Final Project: 7-3-3 Create a Narrative Visualization

CS 498: Data Visualization—Summer Semester 2019

Abstract

Final Project consisting of a Narrative Visualization coded in D3 and javascript, and associated Essay Writeup.

Matthew W. Bennett

Table of Contents

Introduction and Overview

The purpose of this document is to provide details relating to the Narrative Visualization Project for the UIUC CS 498 Data Visualization class for the Summer 2019 term. The goal of the assignment was to create a Narrative Visualization to be implemented through a dynamic and interactive web page using D3 programming with JavaScript. As part of the assignment we were to choose a specific narrative visualization structure ("martini glass", "interactive slideshow", or a "drill down story") to showcase our knowledge of visual communication and what we have learned throughout the class lectures and assignments. Through visual storytelling, our goal was to provide a specific message to the intended audience that walks the viewer through details of a data set while allowing them to interact with the data as little or much as they want while receiving the message.

Throughout this document, the details of my Narrative Visualization will be presented with specific focus on pertinent topics including the intended message, the narrative and visual structure of the visualization, and descriptions of the specific scenes that were created. In addition, I will detail the components that make up the scenes, including annotations, parameters, and the specific triggers that were used. I will conclude with an overall summary of the work that was done and a conclusion with respect to how the assignment specifications were met.

The Interactive Slide-Show narrative visualization that I created for this assignment can be accessed at https://mbennett-wa.github.io./index.html.

Messaging

For this narrative visualization, the specific message that I am try to communicate with the viewer can best be described as the current state of automobile efficiency for gasoline powered engines using data from 2010 through 2017. Please note that automobiles that utilize electricity or diesel to power their engines are specifically excluded from the visualization. The narrative visualization begins with a scatter plot diagram which plots the (Gasoline) Average City Miles Per Gallon (MPG) against Average Highway MPG for the various automobile manufacturers in the United States. This graph is intended to give a quick visual of the breadth of MPG averages for the year 2017. On the graph the viewer is able to see details of the Manufacturer, Average City MPG and Average Highway MPG for each point on the graph via a mouse over. After viewing this scene, the viewer should have a very good understanding of which manufacturers have poor MPG vehicles and which have the best MPG Averages.

To continue with the messaging, the viewer is able to get even more details about each specific manufacturer by clicking on the individual dot points on the graph. When this is done, the viewer is able to investigate the details further by seeing a scene that shows a detailed bar

graph of the breakout of the Highway MPG and City MPG in one place. Should they so desire, the viewer can select individual manufacturers from a drop-down list and the bar graph will filter and update accordingly to show just the specific City and Highway averages for that manufacturer.

Returning from that detailed narrative visualization slide to the main scatter plot slide, the messaging continues by allowing the viewer to go to the final scene where they can get historical details for last eight years on the specifics of the Average combined MPG by manufacturer annually from 2010 through 2017. This final bar graph is also dynamic in that it allows the viewer to drill down into the data by selecting a specific manufacturer from a drop-down list on the page.

When the viewer completes this final scene, the viewer will have received a good understanding of the details of gasoline engine MPG data by year and by manufacturer for the last eight years.

Narrative Structure

For this narrative visualization, I chose the interactive slide-show structure. This type of narrative structure is defined as somewhat of a hybrid presentation structure that starts out by providing the viewer with some baseline information on the message to be communicated. Once this initial messaging is provided, the viewer then has the ability to either continue on to the next predefined scene, or get more insight and information on the current scene by doing some exploring of the details and drill-down options that are provided. The important aspect of this type of visualization is that the viewer has the choice regarding what level of information they want to get out of the presentation. If they want, they can simply run through the high-level scenes to view the data, or, if they are interested in knowing more, they can take advantage of the drill-down features that are provided and actually do some exploring on their own.

In my interactive visualization, the viewer is presented with a starting scene that sets the stage for the overall narrative. On the first slide, the viewer is presented with a graphical representation of the state of gasoline fuel economy by manufacturer for 2017. This visualization provides a lot of information very quickly about automobile manufacturers and how well their cars stack up against the other cars in the US market. There are a number of easily understood data points on this main scene to adequately interest and educate the viewer.

Should the viewer's interest and curiosity be piqued by this initial data however, the viewer is provided with some visual tools which quickly provide more information by simply moving the mouse pointer over the individual plot points. When this is done, a tool-tip becomes visible

which provides the manufacturer's name and the City and Highway mpg numbers associated with that data point.

If the viewer wants even more details, simply clicking on one of the individual points on the scatter plot diagram provides access to a drill-down scene that opens in a new window. This new window provides a detailed bar graph that displays the Average City MPG and the Average Highway MPG in one plot. The power of the drill-down is seen in the addition of a dropdown menu at the top of the graph. In this drop-down is a list of all the manufacturers that were a part of this data, and when the viewer selects one of them from the list, the bar graph dynamically changes to show MPG data that is specific to the selected manufacturer. As with the scatter plot, the viewer is able to access on demand details about the bar charts simply by moving their mouse cursor over each chart, at which point a tool-tip becomes visible with more details on the currently selected MPG values.

When the viewer is satisfied with their lower level data exploration they can return back up to the main level scene with the scatter plot diagram. In following the definition of the interactive slide-show, the viewer then has the ability to continue on with the story by clicking on the button that will take them to the next scene. On this next scene they will find historical data that they can compare against the 2017 data that has already been presented to them. Like the other bar graph in the previous scene, the viewer is able to select from a drop-down list of manufacturers to view historical data on gasoline engine MPG from the last 8 years. Additionally, like the previous bar graph in scene 2, the viewer can get even more details on the graph simply by moving their mouse pointer over each of the bars on the graph. When this is done, a tool-tip is displayed that includes the Average combined MPG for that manufacturer for that particular year. In all of the bar graph scenes, the Y axis values and the redrawing of the bars themselves are animated so that the viewer is able to easily see the changes that exist in the values as they cycle through various manufacturers.

Visual Structure

The visual structure of a narrative visualization is an important aspect since it defines the look and feel for the message as well as the visual methods by which the messaging will be presented. The important aspects of the visual structure include how the scenes are laid out and with what kind of perspective, what items are highlighted to encourage the viewer to focus on the important elements and perhaps want to dig deeper, and how do the scenes transition from one to the next to ensure that the narrative of the story is maintained. All of this is done so that the viewer does not become lost or confused by the presentation, and there is a clear understanding about how everything is connected.

In my interactive slide show narrative, I purposely begin with a wide-angle visual structure that presents to the viewer the highest level of detail for the data. This is done with a more generalized scatterplot diagram that shows how different car manufacturers of gasoline engines are doing with respect to others in terms of MPG. The graph is presented directly in

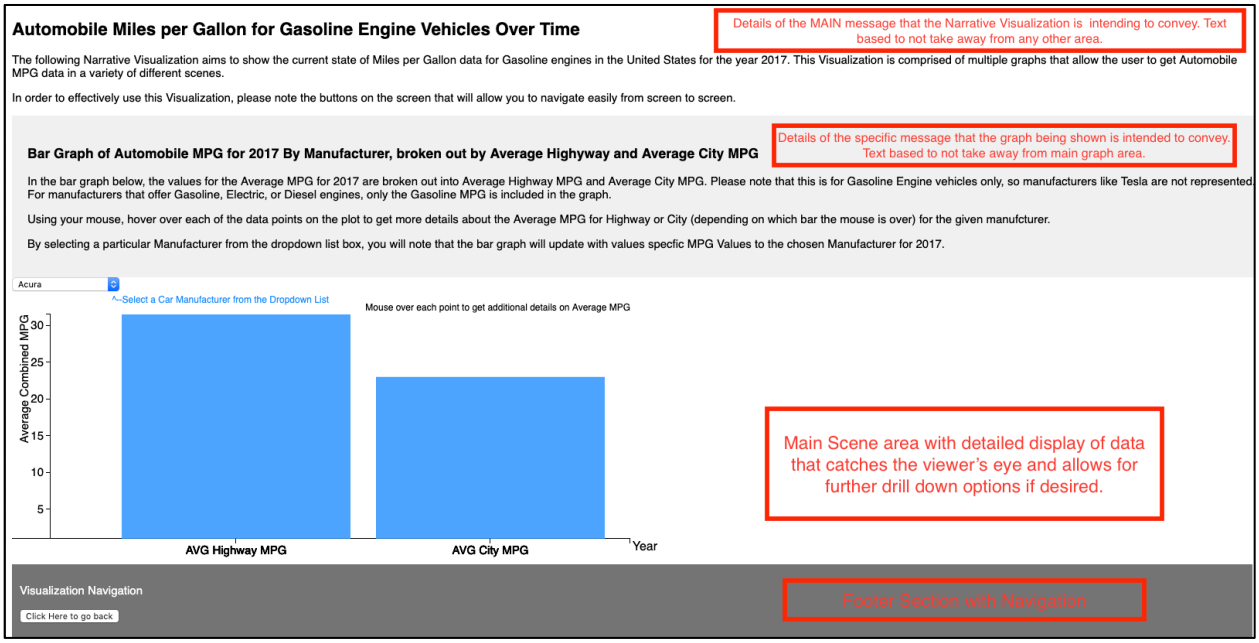
the center of the scene with just a splash of color to draw the viewer's attention directly there. The main graph is not only centered on the page, but it is surrounded by a template of text and muted color that is used throughout the narrative. This repeating template has a section at the top with global information about the narrative and the message that is being presented. By placing it at the top of each scene the viewer never forgets what the narrative visualization is about.

Just under the main message there is another recurring section that is also text based that offers information that is much more specific to the particular scene that is currently being viewed. For each scene there is a new graph with a different (or more detailed) set of data being displayed. This text section describes what is being shown on the current graph and further helps to enhance the message being shared. At the bottom of each scene there is a consistent footer that is also text based and is designed to be the standard place to look for navigation features that allow the viewer to either progress in the slide sequence, return up a level if they have drilled down, or return back to the starting point.

Please note that the visualization template presents valuable information that conveys a message, but it also provides highlights that direct the viewer on how to get more detailed information should the want to see it. In the example above the highlights for drilling down are indicated by the annotations that discuss clicking on plot points to get more refined details.

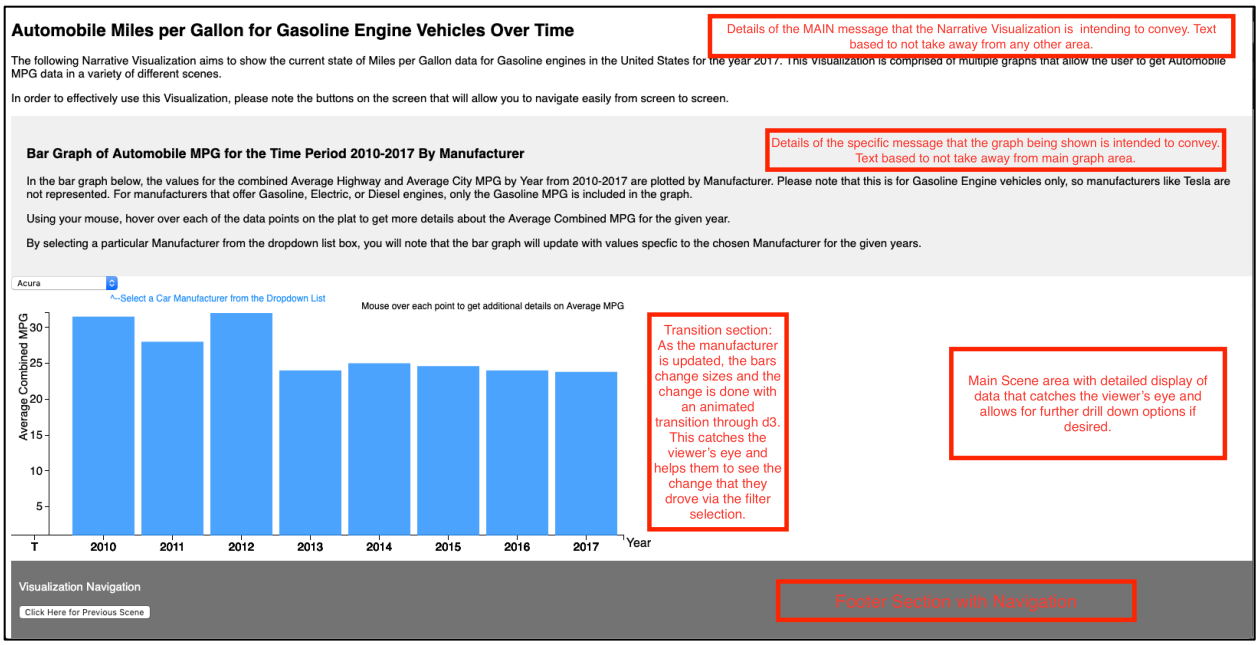
By starting out on the main scene with introducing a standard page template, the viewer quickly becomes comfortable with how to read the messaging. As the viewer progresses through the slide show and subsequent scenes, the use of consistent font type, style, color and size further adds to the understanding by giving the viewer some standards for what each section and font type mean. For example, black font text tells the viewer about tool-tip options while blue text described what happens when a particular visual item is clicked.

In the second scene we see that the visualization layout template remains the same, with the overall message shown at the top, the details of the current graph just below it, the main area displaying the scene's main message (the chart), and the footer for navigation.



The additional item to note in this scene is that from a transition perspective, only the graph area in the middle changed from scene 1 to scene 2. With everything else remaining consistent, the viewer knows what to expect, understands the scene immediately, and is able to focus completely on the message of the narrative visualization without losing focus trying to understand a new layout, new colors, or new fonts.

In the third scene we see that the template continues to remain the same. The information in the graph has changed, as well as the descriptive text associated with the graph, but the overall look and feel of the scene remains constant.

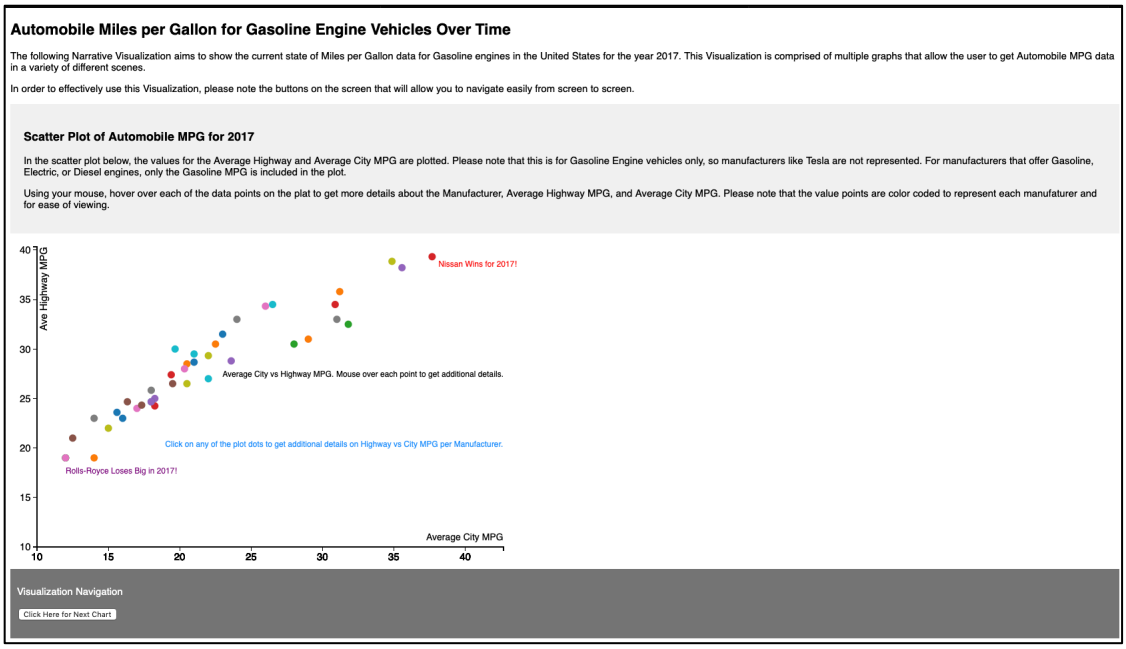


The main item of note for the third scene is the fact that a visual structure relating to transition is front and center for the viewer. As the viewer selects a new manufacturer from the drop-down list, the data is filtered for that vendor and the bar rectangles are automatically redrawn. During this transition point, the transition feature of D3 is used to animate the re-drawing of the bars slowly enough that the view can see them being redrawn dynamically. This transition effect grabs the viewer's attention and encourages more exploration to see what the animation will look like for other manufacturers.

Scenes

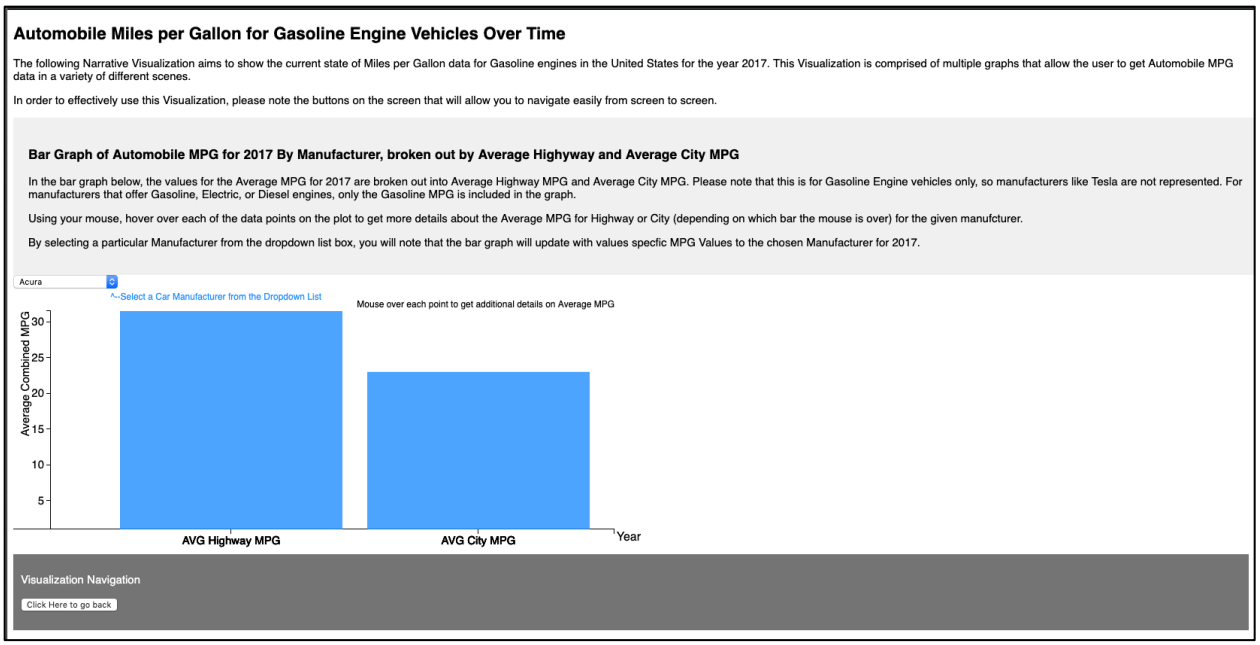
For my narrative visualization, I have created three scenes to convey my message about gasoline engine automobile efficiency. In the first scene, I present the grand overview of the data with a colorful scatter plot diagram which shows how all of the car manufacturers stack up against each other in terms of City and Highway MPG. I chose this as the first slide because I wanted the viewer to see the highest level of the detail at the start. With this opening scene, the viewer gets to know all of the players (manufacturers) at the outset, and how they are ranked against each other for the MPG measure. I also chose to use a variety of color on the graph to catch the viewer's eye immediately and draw them to the important data that the scatter plot represents. By starting at the highest level with this opening scene, I am educating the viewer about what this narrative is all about and I am giving them the background that they need in order to start formulating their own questions and levels of curiosity that can be addressed in the details of the scene and in the drill-down scenes that are to come.

The opening scene for my visualization can be seen below, and you can see how the high-level overview is presented with the option to drill down for more details described in the text annotations next to the data. The goal of this scene is to leave the viewer wanting more.



In the second scene, the viewer's interest is strong enough in the information that was initially provided that they want even more details to satisfy their curiosity. For my second scene, the viewer has decided to click on one of the individual scatter plot dots so that they can see the detailed values that are associated with that point.

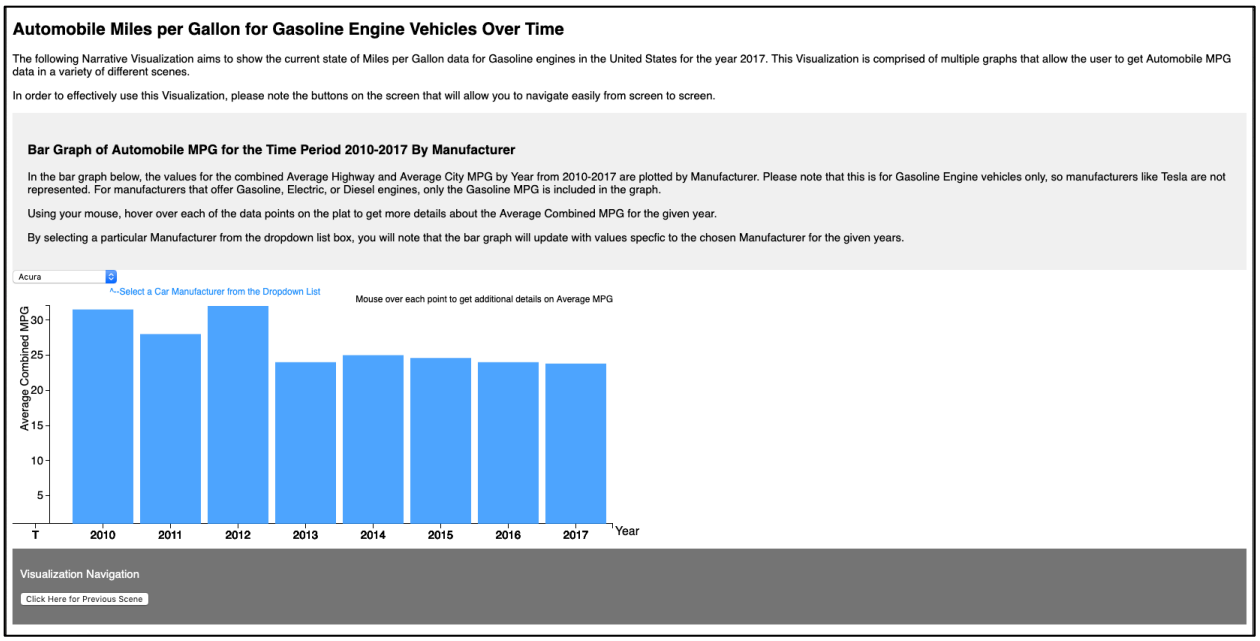
When this happens, the second scene opens and the viewer is presented with a bar graph that shows the Average City MPG and the Average Highway MPG values that determined a particular manufacturer's position in the first scene's scatter plot. This visual representation via a bar graph provides immediate information about these details in a visual format that is eye catching and very easy to understand. Since the viewer had enough interest to drill down into this scene from the first scene, the assumption is that they have enough interest and curiosity in this data that they would want to see more details across all of the manufacturers. In order to satisfy this desire for more information, the second scene provides the ability for the viewer to dynamically change the data being displayed in the bar graph by selecting a new manufacturer from a drop-down list at the top of the screen. By selecting a new manufacturer, the graph is dynamically updated based on the selection, and the data is changed with animation of both the Y axis and the bar graph itself. This is very eye catching to the viewer and is a powerful way to show how the data is being dynamically altered through selection.



Like the graph in scene one, there are a number of on-demand detailed tool-tips that are accessed when the viewer moves the mouse pointer over one of the bars in the graph. Additionally, there are several permanent annotations on the graph area that direct the viewer to utilize the mouse for more details and how to use the drop-down selection list to change the filter for manufacturer.

In the final scene, we want to leave the viewer satisfied with the overall level of the data that has been provided. We also want to make sure that the goal of our initial message has been reached. For my third scene, the viewer has at least reviewed the highest level of data from the first scene, and there is a good chance that they took the opportunity to delve into the details that were provided in the second scene. That being the case, the third scene is designed to provide comparative historical details so that the understanding that has been developed to this point can be viewed in the context of historical data. To accomplish this, my third scene builds on the first two by displaying detailed MPG averages over the last 8 years (from 2010-2017). The assumption is that the viewer has a good understanding about MPG data from 2017, but now wants to know how does that compare to recent year values of the same measurements?

To answer that question, In the third scene the viewer is presented with data in a format similar to scene two but with the added details of side by side yearly data.



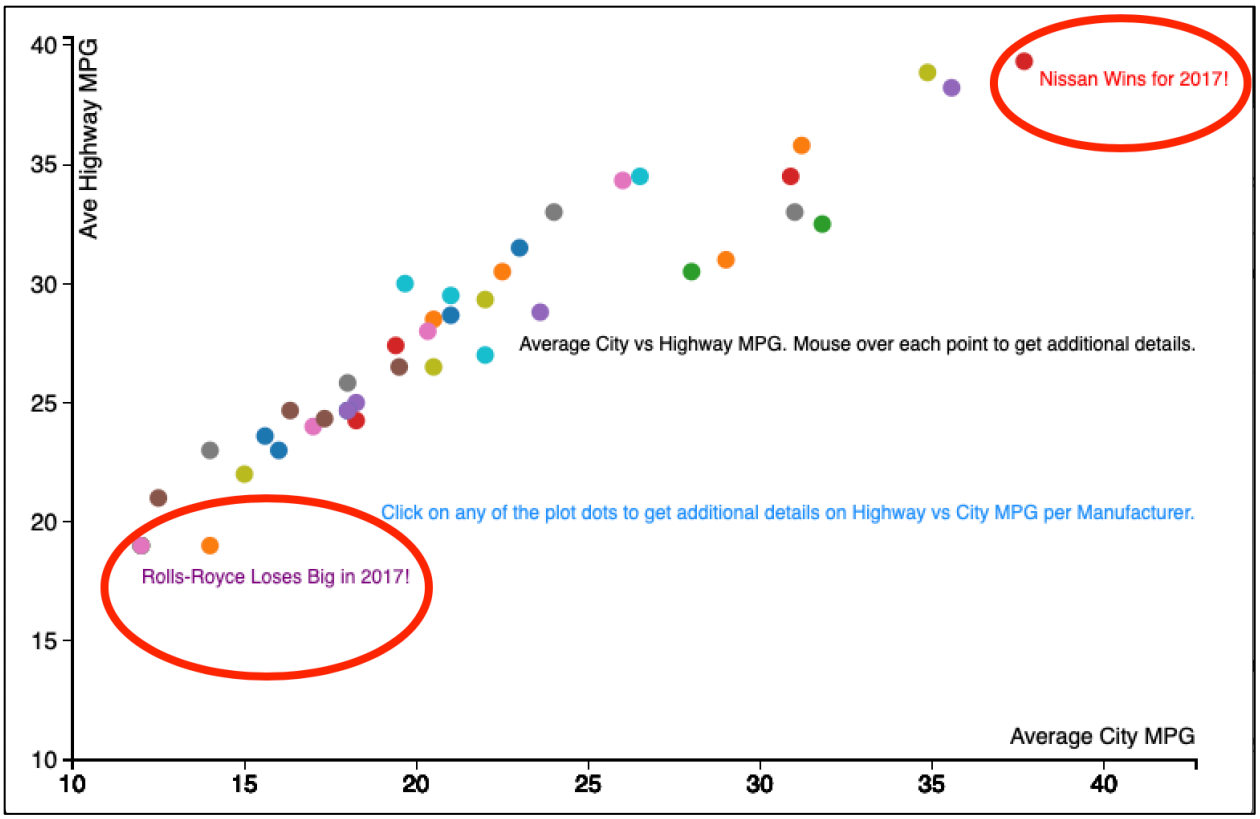
In this final scene, the viewer once again has the option of changing the manufacturer filter from a drop-down list at the top of the page. Like the scenes before it, the third scene provides permanent annotations to direct the viewer to use the mouse over features for more details in a tool-tip , as well as directions about how to select a new filter from the manufacturer drop-down.

Now that the viewer has reached this scene, the original message goal should be achieved and the viewer should now have a solid understanding of gasoline engine MPG averages, as well as a detailed understanding about how this has changed over time.

Annotations

Annotations are an important part of narrative visualization because they offer the viewer additional details about the scene and the level of detail that can be accessed within the implementation. These annotations support the overall messaging of the narrative and are very helpful in acting as instructions for the viewer about how to access the features of the scene in order to get the most out of the data. Additionally, the annotations can provide quick high points that you want the viewer to immediately see and understand so that the overall messaging of the narrative is quickly understood.

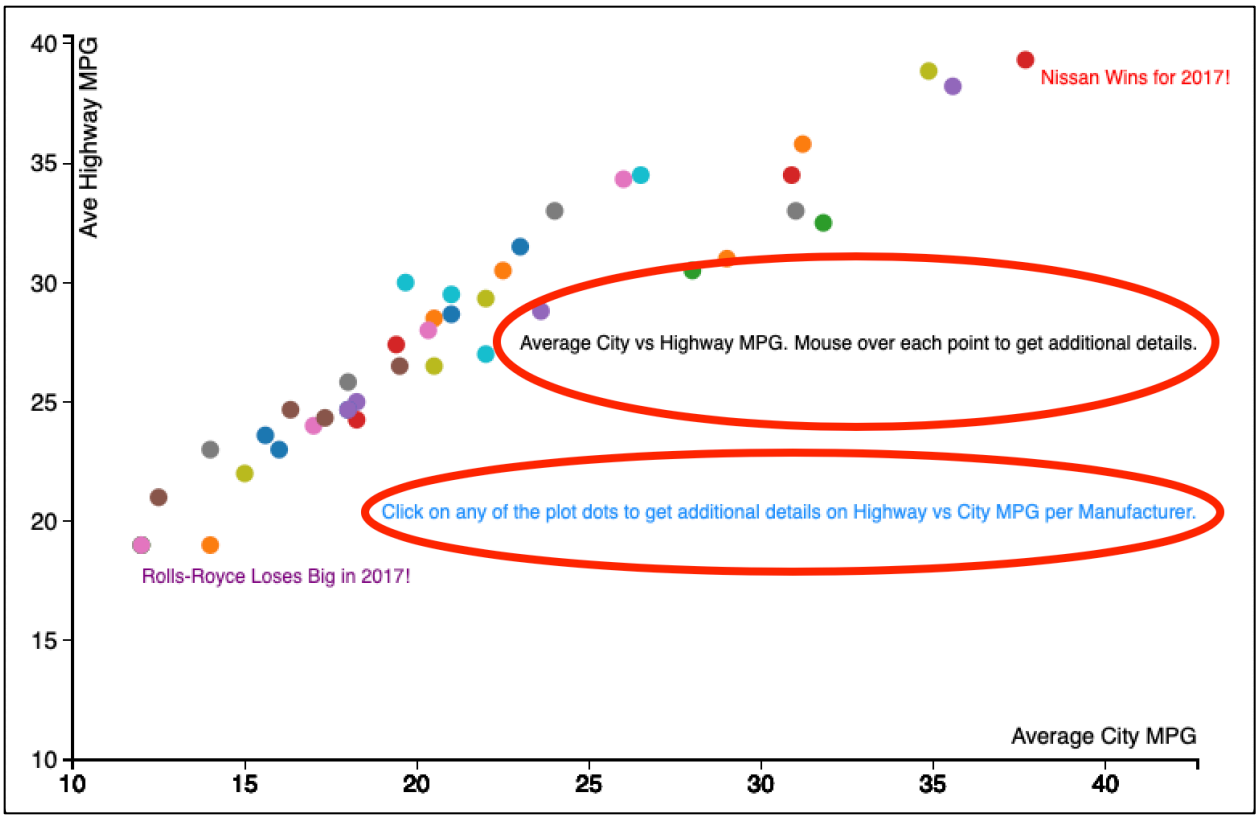
In my interactive slide-show narrative, I utilize annotations in a couple of key areas to assist the viewer in quickly seeing important pieces of information about automobile MPG numbers. In the first scene, there are a number of annotations present that point out to the viewer the low and high MPG manufacturers at a glance so that the viewer doesn't have to spend time highlighting each individual low and high point to understand the data.



As depicted in the screenshot above, the annotations are color coded to match the color associated with the manufacturer, and the annotations quickly draw the viewer's eye and provide immediate details of the data to enhance the overall message.

Annotations can also be used to provide information about how to use the visualization to get more information. These 'instructional' annotations can point out that more information can be had by using the mouse to roll over certain areas of the screen to get more data, or to simply click on a point to have more details appear in a drill-down scene.

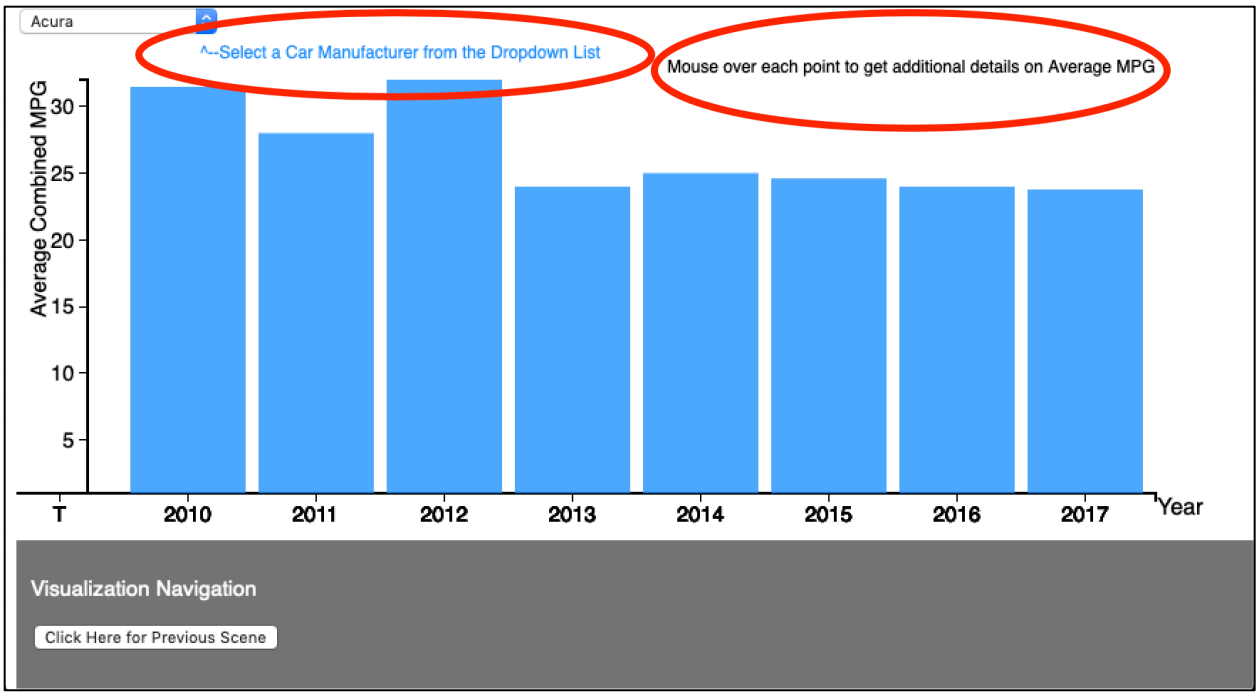
In my visualization, I utilized the 'instructional' annotation concept in a number of areas, including annotations for mouseover tips and where to click to get more detailed information from another scene



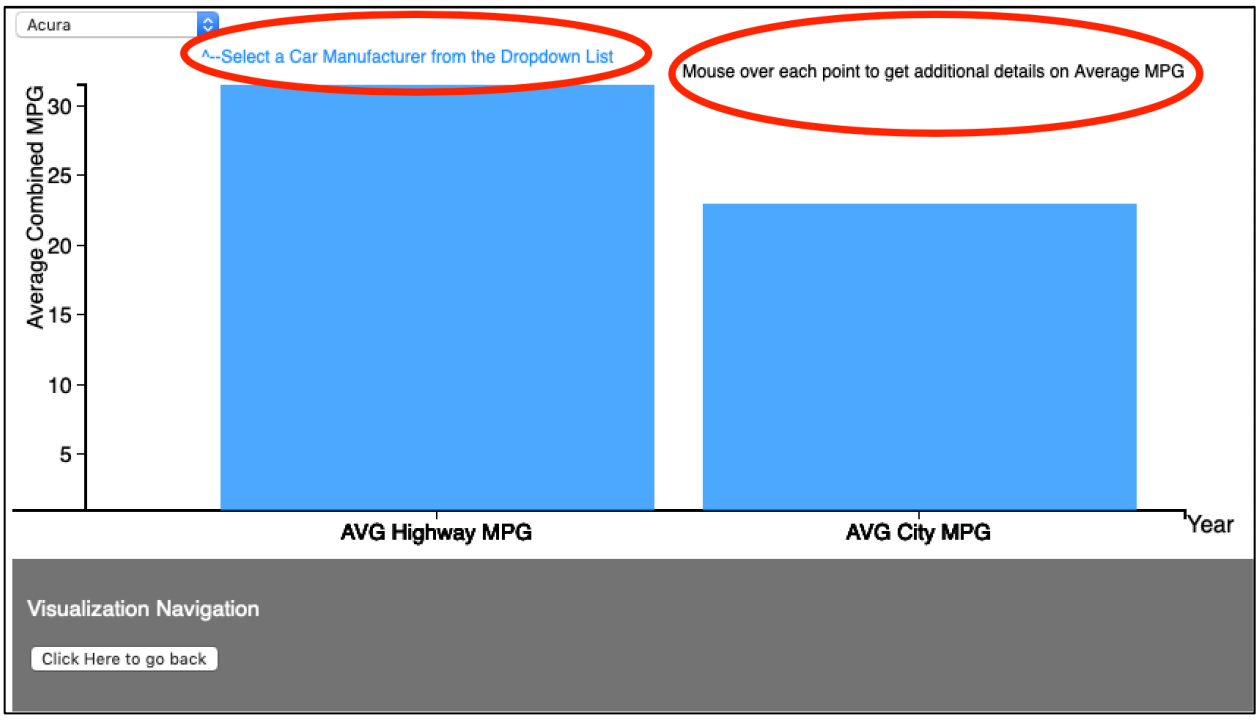
In order to maintain consistency within the visualization, it is important to make sure that there is a template when using annotations within the visualization. If the annotations look and act differently on each scree, the viewer will get confused about what the annotations are trying to convey, and the power of the annotation will be lost.

In my visualization, I was careful to use an annotation template that utilized the same font style and size throughout the scenes. Whenever I placed an annotation on a graph, the annotation was comprised of an informative text field utilizing color to show connections (i.e. the font color being the same as the manufacturer's plot point on the scene 1 graph). In the other scenes, annotations were used to direct the viewer about how to use the drop-down filter list for manufacturers, and how to get additional information via mouse overs on the bar graphs to get tool tips to appear.

By using the same text theme with the same font style and color throughout, the viewer quickly understands what the annotations are and how to use them to get a better understanding of the various scenes within the visualization.



Another example from scene 2 is here:



Parameters

The use of parameters in narrative visualization are what controls the various visual elements of the scenes throughout the slide-show. For example, the messaging for a scene may be to show important information based on a certain year or for a particular entity. As the parameter values change within the scene, either by viewer selection or through the design of the scene itself, the data view of the scene will change accordingly and will give the viewer a different perspective of the message that is being conveyed. The parameter values are also important because they are what drive the D3 and javascript code that are working behind the scenes to process and display the data in the visualization.

In my interactive slide show narrative, I use parameters for all of the scenes. In scene 1, gasoline engine MPG data is being displayed in a scatter plot diagram for each automobile manufacturer for the year 2017. In my original data set, there are records for multiple years and for a variety of engine types including gasoline, diesel, and electric. In this case, since I only wanted my scene message to focus on the year 2017 for gasoline engines, the parameters would be equal to '2017' and 'gasoline'. Now, while I certainly could have implemented the filtering of the data by these parameter values in the D3 and javascript code, for this scene I simply used those parameter values to filter the original data set outside of the code prior to processing it.

In scenes two and three, however, the parameters are changed to utilize a much more dynamic approach. In each one of these scenes, the viewer is able to see more detailed data regarding the breakout of Average City versus Average Highway MPG for 2017 for a particular manufacturer in one scene, and in the last scene to view historical MPG for the last 8 years by manufacturer. In both of these scenes, the primary parameter that directly affects what is being messaged in the scene is the manufacturer value. This parameter value is chosen by the viewer while drilling down into the data via a drop-down list of manufacturer values. As the manufacturer is chosen (and subsequently changes), the D3 and javascript code are alerted to this new value and the data in the bar graphs is dynamically updated for this new parameter value. Within the code, the dataset is filtered prior to being displayed on the screen according to the current selected manufacturer value.

Parameters are an important aspect of setting up the initial state of a scene, and then subsequently modifying that state as the viewer interacts with the scene. In my interactive slide show, the initial state of scene 1 (and the state of the scene throughout) is that the data has been filtered down by year and engine type prior to the viewer interacting with the scene. The state of the scene is purposely set with these initialization parameters so that we can provide a specific starting point for the messaging of the visualization. For the second scene, the initial state is set to show the viewer the breakdown of MPG by City and Highway starting with the manufacturer parameter being equal to the alphabetically first manufacturer on the list (in this case 'Acura'). The state of the scene changes as the viewer selects different manufacturers from the list. Selecting a different manufacturer updates the manufacturer

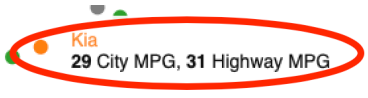
parameter, which causes the D3 and javascript code to redraw the values on the graph for the scene.

The same state action occurs in scene three, with the initial state being that combined MPG data is displayed for the last 8 years for the first manufacturer in the alphabetized list (again, 'Acura'). This initialization state changes once the viewer selects a different manufacturer from the list and the D3 and javascript code redraw the bar graphs based on the new parameter filter.

Triggers

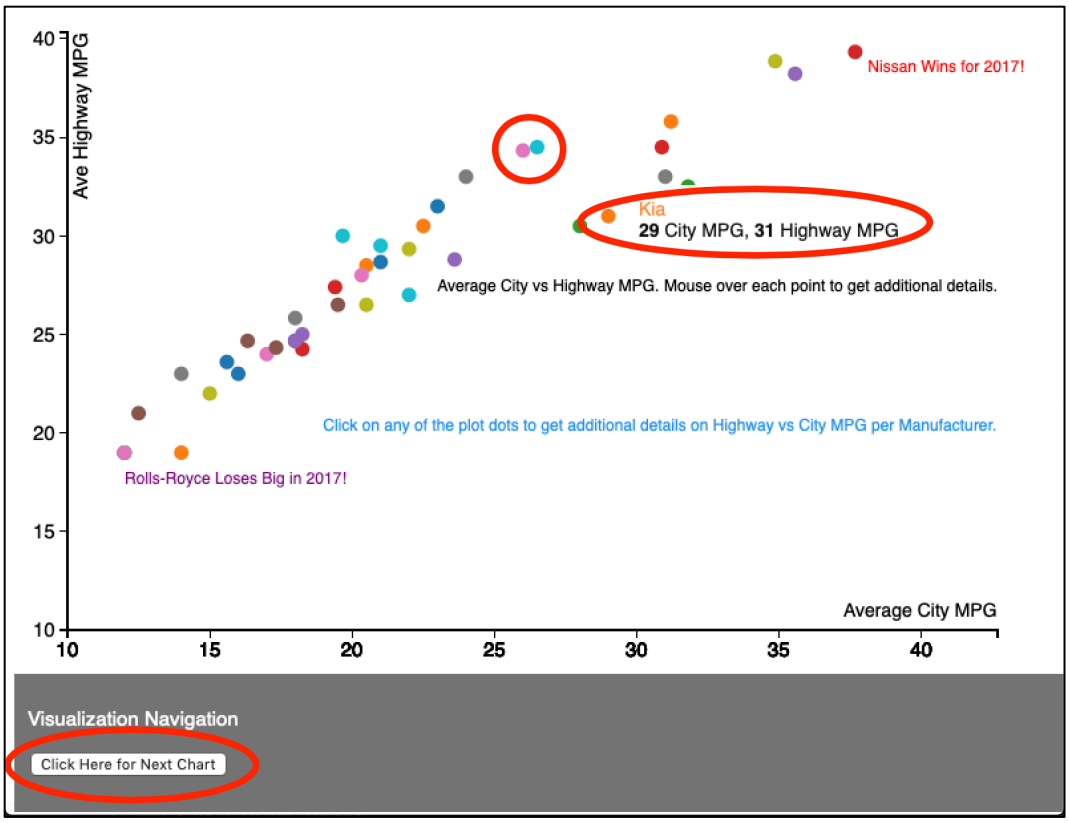
Working in conjunction with the parameter values, triggers are put in place within the code to perform specific tasks based upon viewer actions. For instance, when a viewer changes a value in a drop-down list or rolls their mouse cursor over a specific area on the screen, elements within the code are set off (often referred to as 'callback functions') that cause the narrative to change or the underlying data to be filtered in a certain way. Triggers can also be found within the visualization navigation elements like buttons that, when pressed, transition the viewer to new scenes. The power of the triggers is that it allows the narrative visualization to be dynamic and reactive to what the viewer is acting on within the scene. Triggers are what allow interactive slide shows to actually be interactive.

In my interactive slide show narrative visualization, I use triggers generously in each scene. In scene one, there are a number of triggers associated with the graph in the scene. The main triggers have to do with rolling the mouse over an element on the screen (and the associated rolling the mouse off the element), the clicking of a plot point to activate a drill down scene, and the clicking of a button to advance the slide show to the next scene. In the first case of detecting the mouse location, within the D3 and javascript code there are elements that detect the 'trigger' of the event(s) happening. When the viewer 'mouses over' a plot point, a trigger is fired that performs a callback to a function in the code associated with the mouse over action. In this case the callback function causes the area of the graph to now display an element called a tool-tip which displays more information about the plot point dynamically to the viewer, as in this example:



When the viewer rolls the mouse curser out of the area associated with the trigger, another trigger is fired that performs a callback to a function that removes the tool-tip from the display.

In another example of triggers within scene 1, when the view actually clicks the mouse button while on one of the plot points, a trigger is fired that performs a callback to a function that takes the viewer to the drill down aspect of the interactive visualization, which is scene 2.

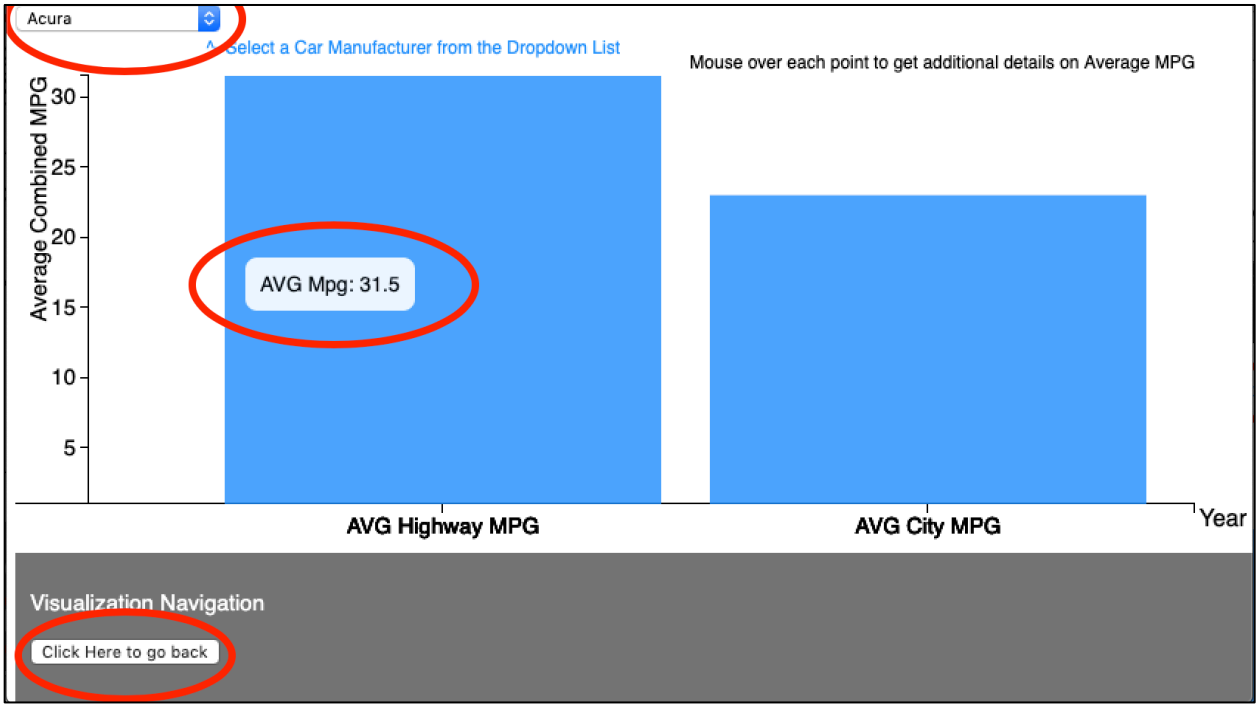


Another example of a trigger in scene 1 is a simple one associated with clicking a navigation button at the bottom of the screen. When the viewer completes this action, the interactive narrative loads the next slide in the presentation.

One of the important elements in narrative visualization is the concept of affordances when talking about triggers. Affordances are visual clues that help the viewer understand when there are actions to be performed on the screen and which elements that make up the scene are actionable. For example, an affordance can be an area on the screen that looks like a button. The affordance aspect is that is appears raised and is in a distinct shape that invites interaction. There may even be text on the button that invites it to be pushed. In my visualization for scene 1, the affordances come in a number of ways, including the use of a standard button shape for navigation. Where the affordance is not so obvious, I include text to notify the viewer that moving the mouse over items will cause things to happen and more drill down elements to be provided. For the click of the plot point affordance, I was unable to use a more traditional beveled looking visualization that would cause a viewer to assume that clicking it would cause an action, so I added additional text to display instructions for usage.

For scenes 2 and 3 of my interactive slide show, there are also a number of triggers that exist to enhance the viewing of the data and/or allow drill-down operations to occur. In scene 2, the viewer has the same capability as described above regarding mousing over areas on the screen and getting tool-tip popups with more data, but the biggest new trigger is associated with the drop-down list for manufacturer. When the viewer selects a different manufacturer from the

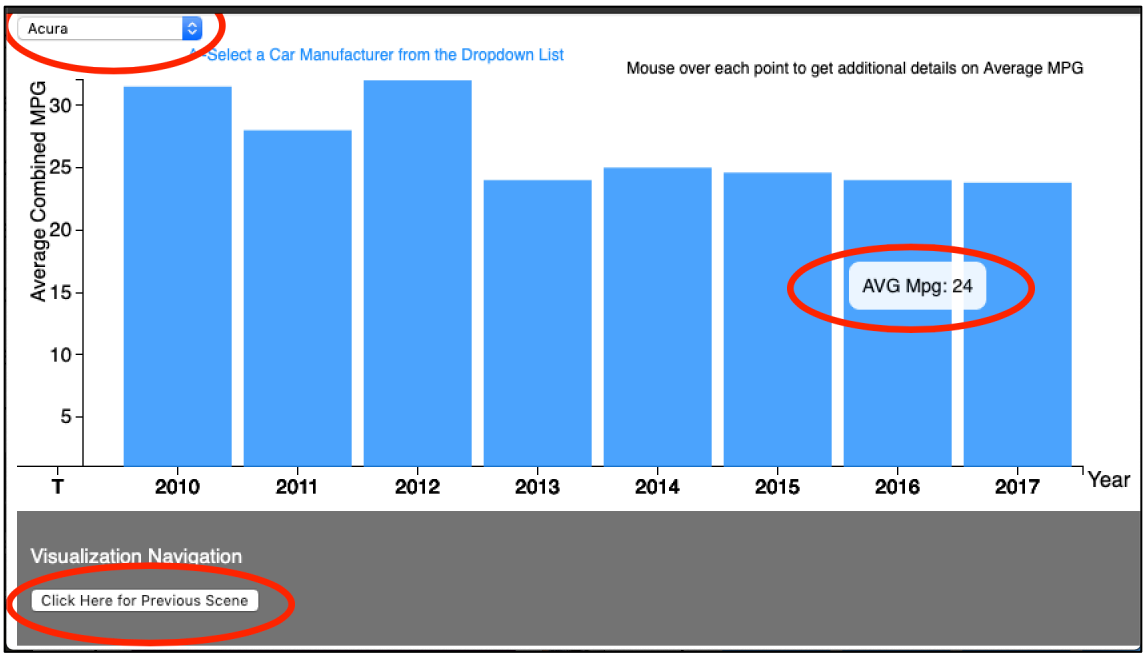
drop-down list, a trigger is fired that performs a callback to a function within the code that performs an update of the graph data. This update performs a filter of the data based on the selected manufacturer, and then redraws the bars to the heights associated with that manufacturer. The function also updates the Y axis scale values.



The final trigger worth mentioning in this scene is the simple one associated with the navigation button at the bottom of the screen that, when pressed, triggers a callback to a function which returns the viewer back to the main scene.

Please note that the affordances for this scene are similar to what was mentioned in scene 1 in that visualization items that look like standard buttons invite the viewer to click. Where it is not obvious, I included text assistance for the affordances to ensure that the viewer did not become confused and miss an opportunity to get more information out of the scene.

For scene 3, the triggers and affordances remain very similar to what was outlined for scene 2. The main difference is that the graph shows more bars associated with historical data for MPG instead of just the two bars for the breakout of City versus Highway MPG. The main trigger of note remains the callback function that occurs when a new manufacturer is selected from the drop-down list. Please note that the discussion points on affordance for this scene remain the same as was outlined for scene 2, and that text cues were added when simple visual affordances were not enough.



Summary and Conclusion

The purpose of this project assignment was to showcase what we have learned through the lectures and weekly homework assignments for this class. The fundamental aspect was to provide a detailed example of a Narrative Visualization utilizing one of the three main types of visual story telling ("martini glass", "interactive slide show", or "drill down story") and then building out the visualization detail elements to showcase our knowledge of narrative and visual structures, scenes, annotations, parameter, and triggers in detail using D3 and javascript code. The overarching goal was to develop a message and associated storyline that was to be presented to the viewer that would lead to knowledge transfer and greater understanding of the topic to be communicated.

To accomplish this task, I selected the topic of fuel efficiency for gasoline powered engines measured in Highway and City Miles Per Gallon. Throughout my interactive slide show I presented the viewer with consecutively more detailed views of the data to further enhance and emphasize my overall message. By taking advantage of all the elements of the visualization narrative, I was able to direct the viewer through the message concepts, while still allowing for flexibility in how the data was viewed according to their desires of when drill down information was important to them. By adding elements like parameters, annotations, and triggers, my overall interactive slide show provided a well-rounded viewing experience while effectively transferring the pertinent information.

Appendix

Data Files Used as Input

In order to make the scenes function dynamically in the D3 and javascript environment, the following data files were use in the creation of the graphs:

- Scene 1—Scatter Plot Diagram for Combined MPG for Each Manufacturer for 2017

File: https://raw.githubusercontent.com/mbennett-WA/mbennett-WA.github.io/master/AvesByCars.csv

- Scene 2—Bar Chart with Average Highway and Average City MPG Filtered by Manufacturer

File: https://raw.githubusercontent.com/mbennett-WA/mbennett-WA.github.io/master/AVG_Each_Make_GB_Fuel.csv

- Scene 3—Bar Chart with Combined MPG Values for the last 8 Years with Filter for Manufacturer

File: https://raw.githubusercontent.com/mbennett-WA/mbennett-WA.github.io/master/AveEachMakeYear.csv