
Drawing with Data

It's time to start drawing with data.

Let's continue working with our simple data set for now:

```
var dataset = [ 5, 10, 15, 20, 25 ];
```

Drawing divs

We'll use this to generate a super-simple bar chart. Bar charts are essentially just rectangles, and an HTML `<div>` is the easiest way to draw a rectangle. (Then again, to a web browser, *everything* is a rectangle, so you could easily adapt this example to use ``span``s or whatever element you prefer.)

Formally, a chart with vertically oriented rectangles is a *column* chart, and one with horizontal rectangles is a *bar* chart. In practice, most people just call them all bar charts, as I'll do from now on.

This `div` could work well as a data bar:

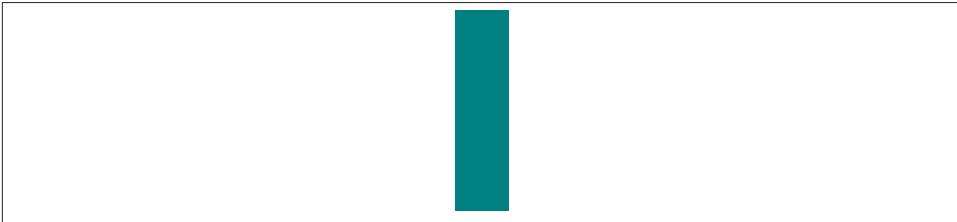


Figure 6-1. A humble div

```
<div style="display: inline-block;
          width: 20px;
          height: 75px;
          background-color: teal;"></div>
```

(Among web standards folks, this is a semantic no-no. Normally, one shouldn't use an empty `div` for purely visual effect, but I am making an exception for the sake of this example.)

Because this is a `div`, its `width` and `height` are set with CSS styles. Except for `height`, each bar in our chart will share the same display properties, so I'll put those shared styles into a class called `bar`:

```
div.bar {
  display: inline-block;
  width: 20px;
  height: 75px; /* We'll override height later */
  background-color: teal;
}
```

Now each `div` needs to be assigned the `bar` class, so our new CSS rule will apply. If you were writing the HTML code by hand, you would write:

```
<div class="bar"></div>
```

Using D3, to add a class to an element, we use the `selection.attr()` method. It's important to understand the difference between `attr()` and its close cousin, `style()`. `attr()` sets DOM attribute values, while `style()` applies CSS styles directly to an element.

Setting Attributes

`attr()` is used to set an HTML attribute and its value on an element. An HTML attribute is any property/value pair that you could include between an element's `<>` brackets. For example, these HTML elements

```
<p class="caption">
<select id="country">

```

contain a total of five attributes (and corresponding values), all of which could be set with `attr()`:

Attribute	Value
class	caption
id	country
src	logo.png
width	100px
alt	Logo

To assign a class of `bar`, we can use:

```
.attr("class", "bar")
```

A Note on Classes

Note that an element's *class* is stored as an HTML attribute. The class, in turn, is used to reference a CSS style rule. This may cause some confusion because there is a difference between setting a *class* (from which styles are inferred) and applying a *style* directly to an element. You can do both with D3. Although you should use whatever approach makes the most sense to you, I recommend using *classes* for properties that are shared by multiple elements, and applying *style* rules directly only when deviating from the norm. (In fact, that's what we'll do in just a moment.)

I also want to briefly mention another D3 method, `classed()`, which can be used to quickly apply or remove classes from elements. The line of code above could be rewritten as:

```
.classed("bar", true)
```

This line simply takes whatever selection is passed to it and applies the class `bar`. If `false` were specified, it would do the opposite, removing the class of `bar` from any elements in the selection:

```
.classed("bar", false)
```

Back to the Bars

Putting it all together with our data set, here is the complete D3 code so far:

```
var dataset = [ 5, 10, 15, 20, 25 ];  
  
d3.select("body").selectAll("div")  
  .data(dataset)  
  .enter()  
  .append("div")  
  .attr("class", "bar");
```



Figure 6-2. Five *div*s masquerading as one

To see what's going on, look at `01_drawing_divs.html` in your browser, view the source, and open your web inspector. You should see five vertical `div` bars, one generated for each point in our data set. However, with no space between them, they look like one big rectangle.

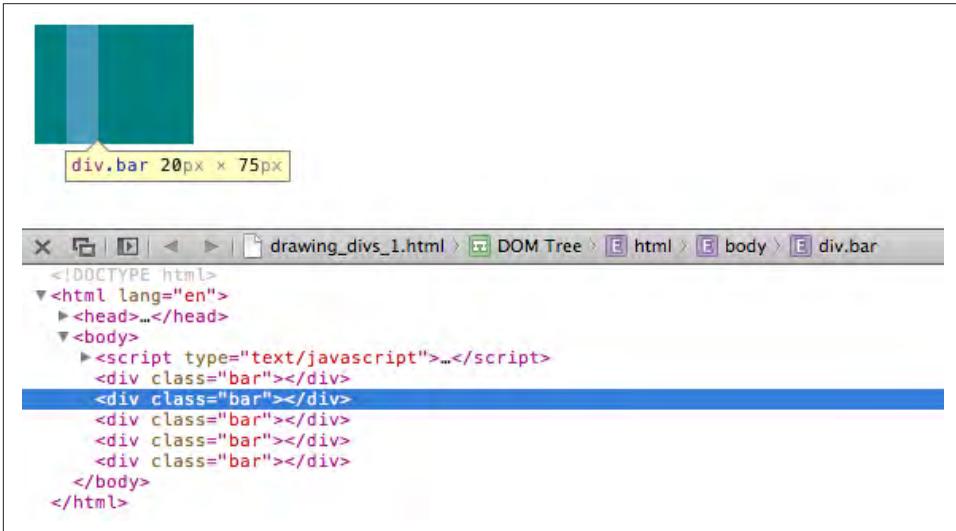


Figure 6-3. Five `div`s masquerading as one, as seen through the web inspector

Setting Styles

The `style()` method is used to apply a CSS property and value directly to an HTML element. This is the equivalent of including CSS rules within a `style` attribute right in your HTML, as in:

```
<div style="height: 75px;"></div>
```

To make a bar chart, the height of each bar must be a function of its corresponding data value. So let's add this to the end of our D3 code:

```
.style("height", function(d) {  
    return d + "px";  
});
```



Figure 6-4. A small bar chart!

See that code in `02_drawing_divs_height.html`. You should see a very small bar chart!

When D3 loops through each data point, the value of `d` will be set to that of the corresponding value. So we are setting a `height` value of `d` (the current data value) while appending the text `px` (to specify the units are pixels). The resulting heights are `5px`, `10px`, `15px`, `20px`, and `25px`.

This looks a little bit silly, so let's make those bars taller

```
.style("height", function(d) {  
    var barHeight = d * 5; //Scale up by factor of 5  
    return barHeight + "px";  
});
```

and add some space to the right of each bar, to space things out:

```
margin-right: 2px;
```

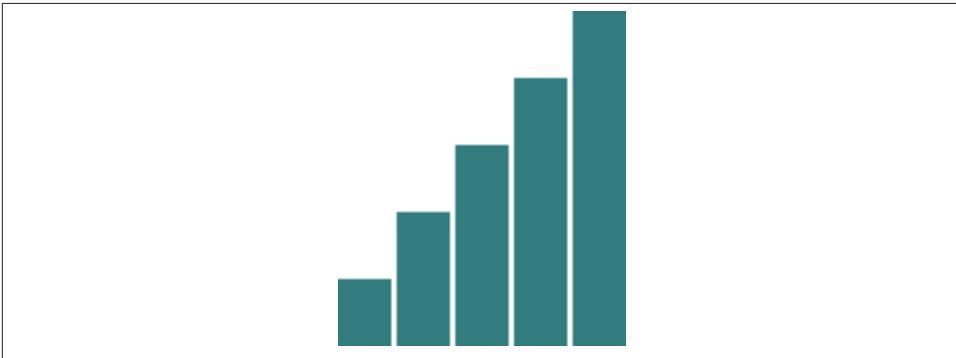


Figure 6-5. A taller bar chart

Nice! We could go to SIGGRAPH with that chart.

Try out the sample code `03_drawing_divs_spaced.html`. Again, view the source and use the web inspector to contrast the original HTML against the final DOM.

The Power of data()

This is exciting, but real-world data is never this clean:

```
var dataset = [ 5, 10, 15, 20, 25 ];
```

Let's make our data a bit messier, as in `04_power_of_data.html`:

```
var dataset = [ 25, 7, 5, 26, 11 ];
```

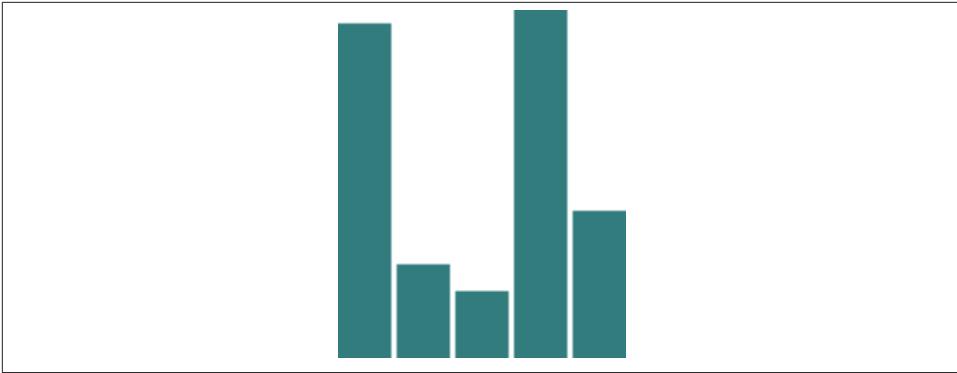


Figure 6-6. New data values

We're not limited to five data points, of course. Let's add many more! (See [05_power_of_data_more_points.html](#).)

```
var dataset = [ 25, 7, 5, 26, 11, 8, 25, 14, 23, 19,
                14, 11, 22, 29, 11, 13, 12, 17, 18, 10,
                24, 18, 25, 9, 3 ];
```

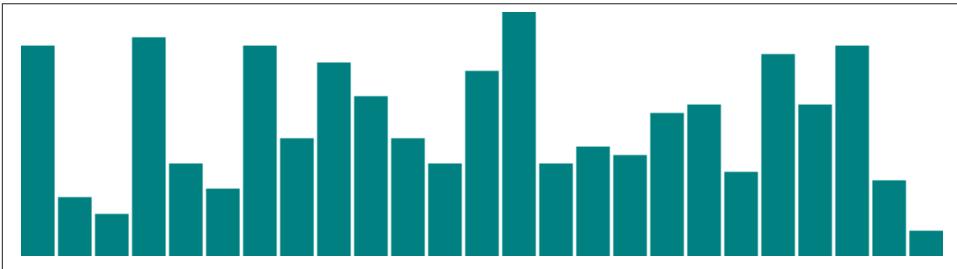


Figure 6-7. Lots more data values

25 data points instead of five! How does D3 automatically expand our chart as needed?

```
d3.select("body").selectAll("div")
  .data(dataset) // <-- The answer is here!
  .enter()
  .append("div")
  .attr("class", "bar")
  .style("height", function(d) {
    var barHeight = d * 5;
    return barHeight + "px";
  });
```

Give `data()` ten values, and it will loop through ten times. Give it one million values, and it will loop through one million times. (Just be patient.)

That is the power of `data()` — being smart enough to loop through the full length of whatever data set you throw at it, executing each method beneath it in the chain, while updating the context in which each method operates, so `d` always refers to the current datum at that point in the loop.

That may be a mouthful, and if it all doesn't make sense yet, it will soon. I encourage you to make a copy of `05_power_of_data_more_points.html`, tweak the `dataset` values, and note how the bar chart changes.

Remember, the *data* is driving the visualization — not the other way around.

Random Data

Sometimes it's fun to generate random data values, whether for testing purposes or just pure geekiness. That's just what I've done in `06_power_of_data_random.html`. Notice that each time you reload the page, the bars render differently.

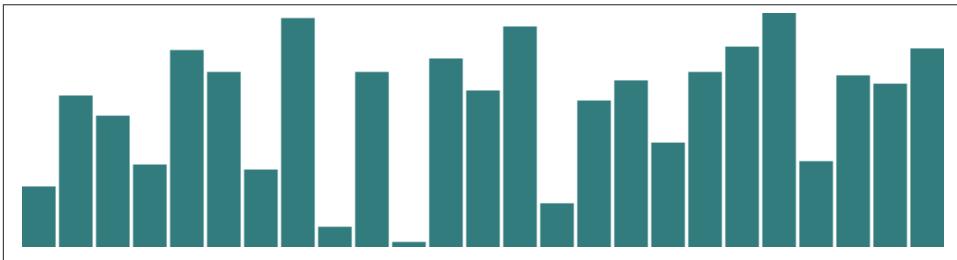
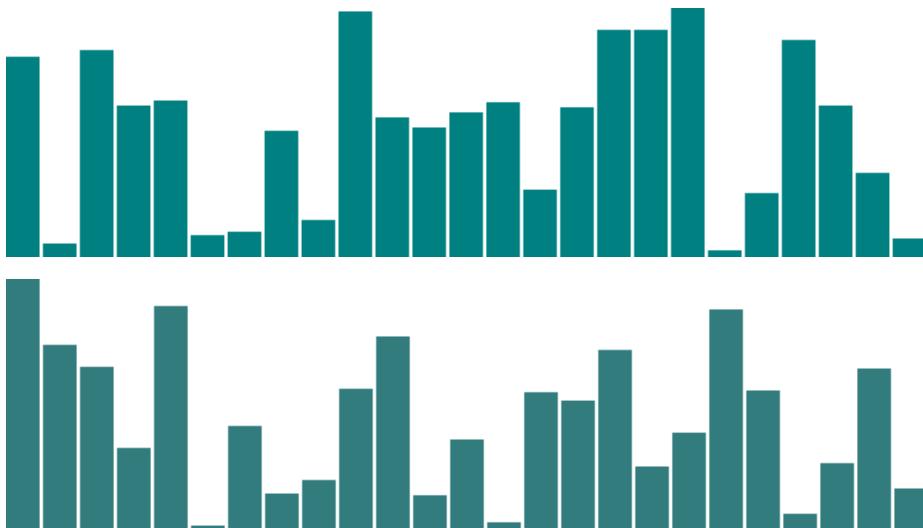


Figure 6-8. Bar charts with random values



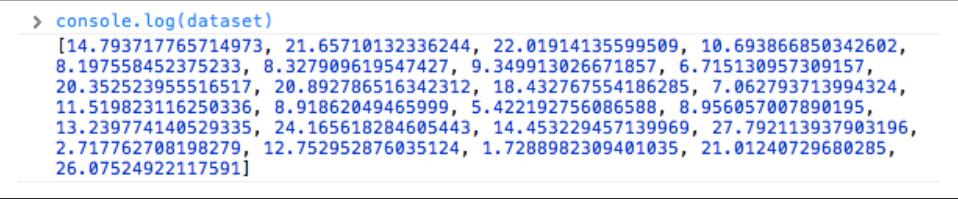
View the source, and you'll see this code:

```
var dataset = []; //Initialize empty array
for (var i = 0; i < 25; i++) { //Loop 25 times
    var newNumber = Math.random() * 30; //New random number (0-30)
    dataset.push(newNumber); //Add new number to array
}
```

That code doesn't use any D3 methods; it's just JavaScript. Without going into too much detail, the code above:

1. Creates an empty array called `dataset`.
2. Initiates a `for` loop, which is executed 25 times.
3. Each time, it generates a new random number with a value between zero and 30.
4. That new number is appended to the `dataset` array. (`push()` is an array method that appends a new value to the end of an array.)

Just for kicks, open up the JavaScript console and enter `console.log(dataset)`. You should see the full array of 25 randomized data values.



```
> console.log(dataset)
[14.793717765714973, 21.65710132336244, 22.01914135599509, 10.693866850342602,
8.197558452375233, 8.327909619547427, 9.349913026671857, 6.715130957309157,
20.352523955516517, 20.892786516342312, 18.432767554186285, 7.062793713994324,
11.519823116250336, 8.91862049465999, 5.422192756086588, 8.956057007890195,
13.239774140529335, 24.165618284605443, 14.453229457139969, 27.792113937903196,
2.717762708198279, 12.752952876035124, 1.7288982309401035, 21.01240729680285,
26.07524922117591]
```

Figure 6-9. Random values in console

Notice that they are all decimal or floating point values (14.793717765714973), not whole numbers or integers (14) like we used initially. For this example, decimal values are fine, but if you ever need whole numbers, you can use JavaScript's `Math.round()` method. For example, you could wrap the random number generator from this line

```
var newNumber = Math.random() * 30;
```

as follows:

```
var newNumber = Math.round(Math.random() * 30);
```

Try it out in `07_power_of_data_rounded.html`, and use the console to verify that the numbers have indeed been rounded to integers:

```
> console.log(dataset)
[23, 19, 18, 16, 24, 29, 1, 5, 13, 4, 29, 23, 11, 9, 16, 10, 15, 4, 28, 23, 13,
19, 20, 20, 27]
```

Figure 6-10. Random integer values in console

That’s about all we can do visually with `div`’s. Let’s expand our visual possibilities with SVG.

Drawing SVGs

For a quick refresher on SVG syntax, see the “SVG” portion of the “Technology Fundamentals” chapter.

One thing you may notice about SVG elements is that all of their properties are specified as *attributes*. That is, they are included as property/value pairs within each element tag, like this:

```
<element property="value"></element>
```

Hmm, that looks strangely like HTML!

```
<p class="eureka">Eureka!</p>
```

We have already used D3’s handy `append()` and `attr()` methods to create new HTML elements and set their attributes. Since SVG elements exist in the DOM, just as HTML elements do, we can use `append()` and `attr()` in exactly the same way to generate SVG images!

Create the SVG

First, we need to create the SVG element in which to place all our shapes.

```
d3.select("body").append(".png");
```

That will find the document’s `body` and append a new `svg` element just before the closing `</body>` tag. While that code will work, may I suggest a slight modification?

```
var svg = d3.select("body").append(".png");
```

Remember how most D3 methods return a reference to the DOM element on which they act? By creating a new variable `svg`, we are able to capture the reference handed back by `append()`. Think of `svg` not as a “variable” but as a “reference pointing to the SVG object that we just created.” This reference will save us a lot of code later. Instead of having to search for that SVG each time — as in `d3.select(".png")` — we just say `svg`.

```
svg.attr("width", 500)
  .attr("height", 50);
```

Alternatively, that could all be written as one line of code:

```
var svg = d3.select("body")
    .append(".png")
    .attr("width", 500)
    .attr("height", 50);
```

See `08_drawing_svgs.html` for that code. Inspect the DOM and notice that there is, indeed, an empty SVG element.

To simplify your life, I recommend putting the width and height values into variables at the top of your code, as in `09_drawing_svgs_size.html`. View the source, and you'll see:

```
//Width and height
var w = 500;
var h = 50;
```

I'll be doing that with all future examples. By *variabilizing* the size values, they can be easily referenced throughout your code, as in:

```
var svg = d3.select("body")
    .append(".png")
    .attr("width", w) // <-- Here
    .attr("height", h); // <-- and here!
```

Also, if you send me a petition to make “variabilize” a real word, I will gladly sign it.

Data-driven Shapes

Time to add some shapes. I'll bring back our trusty old data set

```
var dataset = [ 5, 10, 15, 20, 25 ];
```

and then use `data()` to iterate through each data point, creating a `circle` for each one:

```
svg.selectAll("circle")
    .data(dataset)
    .enter()
    .append("circle");
```

Remember, `selectAll()` will return empty references to all `circle`s` (which don't exist yet), `data()` binds our data to the elements we're about to create, `enter()` returns a placeholder reference to the new element, and `append()` finally adds a `circle` to the DOM. In this case, it appends those `circle`s` to the end of the SVG element, since our initial selection is our reference `svg` (as opposed to the document body, for example).

To make it easy to reference all of the `circle`s` later, we can create a new variable to store references to them all:

```
var circles = svg.selectAll("circle")
    .data(dataset)
    .enter()
    .append("circle");
```

Great, but all these circles still need positions and sizes. Be warned: The following code may blow your mind.

```
circles.attr("cx", function(d, i) {
    return (i * 50) + 25;
})
.attr("cy", h/2)
.attr("r", function(d) {
    return d;
});
```

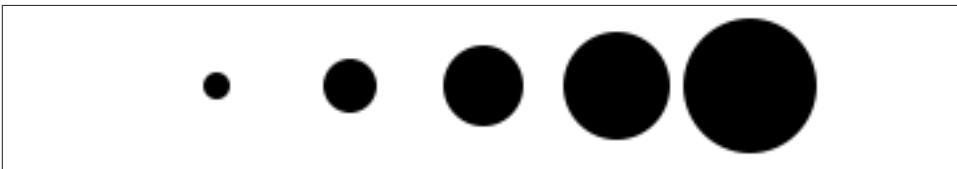


Figure 6-11. Row of data circles

Feast your eyes on the demo, `10_drawing_svgs_circles.html`. Let's step through the code, one line at a time.

```
circles.attr("cx", function(d, i) {
    return (i * 50) + 25;
})
```

This takes the reference to all `circle`s` and sets the ``cx` attribute for each one. (Remember that, in SVG lingo, `cx` is the x position value of the *center* of the circle.) Our data has already been bound to the `circle` elements, so for each `circle`, the value `d` matches the corresponding value in our original data set (5, 10, 15, 20, or 25).

Another value, `i`, is also automatically populated for us. (Thanks, D3!) `i` is a numeric index value of the current element. Counting starts at zero, so for our "first" circle `i == 0`, the second circle's `i == 1` and so on. We're using `i` to push each subsequent circle over to the right, because each subsequent loop through, the value of `i` increases by one.

```
(0 * 50) + 25 //Returns 25
(1 * 50) + 25 //Returns 75
(2 * 50) + 25 //Returns 125
(3 * 50) + 25 //Returns 175
(4 * 50) + 25 //Returns 225
```

To make sure `i` is available to your custom function, you must include it as an argument in the function definition, `function(d, i)`. You must also include `d`, even if you don't use `d` within your function (as in the case above).

On to the next line.

```
.attr("cy", h/2)
```

`cy` is the `y` position value of the center of each circle. We're setting `cy` to `h` divided by two, or one-half of `h`. You'll recall that `h` stores the height of the entire SVG, so `h/2` has the effect of aligning all `circle`s in the vertical center of the image.

```
.attr("r", function(d) {  
    return d;  
});
```

Finally, the radius `r` of each circle is simply set to `d`, the corresponding data value.

Pretty Colors, Oooh!

Color fills and strokes are just other attributes that you can set using the same methods. Simply by appending this code

```
.attr("fill", "yellow")  
.attr("stroke", "orange")  
.attr("stroke-width", function(d) {  
    return d/2;  
});
```

we get the following colorful circles, as seen in `11_drawing_svgs_color.html`:

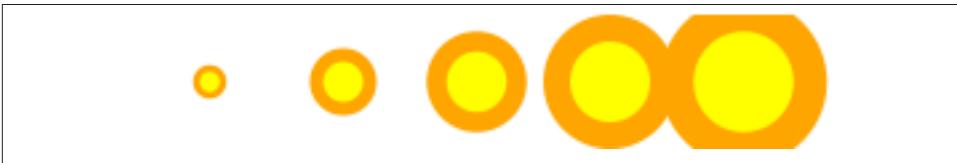


Figure 6-12. Colorful data circles

Of course, you can mix and match attributes and custom functions to apply any combination of properties. The trick with data visualization, of course, is choosing appropriate *mappings*, so the visual expression of your data is understandable and useful for the viewer.

Making a Bar Chart

Now we'll integrate everything we've learned so far to generate a simple bar chart as an SVG image.

We'll start by adapting the `div` bar chart code to draw its bars with SVG instead, giving us more flexibility over the visual presentation. Then we'll add labels, so we can see the data values clearly.

The Old Chart

See the `div` chart, updated with some new data, in `12_making_a_bar_chart_divs.html`:

```
var dataset = [ 5, 10, 13, 19, 21, 25, 22, 18, 15, 13,
               11, 12, 15, 20, 18, 17, 16, 18, 23, 25 ];

d3.select("body").selectAll("div")
  .data(dataset)
  .enter()
  .append("div")
  .attr("class", "bar")
  .style("height", function(d) {
    var barHeight = d * 5;
    return barHeight + "px";
  });
```

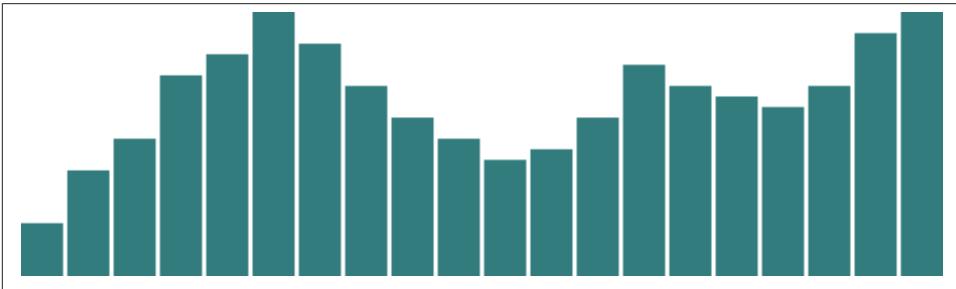


Figure 6-13. Bar chart with `div`s

It may be hard to imagine, but we can definitely improve on this simple bar chart made of `div`'s.

The New Chart

First things first, we need to decide on the size of the new SVG:

```
//Width and height
var w = 500;
var h = 100;
```

(Of course, you could name `w` and `h` something else, like `svgWidth` and `svgHeight`. Use whatever is most clear to you. JavaScript programmers, as a group, are fixated on efficiency, so you'll often see single-character variable names, code written with no spaces, and other hard-to-read, yet programmatically efficient, syntax.)

Then, we tell D3 to create an empty SVG element and add it to the DOM:

```
//Create SVG element
var svg = d3.select("body")
    .append(".png")
    .attr("width", w)
    .attr("height", h);
```

To recap, this inserts a new `<svg>` element just before the closing `</body>` tag, and assigns the SVG a width and height of 500 by 100 pixels. This statement also puts the result into our new variable called `svg`, so we can easily reference the new SVG without having to reselect it later using something like `d3.select(".png")`.

Next, instead of creating `div`s`, we generate ``rect`s` and add them to ``svg`.

```
svg.selectAll("rect")
    .data(dataset)
    .enter()
    .append("rect")
    .attr("x", 0)
    .attr("y", 0)
    .attr("width", 20)
    .attr("height", 100);
```

This code selects all `rect`s` inside of ``svg`. Of course, there aren't any yet, so an empty selection is returned. (Weird, yes, but stay with me. With D3, you always have to first select whatever it is you're about to act on, even if that selection is momentarily empty.)

Then, `data(dataset)` sees that we have 20 values in the data set, so it calls `enter()` 20 times. `enter()`, in turn, returns a placeholder selection for each data point that does not yet have a corresponding `rect` — which is to say, all of them.

For each of the 20 placeholders, `append("rect")` inserts a `rect` into the DOM. As we learned in the “Technology Fundamentals” chapter, every `rect` must have `x`, `y`, `width`, and `height` values. We use `attr()` to add those attributes onto each newly created `rect`.

Beautiful, no?

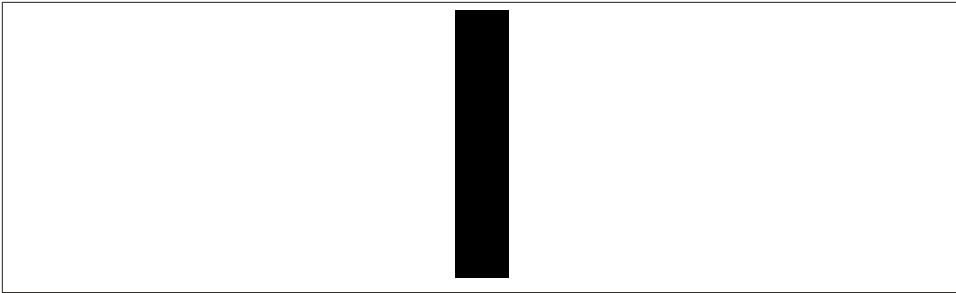


Figure 6-14. One lonely bar

Okay, maybe not. All of the bars are there (check the DOM of `13_making_a_bar_chart_rects.html` with your web inspector), but they all share the same `x`, `y`, `width`, and `height` values, with the result that they all overlap. This isn't a visualization of data yet.

Let's fix the overlap issue first. Instead of an `x` of zero, we'll assign a dynamic value that corresponds to `i`, or each value's position in the data set. So the first bar will be at zero, but subsequent bars will be at 21, then 42, and so on.

```
.attr("x", function(d, i) {  
    return i * 21; //Bar width of 20 plus 1 for padding  
})
```

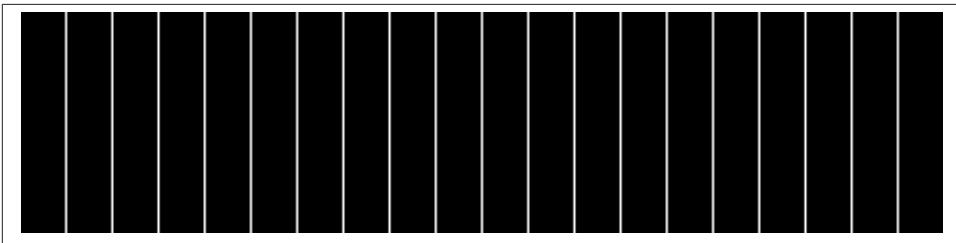


Figure 6-15. Twenty bars

See that code in action with `14_making_a_bar_chart_offset.html`.

That works, but it's not particularly flexible. If our data set were longer, then the bars would just run off to the right, past the end of the SVG! Since each bar is 20 pixels wide, plus 1 pixel of padding, then a 500-pixel wide SVG can only accommodate 23 data points. Note how the 24th bar here gets clipped:

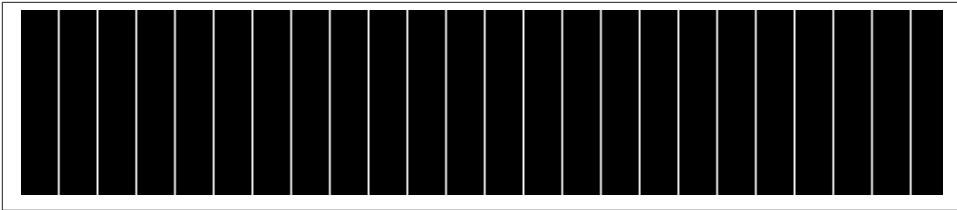


Figure 6-16. Twenty-four bars

It's good practice to use flexible, dynamic coordinates — heights, widths, x values, and y values — so your visualization can scale appropriately along with your data.

As with anything else in programming, there are a thousand ways to achieve that end. I'll use a simple one. First, I'll amend the line where we set each bar's x position:

```
.attr("x", function(d, i) {  
    return i * (w / dataset.length);  
})
```

Notice how the x value is now tied directly to the width of the SVG (`w`) and the number of values in the data set (`dataset.length`). This is exciting, because now our bars will be evenly spaced, whether we have 20 data values

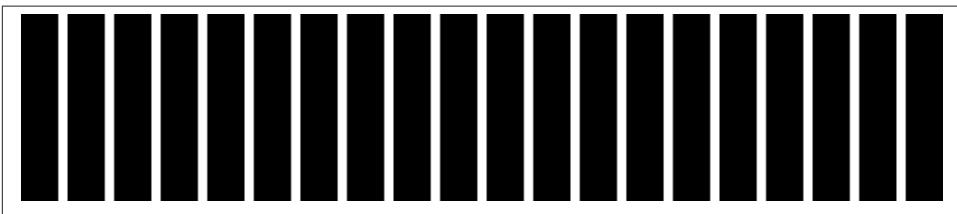


Figure 6-17. Twenty evenly spaced bars

or only five:

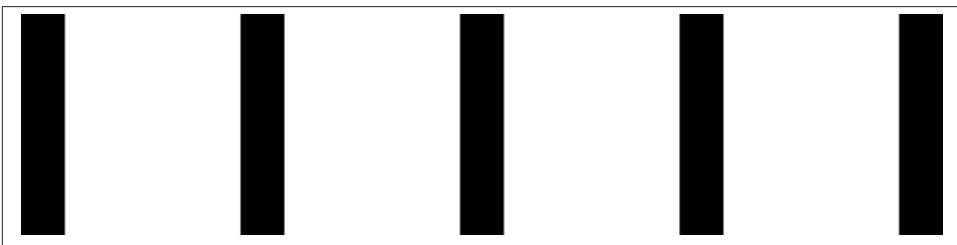


Figure 6-18. Five evenly spaced bars

See that code so far in `15_making_a_bar_chart_even.html`.

Now we should set the bar *widths* to be proportional, too, so they get narrower as more data is added, or wider when there are fewer values. I'll add a new variable near where we set the SVG's width and height

```
//Width and height
var w = 500;
var h = 100;
var barPadding = 1; // <-- New!
```

and then reference that variable in the line where we set each bar's width. Instead of a static value of 20, the width will now be set as a fraction of the SVG width and number of data points, minus a padding value:

```
.attr("width", w / dataset.length - barPadding)
```

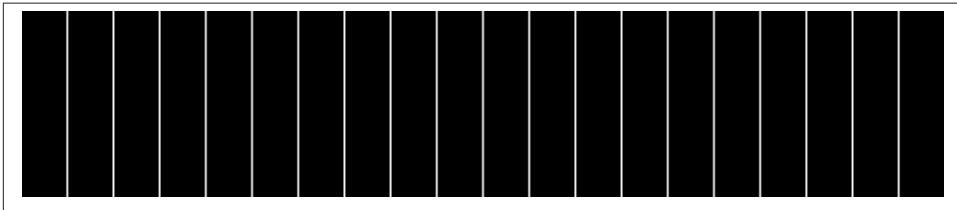


Figure 6-19. Twenty evenly spaced bars with dynamic widths

It works! (See `16_making_a_bar_chart_widths.html`.) The bar widths and x positions scale correctly whether there are 20 points, only five

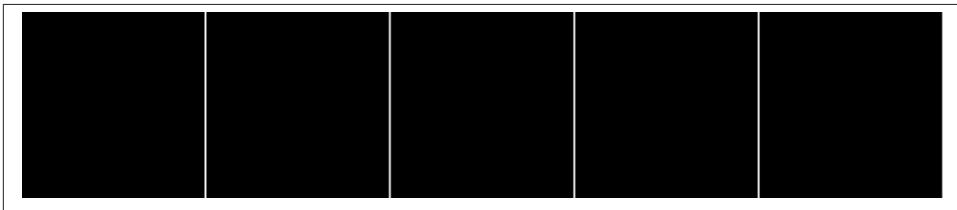


Figure 6-20. Five evenly spaced bars with dynamic widths

or even 100:

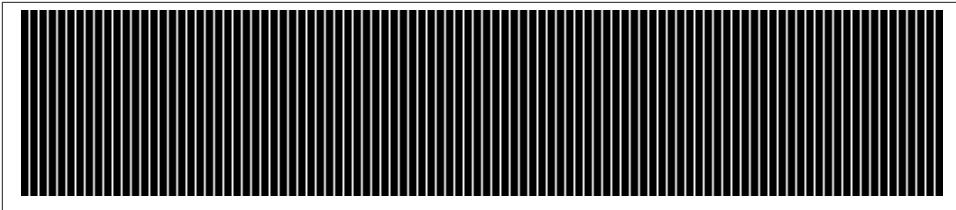


Figure 6-21. One hundred evenly spaced bars with dynamic widths

Finally, we encode our data as the *height* of each bar. You would hope it were as easy as referencing the `d` data value when setting each bar's height:

```
.attr("height", function(d) {
  return d;
});
```

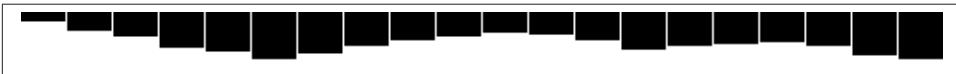


Figure 6-22. Dynamic heights

Hmm, that looks funky. Maybe we can just scale up our numbers a bit?

```
.attr("height", function(d) {
  return d * 4; // <-- Times four!
});
```

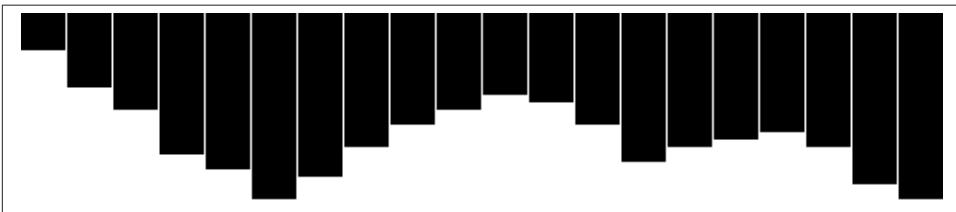


Figure 6-23. Dynamic heights, magnified

Alas, it is not that easy! We want our bars to grow upward from the bottom edge, not down from the top — but don't blame D3, blame SVG.

You'll recall that, when drawing SVG `rect`s`, the `x` and `y` values specify the coordinates of the *upper-left corner*. That is, the origin or reference point for every `rect` is its top-left. For our purposes, it would be soooooo much easier to set the origin point as the bottom-left corner, but that's just not how SVG does it, and, frankly, SVG is pretty indifferent about your feelings on the matter.

Given that our bars do have to “grow down from the top,” then where is “the top” of each bar in relationship to the top of the SVG? Well, the top of each bar could be expressed as a relationship between the height of the SVG and the corresponding data value, as in:

```
.attr("y", function(d) {  
    return h - d; //Height minus data value  
})
```

Then, to put the “bottom” of the bar on the bottom of the SVG, each `rect`'s height can be just the data value itself:

```
.attr("height", function(d) {  
    return d; //Just the data value  
});
```



Figure 6-24. Growing down from above

Let's scale things up a bit by changing `d` to `d * 4`. (Note: Later we'll learn about D3 scales, which offer better ways to accomplish this.)

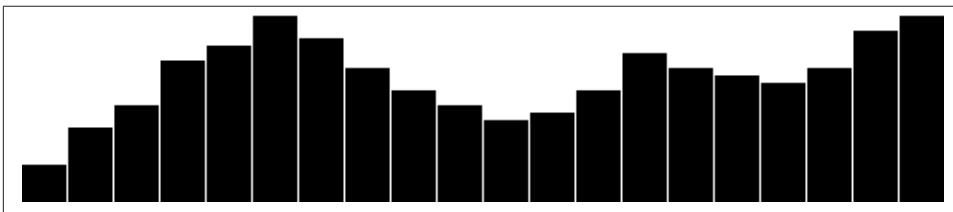


Figure 6-25. Growing bigger from above

The working code for our growing-down-from-above, SVG bar chart is in `17_making_a_bar_chart_heights`.

Color

Adding color is easy. Just use `attr()` to set a fill:

```
.attr("fill", "teal");
```

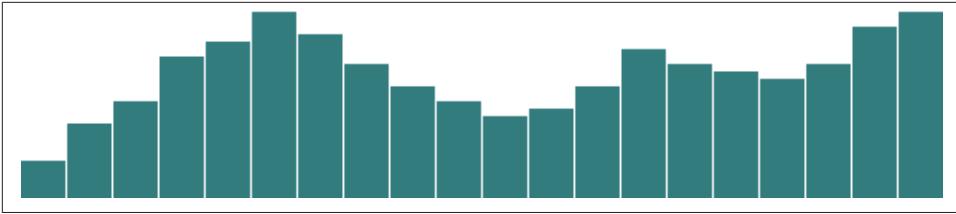


Figure 6-26. Teal bars

Find that all-teal bar chart in `18_making_a_bar_chart_teal`.

Teal is nice, but you'll often want a shape's color to reflect some quality of the data. That is, you may want to *encode* the data values as color. (In the case of our bar chart, that makes a *dual encoding*, in which the same data value is encoded in two different visual properties: both height and color.)

Using data to drive color is as easy as writing a custom function that again references `d`. Here, we replace "teal" with a custom function:

```
.attr("fill", function(d) {
  return "rgb(0, 0, " + (d * 10) + ")";
});
```

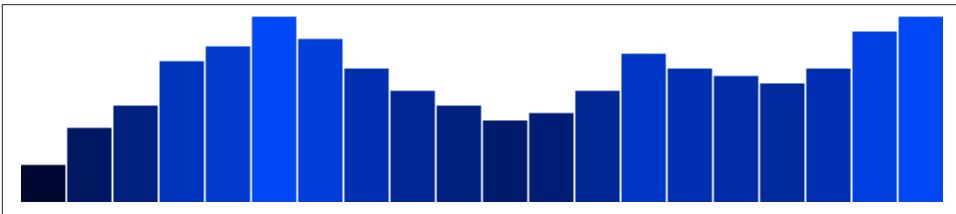


Figure 6-27. Data-driven blue bars

See the code in `19_making_a_bar_chart_blues.html`. This is not a particularly useful visual encoding, but you can get the idea of how to translate data into color. Here, `d` is multiplied by 10, and then used as the blue value in an `rgb()` color definition. So the greater values of `d` (taller bars) will be more blue. Smaller values of `d` (shorter bars) will be less blue (closer to black). The red and green components of the color are fixed at zero.

Labels

Visuals are great, but sometimes you need to show the actual data values as text within the visualization. Here's where value labels come in, and they are very, very easy to generate with D3.

You'll recall from the SVG primer that you can add text elements to an SVG element. We'll start with:

```
svg.selectAll("text")
  .data(dataset)
  .enter()
  .append("text")
```

Look familiar? Just as we did for the `rect`'s, here we do for the `text`'s. First, select what you want, bring in the data, enter the new elements (which are just placeholders at this point), and finally append the new `text` elements to the DOM.

We'll extend that code to include a data value within each text element by using the `text()` method

```
.text(function(d) {
  return d;
})
```

and then extend it further, by including `x` and `y` values to position the text. It's easiest if I just copy and paste the same `x/y` code we used above for the bars:

```
.attr("x", function(d, i) {
  return i * (w / dataset.length);
})
.attr("y", function(d) {
  return h - (d * 4);
});
```

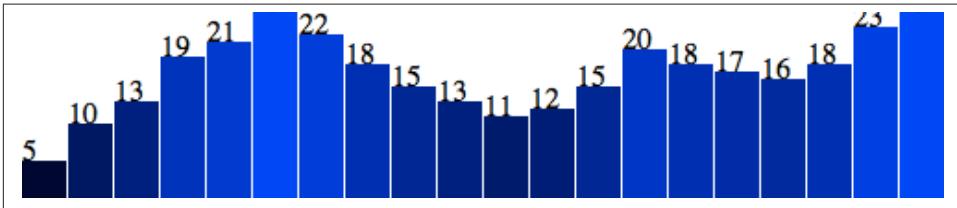


Figure 6-28. Baby value labels!

Aha! Value labels! But some are getting cut off at the top. Let's try moving them down, inside the bars, by adding a small amount to the `x` and `y` calculations:

```
.attr("x", function(d, i) {
  return i * (w / dataset.length) + 5; // +5
})
.attr("y", function(d) {
  return h - (d * 4) + 15; // +15
});
```

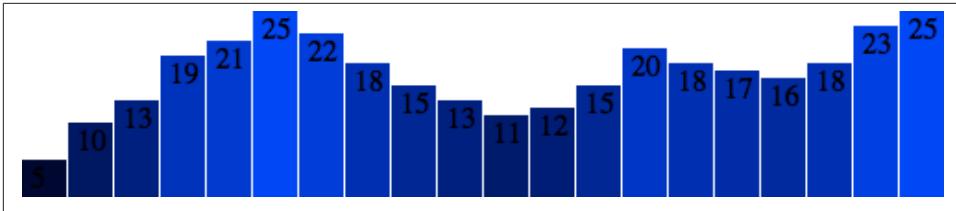


Figure 6-29. In-bar value labels

Better, but not legible. Fortunately, we can fix that:

```
.attr("font-family", "sans-serif")
.attr("font-size", "11px")
.attr("fill", "white");
```

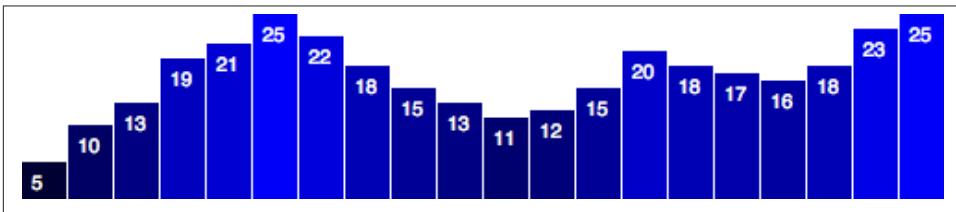


Figure 6-30. Really nice value labels

Fantasti-code! See `20_making_a_bar_chart_labels.html` for that brilliant visualization.

If you are not typographically obsessive, then you're all done. If, however, you are like me, you'll notice that the value labels aren't perfectly aligned within their bars. (For example, note the "5" in the first column.) That's easy enough to fix. Let's use the SVG `text-anchor` attribute to center the text horizontally at the assigned `x` value:

```
.attr("text-anchor", "middle")
```

Then, let's change the way we calculate the `x` position by setting it to the left edge of each bar *plus* half the bar width:

```
.attr("x", function(d, i) {
  return i * (w / dataset.length) + (w / dataset.length - barPadding) / 2;
})
```

And I'll also bring the labels up one pixel for perfect spacing, as you can see in `21_making_a_bar_chart_aligned.html`:

```
.attr("y", function(d) {
  return h - (d * 4) + 14; //15 is now 14
})
```

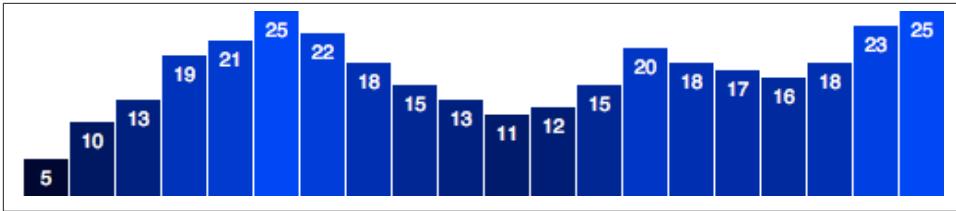


Figure 6-31. Centered labels

Done!

Making a Scatterplot

So far, we’ve drawn only bar charts with simple data — just one-dimensional sets of numbers.

But when you have two sets of values to plot against each other, you need a second dimension. The scatterplot is a common type of visualization that represents two sets of corresponding values on two different axes: horizontal and vertical, x and y.

The Data

As you saw in the “Technology Fundamentals” section on JavaScript, you have a lot of flexibility around how to structure a data set. For our scatterplot, I’m going to use an array of arrays. The primary array will contain one element for each data “point.” Each of those “point” elements will be another array, with just two values: one for the x value, and one for y.

```
var dataset = [
    [5, 20], [480, 90], [250, 50], [100, 33], [330, 95],
    [410, 12], [475, 44], [25, 67], [85, 21], [220, 88]
];
```

Remember, `[]` means array, so nested hard brackets `[][]` indicate an array within another array. We separate array elements with commas, so an array containing three other arrays would look like: `[[[]], [], []]`

We could rewrite our data set with more whitespace so it’s easier to read, like so:

```
var dataset = [
    [ 5,    20 ],
    [ 480,  90 ],
    [ 250,  50 ],
    [ 100,  33 ],
    [ 330,  95 ],
    [ 410,  12 ],
    [ 475,  44 ],
```

```

    [ 25, 67 ],
    [ 85, 21 ],
    [ 220, 88 ]
];

```

Now you can see that each of these 10 rows will correspond to one point in our visualization. With the row [5, 20], for example, we'll use 5 as the x value, and 20 for the y.

The Scatterplot

Let's carry over most of the code from our bar chart experiments, including the piece that creates the SVG element:

```

//Create SVG element
var svg = d3.select("body")
    .append("img")
    .attr("width", w)
    .attr("height", h);

```

Instead of creating `rect`'s, however, we'll make a `circle` for each data point:

```

svg.selectAll("circle") // <-- No longer "rect"
    .data(dataset)
    .enter()
    .append("circle")   // <-- No longer "rect"

```

Also, instead of specifying the `rect` attributes of `x`, `y`, `width`, and `height`, our `circle`'s need `cx`, `cy`, and `r`:

```

    .attr("cx", function(d) {
        return d[0];
    })
    .attr("cy", function(d) {
        return d[1];
    })
    .attr("r", 5);

```

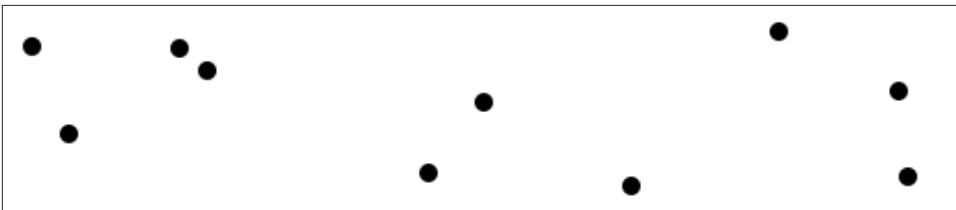


Figure 6-32. Simple scatterplot

See the working scatterplot code in `22_scatterplot.html`.

Notice how we access the data values and use them for the `cx` and `cy` values. When using `function(d)`, D3 automatically hands off the current data value as `d` to your function. In this case, the current data value is one of the smaller, sub-arrays in our larger data set array.

When each single datum `d` is itself an array of values (and not just a single value, like `3.14159`), you need to use bracket notation to access its values. Hence, instead of `return d`, we have `return d[0]` and `return d[1]`, which return the first and second values of the array, respectively.

For example, in the case of our first data point `[5, 20]`, the first value (array position `0`) is `5`, and the second value (array position `1`) is `20`. Thus:

```
d[0] == 5
d[1] == 20
```

By the way, if you ever want to access any value in the larger data set (outside of D3, say), you can do so using bracket notation. For example:

```
dataset[5] == [410, 12]
```

You can even use multiple sets of brackets to access values within nested arrays:

```
dataset[5][1] == 12
```

Don't believe me? Take another look at the scatterplot page `22_scatterplot.html`, open your JavaScript console, type in `dataset[5]` or `dataset[5][1]`, and see what happens.

Size

Maybe you want the circles to be different sizes, so their radii correspond to their `y` values. Instead of setting all `r` values to `5`, try:

```
.attr("r", function(d) {
  return Math.sqrt(h - d[1]);
});
```

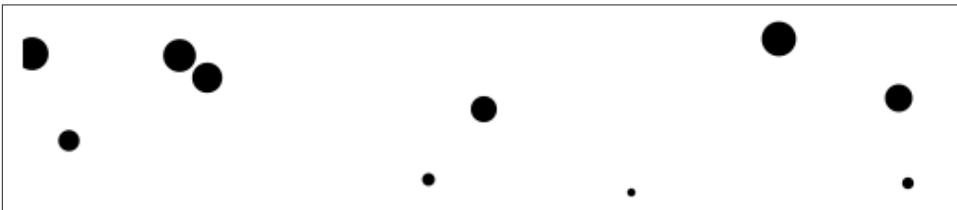


Figure 6-33. Scatterplot with sized circles

See `23_scatterplot_sqrt.html` for the code. `Math.sqrt()` is JavaScript's square root function. Here, we arbitrarily subtract the datum's y value `d[1]` from the SVG height `h`, and then take the square root. The effect is that circles with greater y values (those circles lower down) have smaller radii.

This is neither pretty nor useful, but it illustrates how to use `d`, along with bracket notation, to reference data values and set `r` accordingly.

Labels

Let's label our data points with `text` elements. I'll adapt the label code from our bar chart experiments, starting with:

```
svg.selectAll("text") // <-- Note "text", not "circle" or "rect"
  .data(dataset)
  .enter()
  .append("text")    // <-- Same here!
```

This looks for all `text` elements in the SVG (there aren't any yet), and then appends a new `text` element for each data point. Then use the `text()` method to specify each element's contents:

```
.text(function(d) {
  return d[0] + ", " + d[1];
})
```

This looks messy, but bear with me. Once again, we're using `function(d)` to access each data point. Then, within the function, we're using *both* `d[0]` *and* `d[1]` to get both values within that data point array.

The plus `+` symbols, when used with strings, such as the comma between quotation marks `" , "`, act as *append* operators. So what this one line of code is really saying is: Get the values of `d[0]` and `d[1]` and smush them together with a comma in the middle. The end result should be something like `5,20` or `25,67`.

Next, we specify *where* the text should be placed with `x` and `y` values. For now, let's just use `d[0]` and `d[1]`, the same values that we used to specify the `circle` positions.

```
.attr("x", function(d) {
  return d[0];
})
.attr("y", function(d) {
  return d[1];
})
```

Finally, add a bit of font styling with:

```
.attr("font-family", "sans-serif")
.attr("font-size", "11px")
.attr("fill", "red");
```

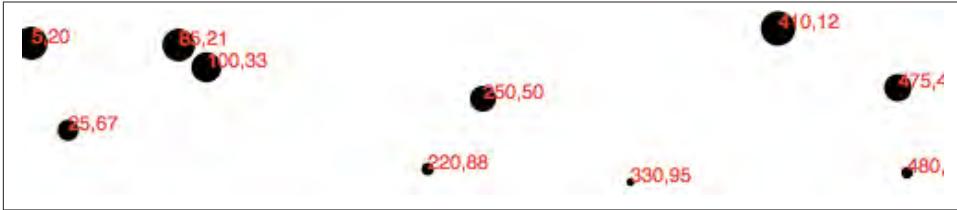


Figure 6-34. Scatterplot with labels

It may not be pretty, but we got it working! See `24_scatterplot_labels.html` for the latest.

Next Steps

Hopefully, some core concepts of D3 are becoming clear: Loading data, generating new elements, and using data values to derive attribute values for those elements.

Yet the image above is barely passable as a data visualization. The scatterplot is hard to read, and the code doesn't use our data flexibly. To be honest, we haven't yet improved on — gag — Excel's *Chart Wizard!*

Not to worry: D3 is way cooler than Chart Wizard (not to mention Clippy), but generating a shiny, interactive chart involves taking our D3 skills to the next level. To use data flexibly, we'll learn about D3's *scales* in the next chapter. And to make our scatterplot easier to read, we'll learn about *axis generators* and axis labels.

This would be a good time to take a break and stretch your legs. Maybe go for a walk, or grab a coffee or a sandwich. I'll hang out here (if you don't mind), and when you get back, we'll jump into D3 scales!