# **Union through Data: Aesthetic Information Visualizations**

Carlon McPherson Csc497 Final Paper

Abstract: For my senior project I created several visualizations relating to various types of data including demographics, race, energy, and waste. The data used was related to Union College. The point of these visualizations is to highlight significant and important aspects of the data. The data is represented through different visual metaphors that are meant to engage the user on a more in-depth level than simply presenting the information in the most easily understandable way.

## 1 **INTRODUCTION:**

Information Visualization is the study of the visual representation of large collections of data. This field is inspired by the fields of Computer Science, Psychology, graphic design, and Art. Originally I wasn't sure how to approach this project; my idea was to use various social statistics related to the country, but I had an extremely difficult time narrowing the scope of the project. I needed data that was somewhat accessible and also something that I thought as an artist, my viewers would be able to relate to. This is how I came to the conclusion of using data about Union College. My hope is that by using data so closely related to us, that my exhibition and pieces will be both informative and interesting to all who attend and have the ability to explore my works. Based on the background of Information Visualization and the development of artistic visualizations this paper will show how, through the use of visual metaphors, visualizations can present data in a way that requires more user immersion. This is the foundation on which I created my own visualizations which will be explained and displayed further in the paper.

### **1.1 INFORMATION VISUALIZATION:**

Information Visualization developed as a field to transform information and data into a visual form so that it could be easily read and understood. By creating effective visual interfaces, visualizations allow us to interact with large collections of data rapidly and effectively to discover hidden patterns and trends [1]. This means that these visualizations became a means of representing data visually in order to promote better understanding and faster evaluation. How can this data be represented visually to accomplish this goal? According to Gaviria, there are two main types of visualizations, functional and aesthetic [2]. The earliest forms of visualizations are

3: 17.	Market Open	ambientul	
Dow .	NASDAQ	S&P 500	
12,687 + 2,	2,439 - 15,	1,425 0,	
СОСЦИ н. +2.0 +15 +1.0			
+.5 0% 5 -1.0 -1.5 -2.0			

Figure 1. Example of a functional visualization. Stock market data (Gavaria 2008).

functional. Functional visualizations are concerned with usability and performance; the visual form of representation matters only in how rapidly it helps communicate the underlying data to the user [2]. As seen in Figure 1, function visualizations are what we are most accustomed to. The practical use and creation of these types of visualizations are seen in objects and places all around us. In this example the stock market data is being

around us. In this example the stock market data is being presented in a handheld device; however, more

importantly, the data is being represented as numbers and a graph showing the up and down trends of the market. For this type of data, this is a standard presentation of the information and through the use of a functional visualization such as this, the data becomes as straight-forward as necessary. According to Gershon et al, since information in these visualizations is abstract data and doesn't have a natural physical representation, one of the research problems is discovering new visual metaphors to represent the information [1]. These visual metaphors, in terms of functional information visualizations, refer to the way in which the data can be presented. In many cases this refers to graphs, and charts, and other contemporary standards for visualizing data; however, over time there began to develop visual metaphors that required more immersion by the user to understand. This is the beginning of aesthetic information visualizations. Rather than aiming for efficiency and understandability, these visualizations attempt to present the data by using abstract visual metaphors, and artistic



Figure 2. Example of an aesthetic visualization. Stock market data displayed using a "spiral" metaphor (Gavaria 2008).

elements. These visualizations tend to be more appealing to the user due to the aesthetical value of them and also they create a deeper level of immersion because they cannot be understood by simply glancing over the visualization.

## **1.2 ARTISTIC VISUALIZATION:**

Artistic visualizations are a fairly new branch of Information Visualizations. It is referred to by many different names, but informative art and ambient art are two of the terms that are used to describe this specific field. According to Otjacques et al, "Holmquist and Skog [2003] restate it as follows: 'informative art is a type of computer applications which borrow their appearance from well-known artistic styles to visualize dynamically updated information' [3]." This is an interesting concept because it defines informative art as not only using artistic elements to present the information; it distinguishes between what kind of data it is. In this case it mentions

dynamically updated information, which means information that is being constantly updated and is constantly changing. Though this is true for much artistic visualization, this definition is not necessarily true because static data can always be used in these visualizations. Ambient art is defined in the same paper as 'the aesthetic presentation of information, using artistic techniques to achieve a pleasing image that also contains hidden depths, where exposure to it over time allows a viewer to understand something about the information sources that it represents' [3]. This definition for ambient art is the definition that applies most to these types of visualizations. The power of artistic visualization lies in its ability to immerse the user differently than presenting data in an efficient manner can. Using the visual metaphors discussed earlier, these visualizations contain 'hidden depths' and by observing them for extending periods of the time, the hope is that these visualizations can present something interesting about the data that it represents. This idea of artistic or aesthetic visualizations has caused much debate and the main question behind this debate is when can visualization be considered art?

### **1.3** When are Information Visualizations Art?

Despite the debate, the significance of the connection between art and information visualizations is clear. There are many researchers who believe that art and creative design can enrich the field of information visualizations. According to A. Vande Moere, "Form Follows Data: The Symbiosis Between Design and Information Visualization", "information visualization should be enriched with the principles of creative design and art, to develop valuable data representations that address the emotional experience and engagement of users, instead of solely focusing on task effectiveness metrics. [4]" This idea is what drives the counterargument to information visualization visualizations not being considered art; they engage the user through visual metaphors and

Page | 4



Figure 3. Two posters from, "If the world were a village of 100 people". World Statistics (infosthetics.com)

interaction. Moere suggests that, "*Metaphors* in visualizations are used to help users understand systems in conceptual terms they already know, by appealing to initial familiarity and experience. [4]" This is the fundamental difference between artistic visualizations and functional ones. Rather than presenting the data for the purpose of easier readability and understanding, by using metaphors and artistic presentation the user is engaged in the visualization in a more exploratory way. They have to search for the understanding of the visualization through the metaphors shown. Moere states, "Data can be presented by very abstract metaphors, which appeal to eye, are user engaging and provoke novel ways of exploration. We propose that such data representations need to be assessed using evaluation criteria that move away from traditional efficiency considerations: maybe users need more time to fully understand them, but enjoy this (longer) time span more, learn complex insights by playing, retain information longer or like to use the application repeatedly. [4]" So his main argument is that creative design enhances the way that the users can engage in data visualizations and because of this deeper level of engagement, they cannot be evaluated based on efficiency. An example can be seen through Figure 3; this figure represents 2 posters in a series of 20 posters that each focus on specific world issues. The twist is that the theme and context of all the images is representing world issues if the world were a village of 100 people. In this case the visualization of the percentages of these issues is simplified into small numbers; however, the visual aesthetic and motifs are generally clever pictures that have some relative connection to the topic which they represent. For example, the first image in Figure 3 is a zebra and the topic is skin color; the numbers and stripes of the zebra correlate to number of non-white and white people in the village of 100. This example is a perfect example of the amount of immersion that aesthetic visualizations allow because as simple as this data is, by imagining the world as 100 people it allows the statistics to be felt on a completely different scale and the visuals lend to an enjoyable, explorative experience. This paper supports the argument that aesthetic data visualizations can be considered art concluding by stating, "design is starting to force information visualization approaches to consider interdisciplinary issues, such as the engagement and mental immersion of users, the use of design cognition insights and the role of aesthetics in the effectiveness of data representations. [4]" This statement shows the ways in which art can enhance visualizations; it takes into account the many different ways that the visualization can affect the user. This idea is at the heart of aesthetic information visualization; creating a product that grabs the user's attention and engages them through the exploration of hidden metaphors. This is the foundation on which this project is based. By creating visualizations using these types of metaphors I hope to highlight interesting aspects of data about Union College. My hope is that by creating a deeper level of user immersion, that after exploring my visualizations, the user can discover what I'm trying to highlight about the data as well as be informed about the content of my data.

# 1.4 MY PROJECT

The focus of my project as stated earlier is to create information visualizations using Union College data. To narrow down what my visualizations are about, here are a few of the types of data that I have worked with: demographics, race, waste, and energy. I created 5 different visualizations using data that I've gathered about Union College. Of these visualizations there are two main implementations that I used to create them, interactive digital visualizations and digital prints. The significance of these visualizations is not only in their aesthetic value as art, but also in their ability to represent data through visual metaphors. By analyzing the data sets chosen, my hope was that I could represent the trends in the data using obscure or abstract metaphors; hopefully this would cause the user to need to spend time and explore the visualizations to decipher the meaning of the data and notice what is trying to be highlighted about the information. The presentation of all of the visualizations will be in the form of an art exhibition.

### 2 MOTIVATION

My project was influenced by several other works in the field of Information Visualization. All of the works that have influenced my project come from a blog of information visualizations called *Infosthetics*. All of the works posted on this site are information visualizations, some are more functional in purpose, but many of them are artistic/aesthetic visualizations. There a variety of topics and themes; however, the ones that stood out the most and influenced this project heavily, were those that focused on world data and represented it in interesting ways. Now I will

Page | 7

outline 3 of the major works that influenced my project and explain how their influence affected my project.

## 2.1 Dynamic Data

Seen in Figure 4 is the perfect example of dynamic data visualization. In this example, there is a large map of America with various "temperature" spots represented by colors ranging from green



Figure 4. US gas price temperature map. Dynamic data map that tracks gas prices in all regions of America (Infosthetics).

to red. The temperature is the metaphor for price of gas, ranging low (green) to high (red). When saying that this is a dynamic visualization, that means that the data is constantly updating and staying current. One can see how this would be a convenient visualization, especially in the past Page | 8 when gas prices were an issue. The visualization itself is impressive because it allows you to type in zip codes and zoom in to any location across America and it seems to contain all of the gas stations in the regions. Visualization on a scale such as this requires a significant amount of



Figure 5-1. Average American consumer spending (Infosthetics).

resources and also some way of mining the data in order for it to remain current. For the purpose of my project these types of visualizations were avoided. Based on the data that I was able to obtain it was not possible to create dynamic visualizations; therefore, all of my visualizations are based on static data.

## 2.2 Static Visualizations

Since this was the way in which the visualizations for this project were created, influence from other works using static data was extremely important. The first example is seen in Figure 5-1, the Average consumer spending of Americans. This piece stood out because of its topic. The



data is separated into its separate categories of spending such as, Apparel, Health care, Housing, Education, Food, Transportation, etc. The major categories of American spending are the larger sections and then in each of these sections it breaks down to specific products and types of each category. This is the most interesting part of this visualization,

Figure 5-2. Detail of Figure 5-1. Zoomed in perspective

not only does it break percentages down to the overall categories of spending, but also within each category the percentages are broken down amongst the specific types. This break down of percentages is shown in Figure 5-2. This visualization has influenced this project because it has found a simple yet extremely attractive way to display

static data. Since this is an interactive piece it allows the user to explore and actively look around the visualization to get the break down of the data, this is the most important aspect of this visualization. From this example I take the importance of interactivity and exploration as a means of immersing the user in the data. The next visualization that has influenced this project is called The Health Visualizer. Seen in Figure 6 below, the Health Visualizer is another interactive visualization; however, it implements a different interaction than the example above. Instead of exploring a visualization that has all of the data spread out in an open format, this visualization implements interaction through comparisons. There are buttons related to three major categories, Demographics, Risk factors, and Diseases and Conditions. Within each of these categories there are several choices, such as Male or Female in Demographics and Diabetes or Stroke in Diseases



Figure 6. Health Visualizer (Infosthetics).

and Conditions. Amongst all of these choices, you can pick any two and the visualization will reorganize the symbols and colors to represent the data as a comparison between the two variables chosen. This visualization is interesting because of its aesthetical quality; also the moving and shifting of the figures give an enjoyable transition between different comparisons. In addition to the aesthetics of the visualization, by focusing on presenting the data through comparisons, it adds a new depth to the visualization by allowing the user to choose what data is presented to them. It is another different, yet effective tool of user immersion because to analyze the data and see what the visualization is trying to represent, the user is forced to invest time in exploring and choosing variables to visualize. From this piece my project gains another option for implementation of interaction. The last example that will be displayed is related to Figure 3 which is about the world visualized as a village of 100 people. In this case, rather than implementing this visualization as posters like the creator of the visualization in Figure 3 did, this version is visualized as a miniature world with individual areas relating to different topics. As seen in Figure 7-1, the world is broken down into 5 zones, Life, Economy, Food, Danger, and

Main. Each of these zones relates to world statistics related to the name of the zone. When one clicks on a specific zone, the user is transported to a larger version of the zone that is inhabited by these legolike people. In each of the zones however, interaction with buildings by clicking



Figure 7-1. If the World were a Village of 100 People (Infosthetics).

or with people by hovering over them creates popup boxes with specific information regarding the zone and relating back to the metaphor of the world as 100 people. For example in Figure 7-2, the detail shows the total world population using the metaphor of the world as 100 people. Therefore, in a pixilated representation of the continents, the visualization shows, using 100 of these small figures, the diversity of the world population. On each of the continents the



Figure 7-2. Detail of the Main Zone. Total world population selected (Infosthetics).

percentages of white (American), African, Asian, and European peoples are shown. The significance of this visualization is that by breaking the numbers down to 100 people, the world statistics have more power on

the user because the statistics are simplified to a very small group. This allows the user to

feel the truth of the statistics much more than they would over a world population of millions. Also this is a highly interactive piece, but more important than the interaction in this case is the metaphor. In this visualization the metaphor and the aesthetic, as always, have the most power over the user because they present information that never could be understood on such basic, simple levels in a highly intriguing way.

These examples are a few of the visualizations that have influenced the way in which my project implements its data. Through a combination of my own aesthetics and metaphors, the implementation of my visualizations will most likely use a combination of skills and techniques used in these examples; more specifically, interaction, heat mapping techniques, and also the concept of "The World as 100". The hope is that like these examples, my own visualizations can portray the information in a way that the user can understand through exploration while also

creating an enjoyable experience that will make the user want to spend time exploring the visualizations.

#### 3 METHOD

For this project the method of creating these visualizations was through programming. The language that was used is the Processing language. The Processing language is a programming language with many similarities to languages like Java; however, the purpose of this language is to create visual output. Because of my programming experience in languages like Java, and the small experience that I've had with Processing prior to beginning this project, using this language was an obvious choice.

#### 4 MY VISUALIZATIONS

The final results of my project were 5 visualizations, 4 which are digital prints and 1 that is an interactive piece. Each of the visualizations is based on data related to Union College. The following sections will display each of the visualizations and explain some of the motivation behind each of the metaphors that were used to represent the data.

## 4.1 UNION AS 100

The first visualization that I created, seen in Figure 8, is based on Union's population and graduation rate data. This visualization is called "Union as 100" because it applies the concept of "The World as 100", as seen in Figure 7.1, to data about Union College. This decision was made based on the idea that by applying this metaphor to data about Union, which is quite a small institution, it would be possible for the data to appear extremely powerful. In my project I was

limited by what data I was able to receive therefore, my decision to use graduation rates data was solely based on the fact that it was the only other data that I had that related to the statistics that I



Figure 8. Union as 100. Total Union population represented as

used to break down Union's population to 100 students. The image is made up of 100 shoes, several of which are

different colors as well as some that are tied and some

that are untied. The colors represent various races, while the laces represent whether or not that individual student graduated. The algorithm for this visualization plots shoes in a grid pattern



shoes.

while determining, based on the data that was read in, what type of shoe to create. It cycles through each race and plots all students who graduated and then those who didn't until all 100

Figure 8-2. Union as 100. Zoomed in detail of shoes.

shoes have been placed. Shoes were chosen to represent people because I didn't want to use an

obvious symbol or image and I also wanted to have an object that naturally had the ability to display several sets of data through its structure. Shoes were a natural choice because they are an article of clothing and therefore something that can be associated with people; in addition to the fact that laces allowed for an interesting metaphor and way to represent the data.

# 4.2 HEAT MAP

This visualization is based on Union's gas and heat data. Using data that provided the cost of heating for each building on campus I comprised a "heat" map of Union College. The motivation behind this visualization is based on influence from other visualizations that use a heat technique



Figure 9. Union Heat Map. Heat indication map of "hottest" buildings on campus.

to differentiate between changes in the data. An example of this can be seen in the temperature map seen in Figure 4. Unlike these visualizations, I took the idea more literally based on the fact that I was actually using heat data and it was one of the most straight-forward ways to represent it. To create the visualization, locations for each heat indicator were hard-coded into the program and the color of the indicator is solely based on the total cost of heating that specific building. Something that I thought about implementing in this visualization was taking into account how large each building is, because obviously it takes more to heat the larger buildings; however, I wasn't able to obtain all the data I needed to make that happen. The most expensive buildings to heat are closer to red while the ones that cost the least on campus are a light shade of blue-green. This visualization represents the data in a much more literal way than any of my other visualizations. Because of this fact, to me, this visualization is the least strong in terms of its consideration as art; however, in designing the visualizations for this project I tried to take into account how informative they were. Some are meant to portray abstracted data so that the viewer has to explore the visualization to completely understand the data; this visualization, on the other hand, is meant to give the viewer an understanding of where heat flows on our campus. It is strictly for an informative purpose, which hopefully in conjunction with the following visualization can be seen.

### 4.3 GAS LIFELINE

This piece is a series of visualizations, 5 in total, that are based on the exact same data that the Heat Map uses. Instead of portraying the data as yearly totals this visualization uses the monthly data for 5 different years to produce 5 visualizations that plot the monthly costs of gas on Union's campus. The algorithm for these visualizations takes in a set a data and plots the peaks based on cost. All peaks (costs) are plotted on the same scale throughout the 5 visualizations; therefore, each peak can be compared to the next. The motivation behind these visualizations is that by seeing the rises and falls of the last 5 years of Union's gas usage, we can begin to understand what the "lifeline" of gas has been at Union during the time. Obviously much is dependent on the type of winters we have but by looking at years like this, you can see clear difference between how much Union spends on a yearly basis. The metaphor of a heart monitor lifeline was used because I thought about this visualization as assessing Union's sustainability

Page | 17

health. In which case, by seeing the years mapped as lines, we could easily determine whether or not Union, in terms of sustainability and more specifically heat, has gotten more or less healthy over the years.

\*\*Figure 10 on following page\*\*



Figure 10. Gas Lifelines. Left to Right are the last 5 academic years starting from 2005-2006 and ending at 2009-2010.

## 4.4 WASTE PILE

This visualization, seen in Figure 11, is a visual representation of the amount of waste Union produces. Using data that provided total weights for Bottles & Cans, Paper, Cardboard, and Solid Waste for the years 2008 and 2009 I created two of these visualizations, one for each year. As you can see the visualization contains 4 sections. Each section in the visualization represents a weight, in tons, of a type of waste. The algorithm takes the total pixel height and designates a percentage of the height for each type of waste. This percentage is determined by assuming that the total waste will take up the entire pixel height of the visualization; therefore, the percent that each type of waste is from the total waste is applied to the pixel height. The types of waste are represented by actual objects that fall into each category. The images have been scanned in and the algorithm random selects which images it will use when creating the visualization. I created this piece because this was data that I was very excited to obtain and use; however, rather than getting total numbers types of waste, I got weights. With these weights I needed an interesting way to represent them without being too obvious; therefore, I came up with the idea that I could represent weights by size. This was the idea that inevitably led to using pixel height to imply weights. This piece when viewed with the other visualization for the following year, should allow the viewer to really have a powerful view of how much we waste on Union's campus. The discrepancy between recycling and waste is incredible, but more importantly with both visualizations the viewer will also be able to see what progress, if any, has been made between the two years.



Figure 11. Waste Piles. These are the visualizations of Waste from 2008 and 2009.

### 4.5 FACT FINDER



Figure 12. Fact Finder

My last visualization is my only interactive piece. The screenshot seen in Figure 12 displays the interface. This visualization takes input from the user and produces output based on what the user has interacted with. The user can draw a line with the mouse and when that line crosses a character in the

visualization, it takes that character and every character that follows that first one in a sentence reveals itself. This means that the visualization is made up of a bunch of different sentences, each which is a random fact/tidbit about Union. The motivation behind creating this piece was that it was the best way I could find to include as much of the data I had gathered as possible in my project. This visualization holds random facts about Union and to discover them, the user must explore the characters in the visualization and interact with the correct ones. This piece works by using a library that applies behavior to text. So each character has behavior attached to it that allows it to become the leader or a follower. A character becomes the leader if it is the first character that is interacted with by the Mouse Path; when this happens all of the remaining letters in the sentence that includes that character, fall into place because they are assigned the follower property. This is the simplified algorithm of this visualization. Overall, this visualization has some quirks to work out, most importantly is trying to implement a way for it to indicate if a statement found in the text is the full statement or just part of it, but it has provided me a means to include as much extra data as I can possibly find; in addition to the fact that it fits my theme perfectly because of its exploratory nature.

### 4 CONCLUSION

This paper has given some background on the field of Information Visualization and its application and significance in the world of artistic visualizations. My own visualizations have been a result of influence that I've taken from works that have motivated me as well as by techniques that I believe have been repeatedly applied to this genre of visualizations. My goal in the beginning of this project was to obtain data related to Union College and use it to make visualizations that would immerse the user and allow for an experience that was both exploratory and informative. I believe that this goal has been accomplished by my visualizations and I anticipate the reception of these pieces when they are displayed in my exhibition. Overall it was quite intriguing working with data about Union and that decision was the one that gave direction to this project. I have learned an immense amount about the institution that I've resided for the last 4 years and to have gotten to work with information regarding this institution was quite rewarding. The 5 visualizations I have created will be presented in an art exhibit during the spring 2010 from May 17-25 in the Atrium Gallery in the Arts building.

# WORKS CITED

- 1. N. Gershon, S. Eick, S. Card, "Information Visualization," ACM interactions, Volume 5 Issue 2, March 1998 [1]
- 2. Andres Ramirez Gaviria, "When Is Information Visualization Art?: Determining the Critical

Criteria," Leonardo, Volume 41, Number 5, October 2008, pp. 479-482 (Article) [2]

**3.** B. Otjacques, F. Feltz, "Redesign of classic information visualization techniques in an artistic computing perspective," ACM, July 2007 [3]

4. A. Vande Moere, "Form Follows Data: The Symbiosis Between Design and Information

Visualization," in Proceedings of CAAD Futures (Vienna: OKK Verlag 2005) pp. 31-40 [4]

## Links to Infosthetics Visualizations

- The World of 100: If the World were a Village of 100 People. (Figure 3). http://infosthetics.com/archives/2009/06/the\_world\_of\_100.html
- **2.** US gas price temperature map. (Figure 4).

http://infosthetics.com/archives/2008/05/us\_gas\_price\_temperature\_map.html

**3.** Average American consumer spending. (Figure 5).

http://infosthetics.com/archives/2008/05/average\_american\_consumer\_spending.html

- 4. Visualizing the Major Health Issues Facing Americans Today. (Figure 6). <u>http://infosthetics.com/archives/2009/05/visualizing\_the\_major\_health\_issues\_facing\_am\_ericans\_today.html</u>
- 5. If the World were a Village of 100 People. (Figure 7).

http://infosthetics.com/archives/2009/01/if\_the\_world\_were\_a\_village\_of\_100\_people\_1.

<u>html</u>