

White Paper

Principles of Data Visualization - What We See in a Visual

Table of Contents

Why We Visualize Data	3
The Importance of Visualizations in Business	3
Two Goals in Data Visualization	4
▶ Explain Data to Solve Specific Problems	4
▶ Explore Large Data Sets for Better Understanding	5
Eyesight and Memory Working in Parallel	6
▶ The Role of Memory in Vision	7
Preattentive Attributes Used by Our Working Memory	8
Forming Analytical Patterns Out of Preattentive Attributes	9
Using the Gestalt Principles to Bring Out Patterns in Visualizations	11
An Example from Recorded Future	12
Gestalt Principles	13
Preattentive Attributes	14
Conclusion	14
About FusionCharts	15

Why We Visualize Data

We visualize information to meet a very basic need - to tell a story. It's one of the most primitive forms of communication known to man, having its origins in cave drawings dated as early as 30,000 B.C., even before written communication, which emerged in 3,000 B.C. Vision is the single most important faculty we use to communicate information.

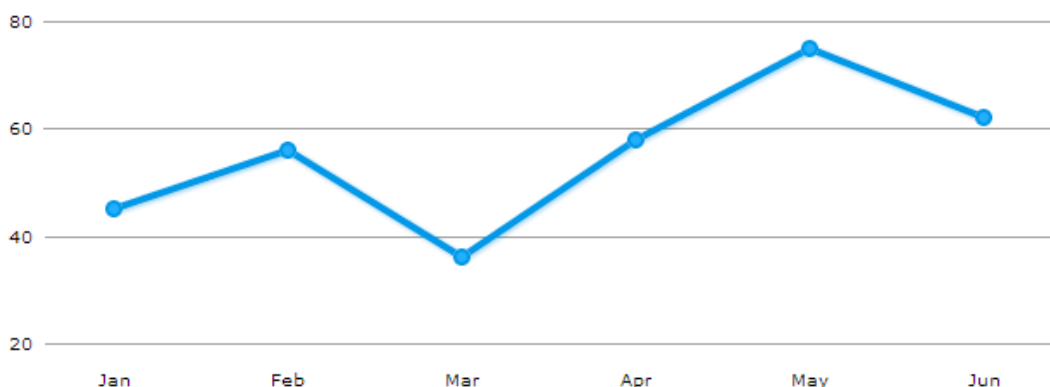
With the passing of time, we've found new ways to visualize information. Today, we're familiar with the basic chart types like the line chart, bar chart, and pie chart. However, we rarely stop to think about why they're more effective than bland tables, text, and numbers. Further, we can't easily spot instances when they're done wrong, and can be improved. This white paper aims to give you a foundational understanding of how we process visual information. You'll be able to use this knowledge when designing your next dashboard, and have an informed opinion on how to communicate even more clearly and powerfully using visualizations. Even if you don't design visualizations yourself, these principles will equip you with skills to better analyze visualizations you come across in the newspapers, or in your daily work reports. Let's start with an understanding of why data visualizations are so important in a business context today.

The Importance of Visualizations in Business

A visual can communicate more information than a table in a much smaller space. This trait of visuals makes them more effective than tables for presenting data. For example, notice the table below, and try to spot the month with the highest sales.

Month	Jan	Feb	Mar	Apr	May	Jun
Sales	45	56	36	58	75	62

This data when visualized gives you the same information in a second or two.



Edward Tufte, a data visualization expert, says 'Graphical excellence is that which gives to the viewer the greatest number of ideas in the shortest time with the least ink in the smallest space.' This trait of visualizations is what makes them vital to businesses.

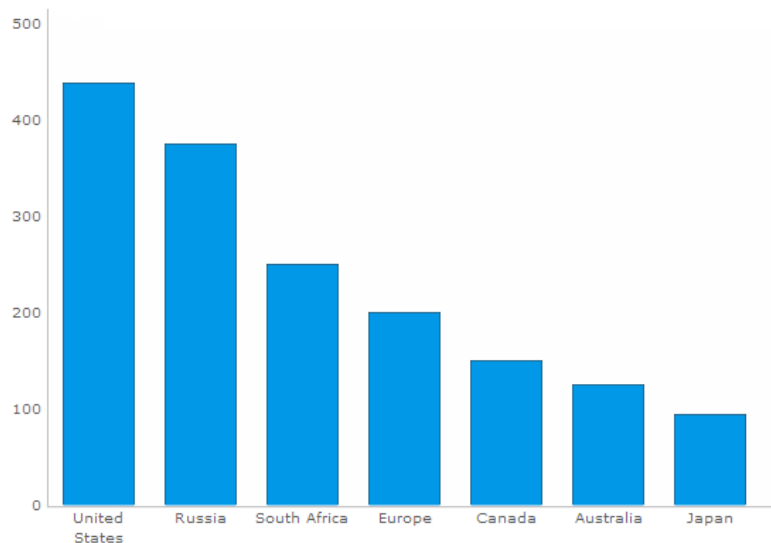
Two Goals in Data Visualization

In a business environment, visualizations can have two broad goals, which sometimes overlap.

- ▶ Explanatory
- ▶ Exploratory

Explain Data to Solve Specific Problems

Visuals that are meant to direct the viewer along a defined path are explanatory in nature. The bulk of business dashboards that we come across in day-to-day scenarios fall in this category.



Here, the viewer starts with a question in mind. For example, 'which country has the highest value?' She then views the visualization, and finds an answer to the question.

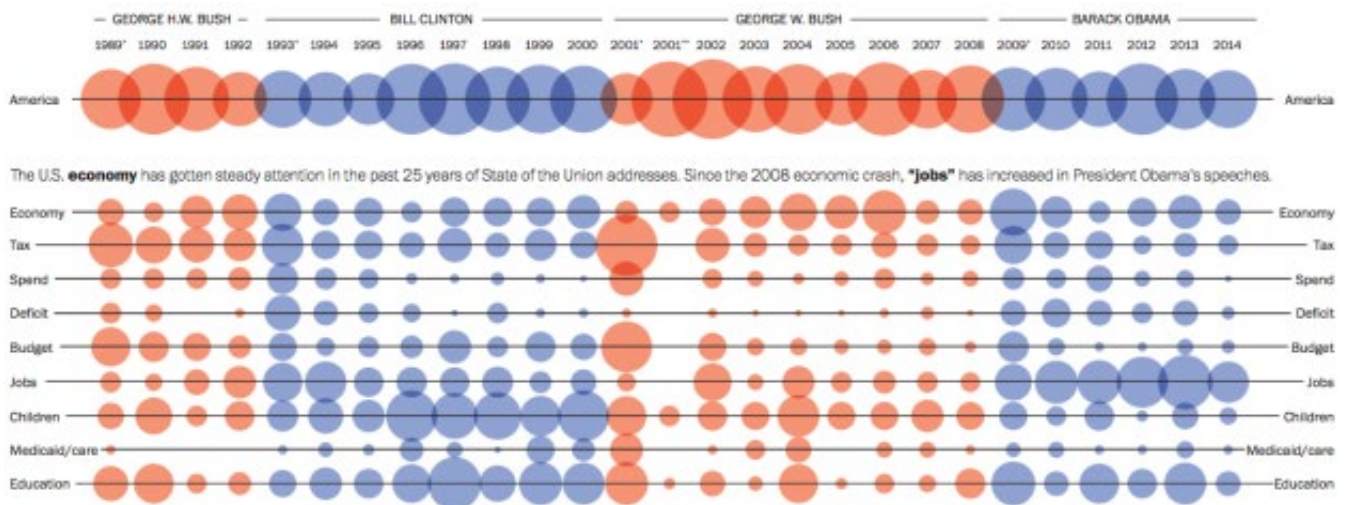


This type of visualization is used in many scenarios for the following tasks:

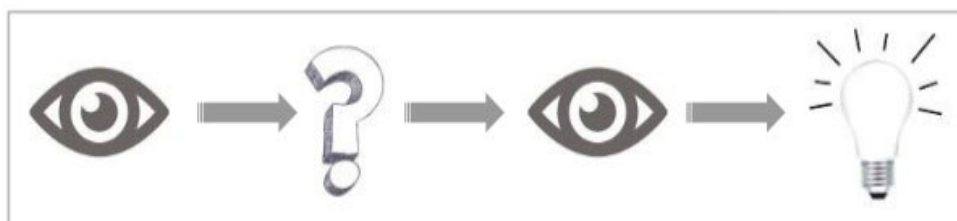
- ▶ Answer a question. E.g., How much sales did we have last quarter?
- ▶ Support a decision. E.g., We need to stock more football jerseys as they were sold out on most days last week
- ▶ Communicate information. E.g., Revenue is on track for this quarter
- ▶ Increase efficiency. E.g., 'Technical specifications' is the most viewed section in the product page. It should be given more visibility.

Explore Large Data Sets for Better Understanding

Exploratory visuals offer the viewer many dimensions to a data set, or compares multiple data sets with each other. They invite the viewer to explore the visual, ask questions along the way, and find answers to those questions. For example, below is a visualization of the State of the Union address of recent Presidents.^[1]



Here, the viewer starts by familiarizing herself with the visualization, then identifies an area of interest. For example, which President spoke more about jobs? She then explores the 'Jobs' section of the visualization, and finds her answer. She could then move on to exploring other areas of the visualization.



Exploratory analysis can be cyclical without a specific end point. Viewers can find many insights from a single visualization, and interact with it to gain understanding rather than make a specific decision. This type of visualization can accomplish the following tasks:

- ▶ Pose new questions
- ▶ Explore and discover

Though they're not as popular as the previous category, exploratory visualizations have gained prominence in recent years with the rise of big data. The high volume of data, and varied data sets that have become common today lend themselves easily to exploratory analysis.

Now that we're aware of what makes visualizations intrinsic to business, let's dive into the mechanics of how we process visual information. We'll understand the role of memory in perceiving visual information, and how to apply that understanding as we work with visualizations.

Eyesight and Memory Working in Parallel

If you're familiar with big data terminology you would have come across the term 'massively parallel processing' (MPP), as most commonly seen in MapReduce technology that powers Hadoop. It breaks data down into small units, and processes each of the units in parallel. A similar process has always been used by us to process information with our vision as well. When we look at a visual, our eyes and our brain work in parallel to take in new information, and break it into small chunks. Then both the eyes and the brain process the chunks in parallel to find meaning. Let's look at an example to better understand this.



Let's say we walk into a supermarket to buy oranges. Our eyes first scan the layout of the supermarket. At the same time, our brain processes the various sections of the layout, and instructs the eyes to zone in on the fruits section. It does this by sending signals about how fruits look from memory. The eyes then break the entire scanned area into parts, and scans each part to spot the fruits section. The same process is repeated till we zero in on the oranges in the fruits section. This process of visualizing information is performed by the eyes and memory working in parallel.

The Role of Memory in Vision

While that's an overview of how we process visual information, let's discuss the vital role of memory in our vision. There are two types of memory that come into play when we process visual information.

- ▶ Long-term memory
- ▶ Working memory

Long-term Memory

This type of memory is what makes us always expect the units to be marked off on the X and Y axes of a chart, or the date range selector at the top of a dashboard. This type of memory is formed by past interactions and experiences. It needs to be considered when designing the layout of a dashboard or visualization. There should be good reason to go contrary to long-term memory when deciding the basic structure of a dashboard or visualization. A discussion of long-term memory is out of the scope of this white paper. Instead, we'll focus on the other type - working memory.

Working Memory

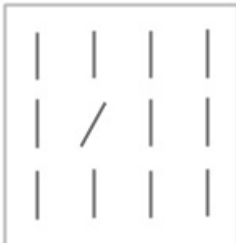
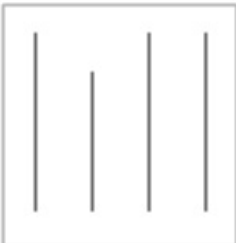


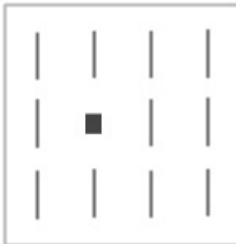





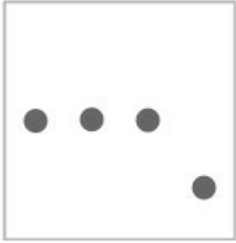
When we look at a line chart, or notice a number in a dashboard, we use our working memory to store just the information we need at the moment. This type of memory breaks the entire visual into small chunks of information, in a process appropriately called 'chunking.' Surprisingly, for all the complexity of our brain, our working memory can hold only about 3 chunks of information at any given time. When perceiving a complex visual, we're constantly replacing the 3 available slots in our working memory. This is why when designing a visual or a dashboard, one of our goals is to limit the number of prominent chunks of information that we want the user to notice, and pack as much information as possible into each chunk. While we want to avoid information overload, a sound understanding of this concept can help us direct the attention of the viewer in a natural way.

Next, Let's discuss working memory in depth. Just as any language has alphabets, which are combined to form words, the language of data visualization consists of certain basic building blocks called preattentive attributes.

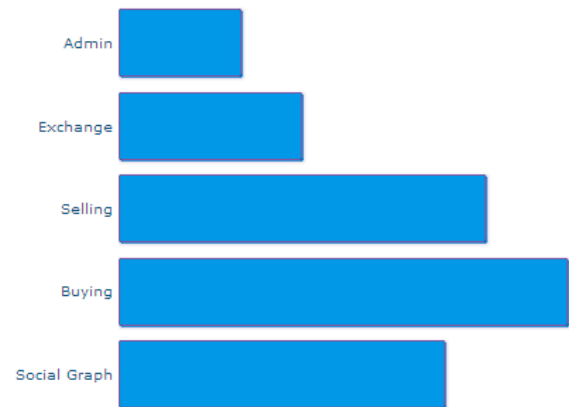
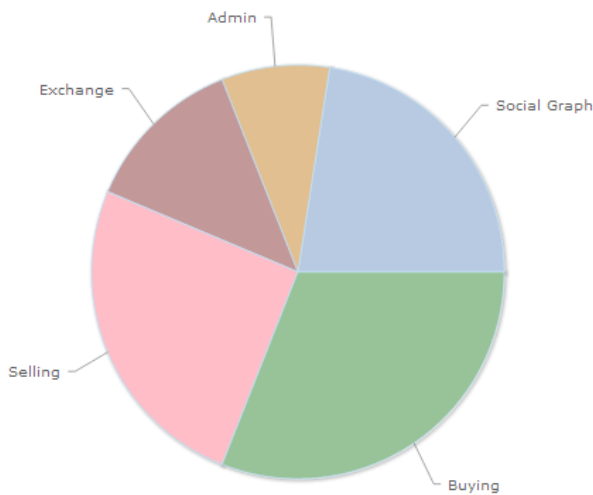
Preattentive Attributes Used by Our Working Memory

Colin Ware, Director of the Data Visualization Research Lab at the University of New Hampshire, terms the basic building blocks of the visualization process as 'Preattentive attributes.' These attributes are what immediately catch our eye when we look at a visualization. They can be perceived in less than 10 milliseconds, even before we make a conscious effort to notice them. Here's a list of the preattentive attributes:

Form

Orientation	Line Length	Line Width	Size
			
Shape	Curvature	Added Marks	Enclosure
			
Color	Spatial Position		
Intensity	Hue	2-D Position	
			


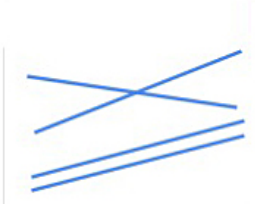






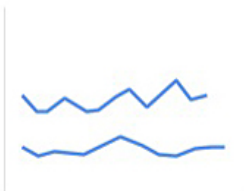

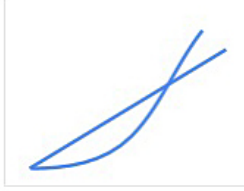

These attributes come into play when we analyze any visualization. Of this list, only Position, and Length can be used to perceive quantitative data with precision. The other attributes are useful for perceiving other types of data such as categorical, or relational data. For example, both the pie chart and the bar chart below show the same data. But you can't easily tell from the pie chart which is the biggest pie. That's more clearly visible in the bar chart as it calls on the preattentive attribute of length.



Considering preattentive attributes can help when deciding which chart type to use for our data. While preattentive attributes are what we immediately identify in a visual, they aren't the only things we notice. We go on to form patterns out of the preattentive attributes.

Forming Analytical Patterns Out of Preattentive Attributes

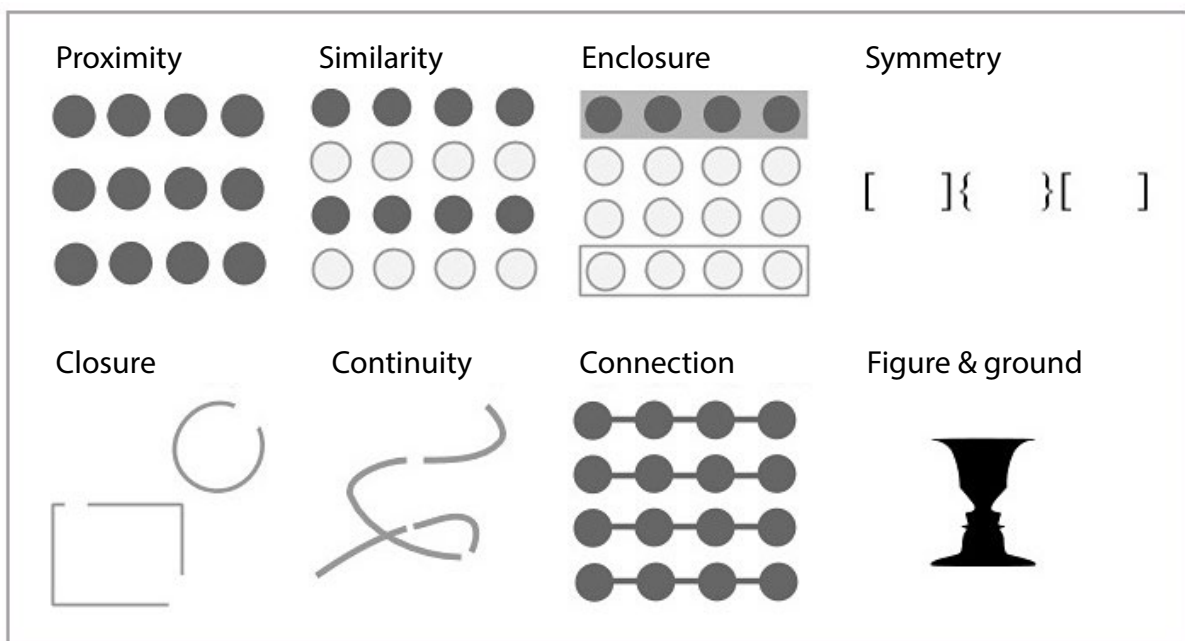
If preattentive attributes are the alphabets of visual language, analytical patterns are the words we form using them. We immediately identify the preattentive attributes in a visualization. We then combine the preattentive attributes to seek out analytical patterns in the visual. Here are the basic analytical patterns that we identify when looking at a visual:

Pattern	Example	Pattern	Example
High, low and in between		Non-intersecting and intersecting	
Going up, going down and remaining flat		Symmetrical and skewed	
Steep and gradual		Wide and narrow	
Steady and fluctuating		Clusters and gaps	
Random and repeating		Tightly and loosely distributed	
Straight and curved		Normal and abnormal	

We've seen how preattentive attributes and patterns enable us to process visual information. However, when crafting visualizations we often want to highlight certain patterns over others. In these cases, Gestalt's Principles come in handy.

Using the Gestalt Principles to Bring Out Patterns in Visualizations

Gestalt principles describe how our mind organizes individual elements into groups. We can use these principles to highlight patterns that are important, and downplay other patterns. The image below illustrates the principles of Gestalt which are relevant to visualization.



Here's what we notice from each of the illustrations:

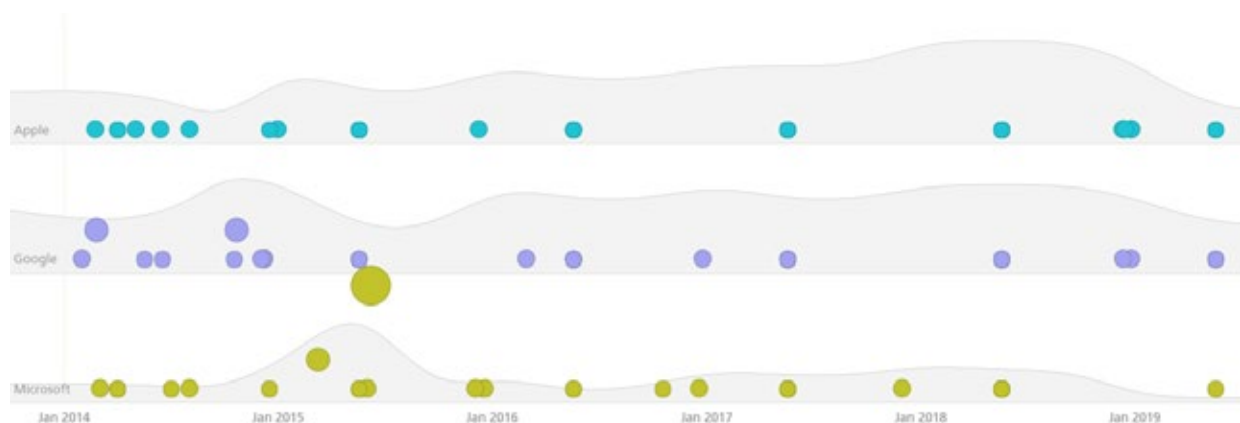
- ▶ **Proximity:** We see three rows of dots instead of four columns of dots because they are closer horizontally than vertically.
- ▶ **Similarity:** We see similar looking objects as part of the same group.
- ▶ **Enclosure:** We group the first four and last four dots as two rows instead of eight dots.
- ▶ **Closure:** We automatically close the square and circle instead of seeing three disconnected paths.
- ▶ **Continuity:** We see one continuous path instead of three arbitrary ones.
- ▶ **Connection:** We group the connected dots as belonging to the same group.
- ▶ **Symmetry:** We see three pairs of symmetrical brackets rather than six individual brackets.
- ▶ **Figure & ground:** We either notice the two faces, or the vase. Whichever we notice becomes the figure, and the other the ground.

These principles can help us perform many tasks such as reduce the noise from charts, choose the ideal aspect ratio, and show relationships between elements more clearly. The Gestalt principles are a powerful tool to have when crafting visualizations.

Now that we have a good understanding of how visualizations work, let's apply these concepts to a real-world example.

An Example from Recorded Future

We'll consider the following visualization from Recorded Future, a web intelligence company.^[2] This visualization compares the mentions of Apple, Google, and Microsoft across the web. The data here is forward-looking, and considers only mentions related to the next five years.



This visualization features two chart types - An area chart, which is grayed out in the background, and a bubble chart, which is color-coded in the foreground. Let's analyze this simple visualization, and identify which elements from this white paper it uses.

Gestalt Principles

Figure & ground

The first thing you notice when looking at this visualization is that the bubbles stand out against the backdrop of the area charts. This shows the Gestalt principle of figure & ground.

Proximity

Zoning in on the bubbles shows 3 distinct groups of bubbles. We can identify this easily because of how close the bubbles are to each other.

Similarity

Further, we notice that the bubbles are of three colors - green, purple, and blue. This similarity brings out the grouping even more clearly.

Analytical Patterns

Going up, going down, remaining flat

This pattern is most visible in the area chart. We notice that the overall trend for Apple is upward, while Google's Microsoft's stays flat.

Wide, narrow

There are many peaks along the area chart, but one of them for Microsoft is particularly noticeable in Jan 2015. This peak is narrower than the other peaks.

Tightly, loosely distributed

We notice the loose distribution of bubbles for all three companies around 2018 and 2019.

Normal, abnormal

Zoning in on the bubbles for Microsoft, we notice an outlier for Microsoft in mid-2015. The other bubbles fall within a more normal range.

Preattentive Attributes

Spatial position

We use the preattentive attribute of position to track the rise and fall of the area chart. Similarly, we notice the abnormal bubble in Microsoft's chart because of its higher position compared to the other bubbles.

Size / Area

The bubbles vary in size. Their size corresponds to the number of web mentions for a particular topic. This makes it easy to spot the important mentions, and explore them in detail.

Hue / Color

As mentioned earlier, the color of the bubbles makes it easy to classify them into three groups. This employs the preattentive attribute of Hue.

Intensity

Finally, the low intensity of the area chart places it in the background, giving priority to the bubbles.

Being aware of the preattentive attributes, analytical patterns, and Gestalt Principles can make a visual come alive to us. Considering we analyze visual information on a daily basis, and how important it is to business, it pays to know how visualizations work.

Conclusion

In this white paper, we dissected the process of visual perception. First, we saw that all visualizations have a goal - explanatory, or exploratory. Second, preattentive attributes are the basic building blocks of a visualization, and are identified by us almost immediately. Third, preattentive attributes lead us to spot patterns in a visual. Finally, when crafting visualizations we could use the Gestalt Principles to prioritize important patterns, and downplay the noisy ones.

In conclusion, it's obvious that we are naturally hard-wired to visualize information in a certain way. Understanding those basic principles of data visualization will help us craft outstanding visualizations, and tell compelling stories, much to the delight our colleagues, and end users.

About FusionCharts

FusionCharts Suite XT is the industry's leading enterprise-grade charting component with delightful JavaScript (HTML5) charts that work across devices and browsers (including IE 6, 7 and 8). Using it, you can create your first chart in under 15 minutes and then add advanced reporting capabilities like drill-down and zoom in a couple of hours after that. It comes with extensive docs, demos and personalized tech support that makes sure implementation is a breeze for you.

22,000 customers and 500,000 developers in 120 countries including organizations like NASA, Microsoft, Cisco, GE, AT&T and World Bank use FusionCharts Suite XT to go from data to delight in minutes.

Learn more about how you can add delight to your products at www.fusioncharts.com



Reference:

1. The Washington Post. History Through the President's Words ^[1]
2. Recorded Future. Apple, Google, and Microsoft during the next five years ^[2]

Further Reading:

1. Tufte, Edward. (2001). The Visual Display of Quantitative Information, 2nd Edition. Graphics
2. Press.Ware, Colin. (2008). Visual Thinking. Morgan Kaufmann.
3. Ware, Colin. (2012). Information Visualization, 3rd Edition. Morgan Kaufmann.
4. Few, Stephen. (2009). Now You See It: Simple Visualization Techniques for Quantitative Analysis. Analytics Press.