tr>
<tr>
<td>1978</td>
<td>10</td>
<td>10/3</td>
<td>60</td>
<td>60/3</td>
<td>30</td>
<td>30/3</td>
<td>100/3</td>
<td>200/3</td>
<td>0.815</td>
</tr>
<tr>
<td>1979</td>
<td>50</td>
<td>50/3</td>
<td>40</td>
<td>40/3</td>
<td>10</td>
<td>10/3</td>
<td>100/3</td>
<td>300/3</td>
<td>1</td>
</tr>
</tbody>
</table>