How to get from idea to execution

Hey, everyone. Welcome back to Video 3 in Module 6. In the last video, we showcased a bunch of examples of data stories. But how do you get from an idea to the final product? In this video, we'll talk about that. We'll look at how The Pudding finds and vets its stories and what makes them special.

So let's jump right in. The first question we often get asked is how do you find stories? And the answer is simple: Question first.

Often people assume that we must spend our day scouring the Internet for cool datasets or finding trends and patterns in everyday life. But it's quite the opposite. Almost all of our stories start with the question.

In fact, if you scroll down through our archive, a lot of the questions reveal themselves to you in the headlines. So how often names are in songs? How much data lies behind the paper you're reading? Why Budapest, Warsaw and Lithuania split themselves in two? How algorithms know what you'll type next. Who is the biggest pop star? And how many high school stars make it in the NBA?

Starting with the question gives you a clear purpose, and it makes it easier to communicate your idea with someone else. The inevitable follow up question that we get asked is, so where do your questions come from? And the bad news is there's no one secret source for inspiration. But the good news is the questions can come from anywhere. I'm a big believer in shower epiphanies, the latest podcast you listened to, something that happened during your commute. Anything can spark a question. The most important thing you can do is make a note of that initial spark of curiosity.

At The Pudding, we all keep a backlog of ideas. It's a place to jot down thoughts, make notes of questions, and remind ourselves of subjects that we should look into further. Even if you don't have a question yet, you might have a subject that really, really interest you. And eventually that question will come along. We also keep a public backlog of our ideas on our Website.

We're a small team of about six to eight journalists, engineers, and we know that we can't possibly produce data stories for every question that comes to mind. So this is a kind of repository, a brainstorm, a dump of ideas, that we want the public and our readers to be able to look at and see if they want to kind of own and take care of any of these questions for us. We're constantly accepting freelance pitches.

So from there, how do you make sure your idea is interesting? There are three main things that we try to do.

One, does your idea spark intense debate or discussion? If you seriously can't stop talking about the idea, you're probably on to something. The idea for Amber and I's piece on pockets came when we were walking around New York City complaining about all the bags we had to carry and how our phones wouldn't fit because we didn't have any functional pockets. Months later, we were still talking about this, and everywhere we looked we saw other women ranting, too.

The second thing you can do is to gently pitch or casually bring up your ideas to your friends or others and gage their response. If one other person is energized by your idea, then it's worth investigating. Don't forget to trust your instincts and follow your passions, too, because your own excitement can be contagious and that can translate and go a long way when you tell the story.

It's equally as important to know your audience. Who are you telling the story to? At The Pudding, we say we write for the general public, so that answer is broad. It's everyone. But we also make sure to ask ourselves the question, if only one person read this story. Who is that person? Nadieh Bremer, who we mentioned in the data arts sphere, has something she calls the dad test. And I use a similar mom test on my stories. I try to tell my mom kind of what's going on in my head. Does it make sense? Gage her reaction. And if she can't understand it, then I probably need to go back to the drawing board.

So what types of stories make for successful data stories? First, let's look at what works well. The number one thing you need for data story is surprise, data. That doesn't mean it has to be clean and perfect, or that it has to be numbers at all, or even that it has to exist yet. You can collect it yourself, but it does mean that you need to be able to structure or categorize that data some way. That raw information needs to be put into a pattern.

The second thing is that it needs to be broadly relatable. People want to see themselves in stories and have connections with the data itself. It also needs to be a topic where visuals can help make sense of the story. Can you illustrate a pattern or show a trend? People really respond well to that kind of order making.

And the final thing, because if you haven't heard it enough and haven't heard me hit it enough yet, your story needs to be question driven. What really sets data stories apart is the question focus. Otherwise, you can end up with a dashboard or a data dump, something that has a load of information but doesn't walk the reader through it in any meaningful way.

On the flip side, here's things that don't work. The first is stories that are highly personal or rely heavily on a human narrative, and those are because there are better avenues to tell those stories and data stories. Often there's documentaries or photo stories or just a really well-thought-out narrative. Stories that are too broad are also difficult to tell as data stories. We always say don't try to boil the ocean. You're doing too much. It's an impossible task. Instead, try to shrink back your idea, scale it back, try to boil a pot on your stove instead of boiling the ocean.

And finally, since data stories are at their best when they're explaining something or answering a question, topics that are too obvious or overly simplistic often don't work. That's the water is wet model.

So when we think about our process of finding the right data stories for The Pudding, we put them through a few paces. The first step of that is that question step, and some of the best questions take you beyond the who, what, where and into the how and why. Those are the types of questions that reveal those universal truths that people really gravitate towards.

The second step is data. After all, we're still talking about data stories, and this is the step that can often trip up a lot of newcomers. Data work is probably the most intimidating part

of the process for a lot of people. But it's important to remember that learning to work with data is exactly as infuriating as it was at the beginning for everyone else. There's no wrong way to solve a problem, no right language to work in, and no limit to how many times you can Google or Stack overflow a question. Even at The Pudding, we all use different languages for data work. And here are just a handful. We'll use are with R with R studio, Python with pandas, Node J.S., QGIS for geographic data, MySQL, and good old-fashioned spreadsheet programs like Google Sheets and Excel. There is no shame in that game. We've even devoted a whole blog post to the topic of data, and you can find that on our site. It's called How to make Dope Shit Part 1 Working With Data.

The third step in the process is storyboarding. And this step like data is also super important because it's how people are going to absorb the information you're giving them. So ask yourself, what's the number one thing that you want people to take away from your story? That's a good place to start. So often at The Pudding our stories end up taking two mirrored pass. There's the small to big approach and the big to small approach.

For the big to small type, that's where the narrative starts with a hyper specific detail or a vignette and then works its way down to reveal a broad picture or overarching trend. I did this with my story that looked at the gender split of gay-borhoods. The first graphic that you scroll down to is a map of New York City. It shows you how I layered on data from Pride March routs, gay bars tagged in Yelp, same sex unmarried partner households, and same sex married joint tax filers to form this idea of a gay-borhood index. But of course, New York City is just one of the 15 cities I looked at in this piece. And that's where that small to big approach comes in. We used New York as a gateway to your larger story to explain those larger trends.

The opposite approach means that you start big and end small, and that means broad picture to small details. Amber and I's piece on women's pockets took this approach. In this piece, the first graphic that you get is a scrolling version showing you the 80 pairs of front pockets in blue jeans that we measured. It moves to the average pocket sizes for men and women, and then tells you our overall findings that women's pockets in jeans just suck.

We also tend to think of our stories in steps or sections, and those are independent from one another and can move around in different puzzle-piece formats. This allows you to rearrange that narrative quickly and alter the story flow. For example that piece on pockets, after the first section you are released in kind of this larger section where you can explore brand by brand and which pieces like iPhones or women's hands can fit in these pockets. At first this section was at the complete end of the story, but because our colleagues told us that people would really enjoy this, we moved it up. And because we were thinking about this as kind of a modular storyboard, that was easy to do.

The fourth step in our process of data storytelling is design. Because now you know exactly how you're going to tell the story, it's important to start designing what it looks like. A few gifted people can sketch a few lines on a napkin, jump straight into code, design as they go and produce something brilliant. But for the rest of us, it's important to consider design and development as completely separate steps.

First, you'll need to figure out what form your data will take. And a great starting point for this is the Financial Times's visual vocabulary chart. It breaks down the chart types by data

relationship. So if you're looking for a correlation, a scatter plot might be the right way to go. Next, you'll want to see how your data looks in the chart type that you've chosen.

To scaffold out data visualizations, our team at The Pudding again uses a variety of tools, but we usually go with whatever where the most familiar with and whatever will produce the quickest results. There's no need to fully code out your charts if you can find an out-of-the-box tool that will give you a good idea of what your data will look like. We've used things like Google Sheets before to create heat maps. We also work in Flourish, which is a tool that gives you a lot of different chart types to choose from and then you're allowed to plug in your data. We use. QGIS when we're working with maps or geographic data.

And if there's not an out-of-the-box tool, sometimes we'll code it out ourselves. And in that case, we'll jump into a tool called Observable where you can code out and prototype D3, which is a data visualization library that attaches to JavaScript.

Just like it's important to start your story with the question, it's also important to consider that question when designing your visualization. On Datawrapper's blog Chartable, Lisa Charlotte Rost writes of a three step process when creating a chart. First, what's your point? What's your original question or hypothesis, and can you reduce that down to a headline? Second, make sure that that one point really comes across in your visualizations. You can use color or type hierarchy to achieve this and make sure that your readers are being directed to the most important point. And finally, does your chart actually show what you want it to show? Is it clear to the reader, or do you need to add additional context?

Finally, you'll want to design your data visualization or visualizations into a larger story experience. At The Pudding, our main tool for a static site design is something called Sigma. It is an online tool that is like Adobe Illustrator and that it operates in kind of vectors. But you can pull everything together more like an InDesign tool so that you can storyboard step by step. Sigma is also great because you can use it collaboratively. People can leave notes. You can design together. And it kind of integrates with the web because you can pull out your own CSS and styles.

Here we're looking at a storyboard that we did for a piece on the structure of stand up comedy, where we focused on Ali Wong's Netflix comedy routine. In this piece, it's kind of like a step by step one screen experience.

And when we look at the actual piece, it really didn't stray far from that original storyboard itself. We'll take a little bit to walk through these steps.

But having a storyboard that is almost identical or very, very close to the final product is a rarity. Most often you'll go through several iterations. And then I want to show you kind of those iterations through a project called RioRun that I made while working at The Guardian US. This was a phone app made for the Olympics that tracked your progress using G.P.S., and put you on this virtual marathon route. So you'd unlock different badges and pass different landmarks the more you ran. This is an article on source where we kind of catalog the entire process of how that came to be. It's how we made it. And there's a big part about design in there. It's called adding sights to sounds.

And here the first kind of step by step mockups we did. The first version was just the test development version. It was completely utilitarian. The colors are very brash and not complimentary at all. But all the functionality was there, and we knew what pieces and parts that we needed. The second version or version 2 is a little bit more polished. We're starting to see everything come together and that all the pieces and all the functionality while the user is going to interact with it. Things are starting to be highlighted. And then looking at version 3, that's the final design. That's the part where we layer on all the color, the style, really kind of make the app have a personality.

We knew that since this was a virtual app that we wanted to capture the vibe of Rio in the app design itself. So the colors are very vibrant, and it's really the only visual that you'll get of the city with this app. It's important to remember that design is a powerful storytelling tool in and of itself. Done effectively, it can make an emotional connection with the reader, but done ineffectively, it can look something like this. This is one of the most common doors that you will see in your day to day life. The sign tells me to pull. But the horizontal means that I should push. And that leaves me feeling like I hate doors.

And the moral of that story is that you should look in your everyday environment and see, are there certain things that are confusing me? And if there are, those are the types of things you don't want to do to your reader. Design should help guide them, not confuse them.

The last step in our process of data stories is development, and because the final front end is what people end up seeing, they assume that this step is the one that takes the longest. In reality, though, it's the behind-the-scenes stuff like data and storyboarding that account for the largest portion of our time.

At The Pudding are stories are built with a HTML, CSS and JavaScript with a heavy reliance on D3 for databases. We work as much as we can in public repositories on our GitHub so that our code is all available and accessible, and we even have a starter template that kind of comes pre-baked with a lot of things that we would use in an average project. And this is good because we don't have to keep reinventing the wheel. We already have these built out, and we can use them each time.

Our starter template includes things like handlebars for HTML templating, and that integrates with our ArchieML, which is a tool that you can write your copy or your text right into Google Docs and then pull that in and inject that into your HTML. Our template is also integrated with Google Sheets, and that makes it easy to pull in live data. We also have things like inline SBG's and a chart template to make quick D3 charts.

When you're designing and developing these last two steps, it's common to hear that you should design mobile first. After all, readers are skewing more mobile. In January, traffic to The Pudding sites hovered around 70 percent mobile. And although we don't always technically build stories mobile first at The Pudding, we do always make sure that they are as delightful on mobile as they are on desktop.

In data visualizations, responsive design is no longer a buzzword. It's not something that you can throw aside. It's simply part of the design and development process. Too often, responsive design means a grand experience on desktop, with features slowly being stripped back as the screen size gets smaller. But it doesn't have to be this way. We're

going to take a look at a couple of projects that The Pudding has used responsive design to make sure that the experience is seamless on both desktop and mobile.

Here's one that Ilia and I did that looks at New York Times headlines from their archives and finds words that appeared most frequently per decade. On the desktop version, we have a kind of a decade by decade navigation on the left side with the text and the word cloud in the middle chart here. And then on the right, there is another kind of aerial line chart that you can interact with. Three column designs are terrible on mobile. So we really had to think about this when we shrunk the down. When we look at this on a smaller screen size, that navigation is still on the left with the decades. The text and the word cloud is on there on the right. And then if you want to see a line chart, interact with the word cloud and it pulls out this drawer that you can then close.

Here's another piece from Amber that takes a look at your local your geo-located data and tells you what the weather would be like on Mars. It's a really cool project in that it sends you postcards as if you were getting them from Mars. On desktop, those postcards look like traditional postcards. They're horizontal and they flip over as you go. When we look at this on mobile, this is what the project looks like. Those cards have just flipped from their horizontal orientation to a vertical orientation. Nothing has been lost in the design. It's all still there.

One tool that we use to go from desktop to mobile seamlessly is called Flexbox. It's used in CSS, and it's kind of this grid-templating system. Basically, you can tell your containers, your divs, or your boxes to either move in a row, which would be a horizontal pattern, or a column, a vertical pattern. And then as your device shrinks, you can switch those back and forth. You can also apply this Flexbox technique to data viz. And Amber, of our team at The Pudding, even has a wonderful blog post on how just to make data viz using boxes and Flexbox on our site.