Perception and Communication

Descriptive Analytics: Data Visualization and Storytelling with Data

The widespread adoption of data analytics poses many challenges in business. While the pace of spending for data analytics has grown exponentially over the years, a majority of companies report that the adoption of data analytics continues to be a major challenge[[1]](#footnote-0). The source of the challenge derives from barriers rooted in communication[[2]](#footnote-1). Communication issues range from poor writing and speaking to delivering ineffective presentations and reports. In descriptive analytics, poor communication tends to stem particularly from:

* Ineffective visuals
* The inability to tell (or sell) the story behind the data

# Effective Visuals

Today’s software libraries and visualization products are sufficiently powerful for visual analytics. Although they cannot cater to every possible scenario, they do allow users to customize individual graphical elements. To create effective visuals, analysts must understand and be able to explain why specific graphical elements must be included, eliminated, or modified. While fine-grained customizations may not be needed for every visualization, they should be employed when needed to communicate the desired information.

## Design and Perception

*Aoccdrnig to rseearch at Cmabrigde Uinervtisy, it deosn’t mttaer in waht oredr the ltteers in a wrod are, the olny iprmoatnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.* — Abeer Hasanin 2015–2016[[3]](#footnote-2)

Visuals are able to convey information effectively because they take advantage of specific characteristics of human perception. Using visuals effectively is more than just transforming raw data into an appealing chart—it requires knowing specifically which kind of visualization is most effective in conveying the desired information. And this knowledge has human psychology as its foundation.

Earlier in the course, we saw how certain questions are best answered by specific visual cue patterns. Underlying these visual cues is a set of principles. Many theories in psychology and cognitive science have categorized these principles as they attempted to explain how the mind interprets visual abstractions.

The most famous psychological perception theories are arguably the *Gestalt principles,*which explore the patterns that the human mind perceives when presented with certain abstract configurations of visual elements. The Gestalt concept espouses the idea that the whole is more than the sum of its parts, which means that the mind creates something more than just what the eye sees as a collection of disparate elements. These principles are studied by UX designers, graphics designers, industrial designers, and artists among others, and should be kept in mind during the construction of data visualizations.

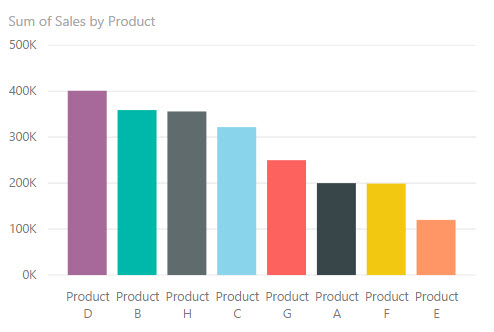
The Gestalt principles most relevant for data v isualizations include the following.

### Law of Simplicity (Pragnanz):

The mind prefers simplicity over complex shapes as it strives to prevent information overload.  
A series of lines depicting a triangle and square on top of each other


*Figure 1: A series of lines depicting a triangle and square on top of each other*

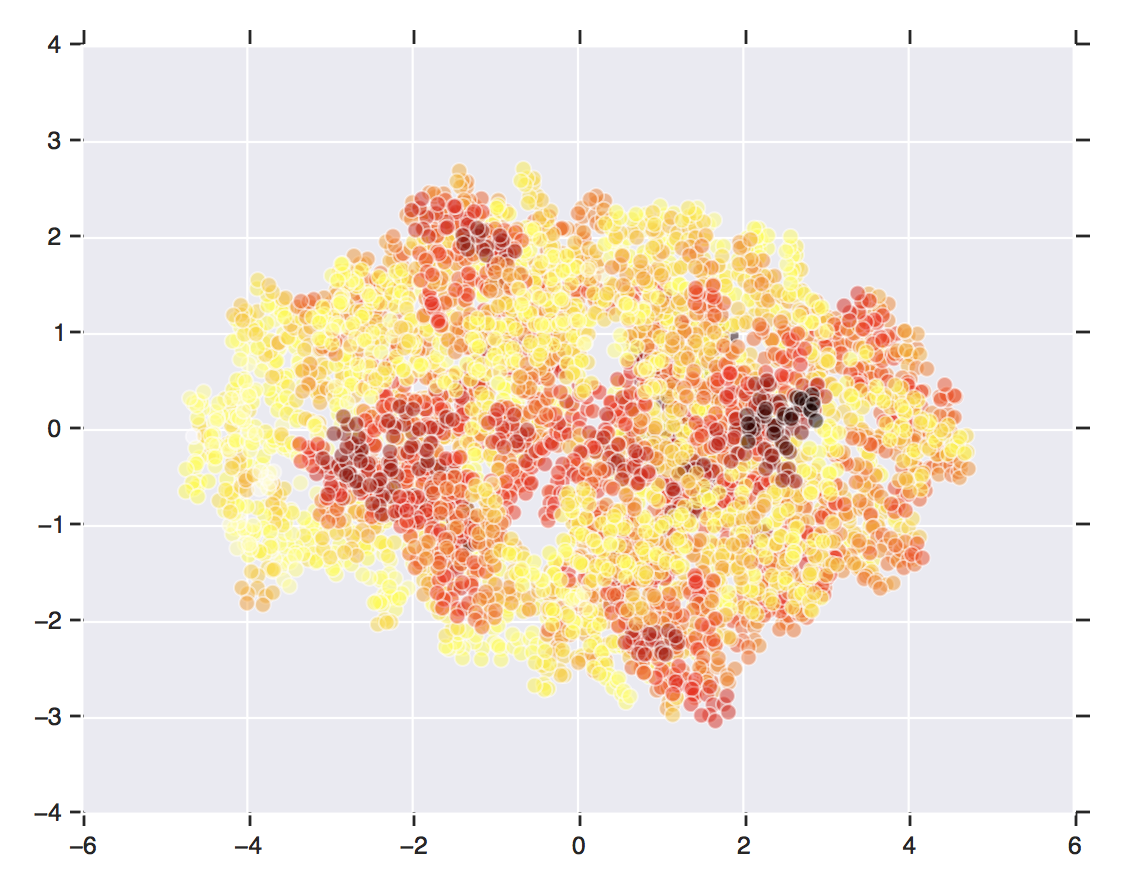
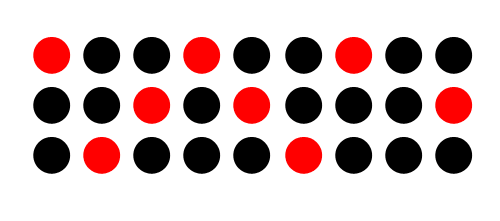
Viewers of Figure 1 see a square and a triangle rather than an assortment of complex shapes—the simple interpretation “wins.” Similarly, consider the bar graphs shown in Figures 2 and 3, which plot sales figures for a set of products. Organizing the graph in decreasing order of sales makes it easier for viewers to see the relationships between sales figures.

Simple sorting of sales in a bar graph makes much more intuitive sense by creating a simple shape that tells the user how the bars are arranged.  
  


*Figure 2: Two bar charts showing sales figures for eight hypothetical products; the top chart is randomly ordered while the bottom chart is in order of decreasing sales*

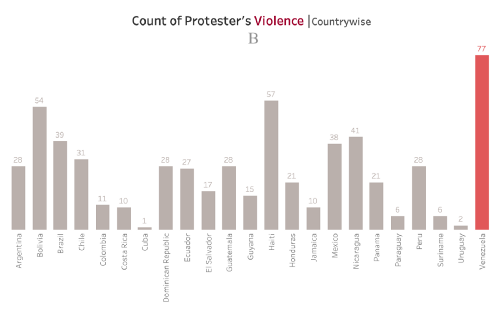
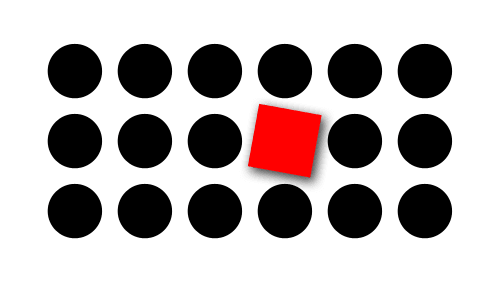
### Law of Similarity

Objects and aesthetics that are similar are considered to be in the same group (Figure 3).



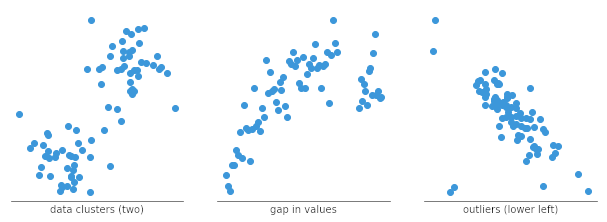
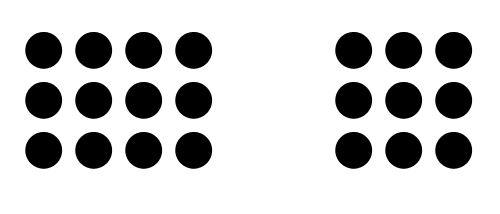
*Figure 3: Two graphs showing sets of dots of different colors; dots having the same or similar colors are perceived to belong to the same group*

### Law of Focal Point

Elements that are not similar to their surroundings are given priority attention (Figure 4). This is related to the law of similarity.

*Figure 4: Two images; the top image has a red square in a field of black dots and draws the viewer’s attention; the bottom image is a bar chart in which one bar is the tallest and drawn in red, which attracts the viewer’s attention*

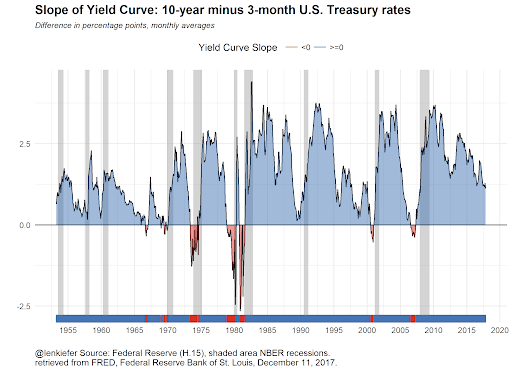
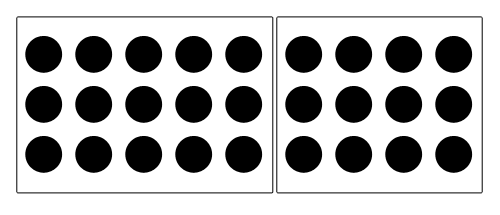
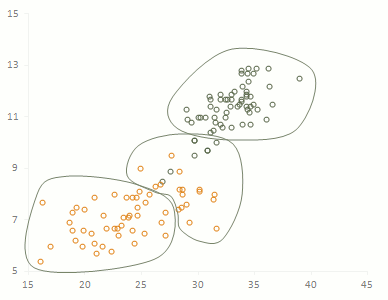
### Law of Proximity

The closer objects are to one another, the more they are perceived as belonging to the same group (Figure 5).  


*Figure 5: Two images of dots in which groups of dots are shown clustered, resulting in the viewer perceiving that the dots in a cluster are members of the same group*

### Law of Common Regions

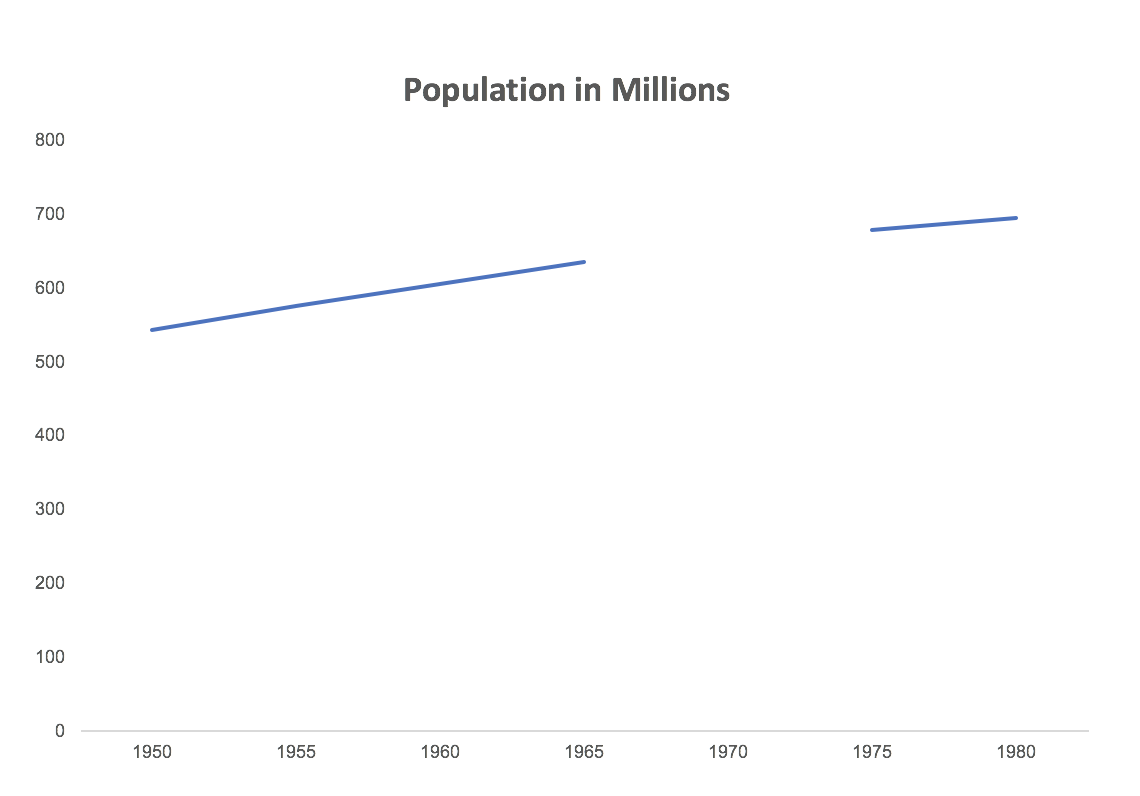
Objects within a boundary structure are considered to be members of the same group (Figure 6).

Different shapes grouped by a different color to visually separate objects by boundary structures that surround them
 

*Figure 6: Four images in which the group membership of separate objects is identified by boundary structures that surround them*

### Law of Closure

In some cases when objects are not entirely visible, the human mind fills in the gaps or imposes patterns and structure from memory. This process, called *reification*, suggests that we don’t need perfect information to find familiar patterns. With closure, we infer form, pattern, and meaning in the absence of complete information.



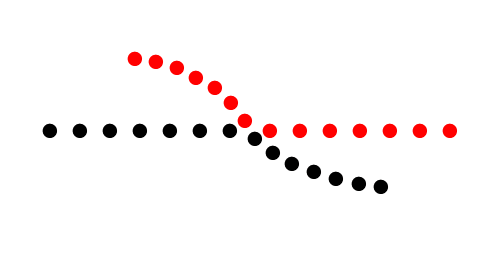




*Figure 7: Four examples of reification, in which the complete shape or structure of several occluded objects or graphs is automatically inferred by the viewer’s mind*

### Law of Continuity

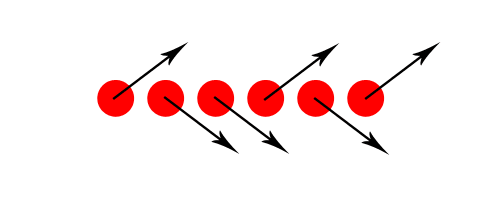
The human mind instinctively follows lines and shapes beyond their ending points. Missing information consisting of points, lines, and boundaries are effectively filled in thereby creating a sense of continuity in line, shape, and pattern. In Figure 8, we see an example of a curve crossing a line rather than two lines that avoid each other or four lines that meet at the center.

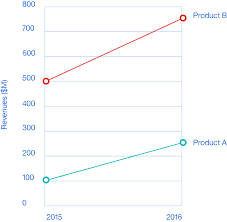


*Figure 8: Two lines made up of large dots appear to cross over, but based on the color, they are actually two lines that only come close to touching each other*

### Law of Common Fate (Synchrony)

Elements that move in the same direction are perceived as more related than elements that are stationary or move in different directions (Figure 9).





*Figure 9: In the series of large red dots shown, those moving upward (as indicated by upward-pointing arrows) are perceived as being related (belonging to one group) and those moving downward (as indicated by downward-pointing arrows) are perceived as being related but belonging to a different group (top image); revenues associated with both Product A and Product B appear to have increased between 2015 and 2016, but without knowing for sure, Product B could have increased substantially while Product A could have decreased substantially*

## Data Density and Visual Density

It’s been said that effective visuals convey the most amount of information with the least amount of ink. This is the principle of *maximizing the data-ink ratio,* and following this principle leads to decluttered visualizations, dashboards, and information layouts. Just as the grammar of graphics allows us to communicate in a visual language, the principle of data-ink maximization keeps visual statements short and to the point.

The *data-ink ratio* is a concept introduced by Edward Tufte, an expert whose work has contributed significantly to the design of effective data presentations. In his 1983 book, *The Visual Display of Quantitative Data*, he stated the following as the primary goal of visualization:

“Above all else show the data” (Tufte, 1983).

Tufte refers to *data-ink* as the non-erasable ink used for the presentation of data. If data-ink is removed from the image, the graphic loses its content. *Non-data-Ink* is the ink used for scales, labels, and boundaries—everything except the information being conveyed. Thus, the data-ink ratio is the proportion of ink that is used to present actual data compared to the total amount of ink used for the entire graphic. (In today’s parlance, we can replace “ink” with “pixels” when thinking about images presented digitally.)

Data-ink

Data-ink ratio = ————————————————

Total ink used to print the graphic

= proportion of a graphic’s ink devoted to the non-redundant display of data-information

= 1.0 — proportion of a graphic that can be erased

*Source:* [*InfoVis Wiki*](https://infovis-wiki.net/wiki/Data-Ink_Ratio)

Tufte then applies this as a principle of design: "Maximize the data-ink ratio, within reason.” Every bit of ink on a graphic requires a reason—and that reason is almost always the presentation of new information.

Let’s look at an example. Consider the chart in Figure 10, which shows market data for a particular stock. While the graph is certainly informative in terms of closing prices over a period of time, it doesn’t readily show whether the most recent market return was positive or negative. There is simply too much “ink” relative to the information we need. If we eliminate the graph altogether, we can simply display the previous day’s return and add a symbol (a plus sign, in this case) to indicate direction. For instance,

**+1.08%**

Or we could use color to denote direction, with green being defined as a positive increase:

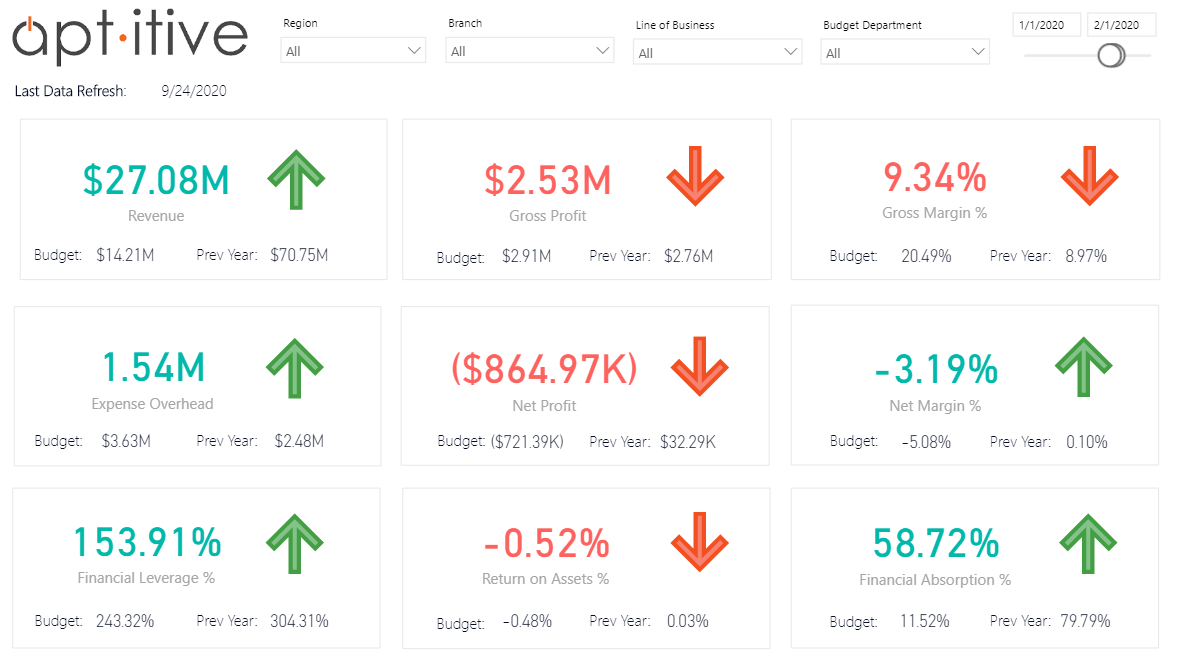
**1.08%**

A plot showing closing prices for a stock with months on the x-axis and price on the y-axis


*Figure 10: A plot showing closing prices for a stock with months on the x-axis and price on the y-axis*

This would be the maximum data-ink representation since we only care about the most recent change. If the only information required is direction, we could combine symbol and color to create an even more simplified representation, such as this upward-pointing green arrow, which reduces the time a viewer needs to process the relevant information.

As the amount of “ink” applied is reduced, the information that remains has a stronger impact as the overall display becomes less cluttered. Dashboards and reports designed with the data-ink ratio in mind tend to have higher data density and lower visual density, which make the results much easier to read at a glance (see Figure 11 for an example of this).



*Figure 11: An example of a dashboard designed with the data-ink ratio in mind. The financial figures shown are represented by numbers, colors (green means up while red means down), and arrows pointing upward or downward indicating whether the figures have increased or decreased.*

## Color

In the grammar of graphics, color can be mapped to specific features of the data you’re displaying. But this is a bit of an oversimplification because color actually has three intrinsic qualities:

* **Hue:** This is often what people think of as “color.” Moving along the scale of hues gives red, yellow, blue, etc.
* **Saturation:** This is the amount of color present, ranging from a neutral gray to a maximum color value.
* **Lightness:** This is intensity, ranging from dark to light.

These qualities can be mapped on continuous scales or discretized so there are only a small number of colors. Choosing between hue, saturation, and lightness, and deciding whether you want to use a continuous or discrete scale can be confusing for individuals who venture away from the default color mappings. Most software applications and libraries apply predefined color palettes to numeric or categorical data and take into account separability, color blindness, and limitations with either displaying or printing graphics. Naturally, you need to take these factors into consideration if you want to create a custom color palette for your charts.

Following are guidelines for choosing color qualities (hue, saturation, lightness) that are appropriate for the type of data you wish to depict.

* **Categorical scales**: Hue is used with categorical data where values do not have an inherent order (such as automobile brands). To account for color blindness, you can vary the lightness of different hues to make them more distinct from each other. You can also use a different hue (or lightness) to call attention to important quantitative values.
* **Quantitative scales**: When your data has an inherent order (such as ages and heights), you can use variations in hue, saturation, or lightness (or a combination of these as shown in Figure 12).
  + *Sequential* variations are used to depict values that range from low to high. They allow viewers to interpret the data without having to refer to a color legend. Typically, sequential variations use gradient-based lightness and saturation scales of one hue.
  + *Diverging* variations in hue are used to depict ranges of ordered values with a clear middle point between the values. They are also used to emphasize extreme values on either end of the scale or to better illuminate differences between values. With a more divergent scale, differences are visually amplified.
  + *Unclassed* variants of sequential scales are used when the data is continuously valued or should be interpreted as continuous. This helps give more detailed or nuanced views of the data in the visualization.
  + *Classed* (otherwise known as stepped, quantized, graduated, binned, or discrete) sequential scales are used when the data is discrete or can be made discrete through bucketing, binning, quantization, etc. Creating data bins often provides more clarity by showing greater separations of extreme values or inviting the viewer to focus on specific value ranges.

Examples of the use of color scales to depict quantitative (sequential and unclassed/classed; diverging and unclassed/classed) and qualitative (categorical) data


*Figure 12: Examples of the use of color scales to depict various kinds of data*

Deciding upon a color scale depends on the nature of the data, the aesthetic mapping, and the intention of the designer. Software can automatically provide the best guess for a color scale based on the data types. However, in some situations, the most effective color scale may not be based on data type, but rather on specific data that the analyst wants to highlight. In this case, the color scale is mapped to the degree of importance rather than to numerical values.

## Associations

Certain shapes and colors often have specific meanings that can vary greatly depending on your audience’s background. For example, in standard U.S. accounting, figures in green indicate an operating profit, figures in black indicate break-even, and figures in red indicate that the revenue is less than the operating expenses. However, the expression “in the black” refers to a company being financially healthy and profitable. This is a carryover from a bygone, pre-digital era when accountants used black and red ink in ledgers to indicate profit and loss. Keep in mind that it’s possible for people within the same country to misunderstand the meaning of colors in financial reports. Table 1 shows the symbolism behind different colors for different countries and regions.

| **International Color Symbolism** | |
| --- | --- |
| Black | * **Western:** death, evil, sin, funerals, mourning, rebellion, elegance, sophistication, formality, power, control * **Eastern:** wealth, prosperity, health, masculinity, knowledge, mystery, evil, mourning * **Europe:** mourning, death, evil, formality, elegance, sophistication * **Middle East:** rebirth, evil, mystery * **Africa:** age, wisdom, maturity, masculinity * **Brazil:** sophistication, authority, mourning, religion, formality * **Cherokees:** problems, death, represents the West * **China:** trust, high quality, color for young boys * **India:** laziness, anger, intolerance, apathy, evil, darkness, negativity, lack of desirability * **Italy:** death, mourning, humility * **Japan:** night, unknown, nonbeing, mystery, anger * **South Korea:** darkness, mystery, mourning * **Thailand:** unhappiness, bad luck, evil * **Australian Aborigines:** ceremonial color |
| White | * **Western:** purity, peace, elegance, wedding color, brides, angels, doctors, antiseptic * **Eastern:** death, mourning, funerals, sadness, purity, age, unhappiness, misfortune * **Europe:** purity, cleanliness, emptiness, neutrality, antiseptic, surrender * **Middle East:** purity, mourning * **Africa:** victory, purity * **Cherokees:** peace, happiness, represents the South * **China:** death, mourning, virginity, purity, humility, age, misfortune * **India:** creation, rebirth, light, serenity, reincarnation, peace, purity, unhappiness, widow, death, funerals * **Ireland:** leisure, sports, peace * **Italy:** virtue, purity * **Japan:** death, mourning * **South Korea:** purity, innocence, morality, birth & death, patriotism * **Thailand:** purity, auspiciousness |
| Yellow | * **Western:** happiness, joy, hope, optimism, creativity, energy, awareness, hazards, warning, weakness, femininity * **Eastern:** sacredness, sovereignty, earth, power, royalty, sun, happiness, masculinity * **Europe:** happiness, joy, quality, weakness, hazard, warning * **Middle East:** happiness, prosperity * **Africa:** wealth, high rank, religion, ceremony, visibility * **Australia:** resurrection, rebirth * **China:** sacredness, sovereignty, nourishment, royalty, honor, masculinity * **Egypt**: soul, sun, happiness, prosperity, mourning * **France:** summer, joy, jealousy * **Greece:** sadness * **India:** merchants, sacredness, auspiciousness, peace, happiness, meditation, mental development * **Japan:** courage, nobility, beauty, refinement, aristocracy, cheerfulness * **Mexico:** mourning * **Netherlands:** food * **Saudi Arabia:** strength, reliability * **Thailand:** Buddhism, royalty, auspiciousness, color for Monday |
| Green | * **Western:** nature, spring, new birth, regeneration, luck, go, environmental awareness, money, jealousy, greed, Saint Patrick's Day * **Eastern:** eternity, new life, regeneration, family, health, prosperity, peace * **Europe:** nature, fertility, confidence, jealousy, inexperience * **Middle East:** Islam, strength, fertility, luck, wealth, prestige * **China:** spring, youth, birth, infidelity, exorcism, desirability, disgrace * **Egypt:** hope, spring, fertility * **India:** Islam, hope, new beginnings, nature, harvest, virtue, happiness * **Indonesia:** a forbidden color * **Ireland:** Ireland – the Emerald Isle, patriotism, nationalism * **Islam:** perfect faith * **Israel:** bad news * **Japan:** eternal life, youth * **South Korea:** youth, energy * **North Africa:** corruption * **Saudi Arabia:** wealth, prestige * **South Africa:** death * **Thailand:** color for Wednesday |
| Blue | * **Western:** trust, loyalty, authority, conservatism, business, peace, calm, depression, sadness, masculinity * **Eastern:** immortality, wealth, self-cultivation * **Europe:** truth, responsibility, fidelity, serenity * **Middle East:** safety, protection * **Belgium:** traditionally used for baby girls * **Cherokees:** defeat, trouble * **China:** immortality, femininity * **Egypt:** truth, justice, virtue, faith, protection, reproduction * **France:** water, reliability, trust * **India:** heavens, love, truth, mercy, strength, bravery, manliness, determination * **Indonesia:** sadness * **Iran:** heaven, spirituality, mourning, immortality * **Israel:** coat of arms * **Italy:** mourning, heaven, purity * **South Korea:** integrity, mourning * **Mexico:** trust, serenity, mourning * **Thailand:** color for Friday * **Turkey:** healing, wealth, repel evil |
| Purple | * **Western:** royalty, spirituality, faith wealth, fame, high-ranking positions of authority, luxury, calm, magic * **Eastern:** wealth, nobility, privilege, sorrow, mourning * **Europe:** royalty, nobility, luxury, power, vanity * **Middle East:** wealth * **Brazil:** death, mourning * **Egypt**: virtue, faith * **France:** freedom, peace * **India:** sorrow, unhappiness * **Italy:** death, mourning, endurance of suffering, nobility * **Japan:** wealth, privilege, power, royalty * **Thailand:** death, mourning, widows, color for Saturday |
| Pink | * **Western:** love, romance, femininity, childhood, babies, nurturing, sweetness, Valentine's Day * **Eastern:** marriage, femininity * **Europe:** baby girls, delicate, flirtation, sensitivity, serenity, femininity * **Belgium:** traditionally used for baby boys * **Japan:** spring, femininity, youth, good health, well-liked by both males & females * **South Korea:** trust * **Thailand:** color for Tuesday |
| Red | * **Western:** stop, danger, anger, blood, energy, excitement, action, adventure, love, passion, Valentine's Day * **Eastern:** prosperity, good fortune, happiness, worn by brides, celebration, communism * **Europe:** visibility, cheapness, loudness * **Middle East:** danger, caution, evil * **Africa:** death, mourning * **China:** good fortune, happiness, celebration, summoning, vitality, prosperity, long life, wedding color * **England:** power, authority, government * **France:** passionate love, lust, virility * **India:** purity, love, sensuality, fertility, beauty, wealth, power, wedding color, married woman, fear, fire * **Italy:** light, fidelity * **Japan:** life, fertility, passion, strength, anger, danger, self-sacrifice * **South Korea:** good fortune, passion * **Netherlands:** nature, government, royalty * **Russia:** communism, revolution, wedding color * **Thailand:** Buddhism, color for Sunday * **Turkey:** death |
| Orange | * **Western:** energy, vitality, excitement, adventure, creativity, caution, construction, harvest, autumn, affordability, Halloween (with black) * **Eastern:** love, happiness, spirituality, humility, good health, immortality * **Europe:** purity, cleanliness, good * **Middle East:** mourning, loss * **Australia:** animals * **Brazil:** environment * **France:** earth * **India:** the most sacred color, purity, courage, sacrifice, religious abstinence, death, rebellion * **Ireland:** religious color for Protestants, appears on the Irish flag along with white for peace & green for Catholics * **Japan:** love, happiness, courage * **Netherlands:** royalty, color of the Dutch royal family * **Thailand:** color for Thursday |
| Brown | * **Western:** comfort, stability, practicality, wholesomeness, dullness, fertility, organic, earthiness, poverty * **Eastern:** earth, mourning * **Europe:** masculinity, earth * **Middle East:** earth, comfort * **Africa:** earth * **Australia:** color of the land * **Brazil:** nature * **China:** earth * **India:** color of mourning * **Indonesia:** earth * **Japan:** earth, strength, durability * **Nicaragua:** disapproval |

*Table 1: International color symbolism Source:* [*PBS*](https://pbs.twimg.com/media/C4s7yX_UoAE8nYN.jpg)

Shapes also have particular meanings. In some cultures, a circled figure in a table indicates a problem with the number. However, in Japan, circles are used to denote correct or supported figures while triangles are used to identify questionable or partially supported figures. In some disciplines, a triangle indicates a change—the number is the difference between two other numbers. Needless to say, the same mark can have different meanings in different contexts and for different audiences.

For reports, presentations, dashboards, and other externally-facing visualizations, it is best to have someone familiar with your target audience’s culture to check for any potential problems that may arise in connection with the colors or symbols you used.

# Storytelling

As the analytical tools available to us become increasingly sophisticated, audiences often find it increasingly challenging to understand the analytical results those tools produce. This makes it concomitantly difficult to foster data literacy, in general. While visualizations can be powerful on their own, audiences are more interested in answering the following questions:

* How does it apply or why does it matter?
* What actions are expected as a result of having this information?

If we can answer these questions clearly, we can begin crafting a “pitch” using storytelling methods designed to make our presentations, reports, and analyses more compelling and likely to garner “buy-in.” Remember that communication in general, whether by visual or oral means, serves to achieve an outcome. This outcome can be immediately actionable or can generate further questions. It can feed the next stage of the analytics pipeline or influence strategic thinking. Most importantly, its assertions must be supported by the data being analyzed.

## Reasoning

In advanced analytics, statistical and computational methods are used to process data and generate predictions, classifications, and regressions. In descriptive analytics, reasoning methods are used to investigate the data, draw hypotheses, and draw conclusions through evidence-based reasoning. Specifically, there are three main types of reasoning:

1. **Deductive**: Deductive reasoning is top-down. We begin with an assertion—a general rule—from which we form a hypothesis. Next, we test this hypothesis against known facts and eventually reach a conclusion. In other words, we reduce facts from a general theory to specific, factual conclusions.

For example, suppose we have a general rule for public companies stating that if company A is about to acquire company B, the price of company A’s stocks will decrease and the price of company B’s stocks will increase. We can then deduce that if Emergent Technologies, Inc. makes a bid to acquire Pacific Telecom, the share price of Emergent will decrease while the share price of Pacific will increase just before the acquisition is completed.

1. **Inductive**: Inductive reasoning is bottom-up. We extrapolate general theories from specific facts. Analytically, we examine a large dataset to determine the likelihood that a particular hypothesis is correct. Unlike deduction, induction does not lead to a logical necessity but instead leads to what could be the consequence of statistical, computational, or even heuristic approaches. The important difference is that induction can be ampliative. It can make predictions or statements about heretofore unobserved phenomena. To apply induction, you must:
   1. Collect data with a specific scope
   2. Draw conclusions based on the data
   3. Generalize the conclusion to all cases or future cases

For example, 80% of restaurants go out of business after 5 years. Restaurant A has been in business for over ten years. Restaurant A most likely *won’t* go out of business this year.

1. **Abductive**: The abductive process generates a hypothesis that explains the relationship between a case and a rule. In other words, it is an assessment of probabilities based on experience and the supposition that an underlying explanatory narrative exists rather than strictly a form of calculation.

For instance, suppose you observed that online user traffic on your site suddenly dropped by millions of users last week. Also, you know that Google updated its search engine algorithm last week. You propose three hypotheses:

* 1. Users are suddenly no longer interested in your site (perhaps due to a boycott or negative social media)
  2. A competitor has hijacked your user traffic
  3. You abduce that the search engine is demoting sites that are not mobile-friendly because your team has not yet moved to a responsive design

Reasoning gives you the method by which you can make your case and sell your approach. The specific method you use depends on the argument you are trying to make, the available data and evidence, and the audience.

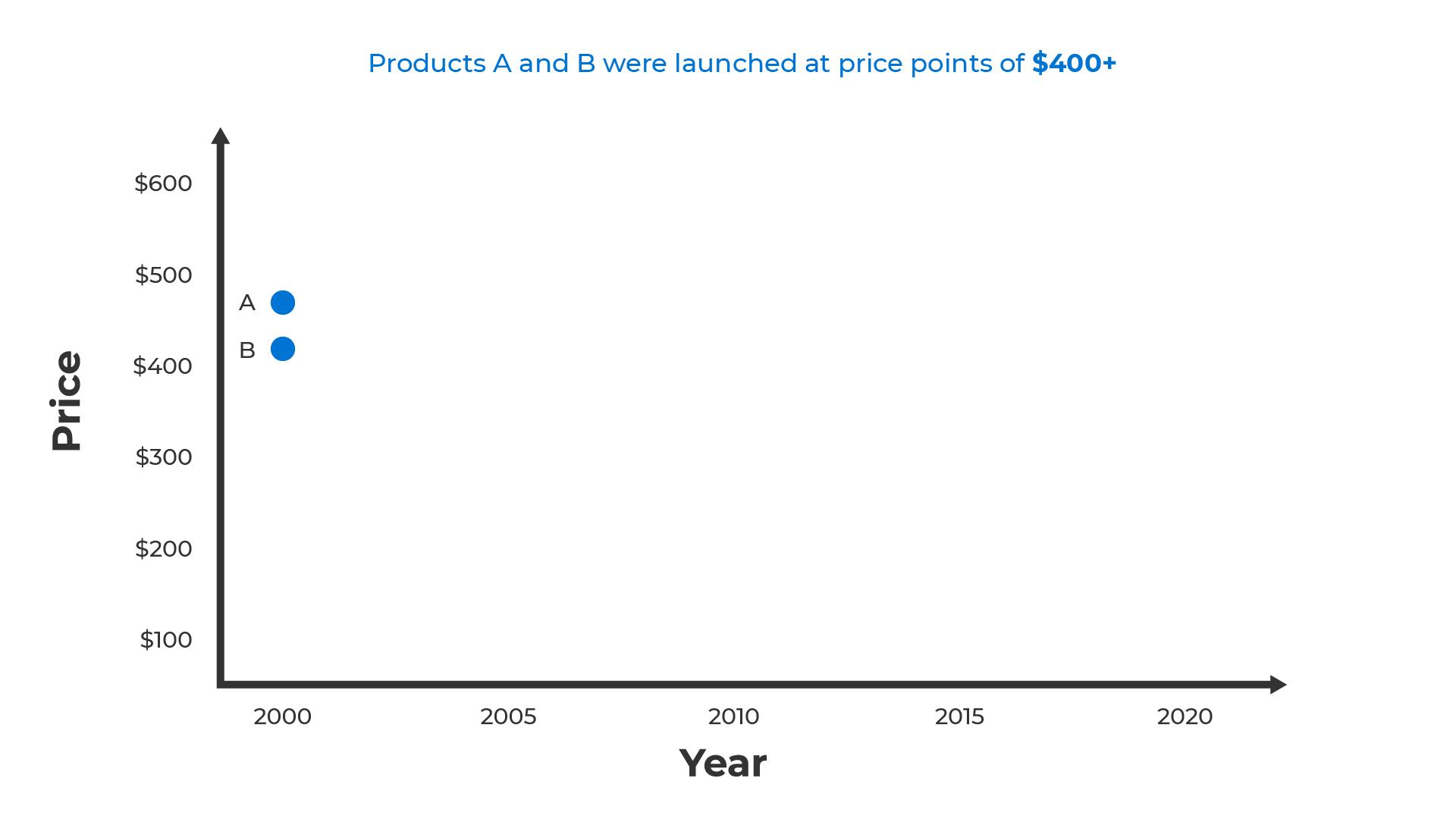
## Structure

Reasoning alone may not “sell” your story because it often fails to convince the audience that (1) a particular issue matters to them, (2) certain conflicts or problems do, in fact, exist, (3) they have options for resolving those conflicts or problems. A good story, like a great play, movie, or book, has a certain structure that can grab (and keep) the audience’s attention. This structure can be broken down into three parts.

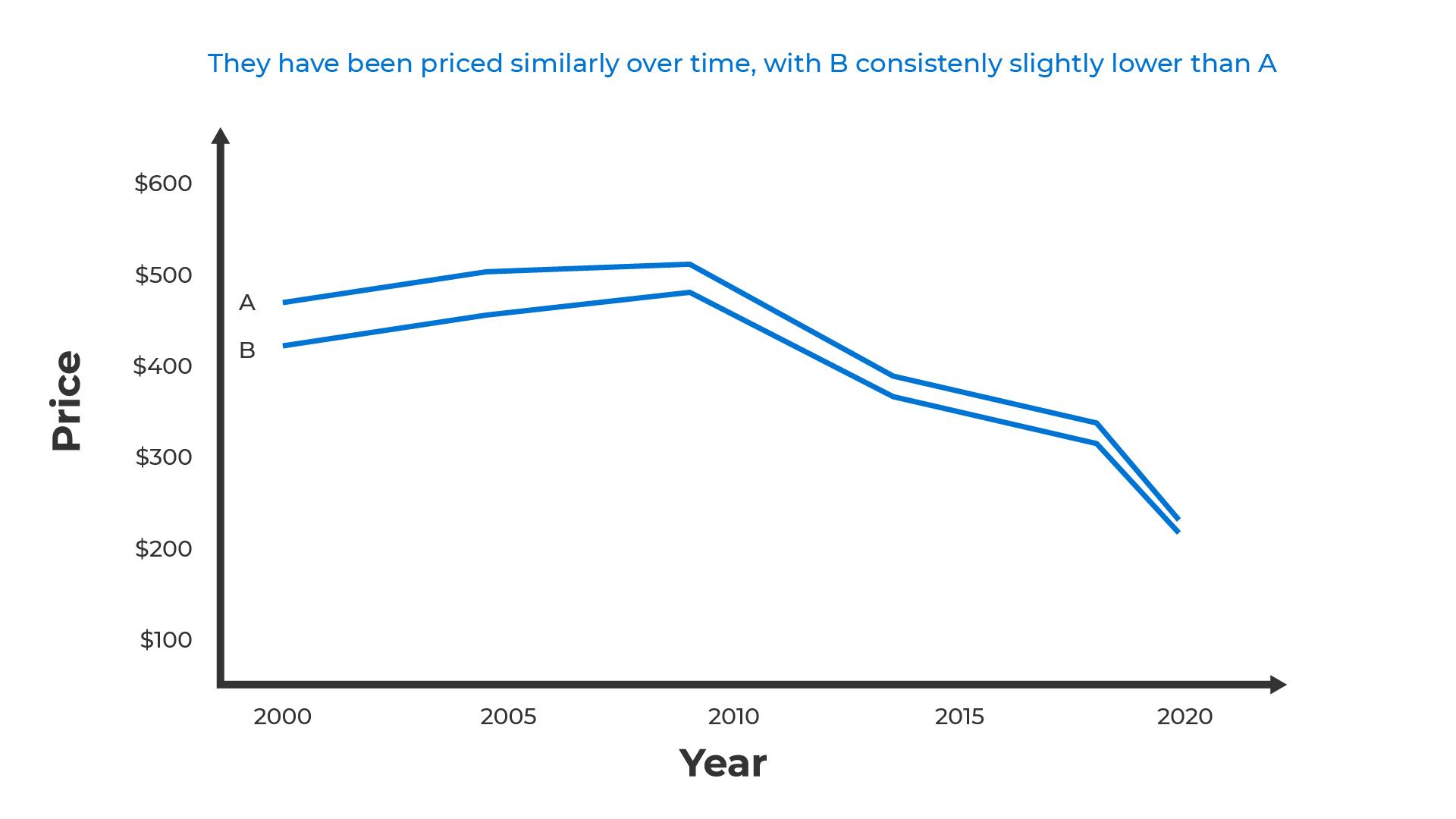
* **First act**: We build context for the audience by explaining why they should care and why the issue being communicated is important. This includes:
  + **Setting, environment, and data context:** This defines the data and feature space, the conditions surrounding the data, and the conceptual mapping from the data space to the business domain.
  + **Focus of attention**: This includes key performance indicators, metrics, and observations containing evidence.
  + **The imbalance**: This consists of conflict or tension. The imbalance could be lower-than-expected KPIs or it could be opportunities like lagging competitor sales rates in a region. This is where reasoning and descriptive analytics observations come into play for making a case.
  + **Desired balance**: What do they want to see happen? What would the ideal outcome be? There should be a substantive difference between the imbalance and the balance. Otherwise, the story is not compelling. A desired balance could consist of a target range for KPIs or it could lead to prediction models for customer behavior.
  + **Solution**: The solution brings about the desired changes. Several possibilities for resolving the conflict or imbalance are introduced at this point. These are presented as options to consider, which leads to the next act.
* **Second act**: Once the audience appreciates the existence of a problem or tension, you need to present possible solutions and explain why they should pick one and proceed with it.
  + Show what will happen if no action is taken.
  + Describe several options for action along with their benefits.
  + Explain the benefits if action *is* taken.
  + Develop motivation for taking action.
* **Third act**: Review the problem and the reasons why finding a solution is urgent. Focus on creating a *call to action*, which urges the audience to engage with the story rather than be a passive observer. Not all descriptive analytics projects end in concrete recommendations for action, but at a minimum, it’s important to take a second or third look at the “what happened” question in preparation for asking the much harder “why did it happen” question. (Or ask the “how can we make it happen” question.)

The combination of reasoning and structure creates a compelling story that is much more effective in reaching and changing the minds of the intended audience than simply providing a deluge of charts and figures. The ability to guide the audience in a step-by-step manner via a narrative consisting of data visualizations helps make the audience feel invested in the solution.

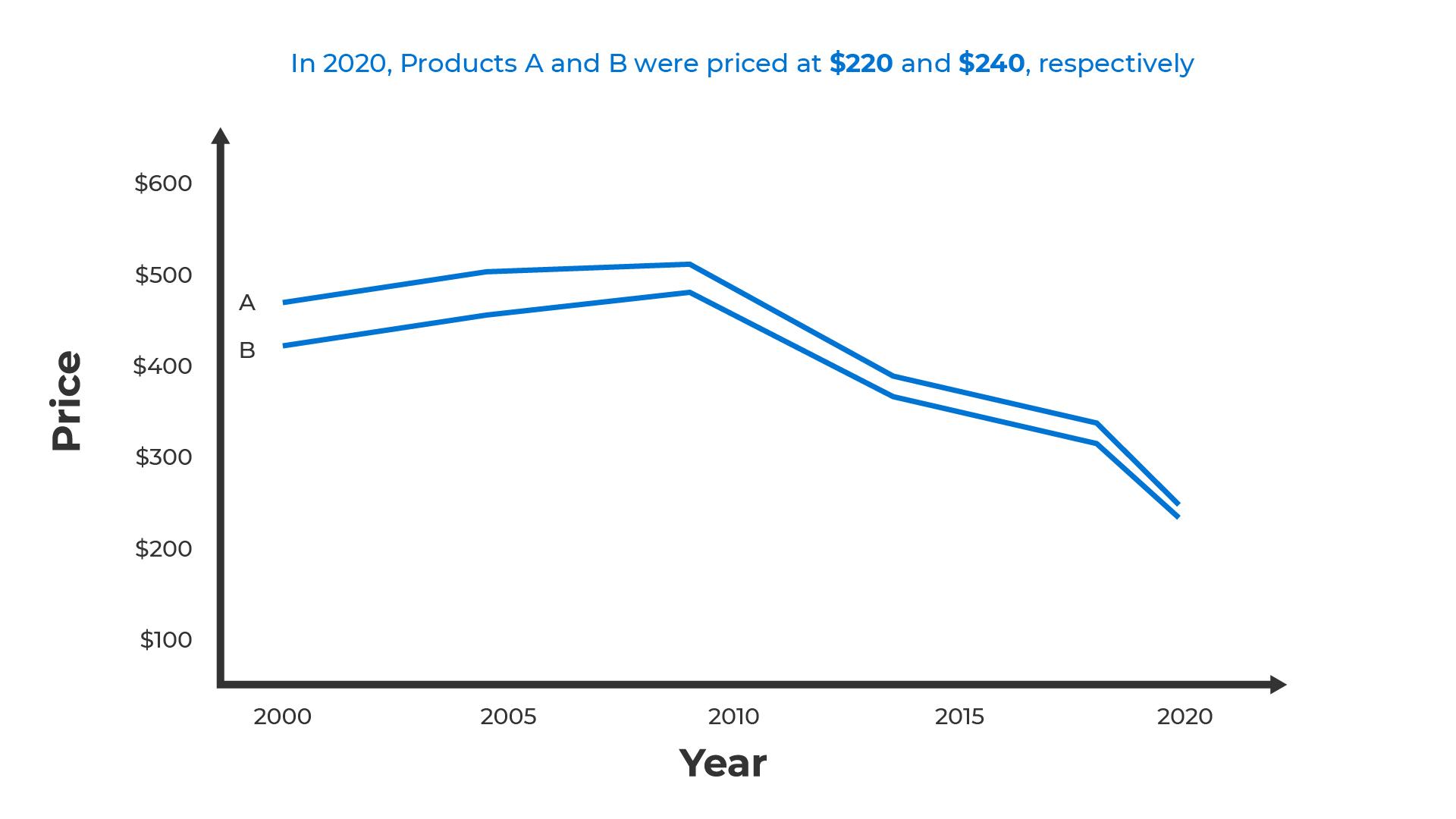
In the following example, you can see how a visual representation of data communicates a narrative of changes over a period of time that can inform decision-making and lead to specific recommendations (Figure 13-20).



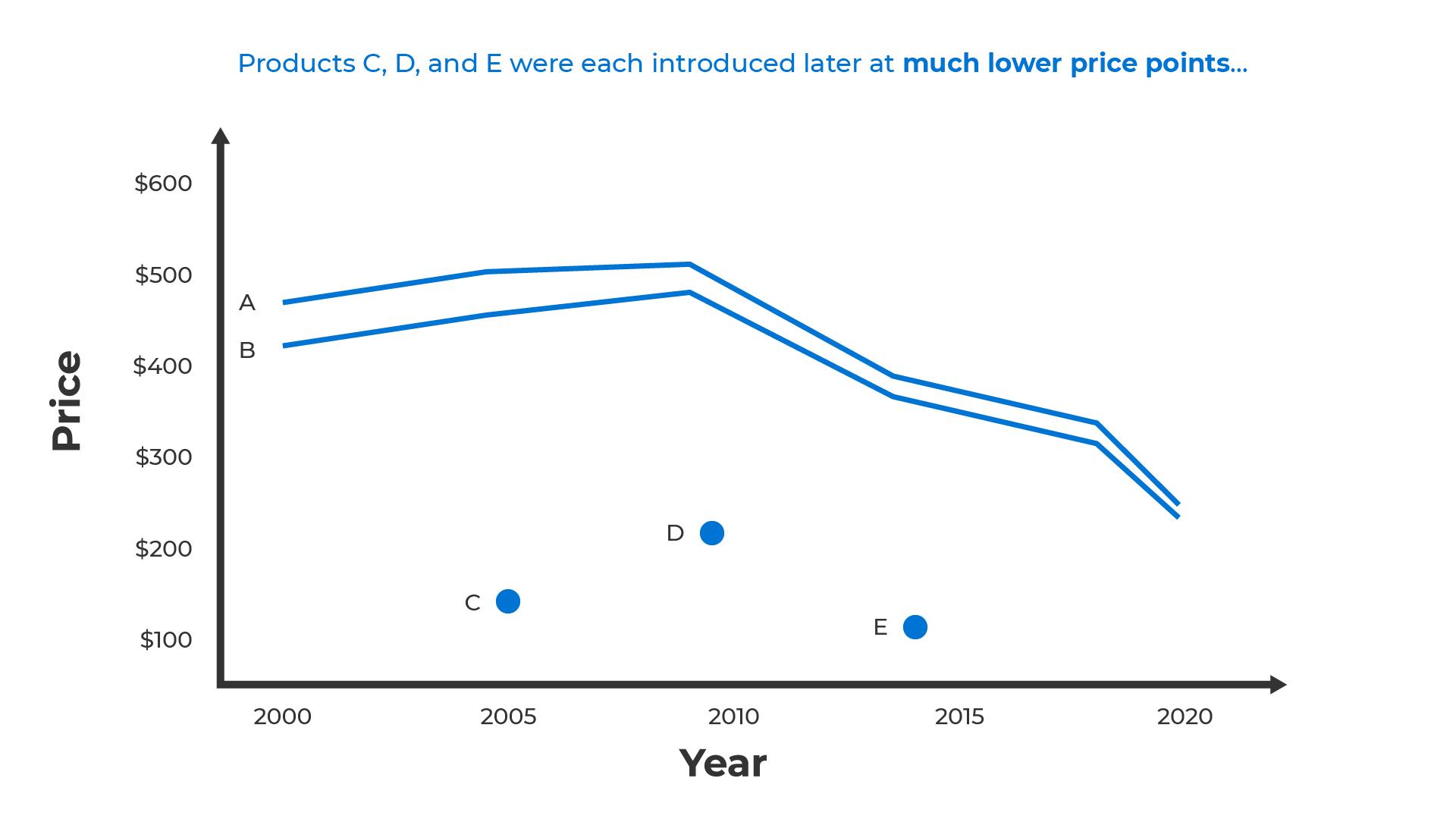
*Figure 13: A graph shows two products (A and B) launched at price points between $400 and $500 in the year 2000.*



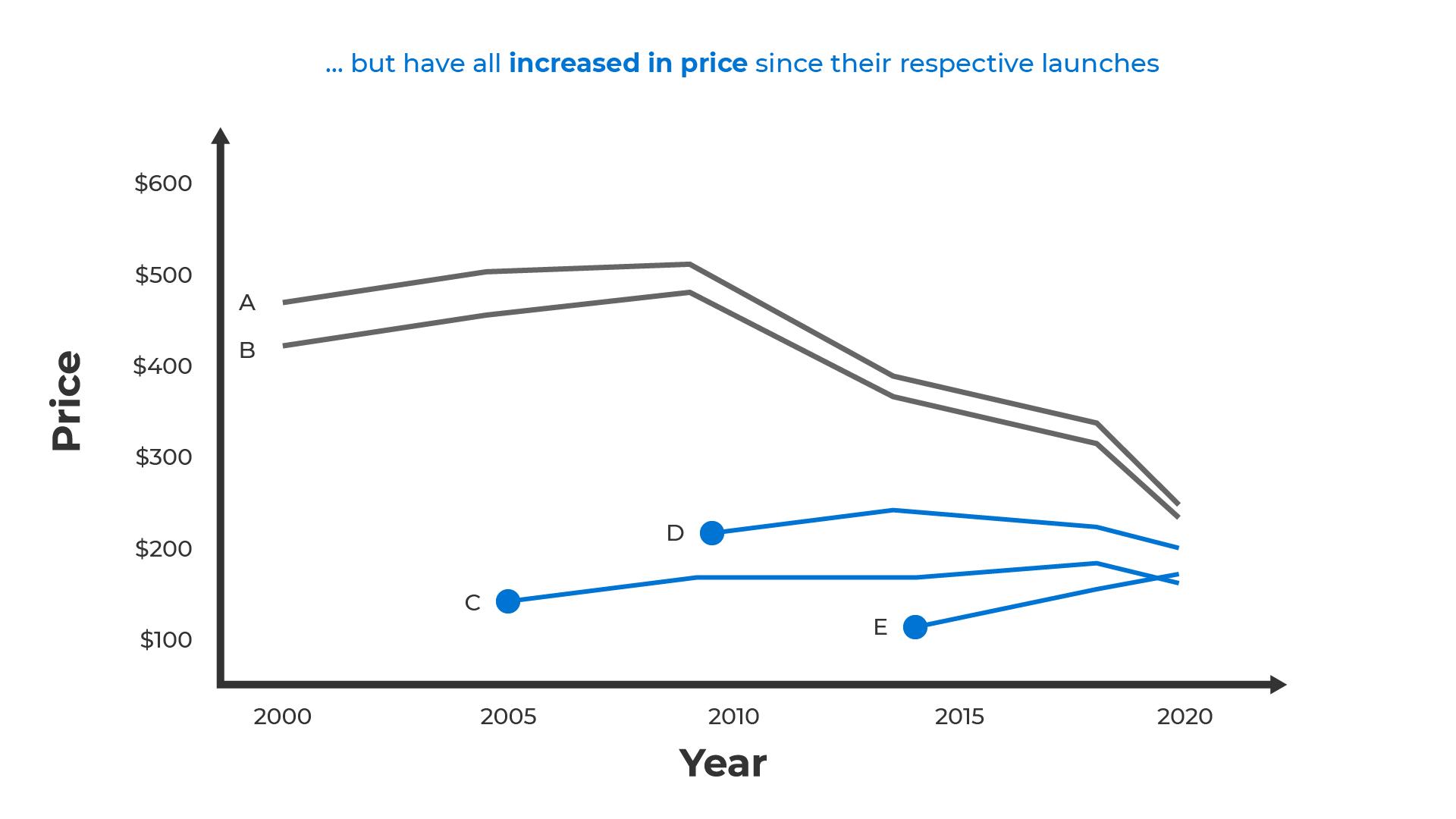
*Figure 14: The same data as shown in Figure 13 illustrating the price of products A and B over time, with B consistently slightly lower than A*



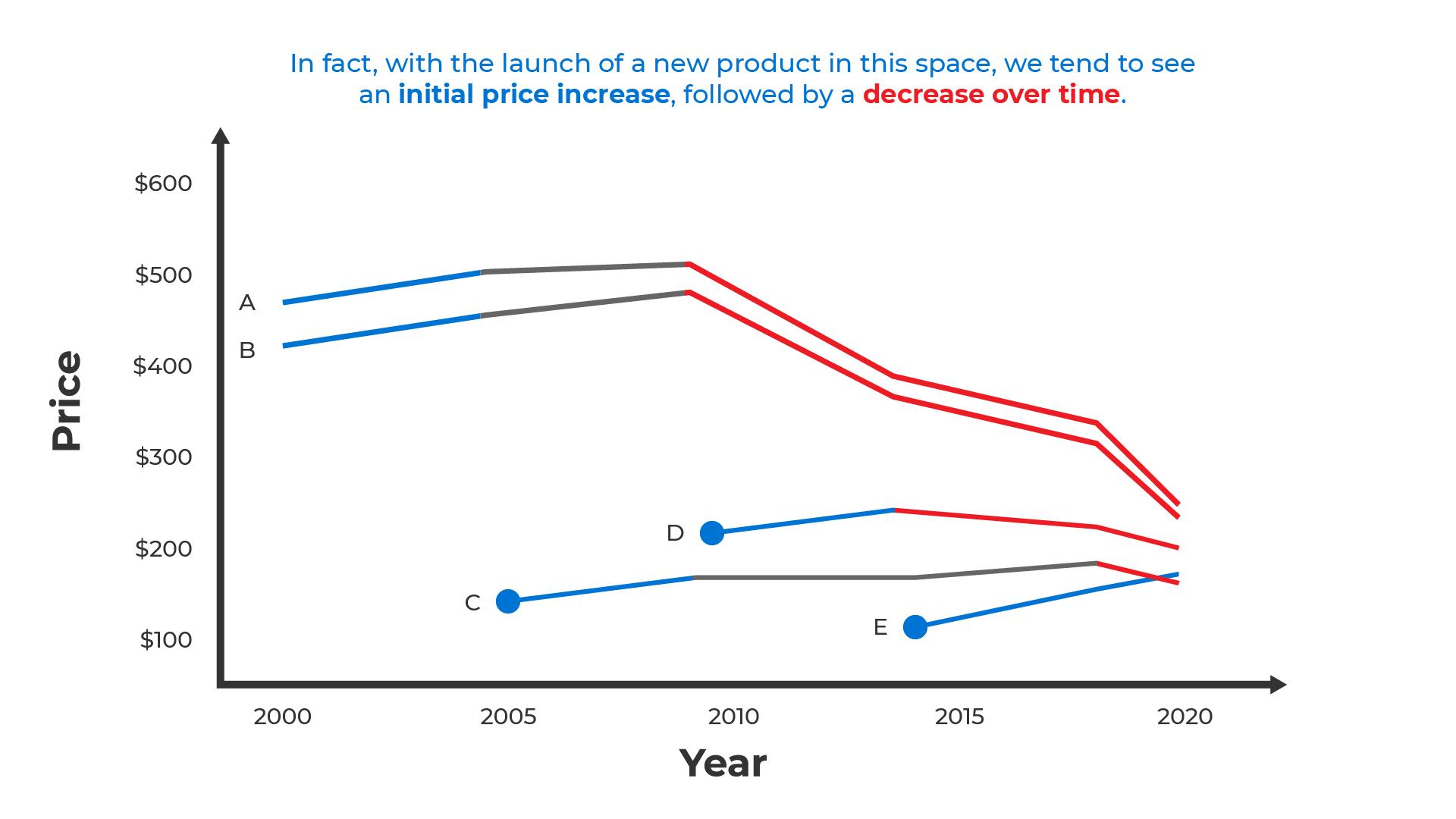
*Figure 15: The same data as shown in Figure 13 illustrating the price of products A and B in 2020 at $220 and $240 respectively.*



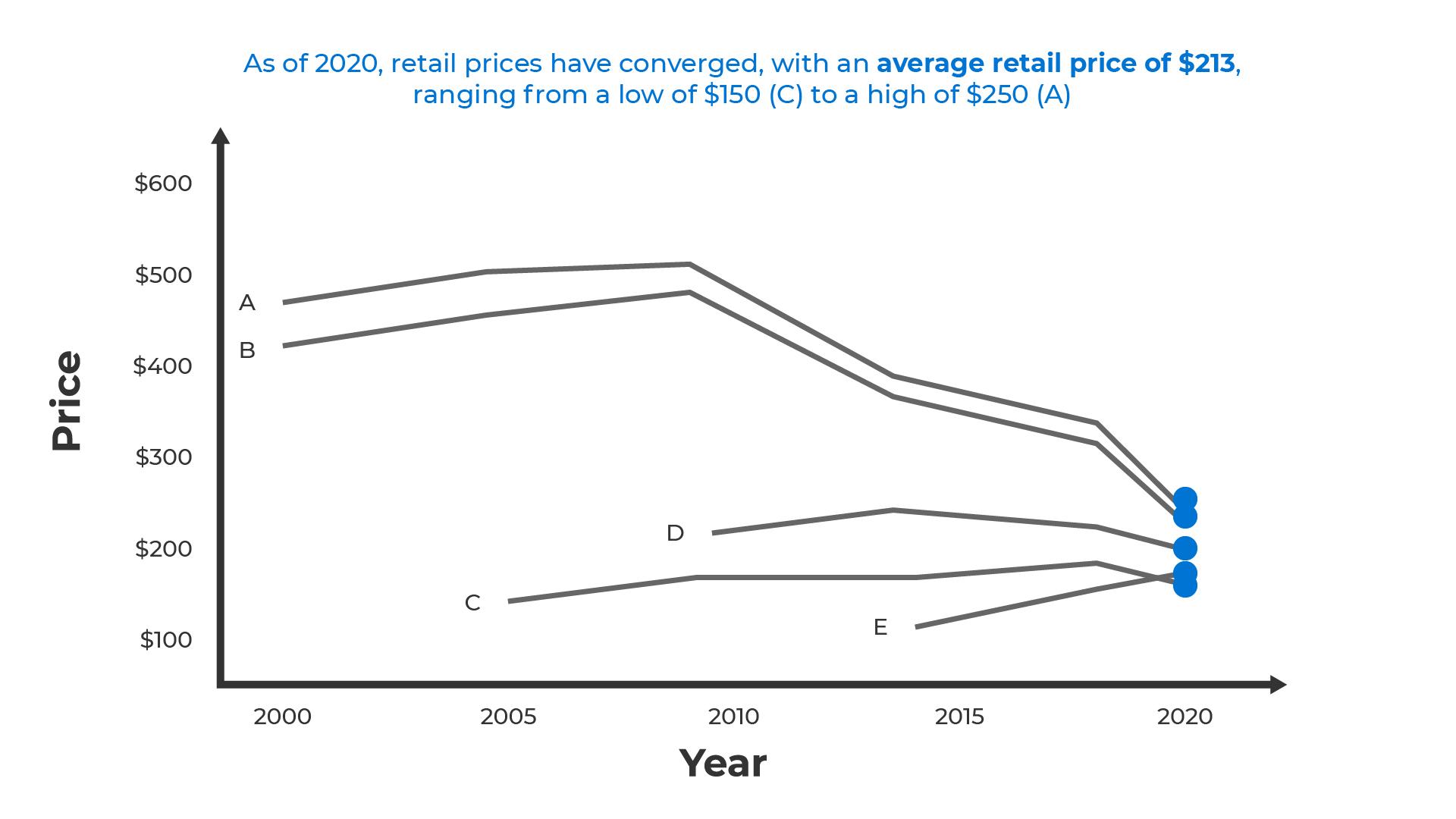
*Figure 16: The same data as shown in Figure 13 introducing other products, C, D, and E at much lower price points*



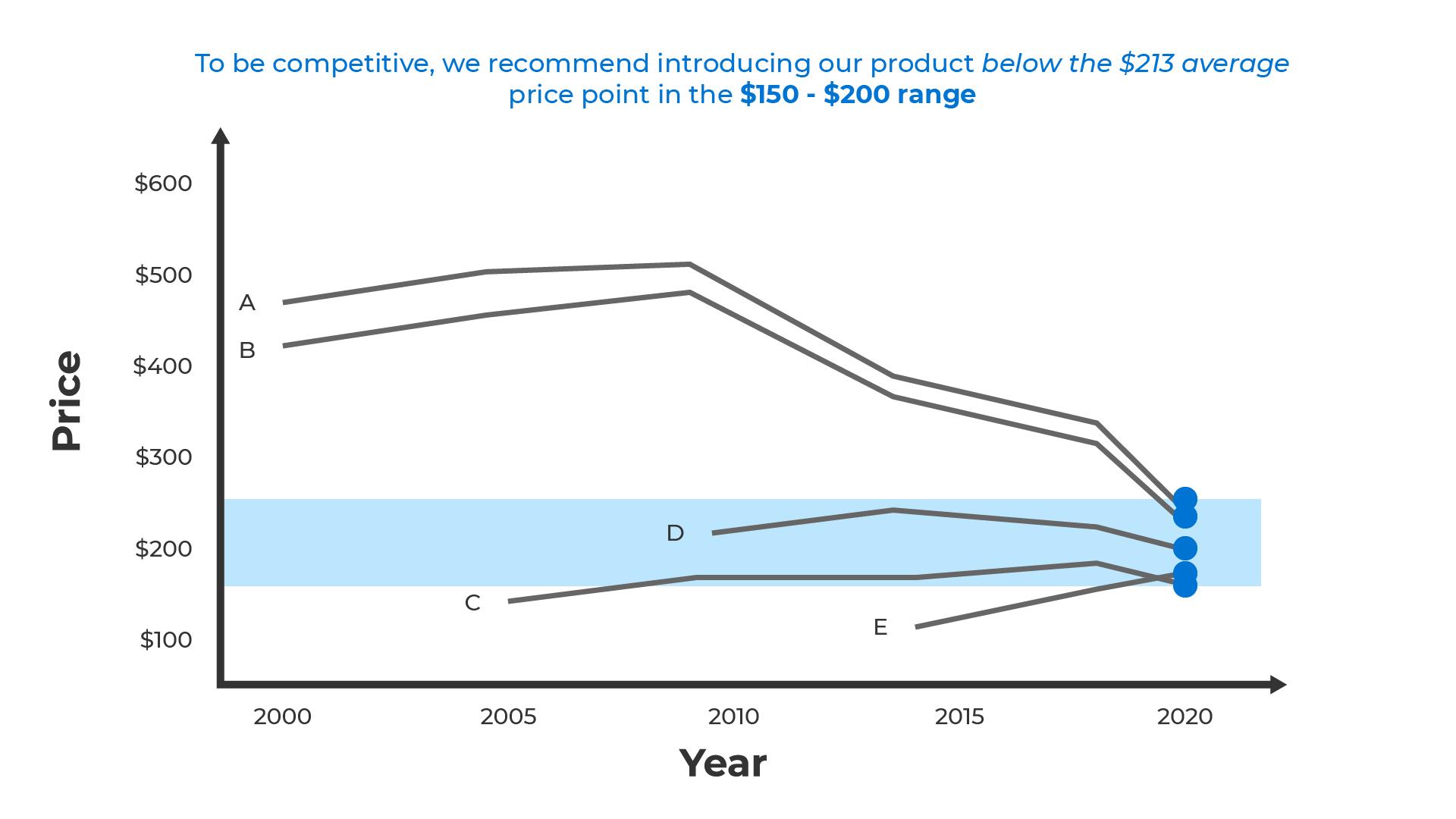
*Figure 17: The same data as shown in Figure 14 illustrating the increases in price for products A, B, C, D, and E since their respective launches*



*Figure 18: The same data as shown in Figure 14 illustrating a trend in price increase for all products (A, B, C, D, and E) after initial launch, followed by a decrease over time*



*Figure 19: The same data as shown in Figure 14 illustrating the average retail price of $223 for all products (A, B, C, D, and E) in 2020, ranging from a low of $150 © to a high of $250 (A)*



*Figure 20: The same data as shown in Figure 14 highlighting a price point in the $150 - $200 range as a recommendation to introduce products below the $213 average*

# Key Points

Communicating effectively through analytics is one of the biggest challenges data analysts face. It’s also one of the main reasons why businesses are not eager to adopt analytics.

However, with some careful consideration communication can be improved by:

* Creating effective visuals:
  + Use the *Gestalt principles* of perception to create optimal visual encodings of the data
  + Maximize the *data-ink ratio* to reduce clutter and improve readability of the visualizations
  + Employ appropriate *color scales* to improve readability of the data and to avoid misinterpretation
* Telling stories:
  + Use reasoning methods like *deduction*, *induction*, and *abduction*to frame arguments and rationales.
  + Apply structure to engage the audience and show them how the information being communicated applies to them, why they should care about it, and what they can do about it. End with a *call to action*.

# Reference

Tufte, E. (1983). *The Visual Display of Quantitative Data*.

1. A survey of 64 C-level technology and business executives representing large corporations like Bloomberg, Bank of China, Sumitomo Bank, Ford, GE, J&J, and American Express was conducted by NewVantage Partners in 2019. 77.1% report that adoption of big data and analytics initiatives remained a major challenge. 95% report cultural (people, process, and communication) issues as the source of the challenge. Only 5% report technology as the source.

   <http://newvantage.com/wp-content/uploads/2018/12/Big-Data-Executive-Survey-2019-Findings-Updated-010219-1.pdf> [↑](#footnote-ref-0)
2. The Gartner study “Fostering Data Literacy and Information as a Second Language” also highlights the fact that the barriers are rooted in poor communication <https://www.gartner.com/en/doc/3860965-fostering-data-literacy-and-information-as-a-second-language-a-gartner-trend-insight-report> [↑](#footnote-ref-1)
3. The illustration of the effect of letter randomization was originally presented in 1976 by Graham Rawlinson in his Ph.D. thesis, “The significance of letter position in word recognition” [↑](#footnote-ref-2)