



Module 2	2
Module 2 Overview	2
Lecture 2-1: Understanding The Growth Of Data	4
Lecture 2-2: Evaluating Methods Of Data Access	112
Lecture 2-3: Focusing Your Communication Journey	22
Lecture 2-4: Planning Your Data Story	29



Module 2

Module 2 Overview

In This Module

Module 2: Working Fast and Thinking Slow

I

Key Concepts:

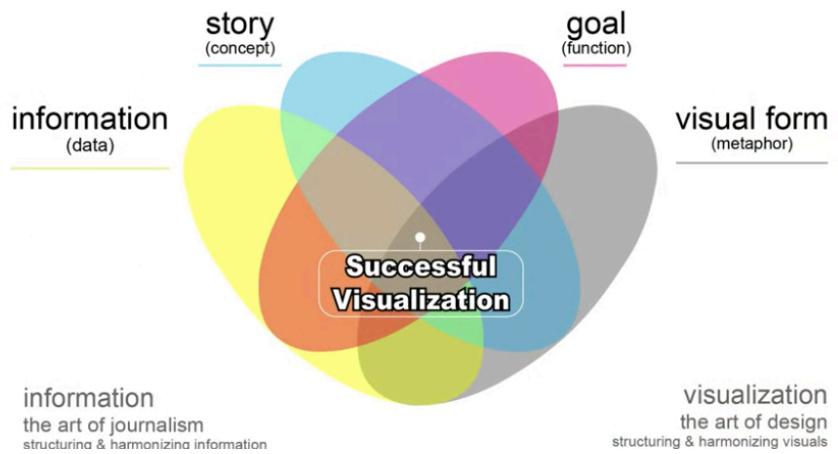
- Understanding the growth of data
- Evaluating methods of data access
- Focusing your communication journey
- Planning your data story
- Case Study: Bellabeat



A couple of concepts we're going to introduce in this module. We will be talking about the unbelievable growth of data. We'll evaluate the different methods that we as analysts have to collecting that data, accessing it. We will talk in terms of focusing our communication journey, particularly around an objective or a goal that we might have. We will talk about planning the data story, really that connective tissue between the data that we've collected and then the goal that we have. Then, we're going to introduce a company named Bellabeat that we will be using as a bit of a case study throughout the lecture videos and returning to them as an example as we introduce new concepts.



McCandless Offers a Thorough Definition of Good Data Stories



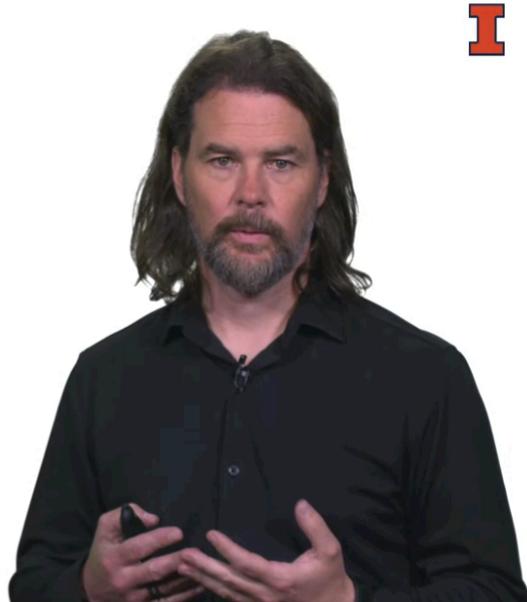
We are working, of course, through this framework, which has all the elements of a good, successful data visualization. In this module and first lesson, we're going to begin by looking at data.



Lecture 2-1: Understanding The Growth Of Data

Understanding the Growth of Data

We are living through the greatest expansion of data the world has ever seen, with no end in sight.



All right. Understanding the Growth of Data. We are truly living through the greatest expansion of data that the world has ever seen. What this has created is an environment where it has never been more important for us to understand how to effectively communicate with data.



90%

of the data in the world today (2018) were generated in the last two years alone.



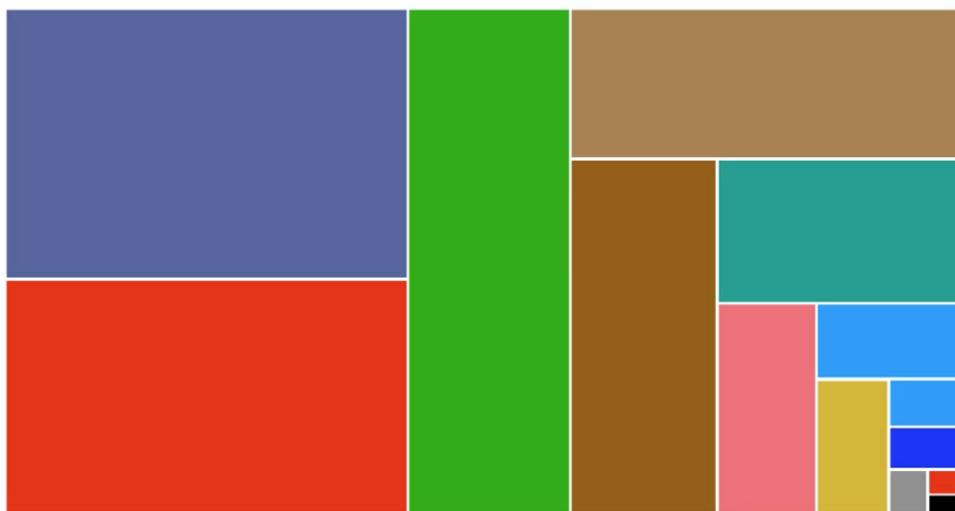
Source: Forbes (2018)

In the last two years alone, we've seen 90 percent of the world's data come to be. It's truly remarkable, the explosion of data all around us. Sometimes we lose context for how much data that really is.



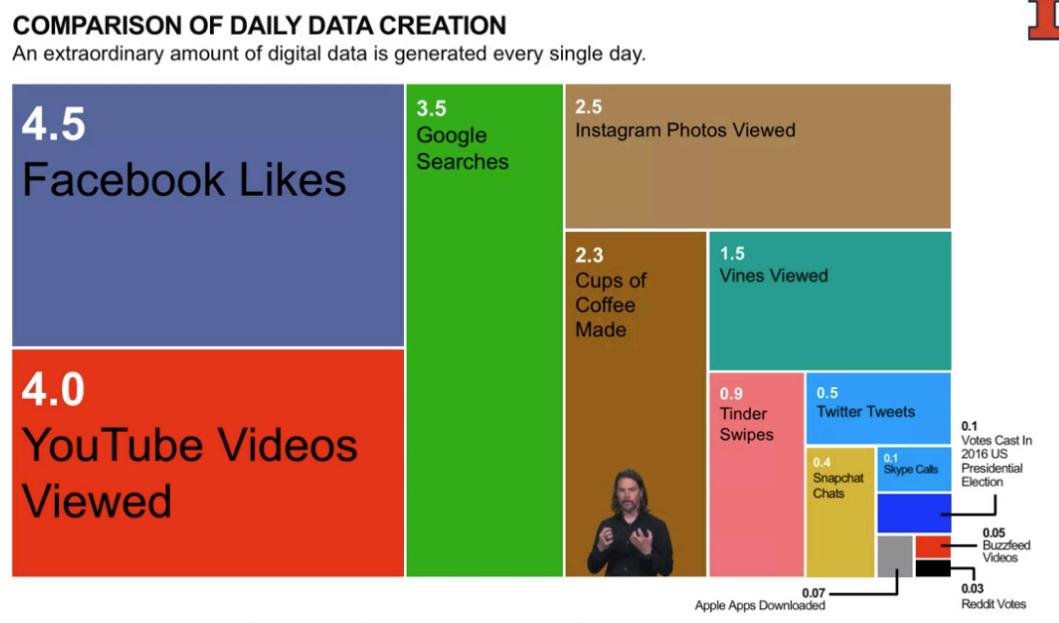
COMPARISON OF DAILY DATA CREATION

An extraordinary amount of digital data is generated every single day.





To that end, I put together this state of a visualization, based on an idea, a concept from David McCandless. What you're looking at is a comparison of daily data creation, and what you'll see is that the extraordinary amount of data being generated every day, this digital data that we can move, that we can, that we can access, that we can analyze and make stories from. So, I basically took data from a number of sources, and then dropped them into this graphic, sizing the boxes to the relative size. It gives us really I think good relative insight into how much data's being created, as well as a point of comparison to some offline data that will show us just how much is being produced.



Source: 2015 figures (unless otherwise noted), Domo, World Development, US Election Atlas

So, here you go. So, every single day in the world, there are 4.5 billion Facebook likes, 4 billion YouTube videos watch, 3.5 billion Google Searches, on and on. All of these things for marketers in particular, tremendously rich and valuable sources of information, allowing us to know what consumers like, what they don't like, and what they want, right? These data that are being generated are providing more insight than we've ever had before in the history of marketing, and say really the history of the world. But how much data really is this? That brown box that I haven't revealed yet? Well, that is the number of cups of coffee consumed every single day. You can literally consume coffee and purchase coffee anywhere in the world. That the rest of the stuff is all generated by the 40 percent or so of the world that is digitally enabled and it's online. That little blue box there in the lower right, well, those were the number of votes cast in the last presidential election. So, that's done once every four years. Everything else is happening every single day. It really does give you a sense of how much information and data is being generated. Useful, valuable, insightful data, from which we can tell



i M B A

University of Illinois at Urbana-Champaign

Introduction to Business Analytics:

Communicating with Data

Prof. Kevin Hartman

great stories, all right? As a result of this, the industries and companies that use and work with this data have also seen a similar path of growth.

I

BRINKER'S MARTECH LANDSCAPE

A crowded space indeed

In August 2011, Scott Brinker of ChiefMartec charted logos providers operating in the nascent business of "marketing technology" as a way to sort out the market. There were 150 of them.



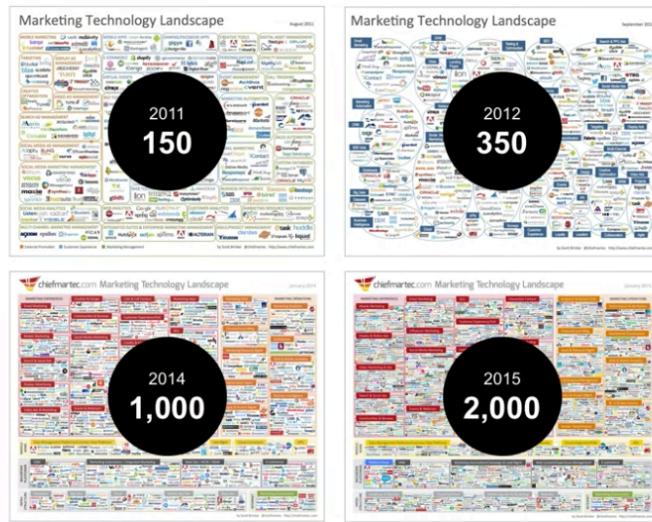
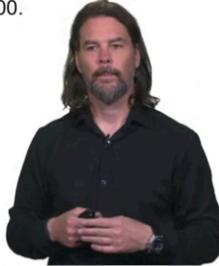
Source: Scott Brinker @chiefmartec www.chiefmartec.com

This is marked very well in an analysis done by Scott Brinker, where in 2011 Scott put together this infographic which became viral and made its way around the data community. Where Scott basically took the logos of all the companies operating in this very new and nascent marketing technology space and categorize them. So, for the first time, you got to see where were these are all the companies in this data space, working with this data that's being generated, making sense of it, processing it, doing something that allows them to make a viable business endeavor. That was in 2011. This was so popular that Scott decided to go back the next year, and then the next, and then the next and continue to revise this number.

BRINKER'S MARTECH LANDSCAPE

A crowded space indeed

By September 2012, when Brinker revisited his infographic, the number of firms operating in marketing technology had grown to 350. Then 1,000. Three years later, that figure would swell to 2,000.



Source: Scott Brinker @chiefmartec www.chiefmartec.com

What he found was pretty surprising. In 2011, there were 150 companies, by 2012, 350, by 2014, 2,000 companies working in this data space.

BRINKER'S MARTECH LANDSCAPE

A crowded space indeed

Brinker's analysis conducted in March of 2016 revealed an astonishing fact: 3,500 firms were operating in the space—almost twice the number just 14 months before.



Source: Scott Brinker @chiefmartec www.chiefmartec.com

By 2015 and 2016 that number had jumped to 3,500.



BRINKER'S MARTECH LANDSCAPE

A crowded space indeed

By 2017, the number of companies operating in the marketing technology space reached 5,000, growing more than 40% from 2016's total.



The one that was produced in 2016 showed 5,0000 companies.



BRINKER'S MARTECH LANDSCAPE

A crowded space indeed

In 2018, the number of companies operating in the marketing technology space has grown to nearly seven thousand.



Source: Scott Brinker @chiefmartec www.chiefmartec.com

The most recent in 2018 graphic that Scott generated shows nearly 7,000 companies now working in the marketing technology space, utilizing this data that is being created. Truly, there has never been a time where understanding that data and knowing how to communicate with that data has ever been as important as it is today.



1%

of the change technology will introduce to our lives and businesses is realized today.



But even with all these numbers, it might not be the most important number that I want to share. That number maybe this one, a simple one, but one with great meaning, one percent. This comes from a recent Mackenzie study, which found that the amount of change that technology will introduce into our personal and professional lives, we've but experienced one percent of that disruption. We have not even begun. So, with all the numbers and data that we have generated and can work with now, we're at the very beginning of this journey. Truly, if there's ever a time to understand how to use data to communicate, now is the time to start to learn these tasks and these skills.

Communicating with Data Has Never Been More Important



We are living in the midst of a data explosion

Extraordinary amounts of useful, portable digital data are being generated each day.

New companies (and industries!) working with data are emerging at a rapid pace.

We have only just begun to see the impact technology will have on our personal and professional lives.

So, as I've said communicating with data is never been more important. We are living in the midst of an enormous data explosion, and extraordinary amounts of this data as we've seen, is being generated every single day. Useful, rich, vital information that can tell great stories. New companies, new industries are springing out of this data, because it is just so valuable and so rich. Yet, we have only seen the very beginning of the revolution that we're living through. This is going to continue on and we will see much greater change over time.



Lecture 2-2: Evaluating Methods Of Data Access

Evaluating Methods of Data Access

Technology growth has led to new online access points for data, making digital analysis more robust and accessible.



All right. Evaluating the methods of data access. So, we've talked about the importance of data in our data communication story. Right? We've talked about the explosion and growth of data. So, what are the ways we can access that data? Let's look at a few of them and give each of them a little bit of an evaluation.

Technology Growth Has Led to New Online Access Points



Bulk downloads



APIs



Web scraping

I



There really are three primary ways that we as analysts will access the data that we use in our Data Communication story. We'll call those bulk downloads, APIs which is application programming interfaces and then web-scraping. Let's look at each of them individually. Bulk downloads is the idea that someone, a data owner, has provided a place for you as an analyst to go and collect that data.

Bulk Is a Controlled Way for Data Owners to Offer Data Access

Data tables are accessed through a GUI built by a data owner that directs query parameters.

Accessible data is typically stripped of sensitive information.

Download limits, user-based access can be enforced by data owners as well.

Data access can be simplified, making data more broadly attainable.

These are very controlled releases of data. Some of these, I would even think of as your internal company data. Someone is the owner of that data, you are accessing it for analysis purposes, that is what I would classify as a bulk download just as sure as any third party download sites. So, these bulk download areas are typically data tables that are accessed through some graphical user interface, or GUI, which is built for the purpose of giving you that data. The people who own this data want you to have in the set up a bridge for you to get to that data. These gives the data owner the ability to determine what they give out and what they don't to hold back the sensitive information. It also gives the data owner the ability to provide limits to how much of the data you can get and to differentiate access based on who you are as someone who has declared through a login who you are, what data you get, what data you don't. It is the easiest way to access data bar none.



Use Bulk Downloading to Access Large Data Sets via GUIs



Right. A good example of this is the census data here in the US.



They've constructed a site called American FactFinder. You can go to American FactFinder, there's great GUI and very simply and easily access all of or most of the US censuses data, at least the data they want you to get access to.

The screenshot shows a web browser window with the URL www.imdb.com/interfaces. The page title is "Alternative Interfaces". It features a central text box with instructions for accessing IMDb data locally via plain text files or Unix command line programs. On the left, there's a small image of a Samsung Galaxy Note 4 smartphone displaying the time "12:45". On the right, there's a large image of a man with a beard, wearing a black shirt, holding a small device and gesturing while speaking. The browser's address bar shows various tabs and links, including "How is 'beidh me le'" and "Google Lifestyle Per...".

Now, these things don't always look great. IMDb has what I would call a bulk download source as well, it's a little bit hidden, a little harder to find. But all the same, they have set up a place for you to go access data as an individual.



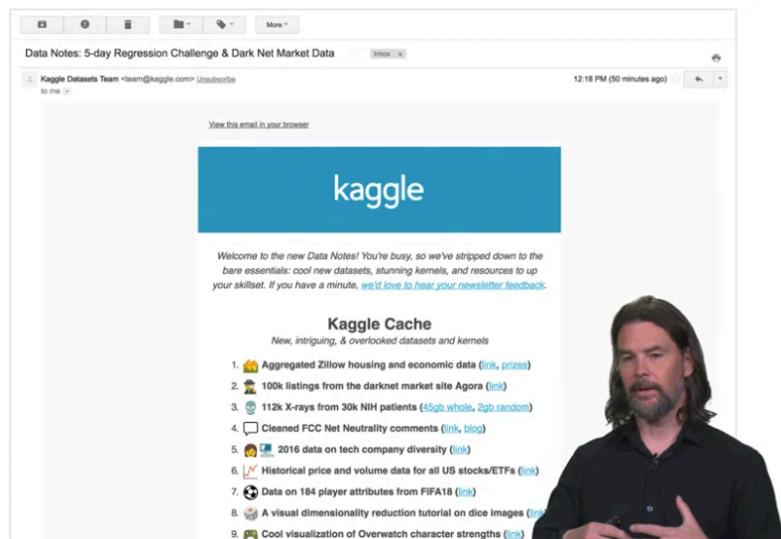
I

**DATA SO PLENTIFUL,
DELIVERED TO YOUR INBOX**

Kaggle and others proactively push data sets to analysts

Data have become such an inexpensive and important resource that services whose business models depend upon engagement can use it as a free offering to entice analysts' participation.

Once a hard-earned asset whose value was driven by its uniqueness, data have become an effective loss leader.



We live in an era of such plentiful data that we really need to do little more than sign up for a newsletter to get tremendous amounts of information. Interesting, vital datasets sent our way. Kaggle does this weekly, we'll send you a list of interesting new databases that you can access through a prescribed dedicated access link, that is again this idea of bulk downloads.

APIs Provide Instant, Automatic Access to Data

I

Typically, just a URL and API works like the way your browser makes a call.

API operation is a machine-to-machine process of data exchange.

Scripting language, set of commands that meet data host's vernacular required to access data.

API keys, access tokens are used to manage data access.



A second way that we can access data is what we call API's application programming interfaces. Truly the idea of an API is to let a computer talk to a computer. It is a machine to machine connection. Doing this accessing and building an API connection for yourself would require some understanding scripting language, a set of commands for you to use, there are tokens and keys that access security. Again, the design is meant to be Machine to Machine and taken care of automatically. Well, you can, as an analyst, build your own API access tool to go out and tap this data. Typically, data that we want to have access to would be better found through a bulk download. But APIs can work.

Example

The screenshot shows a presentation slide with a light gray background. On the left, there is a code editor window containing an HTML script for a Google Map. The code includes a DOCTYPE declaration, an HTML tag, a head section with a script tag pointing to the Google Maps API, and a body section with a div element for the map. Below the code editor is a green button labeled "Try it yourself ». On the right side of the slide, there is a video player window showing a man with a beard and dark hair, wearing a black shirt, gesturing with his right hand while speaking. Above the video player is a red stylized letter 'T' icon.

```
<!DOCTYPE html>
<html>
<head>
<script src="http://maps.googleapis.com/maps/api/js"></script>
<script>
function initialize() {
  var mapProp = {
    center:new google.maps.LatLng(51.508742,-0.120850),
    zoom:5,
    mapTypeId:google.maps.MapTypeId.ROADMAP
  };
  var map=new google.maps.Map(document.getElementById("googleMap"),mapProp);
}
google.maps.event.addDomListener(window, 'load', initialize);
</script>
</head>

<body>
<div id="googleMap" style="width:500px;height:380px;"></div>
</body>
</html>
```

Try it yourself »

This is an example of some of that scripting, this is the API for Google Maps. Rather than going out and running your own API, you can utilize sites that are really fancy user-friendly markups of APIs.



The screenshot shows a Twitter analysis tool named "Tweet Binder". At the top, it displays a collection for "@realdonaldtrump" with 1,983 tweets from June 19, 2018, at 8:32:28 PM. It offers options to "NEW COLLECTION", "Collections", "Instagram reports", "Events", "Plans and pricing", "Contact us", "Help", "SHARE ON TWITTER", "DOWNLOAD EXCEL", and "INFOGRAPHIC REPORT". Below this, the main area is titled "Activity" and features a line chart showing the volume of tweets over time. The chart has a yellow background and shows several peaks. To the left of the chart, there's a note: "You could get up to 30 days old tweets" and a button "GET A HISTORICAL REPORT". To the right, there are four colored circles representing different metrics: 64 original (dark blue), 1,113 retweets (light blue), 720 replies (green), and 186 links & pics (pink). There's also a note: "Activate real time streaming to continue capturing tweets now" and a button "GET A LIVE REPORT". At the bottom, it says "Contributor rankings". A red UIUC logo is visible in the top right corner.

Let me show you an example, Tweet Binder is one of those. Tweet Binder will give you great access and insight into Twitter data. Basically, what Tweet Binder is, it's an API connection to the Twitter API that takes in through its nice graphical interface, some information that then pops into their API query, goes off pings the Twitter API, grabs data, brings it back all cleaned up for you. So, in this case it's cheating, because you're using an API but you're accessing it through someone else's GUI. That would be in my estimation not as effective as finding a bulk download source, it is still a better alternative to writing your own API scripts that take a little more time and a little more expertise.

Technology Growth Has Led to New Online Access Points



Web scraping



A third way that we can access data is through this practice that's called Web Scraping. Now, in the early days of the web, this was the black hat access to data, this is the sneaky way to go about and get whatever data you need. It's frankly fallen out of practice, and that's because we live again with so much information around us, and so many organizations and data owners are willing to give you that data. There's very little reason for you to go on and scrape data that you're not intended to have.

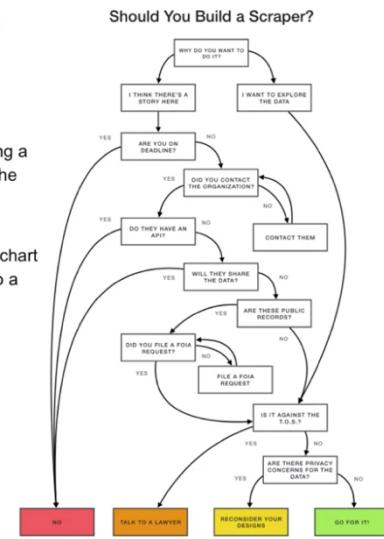
TO SCRAPE OR NOT TO SCRAPE, THAT IS THE QUESTION

While tempting, scraping data from websites is problematic

Sophie Chou is a graduate student at the MIT Media Lab focusing on machine learning and journalism. After attending a session on web scraping at the 2016 NICAR conference, she wrote an enlightening piece on the technical (and ethical) challenges surrounding web scraping.

Ms. Chou's point of view was well illustrated through a flowchart she developed for her article, where nearly all roads lead to a decision not to scrape.

(Chou & Storybench.org, 2016)



In fact, in an analysis an interesting one was done by Sophie Chou who put together this thought matrix to say through a decision tree, whether you should or should not build a Web scraper, and almost every single path you went down said, "Don't build a Web scraper." Why would you scrape this data?



Webpage Scraping Grabs HTML Formatted Data off Web

I

Data are intended for presentation, not necessarily download.

Unlike APIs and bulk download sites, data exchange is not facilitated by data host.

Data can often be governed by copyrights and/or needs for special permissions.

Access to data ranges from simple (highlight, right-click, “save as”) to complicated (python scripts).

There really are a number of good reasons why you would not. If you need to scrape data from an html webpage or any source, that's only because it hasn't been packaged for you to collect, which means that unlike through an API or bulk download site, that data owner really probably doesn't want you to have that data. If they did, they would make it available. Right. There are usually copyrights and other licensing issues that govern much of that data, and it can be at times a very easy case of simply just copying and pasting data from an HTML page, can require pre-sophisticated Python programming or other language use to go out and grab the data that you want. Again, data that probably was never intended for you to get anyways.

Technology Growth Has Led to New Online Access Points

I



Bulk downloads



APIs



Web scraping



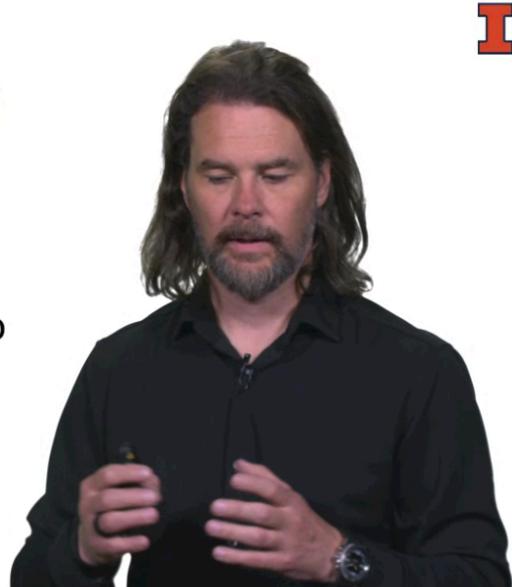
So, those are the three ways that we as analysts can access all this great data that's being created. Again, in my opinion the bulk download area on site is the one that should form the vast majority of the data that we're using in our analysis, either publicly available data that we access through that means or our own company internal data that we access through that means. APIs can in some cases work well and Web scraping is something that I think as an analyst, we're better just to leave behind.



Lecture 2-3: Focusing Your Communication Journey

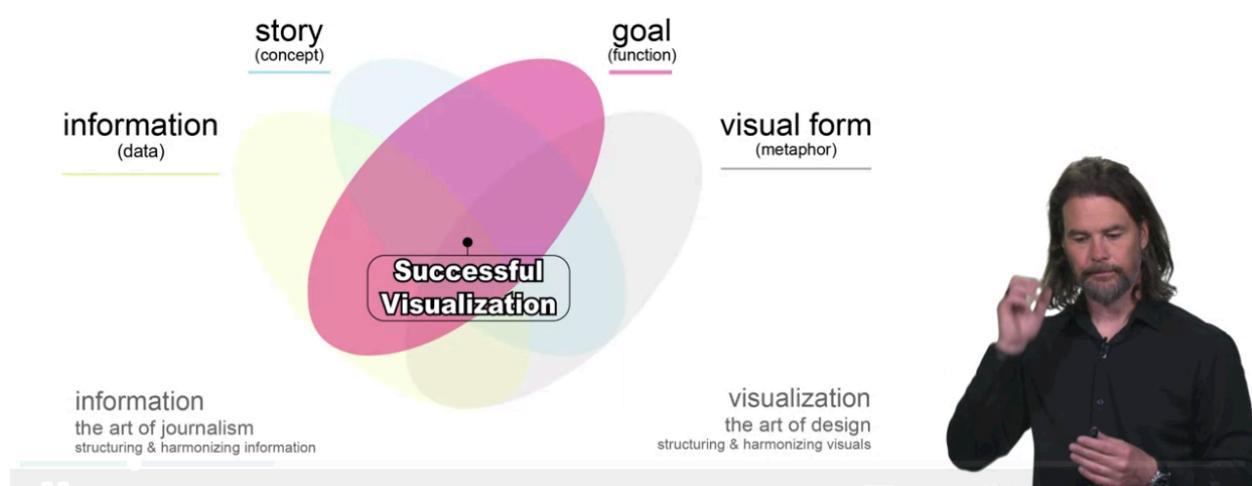
Focusing Your Communication Journey

Clearly articulating a goal helps the analyst center the story they plan to tell, thereby focusing their approach.



Focusing your communication journey, let's talk a little bit now about establishing a goal for your communication journey, for your ultimate data visualization. What that goal would do is allow you to again be focused in the data that you're collecting. It'll give purpose to that narrative, the story that you're building, that's going to connect the data to that objective and goal, and ultimately save you loads of time and make you much more efficient as an analyst.

McCandless Offers a Thorough Definition of Good Data Stories





This goal, as we saw, came as a crucial element of the framework we're using to evaluate a good and successful data visualization. Now, there are a number, a whole host of goals and objectives that you could have as an analyst.

Being Clear and Singular in Your Objective Is Paramount



Build Awareness

Influence Consideration

Improve Sales Process

Do consumers recall and recognize my brand?

Do my products satisfy consumers' needs?

Do my sales efforts result in wins for my brands?

Reposition the Brand

Grow Loyalty

Do the experiences I deliver fulfill customer expectations?

Do consumers advocate for my brand?



If you are working as a business analysts, I would argue that there are five broad objectives that any company has. These five objectives, I would also argue, are what we would call MECE, M-E-C-E, Mutually Exclusive, Collectively Exhaustive. Each one is distinct from the other, and taken together, they comprise the entire universe of goals that we would want. Now, they are pretty high level, but when you boil all the business objectives in the world down to just five, you need to stay a little high level. The thing I love about these objectives, building awareness, influencing consideration, improving sales process, reposition the brand, or growing loyalty, is that they can be really easily distinguished, and that is through a simple question. Ask yourself this question, do consumers recall and recognize my brand? If they don't to the extent that you want, well, you have an issue with building awareness. You should make an objective of your company to build better awareness, and on and on. You can do this for each of those five objectives. Those five objectives then will serve as the focal point of any analysis that we're pulling together. We'll also want to focus on a single one, because any company in the world could probably look at these five and say, "Yeah, I would like to get better." Probably, very few companies in the world have any of these completely linked to the point that they think, "I don't need to even worry about that." But when we are putting together a data story and a narrative and trying to stay focused, we would want to choose just a singular one. Doesn't mean that the others aren't relevant, it just



means that we are focusing on a single. Now, because we have this goal, we still need to make sense of the data that we're collecting. Again, all this vast, rich data that we've pulled in. In that way, a framework, some organizing philosophy, will go a long way.

The Consumer Decision Journey

Change in consumer behavior brings change to marketing.

Then: The Funnel Metaphor

For years, marketers assumed that consumers started with a large number of potential brands in mind and methodically winnowed their choices until they'd decided which one to buy. After purchase, their relationship with the brand typically focused on the use of the product or service itself.



Now: The Consumer Decision Journey

New research shows that rather than systematically narrowing their choices, consumers add and subtract brands from a group under consideration during an extended evaluation phase. After purchase, they often enter into an open-ended relationship with the brand, sharing their experience with it online.



Source: McKinsey & Company, David C. Edelman and Marc Singer

When I am working in marketing data or towards a business objective and I'm looking at a consumer, the framework that I love is what we call the consumer decision journey. This comes out of some research from Mackenzie, which really sought to revise and revitalize the old marketing funnel. This funnel idea was that consumers start with all these companies and they whittle their way down to the one that they finally choose before they make a purchase. What we've learned now with the advent of digital and the way that consumers have access to so much information, there's a much more fluid process. In fact, looking at it from the consumer's perspective, that idea, the funnel is really rejected to more of a looped process that we call the consumer decision journey.



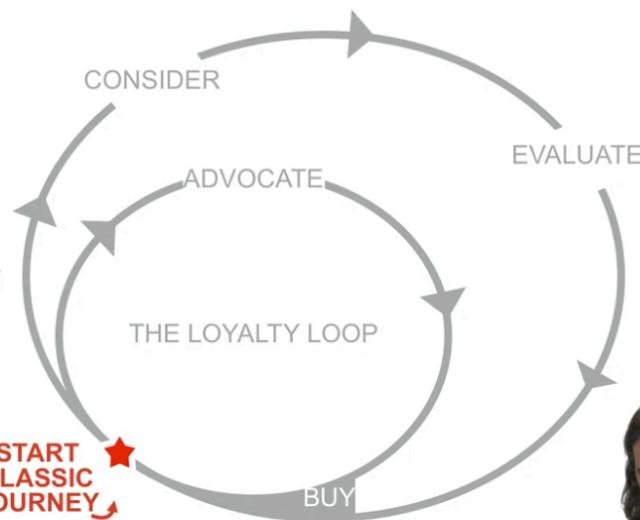
THE CUSTOMER DECISION JOURNEY

Adapting the marketer's POV to the consumer's perspective

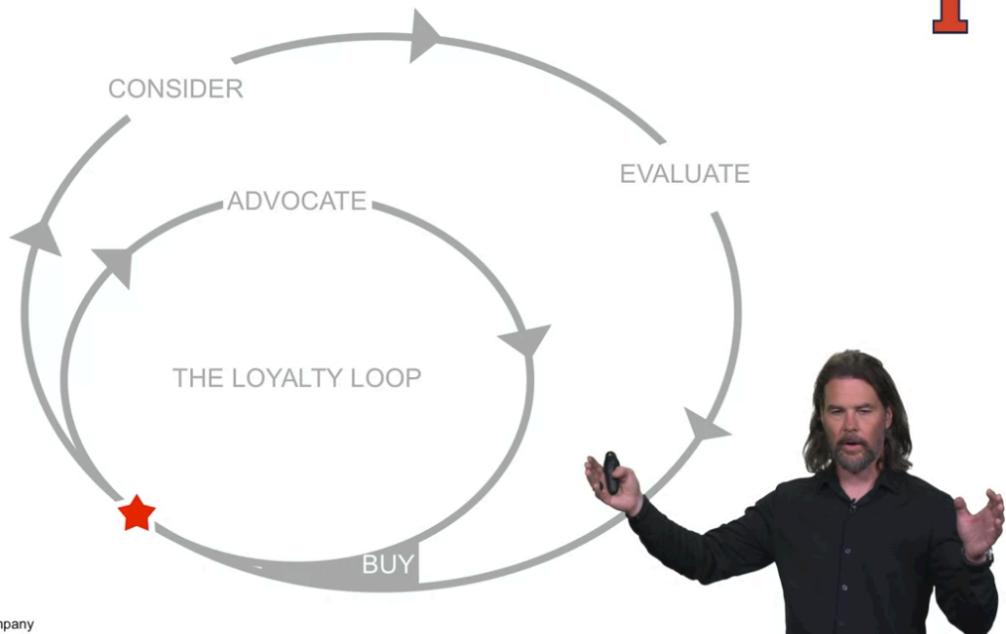
New research shows that rather than systematically narrowing their choices, consumers add and subtract brands from a group under consideration during an extended evaluation phase.

After purchase, they often enter into an open-ended relationship with the brand, sharing their experience with it online.

START CLASSIC JOURNEY



This consumer decision journey starts with some kind of trigger. There's something that says to consumer, "You need a product. You grow out of the shoes you have. You break a computer that you own. Your car breaks down." Whatever it is, something has happened that now has put you into market to buy a new product. You will enter through a number of phases. The first phase is the consider phase. This is where you are looking at what Mackenzie terms the initial consideration set. There are a number of brands that pop to mind immediately. If I'm looking for new sneakers, I think Adidas, I think Nike, I think maybe Puma, some others. Some of those come immediately to mind. It doesn't mean I'm going to buy any of those. It just means that they have the greatest awareness, and they are in my initial consideration set. I then as a consumer enter into what we would call the evaluation phase. In this case, I start evaluating brands against each other, and new brands frankly can come in, things that I hadn't thought of before. So, I am now working through an enormous set or a growing and contracting set of brands as I learn new information through the evaluation phase. I finally then do arrive at one and make a purchase. I am then in, what Mackenzie calls, the post-purchase experience. Do I like this? Did it live up to my expectations, the product that I purchased? If it has and fits the needs and becomes a product that I do indeed love, the objective of the brand is to get me as a consumer into, what's called, the loyalty loop. The next time that I have that need, I no longer go through the consideration or evaluation phases. I just go right to the source and buy that product again. I am now an advocate for that brand.



This framework works beautifully. It works in any purchase decision as large as a house or an automobile, down to a pack of gum. You're still going through these phases. You might do them more quickly with something that is less of an investment for you, but it does work very well to explain the way that consumers think about the purchase decision.



Now, the way this fits back with our goals, is that you can literally take those five questions that I introduced, each one aligned to a goal, and layer them into this consumer decision journey. If I indeed do not find that consumers recall or recognize my brand and I need to build awareness, I need to get to them during that consideration phase. If I'm finding that the products that I produce actually don't fit the needs of the consumers that I'm advertising to, well, I have to evaluate or I have to improve my position and their evaluation. I need to influence consideration there a little better on and on. You can see that each one then lines up and it points me to a direction of the consumer journey that I will focus on given whatever goal I find most relevant. These things together focus not only the data that we're collecting, but also the analysis that we're going to do and ultimately the story that we will tell as we focus in on a singular goal, that then ties to a very discrete and distinct point of the consumer decision journey.



Organizing Data Around a Goal Is Critical **I** for Effective Dataviz

Huge volume of available data means some organizing principle is needed to bring comprehension.

Frameworks help focus collected data with purpose around your goal.

Goals give meaning to collected data by setting context for the story you wish to tell.

McKinsey's CDJ is an exceptional framework for nearly any marketing-related analysis.

So, this practice is really crucial for effective data visualization and communicating with data, because we are surrounded with so much data. We need some sort of organizing framework to make sense of it, and we need some sort of focus just to keep us from going off track. The high volume of data that is around us means that we need these things in place. Those frameworks really can help us make sense of very complicated data stories and data collection processes. So, to categorize them and get them into a spot that we can ultimately tell a story from. The goal that we select will start to provide context for the story that we want to tell. As I said, the consumer decision journey from Mackenzie, I've always found to be an exceptional framework for anything that is marketing or really business-related. Anything that involves a consumer, it becomes my go-to framework to organize that objective and point me in the right direction for my analysis.



Lecture 2-4: Planning Your Data Story

Planning Your Data Story

I

Even the best designed dataviz fall flat if the broader narrative lacks structure. Organize yours through a few simple guides.



Planning your dataviz story. So, let's talk about how we develop that connective tissue between the information and the data we've collected and the objective that we're working against. This is our story.

McCandless Offers a Thorough Definition of Good Data Stories

I



0:22 / 9:31





The story we know is a vital part of creating a successful visualization, and does just that. It really does become that connective tissue between our data and our goal. Now, organization is of utmost importance here. We need in the story we tell to be organized so that we can communicate effectively to our audience, and lead them through a narrative efficiently and effectively.

“

Most people have a pretty good idea of what they intend to write about but no specific plan for what to say or how to phrase it.

”

—Barbara Minto

I

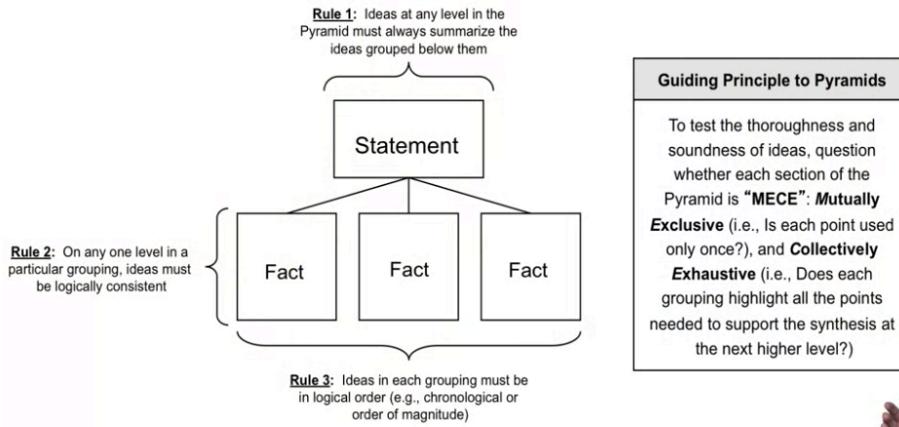


To that end, I want to introduce you to Barbara Minto. Barbara Minto was a Mackenzie consultant who developed a brilliant set of concepts that she called the Memento pyramid principal.



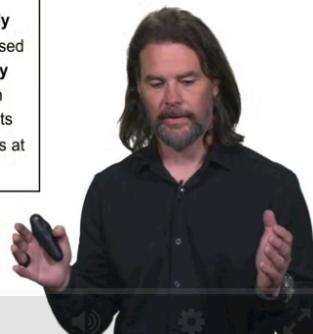
I

Minto's Rules Ensure Story Coherence and Completeness



Guiding Principle to Pyramids

To test the thoroughness and soundness of ideas, question whether each section of the Pyramid is "MECE": **Mutually Exclusive** (i.e., Is each point used only once?), and **Collectively Exhaustive** (i.e., Does each grouping highlight all the points needed to support the synthesis at the next higher level?)



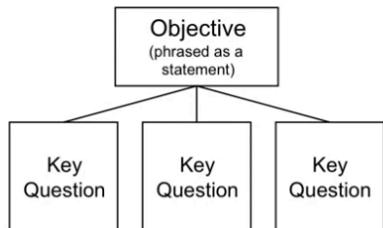
Source: Adapted from Barbara Minto "The Pyramid Principle"

This principal basically said that any statement that is made needs to be supported by facts, and the statement must be a perfect summarization of those facts. She had a couple of other rules. The first one is that, the most important that a statement we make must be a good summarization of the elements that we're using below it to support that statement. Those elements, the facts that we're using to support that statement must all be similar in some way, like a group, right? We want apples here. Not apples and oranges. Then the ordering of the facts as we display them should be something that is unsurprising to our audience. We either order them in time of occurrence from earliest to latest, from biggest to smallest, something that will make sense to them. Now, I use this term before but this is where the idea of MECE, mutually exclusive collectively exhaustive comes from. As Minto pushes us here, the facts that we are using on a single line should be MECE. They should be mutually exclusive from one another as well as collectively exhaustive. When we have that, we know we have a robust well-organized story. This is the way we should think about organizing the story we're telling, and not in a single pyramid, you can use as many as you want depending on how many chapters you have or points you have in your story. But this kind of organization can be very important. What it can also do is really lead us as we begin to plan for how we want to collect data and use data.



Applying Minto's Ideas to Data Pursuit Ensures a Sound Approach

I

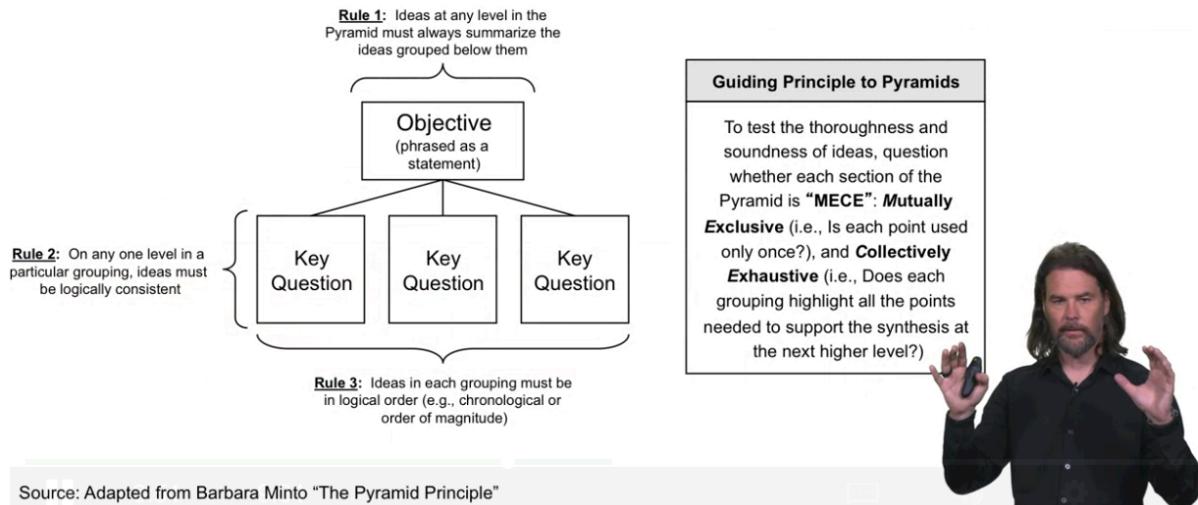


Source: Adapted from Barbara Minto "The Pyramid Principle"

In that way, if we don't have a statement yet or a set of facts, we can change that statement to an objective. That we state and we phrase that in terms of a statement, right? So, the objective that we talked about in the previous lesson becomes that objective that we're after. We don't have facts yet but we do have hypotheses, we do have key questions, so we can start to think about what those are and place those together. This will help us if we can articulate these questions correctly and ensure that this whole thing is MECE. It will help us develop a robust and solid story.

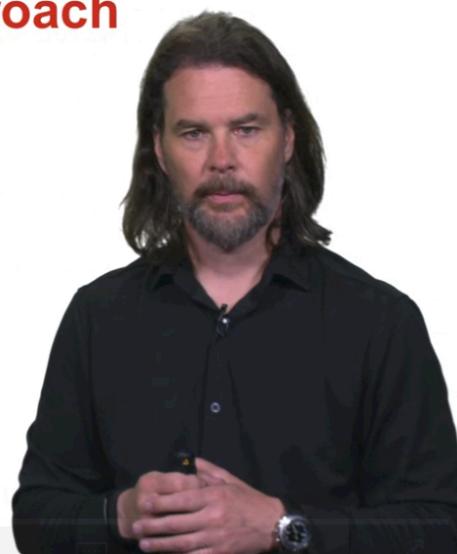
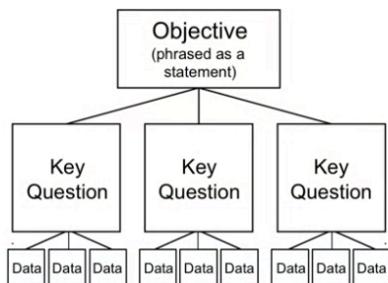


Applying Minto's Ideas to Data Pursuit Ensures a Sound Approach



Again, we will want to apply the same rules that we would to a finished statement and finished story to our planning process. This will point us in the right direction and move us toward the right story that we ultimately want to produce.

Applying Minto's Ideas to Data Pursuit Ensures a Sound Approach



Source: Adapted from Barbara Minto "The Pyramid Principle"



Each of those key questions, the hypotheses that we develop, we want to support with some data. In this way then, we can develop a plan to go out and collect information, collect data to prove or disprove a hypothesis, answer the key question that we've developed, and then ultimately roll up to some statement, and these things all fit together nicely in a well organized story. That's complicated, so, let me try to show you an example.



To do this, I want to introduce a company named Bellabeat. This is a real company but we're going to use them illustratively here. What Bellabeat produces is high-tech, very beautifully designed instruments primarily for women who are pregnant. What you're seeing here is the device called the shell, which allows a pregnant woman to hear her baby's heartbeat. It can also be turned around and send music to the child. They have a couple of other well-designed products as well.



LEAF



BALANCE



SHELL



One on the left here called the leaf, which is an activity tracker that looks more like a piece of jewelry, it's a beautiful piece. Then on the right is something that's called balance, which is a Bluetooth enabled scale. All of these products are well-designed and incorporate great technology. But, most of them are not known for sure. So, Bellebeat, whether you're talking about leaf, balance, or shell are facing an awareness problem. Consumers particularly women carrying children do not know of their existence where these products would probably fit many of the needs that they want, they just don't know about it.



Applying Minto's Ideas to Data Pursuit Ensures a Sound Approach

I

Source: Adapted from Barbara Minto "The Pyramid Principle"

So, we can construct based on that objective this design a Memento pyramid scheme. At the very top, we put that objective of building awareness, and we're doing this in the form of a statement, right? To increase sales which is the ultimate output and objective of Bellabeat. They first must grow awareness. So, there we have captured our goal, growing awareness in a statement. Now, underneath that, we start to look at some key questions that will help us determine the plan or the approach that Bellabeat should take to growing awareness. To ensure that we have a MECE set of questions, one of the things I like to do is to begin one question with the word what, the second with the word how, the third with the word why. If I can answer a what, how, and why question about the statement I've made, I know that I have a certainly a mutual exclusive set of questions that I'm after, but also a very robust and complete view of that statement. So, in this case, I will say what is the level of awareness for Bellabeat today? Certainly an important thing to know. I want to know how has the awareness for Bellabeat changed over time. I will want to know as why? Why is awareness critical to Bellabeat's product adoption. Just to prove out that, right? And then under each one, you can see I can articulate a data source that will help me answer that question. In this way, I have formed a visual outline for the communication journey that I'm about to set on using data that I know will answer a question, that I know will support an ultimate goal and objective. You can see how all of these things from story-wise would fit together very well.



Minto Offers an Excellent Way to Organize Data Stories

Good stories are built around statements supported by facts.

MECE collections of facts communicate thorough, sound ideas.

Minto's guidelines provide exceptional structure for any story.

Application of Minto's concept to data collection ensures your approach will be sound.

So, that Minto idea really does offer an excellent way for us to organize data stories. It ensures that the statements that we make are all supported by facts. It introduces this concept of mutual exclusivity and things being taken together collectively exhaustive, ensuring that we are communicating sound ideas. The guidelines provide excellent structure for any story, organized, clear, and concise. This will also ensure that the approach we take in collecting data and the analysis that we do will be sound, as this becomes the plan that we will begin with.

In This Module

Module 2: Working Fast and Thinking Slow

Key Concepts

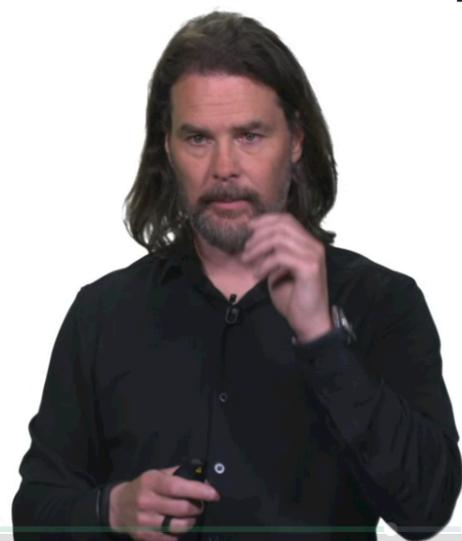
Understanding the growth of data

Evaluating methods of data access

Focusing your communication journey

Planning your data story

Case Study: Bellabeat





Okay, in this module, we looked at a number of things. We took a look at this explosive growth of data. Really seeing that there is no better time than now to understand how to communicate effectively with data. We evaluated methods of accessing data. How do we get our hands on this analysts the data? What are the right paths and avenues? What are the wrong ones? We looked at an approach to focusing our communication journey through the use of a goal and an objective. Then, we looked through the Minto pyramid principle. This idea of how we can start to plan our data story, and layout in visual form a good outline. We used again the case study for Bellebeat to really illustrate those concepts. We will continue on with Bellebeat as we move forward in this class.