Visualization Of Data Science

What makes a Good Visualization?

1. In the perception article:

A). What is perception?

Perception, in a sense, makes user experience design possible. If no one is viewing, perceiving or experiencing your product, does the product exist?
Perception is in play from the moment your customers learn about a new product to the moments after they purchase it and evaluate how they feel about it.
Stimuli, emotions and opinions all impact and affect perception. Therefore, it is important for the UX designer to be familiar with perception and related concepts.

B). What does vision do?

Vision begins in the eye which receives the inputs, in the form of light, and finishes in the brain which interprets those inputs and gives us the information we need from the data we receive. The components of the eye are pictured below.

The eye focuses light on to the retina. In the retina there is a layer of photoreceptor (light receiving) cells which are designed to change light into a series of electrochemical signals to be transmitted to the brain. There are two types of photoreceptor – rods and cones

C). What do rods do in the eye?

Rods tend to be found in the peripheral areas of the retina and are designed to respond to low levels of light. They are responsible for our night vision and because of where they are placed on the retina – you can improve your night vision by learning to focus slightly to the side of whatever you are looking at; allowing the light to reach the rod cells most successfully.

D). What do cones do in the eye?

Cones cells are found in the fovea (the center of the retina); cone cells handle the high acuity visual tasks such as reading and color vision. Cone cells respond to red, green or blue light and by combining the signals from these three receptors we can perceive a full range of color.

E). Where does visual perception takes place in the brain?

Visual perception takes place in the cerebral cortex and the electrochemical signal travels through the optic nerve and via the thalamus (another area of the brain) to the cerebral cortex.

F). How much of the data do we process if visual?

It is estimated that 70% of all the data we process is visual but again this is not "hard fact" but based on our understanding of how data works in computer systems.

G). What are the two challenges associated with visual perception?

Visual Stress

Visual stress is a peculiar phenomenon that affects a small, but significant, percentage of the population. When striped patterns (at about 3 cycles per degree) are shown at a flicker rate of about 20 Hz (cycles per second) they can cause seizures in people susceptible to visual stress.

Visual stress is sometimes known as "pattern induced epilepsy" and while this is the most extreme manifestation of visual stress; it's worth noting that visual stress can be induced at milder levels by striped patterns in most people.

Color Blindness

Color blindness is mislabeled. It's not blindness but rather a deficiency in color vision. It is the inability (or sometimes decreased ability) to see certain colors, or perceive color contrasts in normal light.

H). In terms of color blindness, who suffer more? Men or women?

For some reason men suffer from color blindness more often than women. 1 in 12 men have color blindness compared to 1 in 200 women. Color blindness is normally genetic and the trait is inherited from the mother but in some cases it may be induced by disease or ageing.

I). The most common form of color blindness is?

The most common form of color blindness is red/green color blindness – this doesn't mean that the person cannot see red or green but rather that they confuse colors which have some elements of red or green within them. There are other less common forms of color blindness which affect different pairings of colors too. There are many tests for color blindness (some are pictured below) but the condition should not be self-diagnosed but rather diagnosed by an optician or medical professional

2) In Perception in Visualization article:

What makes for a good visualization? (Animation, Color, Composition)

2a) Examples of preattentive processing

Preattentive Processing

For many years vision researchers have been investigating how the human visual system analyses images. An important initial result was the discovery of a limited set of visual properties that are detected very rapidly and accurately by the low-level visual system. These properties were initially called ***preattentive,*** since their detection seemed to precede focused attention. We now know that attention plays a critical role in what we see, even at this early stage of vision. The term preattentive continues to be used, however, since it conveys an intuitive notion of the speed and ease with which these properties are identified.

Typically, tasks that can be performed on large multi-element displays in less than 200 to 250 milliseconds (msec) are considered preattentive. Eye movements take at least 200 msec to initiate, and random locations of the elements in the display ensure that attention cannot be prefocused on any particular location, yet viewers report that these tasks can be completed with very little effort. This suggests that certain information in the display is processed in parallel by the low-level visual system.

A simple example of a preattentive task is the detection of a red circle in a group of blue circles (Fig. 1). The target object has a visual property "red" that the blue distractor objects do not (all non-target objects are considered distractors). A viewer can tell at a glance whether the target is present or absent.

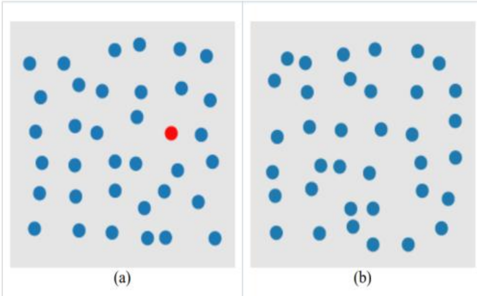


Fig. 1: An example of searching for a target red circle based on a difference in hue: (a) target is present in a sea of blue circle distractors; (b) target is absent

In Fig. 1 the visual system identifies the target through a difference in hue, specifically, a red target in a sea of blue distractors. Hue is not the only visual feature which is preattentive. In Fig. 2 the target is again a red circle, while the distractors are red squares. As before, a viewer can rapidly and accurately determine whether the target is present or absent. Here, the visual system identifies the target through a difference in curvature (or form).

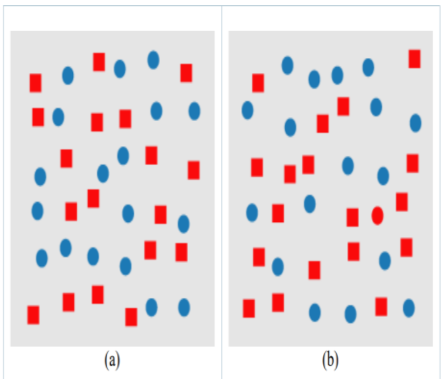


Fig. 3: An example of a conjunction search for a target red circle: (a) target is absent in a sea of red square and blue circle distractors; (b) target is present

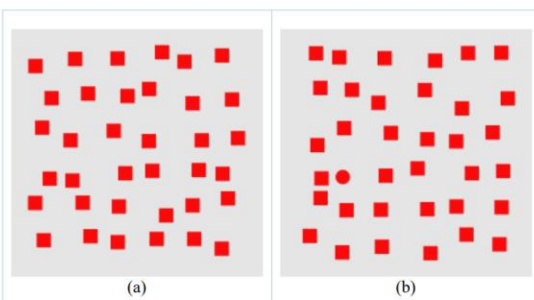


Fig. 2: An example of searching for a target red circle based on a difference in curvature: (a) target is absent in a sea of red square distractors; (b) target is present

A unique visual property in the target (***e.g.,*** a red hue in Fig. 1, or a curved form in Fig. 2) allows it to "pop out" of a display. A target made up of a combination of non-unique features (a ***conjunction*** target) normally cannot be detected preattentively. Fig. 3 shows an example of conjunction search. The red circle target is made up of two features: red and circular. One of these features is present in each of the distractor objects (red squares and blue circles). This means the visual system has no unique visual property to search for when trying to locate the target. If a viewer searches for red items, the visual system always returns true because there are red squares in each display. Similarly, a search for circular items always sees blue circles. Numerous studies have shown that this target cannot be detected preattentively. Viewers must perform a time-consuming serial search through the displays to confirm its presence or absence.

2b)Four theories of preattentive processing

Theories of Preattentive Processing

A number of theories have been proposed to explain how preattentive processing occurs within the visual system. We describe four well-known models: feature integration theory, texton theory, similarity theory, and guided search theory. We also discuss briefly the phenomena of postattentive vision, which shows that prior exposure to an scene does not help a viewer answer questions about the content of the scene.

Feature Integration Theory

Anne Treisman was one of the original researchers to document the area of preattentive processing. She provided important insight into this phenomena by studying two important problems. First, she tried to determine which visual properties are detected preattentively [Treisman 91, Treisman & Gormican 88, Treisman & Souther 86]. She called these properties "preattentive features" [Treisman 85]. Second, she formulated a hypothesis about how the human visual system performs preattentive processing [Treisman & Gelade 80].

Texton Theory

Bela Julész was also instrumental in expanding our understanding of what we "see" in an image. Julész's initial investigations focused on statistical analysis of texture patterns [Julész 71, Julész 75, Julész 81b, Julész et al. 73, Julész et al. 78]. His goal was to determine whether variations in a particular order statistic were seen (or not seen) by the low-level visual system. Examples of variations in order statistics include contrast (a variation in a texture's first-order statistic), orientation and regularity (a variation of the second-order statistic), and curvature (a variation of the third-order statistic). Unfortunately, Julész's results were inconclusive. First-order variations were detected preattentively. In addition, some (but not all) second-order variations were also preattentive, as were an even smaller set of third-order variations.

Similarity Theory

Some researchers do not support the dichotomy of serial and parallel search modes. Initial work in this area was done by Quinlan and Humphreys [87]. They investigated conjunction searches by focusing on two factors. First, search time may depend on the number of items of information required to identify the target. Second, search time may depend on how easily a target can be distinguished from its distractors, regardless of the presence of unique preattentive features. Treisman addressed this second factor in her later work [Treisman 88]. Quinlan and Humphreys found that Treisman's feature integration theory was unable to explain the results they obtained from their experiments. Duncan and Humphreys proceeded to develop their own explanation of preattentive processing. Their model assumes that search ability varies continuously, depending on both the type of task and the display conditions [Duncan 89a, Duncan 89b, Müller 90]. Search time is based on two criteria: T-N similarity and N-N similarity. T-N similarity is the amount of similarity between the targets and nontargets. N-N similarity is the amount of similarity within the nontargets themselves. These two factors affect search time as follows:

- as T-N similarity increases, search efficiency decreases and search time increases,
- as N-N similarity decreases, search efficiency decreases and search time increases, and
- T-N similarity and N-N similarity are related (Fig. 7); decreasing N-N similarity has little effect if T-N similarity is low; increasing T-N similarity has little effect if N-N similarity is high.

Guided Search Theory

More recently, Jeremy Wolfe has suggested a visual search theory that he calls "guided search" [Wolfe 94, Wolfe & Cave 89, Wolfe et al. 89]. He hypothesized that an activation map based on both bottom-up and top-down information is constructed during visual search. Attention is drawn to peaks in the activation map that represent areas in the image with the largest combination of bottom-up and top-down influence.

As with Treisman, Wolfe believes early vision divides an image into individual feature maps (Fig. 8). In his theory, there is one map for each feature type (***e.g.***, one map for colour, one map for orientation, and so on). Within each map a feature is filtered into multiple categories. For example, in the colour map there might be independent representations for red, green, blue, and yellow. Wolfe had already found evidence to suggest that orientation is categorized into steep, shallow, right, and left [Wolfe 92]. The relationship between values within a feature map is different than the relationship between values from different maps (***i.e.***, the relationship between "red" and "blue" is different than the relationship between "blue" and "shallow").

2c)Which figure illustrates some aspect of preattentive processing?

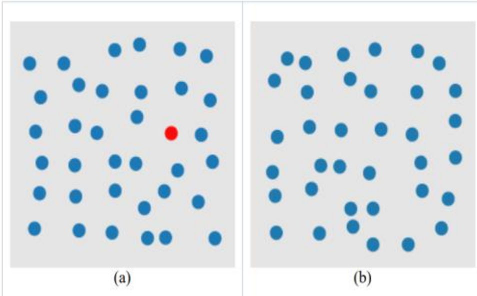


Fig. 1: An example of searching for a target red circle based on a difference in hue: (a) target is present in a sea of blue circle distractors; (b) target is absent

In Fig. 1 the visual system identifies the target through a difference in hue, specifically, a red target in a sea of blue distractors. Hue is not the only visual feature which is preattentive. In Fig. 2 the target is again a red circle, while the distractors are red squares. As before, a viewer can rapidly and accurately determine whether the target is present or absent. Here, the visual system identifies the target through a difference in curvature (or form).

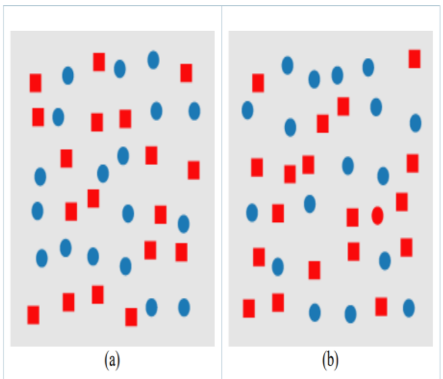


Fig. 3: An example of a conjunction search for a target red circle: (a) target is absent in a sea of red square and blue circle distractors; (b) target is present

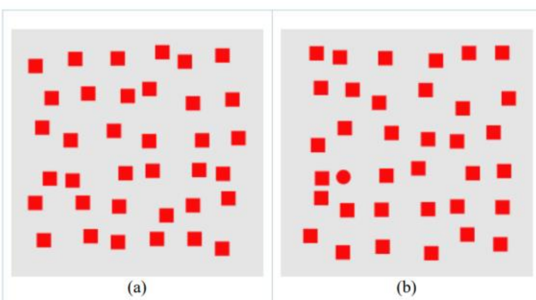


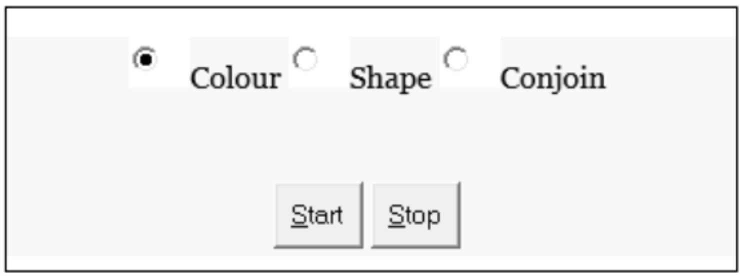
Fig. 2: An example of searching for a target red circle based on a difference in curvature: (a) target is absent in a sea of red square distractors; (b) target is present

A unique visual property in the target (***e.g.,*** a red hue in Fig. 1, or a curved form in Fig. 2) allows it to "pop out" of a display. A target made up of a combination of non-unique features (a ***conjunction*** target) normally cannot be detected preattentively. Fig. 3 shows an example of conjunction search. The red circle target is made up of two features: red and circular. One of these features is present in each of the distractor objects (red squares and blue circles). This means the visual system has no unique visual property to search for when trying to locate the target. If a viewer searches for red items, the visual system always returns true because there are red squares in each display. Similarly, a search for circular items always sees blue circles. Numerous studies have shown that this target cannot be detected preattentively. Viewers must perform a time-consuming serial search through the displays to confirm its presence or absence.

2d)Target Detection?

Target detection: users rapidly and accurately detect the presence or absence of a "target" element with a unique visual feature within a field of distractor elements.

The Javascript applet below **(Click on Hyperlink image)** will let you experiment with the three different target detection searches: colour, shape, and conjunction. As in the figures above, the target is a red circle. Background elements are either blue circles (during colour searches), red squares (during shape searches), or blue circles and red squares (during conjunction searches). The "Exposure Duration:" slider lets you control how long each display is shown (anywhere from 100 to 1000 msec). The "Elements per Display:" slider lets you control the total number of elements in each display (from a minimum of 10 to a maximum of 70). The "Number of Trials:" slider lets you control how many displays to run.

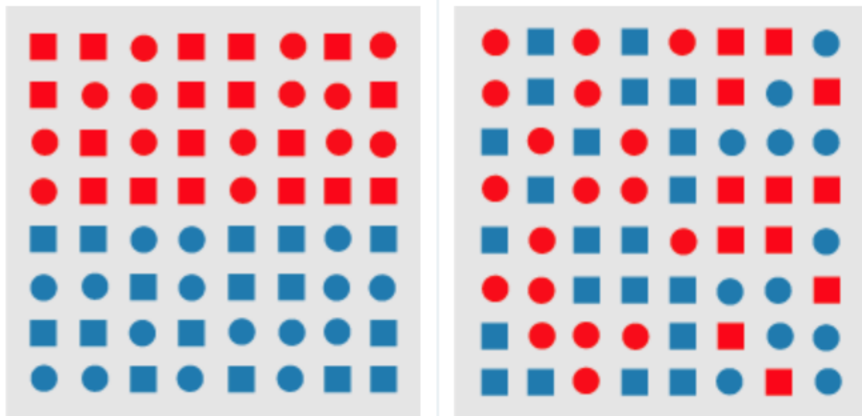


After each display, the applet will tell you whether the target was present or absent. This allows you to compare your answer with the correct answer for each display you see.

If the low-level visual system can be harnessed during visualization, it can be used to draw attention to areas of potential interest in a display. This cannot be accomplished in an ad-hoc fashion, however. The visual features assigned to different data attributes (the *data-feature mapping*) must take advantage of the strengths of our visual system, must be well-suited to the analysis needs of the viewer, and must not produce that visual interference effects (*e.g.*, conjunction search) that could mask information in a display.

2e) What is boundary Detection?

Treisman ran experiments using target and boundary detection to classify preattentive features. For target detection, subjects had to determine whether a target element was present or absent in a field of background distractor elements .Boundary detection involved placing a group of target elements with a unique visual feature within a set of distractors to see if the boundary could be preattentively detected .



(a) (b)

Fig. 4: An example of a boundary detection from Treisman's experiments: (a) a boundary defined by a unique feature hue (red circles and red squares on the top, blue circles and blue squares on the bottom) is preattentively classified as horizontal; (b) a boundary defined by a conjunction of features (red circles and blue squares on the left, blue circles and red squares on the right) cannot be preattentively classified as vertical

---

• *boundary detection:* users rapidly and accurately detect a texture boundary between two groups of elements, where all of the elements in each group have a common visual property .

2F) What is region tracking?

Region tracking: users track one or more elements with a unique visual feature as they move in time and space, and

Extra:

• *counting and estimation:* users count or estimate the number of elements with a unique visual feature.

2G) Texton Theory

Texton Theory

Bela Julész was also instrumental in expanding our understanding of what we "see" in an image. Julész's initial investigations focused on statistical analysis of texture patterns [Julész 71, Julész 75, Julész 81b, Julész et al. 73, Julész et al. 78]. His goal was to determine whether variations in a particular order statistic were seen (or not seen) by the low-level visual system. Examples of variations in order statistics include contrast (a variation in a texture's first-order statistic), orientation and regularity (a variation of the second-order statistic), and curvature (a variation of the third-order statistic). Unfortunately, Julész's results were inconclusive. First-order variations were detected preattentively. In addition, some (but not all) second-order variations were also preattentive, as were an even smaller set of third-order variations.

Based on these findings, Julész modified his theory of how preattentive processing occurs. He suggested that the early visual system detects a group of features called *textons* [Julész 81a, Julész 81b, Julész & Bergen 84]. Textons can be classified into three general categories:

1. Elongated blobs (*e.g.*, line segments, rectangles, ellipses) with specific properties such as hue, orientation, and width.
2. Terminators (ends of line segments).
3. Crossings of line segments.

Julész believed that only a difference in textons or in their density can be detected preattentively. No positional information about neighbouring textons is available without focused attention. Like Treisman, Julész suggested that preattentive processing occurs in parallel and focused attention occurs in serial.

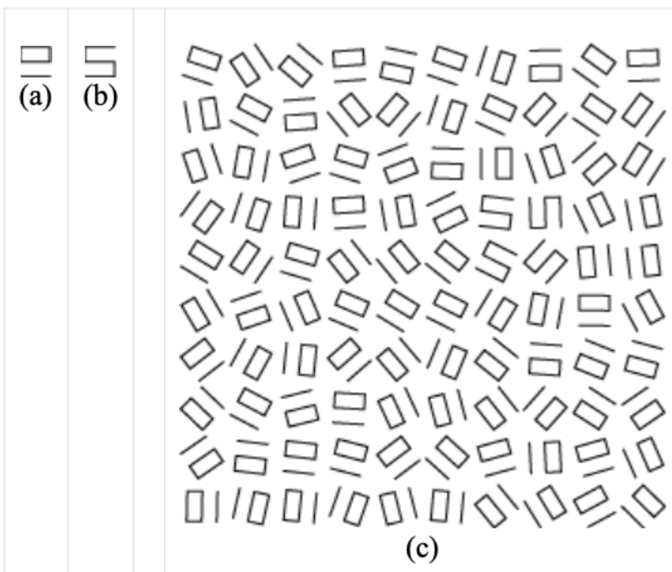


Fig. 6: An example of textons: (a,b) two textons (A and B) that appear different in isolation, but have the same size, number of terminators, and join points; (c) a target group of B-textons is difficult to detect in a background of A-textons when a random rotation is applied

Julész used texture segregation, the task of locating groups of similar objects and the boundaries that separate them, to demonstrate his theory (other researchers like Treisman also used this type of task, for example, identifying the orientation of the boundary between groups of common elements in Fig. 4). Fig. 6 shows an example of an image that supports the texton hypothesis. Although the two objects look very different in isolation, they are actually the same texton. Both are blobs with the same height and width. Both are made up of the same set of line segments and each has two terminators. When oriented randomly in an image, one cannot preattentively detect the texture boundary between the target group and the background distractors.

Visualization Of Data Science

Similarity Theory

Some researchers do not support the dichotomy of serial and parallel search modes. Initial work in this area was done by Quinlan and Humphreys [87]. They investigated conjunction searches by focusing on two factors. First, search time may depend on the number of items of information required to identify the target. Second, search time may depend on how easily a target can be distinguished from its distractors, regardless of the presence of unique preattentive features. Treisman addressed this second factor in her later work [Treisman 88]. Quinlan and Humphreys found that Treisman's feature integration theory was unable to explain the results they obtained from their experiments. Duncan and Humphreys proceeded to develop their own explanation of preattentive processing. Their model assumes that search ability varies continuously, depending on both the type of task and the display conditions [Duncan 89a, Duncan 89b, Müller 90]. Search time is based on two criteria: T-N similarity and N-N similarity. T-N similarity is the amount of similarity between the targets and nontargets. N-N similarity is the amount of similarity within the nontargets themselves. These two factors affect search time as follows:

- as T-N similarity increases, search efficiency decreases and search time increases,
- as N-N similarity decreases, search efficiency decreases and search time increases, and
- T-N similarity and N-N similarity are related (Fig. 7); decreasing N-N similarity has little effect if T-N similarity is low; increasing T-N similarity has little effect if N-N similarity is high.

Fig. 7: Example of N-N similarity affecting search efficiency for a target shaped like the letter L: (a) high N-N (nontarget-nontarget) similarity allows easy detection of target L; (b) low N-N similarity increases the difficulty of detecting the target L

Treisman's feature integration theory has difficulty explaining the results of Fig. 7. In both cases, the distractors seem to use exactly the same features as the target, namely oriented, connected lines of a fixed length. Yet experimental results show displays similar to Fig. 7a produce an average search time increase of 4.5 msec per additional distractor, while displays similar to Fig. 7b produce an average search time increase of 54.5 msec per additional distractor. In order to explain the above and other search phenomena, Duncan and Humphreys proposed a three-step theory of visual selection.

1. The visual field is segmented into structural units. Individual structural units share some common property (*e.g.*, spatial proximity, hue, shape, motion). Each structural unit may again be segmented into smaller units. This produces a hierarchical representation of the visual field. Within the hierarchy, each structural unit is described by a set of properties (*e.g.*, spatial location, hue, texture, size). This segmentation process occurs in parallel.

2I) Guided Search Theory

Guided Search Theory

More recently, Jeremy Wolfe has suggested a visual search theory that he calls "guided search" [Wolfe 94, Wolfe & Cave 89, Wolfe et al. 89]. He hypothesized that an activation map based on both bottom-up and top-down information is constructed during visual search. Attention is drawn to peaks in the activation map that represent areas in the image with the largest combination of bottom-up and top-down influence.

As with Treisman, Wolfe believes early vision divides an image into individual feature maps (Fig. 8). In his theory, there is one map for each feature type (*e.g.*, one map for colour, one map for orientation, and so on). Within each map a feature is filtered into multiple categories. For example, in the colour map there might be independent representations for red, green, blue, and yellow. Wolfe had already found evidence to suggest that orientation is categorized into steep, shallow, right, and left [Wolfe 92]. The relationship between values within a feature map is different than the relationship between values from different maps (*i.e.*, the relationship between "red" and "blue" is different than the relationship between "blue" and "shallow").

Bottom-up activation follows feature categorization. It measures how different an element is from its neighbours. Differences for each relevant feature map are computed and combined (*e.g.*, how different are the elements in terms of colour, how different are they in terms of orientation?) The "metrics" used to measure differences in each feature map are still being investigated.

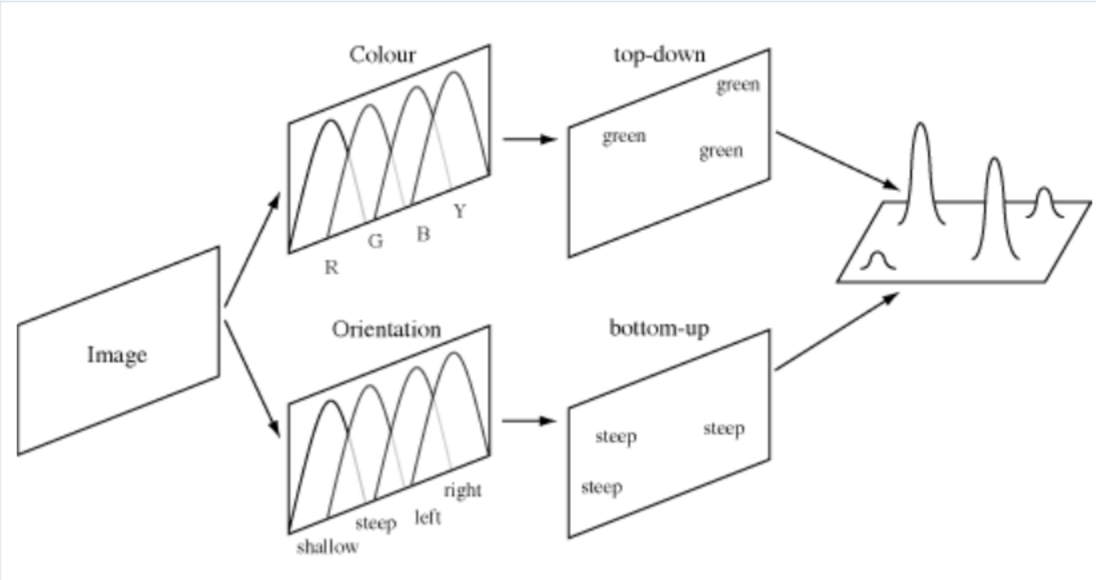


Fig. 8: Framework for guided search, user wants to find a green steep target; image is filtered into categories for each feature map, bottom-up and top-down activation "mark" regions of the image; an activation map is built by combining bottom-up and top-down information, attention is draw to the highest "hills" in the map

Top-down activation is a user-driven attempt to find items with a specific property or set of properties. For example, visual search for a blue element would generate a top-down request that activates "blue" locations. Previous work suggests subjects must specify requests in terms of the categories provided by each feature map [Wolfe & Franzel 88, Wolfe 92]. Thus, subjects could search for "steep" or "shallow" elements, but not for elements rotated by a specific angle. Obviously, subjects should pick the category that best differentiates the target from its distractors. Finding the "best" category is often nonintuitive, however. Wolfe suggests this might explain cases where subjects' performance for a task improves over time.

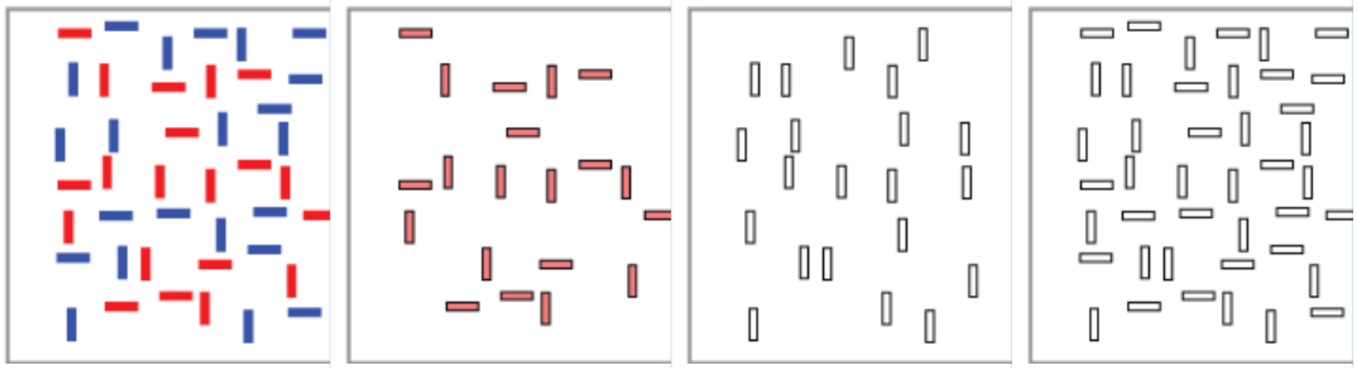
The activation map is a combination of bottom-up and top-down activation. The weights assigned to these two values are task dependent. A conjunction search would place priority on top-down information, since bottom-up results are, in essence, useless. Search for a target with a unique feature would assign a high weight to bottom-up activation. Hills in the activation map mark regions that generated a relatively large amount of bottom-up or top-down influence. There is no information in the activation map about the source of a hill. High activation from a colour map looks exactly the same as high activation from an orientation map. A subject's attention is drawn from hill to hill in order of decreasing activation.

Wolfe's theory easily explains traditional "parallel" visual search. Target elements produce the highest level of activation, regardless of the number of distractor elements. This causes the target to "pop-out" of the scene in time independent of the number of distractors. This also explains Duncan and Humphreys' similarity theory results. Low N-N similarity causes distractors to report higher bottom-up activation, since they now differ from their neighbours. High T-N similarity causes a reduction in the target elements' bottom-up activation. Moreover, guided search also provides a possible explanation for situations where conjunction search can be performed preattentively [Nakayama & Silverman 86, Wolfe 89, Wolfe 90]. User-driven top-down activation may permit efficient searching for conjunction targets.

2J) Boolean Map Theory

A more recent model of low-level vision has been presented by Huang et al. [Huang and Pashler 2007, Huang et al. 2007]. This theory carefully divides a visual search task into two parts: *selection* and *access*. Selection involves choosing a set of objects from a scene. Access determines what properties of the selected objects a viewer can apprehend. Although both operations are implicitly present in previous theories, they are often described as a whole and not as separate steps.

Huang et al. propose that the visual system is capable of dividing a scene into exactly two parts: selected elements and excluded elements. This is the "boolean map" that underlies their theory. The visual system can then access certain properties of the selected elements in the map.



(a) (b) (c) (d)

Fig. 9: Boolean maps: (a) original scene with red and blue, vertical and horizontal elements; (b) boolean map on selection "red", colour label is red, orientation label is undefined; (c) boolean map on selection "vertical", orientation label is vertical, colour label is undefined; (d) boolean map on selection "all locations", colour label is undefined, orientation label is undefined

How are elements selected? That is, how is a boolean map created by the visual system? One way is for a viewer to specify a single value of an individual feature. All objects that contain the feature value are then selected. For example, a viewer could look for red objects, or vertical objects, or big objects. An important distinction between feature integration and boolean maps is that, in feature integration, presence or absence of a feature is available preattentively, but no information on location is provided. A boolean map, on the other hand, encodes the specific spatial locations of the elements that are selected.The boolean map contains feature labels to define properties of the selected objects. If, for example, a viewer selected red objects (Fig. 9b), the colour feature label for the resulting boolean map would be "red". Labels for other features (e.g. orientation, size) would be undefined, since they have not (yet) participated in the creation of the map.

A second method of selection is for a viewer to choose a set of elements at specific spatial locations. In this scenario, the boolean map would provide information about the spatial layout of the selected objects, but all feature labels would be undefined, since no specific feature value was used to identify the selected elements. Fig. 9 shows an example of a simple scene, and the resulting boolean maps for selecting red objects, selecting vertical objects, or selecting all locations.

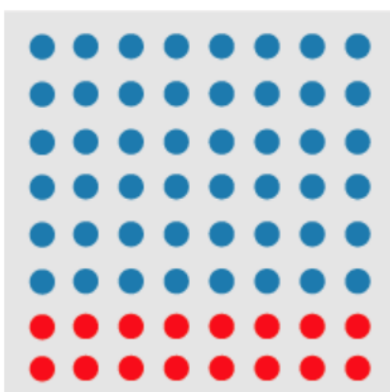
Once a boolean map is available, what information can a viewer access? Two properties of a boolean map are available to a viewer: the label for any feature in the map, and the spatial location of the selected elements. For example, in Fig. 9b where red objects have been created, a viewer can access colour of any object as red, and can also identify the spatial locations of the selected objects. A viewer cannot, however, determine the orientation of a particular object in the map, because the orientation label is undefined. In order to locate, for example, vertical objects, a new boolean map based on this feature value must be created.

2K) Feature Hierarchy

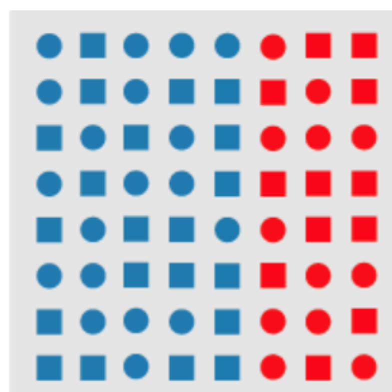
Feature Hierarchy

Based on our understanding of low-level human vision, one promising strategy for multidimensional visualization is to assign different visual features to different data attributes (*i.e.*, building a *data-feature mapping* that maps data to a visual representation). This allows multiple data values to be shown simultaneously in a single image. One key requirement of this method is a data-feature mapping that does not produce visual interference. Interactions between different visual features hide or mask information in a display. Obviously, we want to avoid this situation during visualization. One simple example of visual interference is the conjunction search shown in Fig. 3. If we want to search rapidly for combinations of data values, care must be taken to ensure the resulting combinations contain at least one unique feature for the visual system to cue on.

Other types of visual interference can also occur. An important type of interference results from a *feature hierarchy* that appears to exist in the visual system. For certain tasks the visual system seems to favour one type of visual feature over another. For example, during boundary detection researchers have shown that the visual system favours colour over shape (Fig. 14). Background variations in colour interfere with a viewer's ability to identify the presence of individual shapes and the spatial patterns they form [Callaghan 90]. If colour is held constant across the display, these same shape patterns are immediately visible. The interference is asymmetric: random variations in shape have no effect on a viewer's ability to see colour patterns.



(a)



(b)

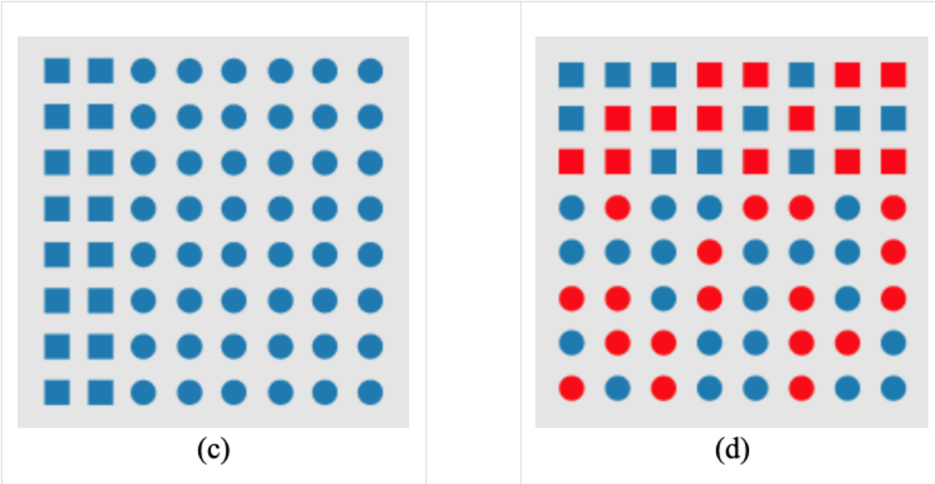


Fig. 14: An example of hue-on-form feature hierarchy: (a) a horizontal hue boundary is preattentive identified when form is held constant; (b) a vertical hue boundary is preattentively identified when form varies randomly in the background; (c) a vertical form boundary is preattentively identified when hue is held constant; (d) a horizontal form boundary cannot be preattentively identified when hue varies randomly in the background

---

Callaghan also documented a luminance-on-hue preference during her experiments [Callaghan 84, Callaghan 89]. More recently, a hue-on-texture interference has been shown to exist [Healey & Enns 98, Healey & Enns 99, Snowden 98, Treisman 85]; random variations in hue interfere with the identification of texture patterns, but not vice-versa. These hierarchies suggest the most important data attributes (as defined by the viewer) should be displayed with the most salient visual features, if possible. The data-feature mapping should avoid situations where the display of secondary data values masks the information the viewer wants to see.

The Javascript applet below **(Cliick on the image)** will let you experiment with the two different boundary identifications: colour and shape. As in Fig. 14, the boundary is either horizontal or vertical, defined by either colour or shape. The "Interference" checkbox enables interference; for colour boundaries, this means that the brightness of each element varies randomly, for shape boundaries the colour varies randomly. The "Exposure Duration:" slider lets you control how long each display is shown (anywhere from 100 to 1000 msec). The "Number of Trials:" slider lets you control how many displays you want to try. After each display, the applet will tell you whether the boundary was horizontal or vertical. This allows you to compare your answer with the correct answer for each display you see.

2 L) Change blindness

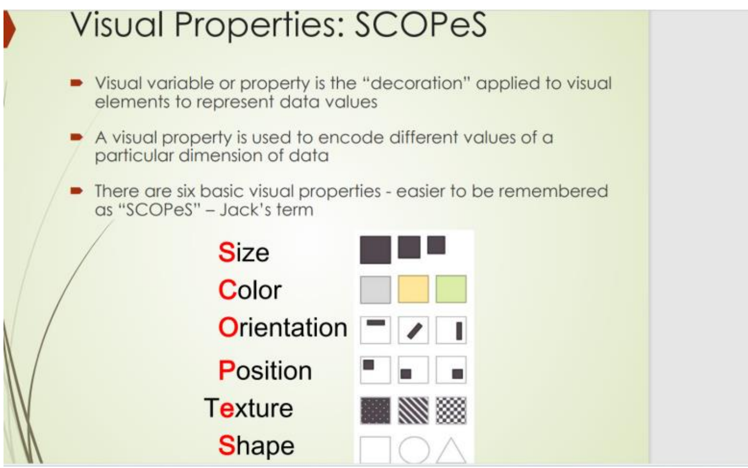
Change Blindness

Recent research in visualization has explored ways to apply rules of perception to produce images that are visually salient [Ware 2000]. This work is based in large part on psychophysical studies of the low-level human visual system. One of the most important lessons of the past twenty-five years is that human vision does not resemble the relatively faithful and largely passive process of modern photography [Pomerantz & Pristach 89, Treisman 85, Treisman & Gormican, Wolfe 94, Wolfe et al. 2000]. The goal of human vision is not to create a replica or image of the seen world in our heads. A much better metaphor for vision is that of a dynamic and ongoing construction project,

where the products being built are short-lived models of the external world that are specifically designed for the current visually guided tasks of the viewer [Egeth & Yantis 97, Mack & Rock 98, Rensink 2000, Simons 2000]. There does not appear to be any general purpose vision. What we "see" when confronted with a new scene depends as much on our goals and expectations as it does on the array of light that enters our eyes.

These new findings differ from one of the initial ideas of preattentive processing, that only certain features in an image are recognized without the need for focused attention, and that other features ***cannot*** be detected, even when viewers actively search for these exact features. More recent work in preattentive vision has presented evidence to suggest that this strict dichotomy does not hold. Instead, "visible" or "not visible" represent two ends of a continuous spectrum. Issues like the difference between a target's visual features and its neighbours' features, what a viewer is searching for, and how the image is presented can all have an effect on search performance. For example, Wolfe's guided search theory assumes both bottom-up (***i.e.,*** preattentive) and top-down (***i.e.,*** attention-based) activation of features in an image [Wolfe & Cave 89, Wolfe et al. 89, Wolfe 94]. Other researchers like Treisman have also studied the dual effects of preattentive and attention-driven demands on what the visual system sees [Treisman 91, Treisman & Souther 86]. Wolfe's discussion of postattentive vision also points to the fact that details of an image cannot be

4)What is Scopes?



5)What are the steps in the WK05 NotesB2 Visual Analysis Guidebook?

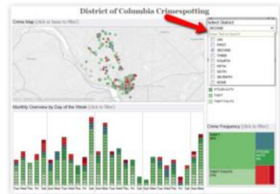
Filters

Filters let you slice data from different angles or drill down to a more detailed level. They are great ways to enable multi-level data exploration and user-driven data analysis. Tableau provides many options for building powerful filters right into your dashboards; however, yet there is always the potential to confuse your audience if you do not use filters properly. The following steps will help you create effective filters:

1. **Think about what you want your filters to do.** Before you start adding filters to your dashboard, it is often useful to ask yourself these questions: What are your views? How much flexibility do you want your users to have? Which filters will bring maximum value to your views? Are these filters already part of your views? What happens when you apply these filters? Are these filters necessary for your viewers to obtain certain information? How do your filters work together with your highlight actions? Do you want your filters to apply to one view, a few selected views, or to all of your views?

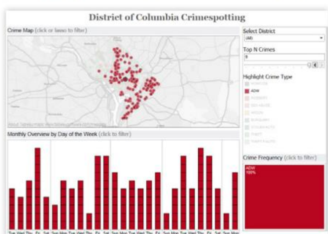
2. **Determine which types of filters to add.** Tableau provides four ways to add filters to your dashboards. Determining which types of filters to add usually depends on the answers to your previous questions. Below is a summary of the different filters and how they work:

- **Quick filters:** Applying a quick filter is the easiest way to add filters to your views. Simply right click a field or drag a field onto the filters shelf in your worksheet and then bring them up on the dashboard. In this example, notice that the filter is a sub-area of the default dashboard view –the filter is District within DC and the default geographical area is the entire DC. In other words, this filter adds great value to the dashboard by enabling users to drill down views from the entire District of Columbia into a subarea of the district. Also note that the default scope of this filter is to one view only. To apply it to additional views, you will need to manually adjust the filter.

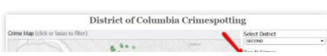


- **Use a view as a filter:** This is another easy way to add filters to your dashboard. Unlike with the quick filter, though, the default for this filter is to apply to all views and fields on the dashboard. Below is an example of what you might see if you select the "Use as Filter" option on the map view; in this case, the other views filter down by crime and day of the week when you select a data point on the map. By seeing this result, we realize that the value of seeing each crime on each day of the week is less informative than seeing each crime on all of the day of week. This is where the next filter type comes in handy.

- **Filter action:** This option gives you great control and flexibility over your filters; you can choose your source views, target views, which fields to filter by, how your users activate the filters, and what you want to happen after you clear the filters. You can either create a filter action from the dashboard action menu or just edit the action that is automatically generated with the "Use as Filter" option. Below is the new and more informative view that you would see after changing your existing map filter into a filter by crime type.



- **Filter with a parameter**: Compared to the previous three filter types, this option is the most powerful. Filtering with a parameter gives you the ability to make more flexible and interesting filters which can't be achieved by other type of filters. It even lets you filter across different data sources. Below is a simple example of a filter with a parameter; in this view, users can choose a top number of crimes to filter the view by.



6).

A). Name 5 ways to create effective views

I. Emphasize the most important data (A rule of thumb is to put the most important data on the X- or Y- axis and less important data on color, size, or shape)

II. Orient your views for legibility (If you find yourself with a view that has long labels that only fit vertically, try rotating the view.)

III. Organize your views (Bullet chart combines a bar chart with reference lines to create a great visual comparison between actual and target numbers.)

IV. Avoid overloading your views (Instead of stacking many measures and dimensions into one condensed view, break them down to small multiples)

V. Limit the number of colors and shapes in a single view (Limit the number of colors and shapes in one view to 7-10 so that you can distinguish them and see important patterns)

B). What are the five ways to create holistic dashboard?

I. General guidelines (Use interactive views only when it is necessary: you need to guide a story, encourage user exploration, or there is too much detail to show all at once)

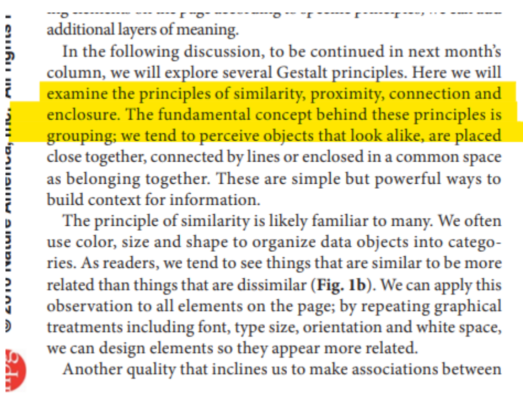
II. Highlighting (Highlights let you quickly show relationships between values in a specific area or category—even across multiple views.)

III. Filters(Filters are great ways to enable multi-level data exploration and user driven data analysis)

IV. Hyperlinking and using the power of the web(You can use URL actions to link to information hosted outside of your data source.)

V. Sizing: Making sure your visualization is visible (No matter where you are publishing, be sure to construct your visualization at the size that you will eventually publish at— and use the Range sizing feature to avoid scrollbars or scrunched views)

7)What are some Gestalt principles?

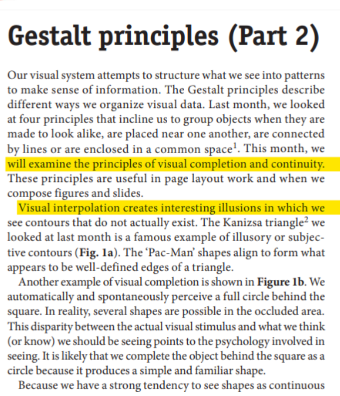


additional layers of meaning.

In the following discussion, to be continued in next month's column, we will explore several Gestalt principles. Here we will examine the principles of similarity, proximity, connection and enclosure. The fundamental concept behind these principles is grouping; we tend to perceive objects that look alike, are placed close together, connected by lines or enclosed in a common space as belonging together. These are simple but powerful ways to build context for information.

The principle of similarity is likely familiar to many. We often use color, size and shape to organize data objects into categories. As readers, we tend to see things that are similar to be more related than things that are dissimilar (**Fig. 1b**). We can apply this observation to all elements on the page; by repeating graphical treatments including font, type size, orientation and white space, we can design elements so they appear more related.

Another quality that inclines us to make associations between



Gestalt principles (Part 2)

Our visual system attempts to structure what we see into patterns to make sense of information. The Gestalt principles describe different ways we organize visual data. Last month, we looked at four principles that incline us to group objects when they are made to look alike, are placed near one another, are connected by lines or are enclosed in a common space[1]. This month, we will examine the principles of visual completion and continuity. These principles are useful in page layout work and when we compose figures and slides.

Visual interpolation creates interesting illusions in which we see contours that do not actually exist. The Kanizsa triangle[2] we looked at last month is a famous example of illusory or subjective contours (**Fig. 1a**). The 'Pac-Man' shapes align to form what appears to be well-defined edges of a triangle.

Another example of visual completion is shown in **Figure 1b**. We automatically and spontaneously perceive a full circle behind the square. In reality, several shapes are possible in the occluded area. This disparity between the actual visual stimulus and what we think (or know) we should be seeing points to the psychology involved in seeing. It is likely that we complete the object behind the square as a circle because it produces a simple and familiar shape.

Because we have a strong tendency to see shapes as continuous

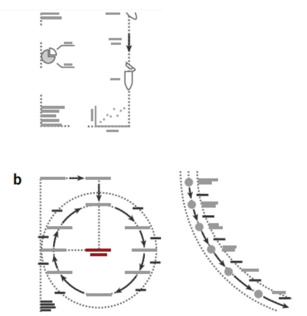
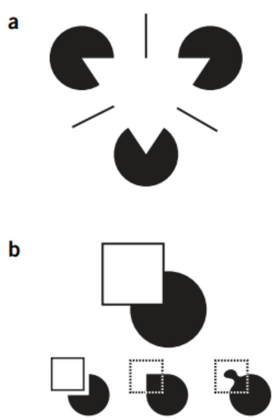


Figure 2 | Alignment. (**a**) Graphics and text used as vertices and edges of geometric shapes. (**b**) Geometric and curvilinear shapes used as flexible guides to align content.

8)What is the Kanizsa triangle?

Visual interpolation creates interesting illusions in which we see contours that do not actually exist. The Kanizsa triangle[2] we looked at last month is a famous example of illusory or subjective contours (**Fig. 1a**). The 'Pac-Man' shapes align to form what appears to be well-defined edges of a triangle.

Another example of visual completion is shown in **Figure 1b**. We automatically and spontaneously perceive a full circle behind the square. In reality, several shapes are possible in the occluded area



. | Visual completion. (**a**) The Kanizsa triangle and illusory contour. ıtaneous and automatic completion of occluded surfaces as a simple iliar circle.

Story telling with data:

9) What's Data-Driven Storytelling?

Data storytelling is the practice of blending hard data with human communication to craft an engaging

narrative that's anchored by facts. It uses data visualization techniques (e.g., charts and images) to help

convey the meaning of the data in a way that's compelling and relevant to the audience.

Data-driven stories are created through the process of analyzing and filtering large datasets to uncover

insights and reveal new or different ways to understand the information. They're tailored to a specific

audience and the context in which they're consumed. This can help communicate information or a point-of-

view most effectively while generating the least cognitive load, which affects the mental energy the audience

needs to spend on grasping your message and therefore, impact how well it's received.

10) Name 5 Benefits of Data-Driven Storytelling?

The Benefits of Data-Driven Storytelling

Here are some benefits of using data storytelling as a communication tool:

• Data stories add value by assigning data meaning and context so the audience can connect the dots

and turn numbers into insights that they can use. The insights, in turn, facilitate decision-making

and spur actions.

• By using numbers and facts, you can anchor your claims and increase the credibility of your content.

This builds trust with your audience and increases the likelihood that they'll be convinced by your

point-of-view.

• Data stories crafted with internal and proprietary data help you stand out and capture attention. The

original insights, useful perspectives and unexpected angles allow you to cut through the clutter in a

world full of regurgitated content.

• The graphic elements are attractive to the media, increasing the chances that your content gets

picked up by high-profile publications or influencers. This will help you generate brand awareness,

reach new audiences and position your brand as a thought leader.

• The combination of narrative and visual elements activates both sides of the brain, delivering an

experience that's analytical and emotional at the same time to help your audience cement the information through comprehension, retention and appeal.

• The use of various techniques, such as interactive data visualization, can further foster engagement

with the audience. For example, by guiding them to arrive at a conclusion or exploring the part of

the data story that's most relevant to them.

• Data-driven storytelling is versatile. It can be incorporated and repurposed in many external and

internal communication channels, such as annual reports, brochures, case studies, presentations,

videos, website content, white papers, social media posts and more.

11) Give 5 examples of How to tell an Effective Data Story with Data Visualization Best Practices

1. Define Your Objectives and Know Your Audience
2. Identify a Compelling Narrative
3. Incorporate Key Elements of Analysis Storytelling
4. Be Objective and Transparent
5. Choose the Right Data Visualization Method
6. Follow Visual Design Best Practices
7. Use the Right Data Visualization Tools
8. Make Your Data Story Human and Insightful
9. Create Synergies Between Data Storytelling and Content

12. What is the difference between data visualization and story telling with data?

Data visualization is simply the visual representation of data. This might be basic charts and tables that are generated from a spreadsheet. Or, it could go well beyond those modalities to include any use of shapes, color, and sizing to draw visual focus to data findings. Bottom line, data visualization is about communicating the substance of your metrics in a visual way. Data visualization can certainly be used to tell a story at the slide level. It can:

- Provide context
- Elevate and draw attention to key insights (and visually subdue the others)
- Lead to action (AKA: the "ask")

However, the real magic happens when data visualization is driven by storytelling at the slide level and story level.

Storytelling with data differs from data visualization because it requires communicators to offer a larger, holistic, view of their message. You must focus first on your audience and structure a larger message before any visuals are rendered. You must identify from the start:

- What do I want my audience to know or do with the data I am presenting?
- How will I structure a narrative that leads to desired action?
- How is my data helping drive a decision?

There is no understating how important it is for all presented data to have a purpose. Every piece of data you include should further this purpose – or it should be left out.

13. From storytelling that moves people, what is the root of all the great storytelling?

Self-knowledge is the root of all great storytelling. A storyteller creates all characters from the self by asking the question, "If I were this character in these circumstances, what would I do?" The more you understand your own humanity, the more you can appreciate the humanity of others in all their good-versus-evil struggles. I would argue that the great leaders Jim Collins describes are people with enormous self-knowledge. They have self-insight and self-respect balanced by skepticism. Great storytellers—and, I suspect, great leaders—are skeptics who understand their own masks as well as the masks of life, and this understanding makes them humble. They see the humanity in others and deal with them in a compassionate yet realistic way. That duality makes for a wonderful leader.

14. According to Pixar, what are the 6 rules of great story telling?

1. Great stories are universal - Great storytelling is about taking a piece of the human condition (so things like birth, growth, emotionality, aspiration, conflict) and conveying it in a unique situation.

2. Great stories have a clear structure and purpose

3.Great stories have a character to root for (an underdog)

4. Great stories appeal to our deepest emotions

5. Great stories are surprising and unexpected

6. Great stories are simple and focused

Data Visualization Tools: Tableau

15. Dashboard, worksheet, measures, dimensions, shelf

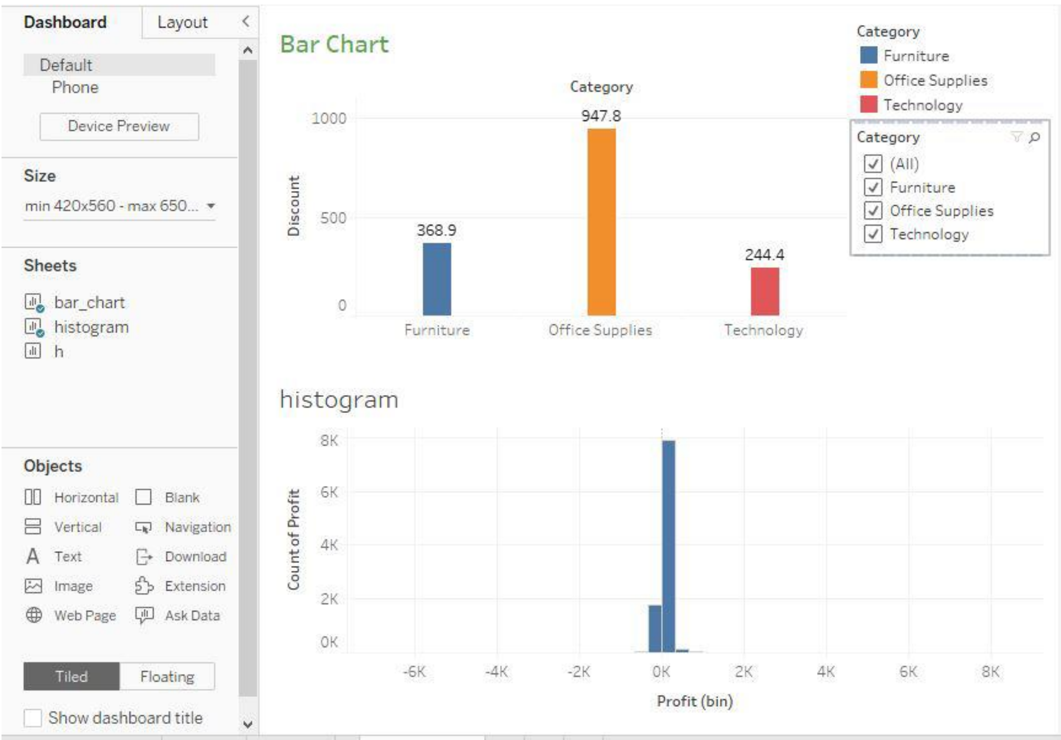
Create a Dashboard in Tableau

1. Select Dashboard and New Dashboard from the toolbar.

2. Under the dashboard and sheets pane, the existing charts/sheets appear.

3. Drag the drop the relevant sheet onto the dashboard.

4. To add filters to the dashboard – Select Analysis from the toolbar, under it select Filter, all relevant filters appear. Go ahead and choose the one appropriate for the current dashboard.



Save the dashboard to tableau public: Tableau desktop provides a variety of options to save a dashboard, most commonly used are twb(tableau workbook file) and twbx(tableau packed workbook), the former is the XML version of the dashboard and is intended for collaboration and doesn't enclose the data, the latter dashboard is inclusive of data and it's intended for sharing.

In Tableau public, all dashboards including the data need to be saved in the public domain

In Tableau, when we connect to a new data source, each field in the data source is either mapped as measures or dimensions. These fields are the columns defined in the data source. Each field is assigned a dataType (integer, string, etc.) and a role (discrete dimension or continuous measure).

Measures contain numeric values that are analyzed by a dimension table. Measures are stored in a table that allows storage of multiple records and contains foreign keys referring uniquely to the associated dimension tables.

While Dimensions contain qualitative values (name, dates, geographical data) to define comprehensive attributes to categorize, segment, and reveal the data details.

Tableau worksheets contain various named elements like columns, rows, marks, filters, pages, etc. which are called shelves. You can place fields on shelves to create visualizations, increase the level of detail, or add context to it. The largest area of the workspace is the view(worksheet), the area where the visualization is visible.

A Dashboard *in Tableau is a collection of views (worksheets) arranged on a single canvas. It allows users to combine multiple visualizations and related information into a single, interactive interface. Dashboards are often used to present a comprehensive overview of data and facilitate data analysis.*

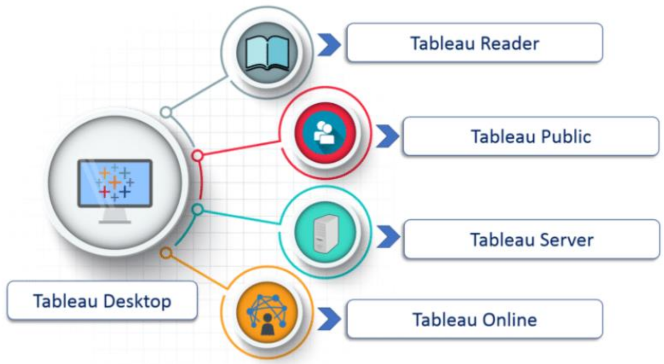
A Worksheet *is a single tab within Tableau where you build and design visualizations. It's the canvas where you drag and drop fields to create charts, graphs, and other visual representations of your data. Each worksheet is associated with a specific data source and contains shelves for Columns and Rows, as well as other shelf areas for Filters, Marks, and Pages.*

Measures *in Tableau are the numeric fields that contain the quantitative data you want to analyze. Examples of measures include sales revenue, profit, quantity sold, or any other numerical data point. Measures are typically used to create visualizations like bar charts, line charts, and scatter plots. They represent the "how much" aspect of your data.*

Dimensions *in Tableau are categorical fields that provide context to your data. Examples of dimensions include date, geographic location, product categories, or any other qualitative data point. Dimensions are used to break down and categorize measures, and they represent the "what" aspect of your data. Dimensions are often used to define the structure of your visualization, such as the categories on the x-axis or y-axis of a chart*

---

16. what are the different tableau products and what’s is the latest tableau version of tableau?



1. Tableau Desktop: Desktop product is used to create optimized queries out from pictures of data. Once the queries are ready, you can perform those queries without the need to code. Tableau desktop encompasses data from various sources into its data engine and creates an interactive dashboard.
2. Tableau Server: When you have published dashboards using Tableau Desktop, Tableau servers help in sharing them throughout the organization. It is an enterprise-level feature that is installed on a Windows or Linux server.
3. Tableau Reader: Tableau Reader is a free feature available on Desktop that lets you open and views data visualizations. You can filter or drill down the data but restricts editing any formulas or performing any kind of actions on it. It is also used to extract connection files.
4. Tableau Online: Tableau online is also a paid feature but doesn't need exclusive installation. It comes with the software and is used to share the published dashboards anywhere and everywhere.
5. Tableau Public: Tableau public is yet another free feature to view your data visualizations by saving them as worksheets or workbooks on Tableau Server.

17. What are the different types of joins in Tableau?

Tableau is pretty similar to SQL. Therefore, the types of joins in Tableau are similar:

1. Left Outer Join: Extracts all the records from the left table and the matching rows from the right table.
2. Right Outer Join: Extracts all the records from the right table and the matching rows from the left table.
3. Full Outer Join: Extracts the records from both the left and right tables. All unmatched rows go with the NULL value.
4. Inner Join: Extracts the records from both tables.

18. **How many maximum tables can you join in Tableau?**

You can join a maximum of 32 tables in Tableau.

19. **What are the supported datatypes in tableau?**

Tableau supports various data types to accommodate different types of data. Here are the main data types in Tableau:

1. **Integer:** Whole numbers without decimal points, used for counting or ordinal values.
2. **Float:** Numeric values with decimal points, used for continuous numerical data.
3. **String (Text):** Alphanumeric characters, used for text data, names, labels, and categorical values.
4. **Boolean:** Logical values representing True or False, used for binary data or conditional statements.
5. **Date:** Dates without time information, used for date-based analysis and time series.
6. **Datetime:** Dates with time information, used for more granular time-based analysis.
7. **Geographic Role:** Special data type for geographical data, such as latitude and longitude, used for creating maps.
8. **Duration:** Represents time duration, used for calculating time differences.
9. **Currency:** Used to display currency symbols and values with proper formatting.
10. **Percentage:** Used for percentage values with formatting to show percentage symbol.
11. **Whole Number (Discrete):** Similar to Integer but treated as a discrete (categorical) dimension.
12. **Decimal Number (Discrete):** Similar to Float but treated as a discrete (categorical) dimension.
13. **Ordinal:** Categorical data with a specific order, used for data with inherent ranking.
14. **Measure Names:** Special data type used for representing the names of measures in Tableau.
15. **Measure Values:** Special data type used for representing the values of measures in Tableau.

20. **What are sets?**

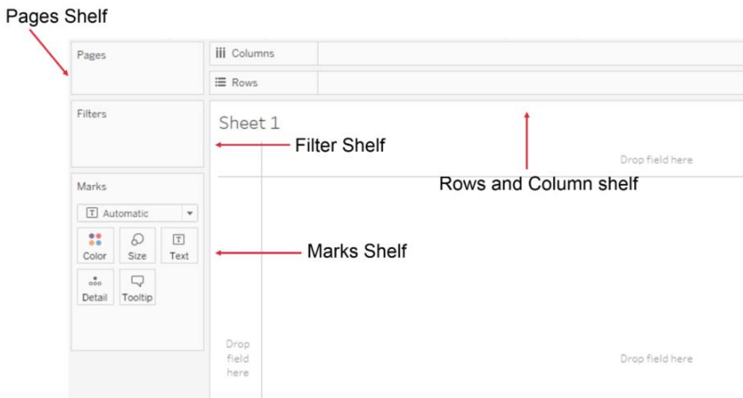
Sets are custom fields that define a subset of data based on some conditions. A **set** can be based on a computed condition, for example, a **set** may contain customers with sales over a certain threshold. Computed **sets** update as your data changes. Alternatively, a **set** can be based on specific data point in your view.

21. **What are groups in tableau?**

A group is a combination of dimension members that make higher level categories. For example, if you are working with a view that shows average test scores by major, you may want to group certain majors together to create major categories.

22. **What are shelves?**

They are Named areas to the left and top of the view. You build views by placing fields onto the shelves. Some shelves are available only when you select certain mark types.



23. **What is data blending in Tableau?**

In Tableau, data blending is the process of combining data from different sources or links to make a single view that can be used for analysis. It is used when there is no direct link between the data sources or a key that they both share. Tableau finds similar dimensions and makes sensible connections between the main and secondary data sources. This lets users work with complicated sets of data and get insights from data that can't be easily put together using standard joins. But there may be some limits to data blending, such as the possibility of data duplication and speed issues. Still, it is a powerful Tableau tool that lets you analyze many different datasets in depth.

24. ** NOT AVAILABLE**

25. **What is a dual axis?**

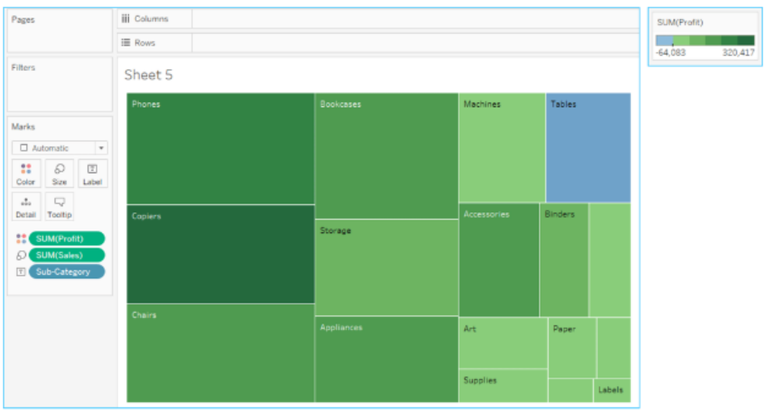
Dual Axis is an excellent phenomenon supported by Tableau that helps users view two scales of two measures in the same graph. Many websites like Indeed.com and other make use of dual axis to show the comparison between two measures and their growth rate in a septic set of years. Dual axes let you compare multiple measures at once, having two independent axes layered on top of one another. This is how it looks like:

26. **What is the difference between a tree map and heat map?**

A heat map can be used for comparing categories with color and size. With heat maps, you can compare two different measures together.

Heat Map

A tree map also does the same except it is considered a very powerful visualization as it can be used for illustrating hierarchical data and part-to-whole relationships.



Tree Map

27. **Explain when would you use Joins vs. Blending in Tableau?**

While the two terms may sound similar, there is a difference in their meaning and use in Tableau:

While Join is used to combine two or more tables within the same data source.

Blending is used to combine data from multiple data sources such as Oracle, Excel, SQL server, etc.

Explain when would you use Joins vs. Blending in Tableau?

If data resides in a single source, it is always desirable to use Joins. When your data is not in one place blending is the most viable way to create a left join like the connection between your primary and secondary data sources.

28. What is the Rank Function in Tableau?

Rank function is used to give positions (rank) to any measure in the data set. Tableau can rank measure in the following ways:

- Rank: The rank function in Tableau accepts two arguments: aggregated measure and ranking order (optional) with a default value of desc.
- Rank_dense: The rank_dense also accepts the two arguments: aggregated measure and ranking order. This assigns the same rank to the same values but doesn't stop there and keeps incrementing with the other values. For instance, if you have values 10, 20, 20, 30, then ranks will be 1, 2, 2, 3.
- Rank_modified: The rank_modified assigns the same rank to similar values.
- Rank_unique: The rank_unique assigns a unique rank to each and every value. For example, If the values are 10, 20, 20, 30 then the assigned ranks will be 1,2,3,4 respectively.

29. What is a hierarchical field?

A hierarchical field in tableau is used for drilling down data. It means viewing your data in a more granular level.

30. **What are the different filters in Tableau and how are they different from each other?**

In Tableau, filters are used to restrict the data from database.

The different filters in Tableau are: Quick , Context and Normal/Traditional filter are:

- ***Normal Filter*** is used to restrict the data from database based on selected dimension or measure. A Traditional Filter can be created by simply dragging a field onto the 'Filters' shelf.
- ***Quick filter*** is used to view the filtering options and filter each worksheet on a dashboard while changing the values dynamically (within the range defined) during the run time.
- ***Context Filter*** is used to filter the data that is transferred to each individual worksheet. When a worksheet queries the data source, it creates a temporary, flat table that is uses to compute the chart. This temporary table includes all values that are not filtered out by either the Custom SQL or the Context Filter.

31. How to view underlying SQL Queries in Tableau?

Viewing underlying SQL Queries in Tableau provides two options:

- **Create a Performance Recording** to record performance information about the main events you interact with workbook. Users can view the performance metrics in a workbook created Tableau.
  Help -> Settings and Performance -> Start Performance Recording
  Help -> Setting and Performance -> Stop Performance Recording.

- **Reviewing the Tableau Desktop Logs** located at C:UsersMy DocumentsMy Tableau Repository. For live connection to data source, you can check log.txt and tabprotosrv.txt files. For an extract, check tdeserver.txt file.

32. **Name the components of a Dashboard.**

- **Horizontal –** Horizontal layout containers allow the designer to group worksheets and dashboard components left to right across your page and edit the height of all elements at once.
- **Vertical –** Vertical containers allow the user to group worksheets and dashboard components top to bottom down your page and edit the width of all elements at once.
- **Text –** All textual fields.
- **Image Extract –** A Tableau workbook is in XML format. In order to extracts images, Tableau applies some codes to extract an image which can be stored in XML.
- **Web [URL ACTION] –** A URL action is a hyperlink that points to a Web page, file, or other web-based resource outside of Tableau. You can use URL actions to link to more information about your data that may be hosted outside of your data source. To make the link relevant to your data, you can substitute field values of a selection into the URL as parameters.

34. **What is story in Tableau?**

A story is a sheet that contains a sequence of worksheets or dashboards that work together to convey information. You can create stories to show how facts are connected, provide context, demonstrate how decisions relate to outcomes, or simply make a compelling case. Each individual sheet in a story is called a story point.

35. What is the difference between discrete and continuous in Tableau?

There are two types of data roles in Tableau – discrete and continuous dimension.

- Discrete data roles are values that are counted as distinct and separate and can only take individual values within a range. Examples: number of threads in a sheet, customer name or row ID or State. Discrete values are shown as blue pills on the shelves and blue icons in the data window.
- Continuous data roles are used to measure continuous data and can take on any value within a finite or infinite interval. Examples: unit price, time and profit or order quantity. Continuous variables behave in a similar way in that they can take on any value. Continuous values are shown as green pills.

36. Mention what is the difference between published data sources and embedded data sources in Tableau?

The difference between published data source and embedded data source is that,

- **Published data source**: It contains connection information that is independent of any workbook and can be used by multiple workbooks.
- **Embedded data source**: It contains connection information and is associated with a workbook.

37. What are different Tableau files?

- **Workbooks:** Workbooks contain one or more worksheets and dashboard elements.
- **Bookmarks:** Contains a single worksheet that is easier to share.
- **Packaged Workbooks:** Contains a workbook along with supporting local file data and background images.
- **Data Extraction Files:** Extract files that contain a subset of data.
- **Data Connection Files:** Small XML file with various connection information.

38. What is the maximum no. of rows Tableau can utilize at one time?

The maximum number of rows or columns is indefinite because even though Tableau contains petabytes of data, it intelligently uses only those rows and columns which you need to extract for your purpose.