



The Role of Data in Decision Making

Jennifer Lewis Priestley, Ph.D.
Associate Dean and Professor of Statistics
Executive Director, Analytics and Data Science Institute
Kennesaw State University

Data Science: (United States)

100

90

80

70

60

50

40

30

20

10

0

2004-01

2004-04

2004-07

2004-10

2005-01

2005-04

2005-07

2005-10

2006-01

2006-04

2006-07

2006-10

2007-01

2007-04

2007-07

2007-10

2008-01

2008-04

2008-07

2008-10

2009-01

2009-04

2009-07

2009-10

2010-01

2010-04

2010-07

2010-10

2011-01

2011-04

2011-07

2011-10

2012-01

2012-04

2012-07

2012-10

2013-01

2013-04

2013-07

2013-10

2014-01

2014-04

2014-07

2014-10

2015-01

2015-04

2015-07

2015-10

2016-01

2016-04

2016-07

2016-10

2017-01

2017-04

2017-07

2017-10

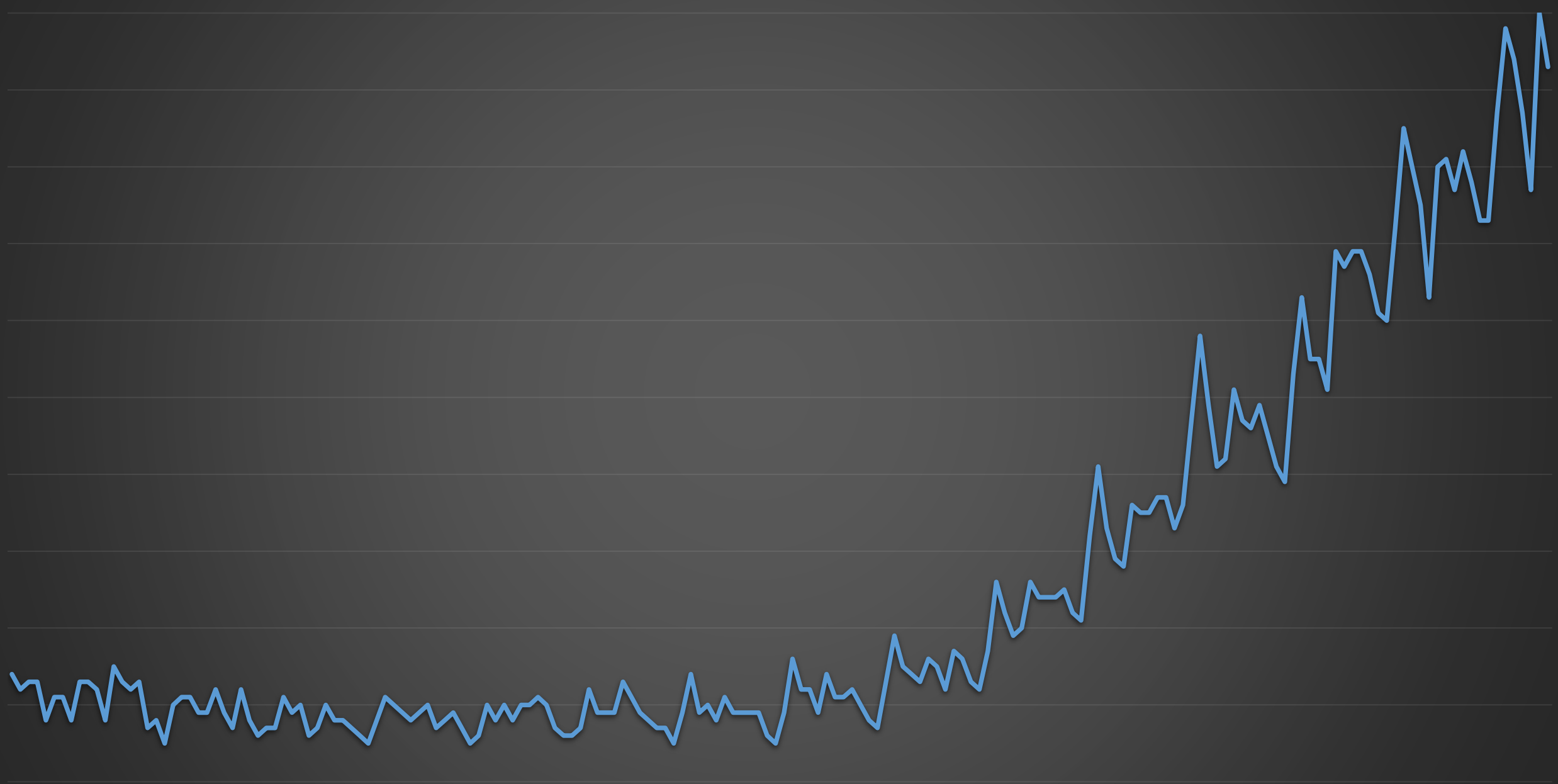
2018-01

2018-04

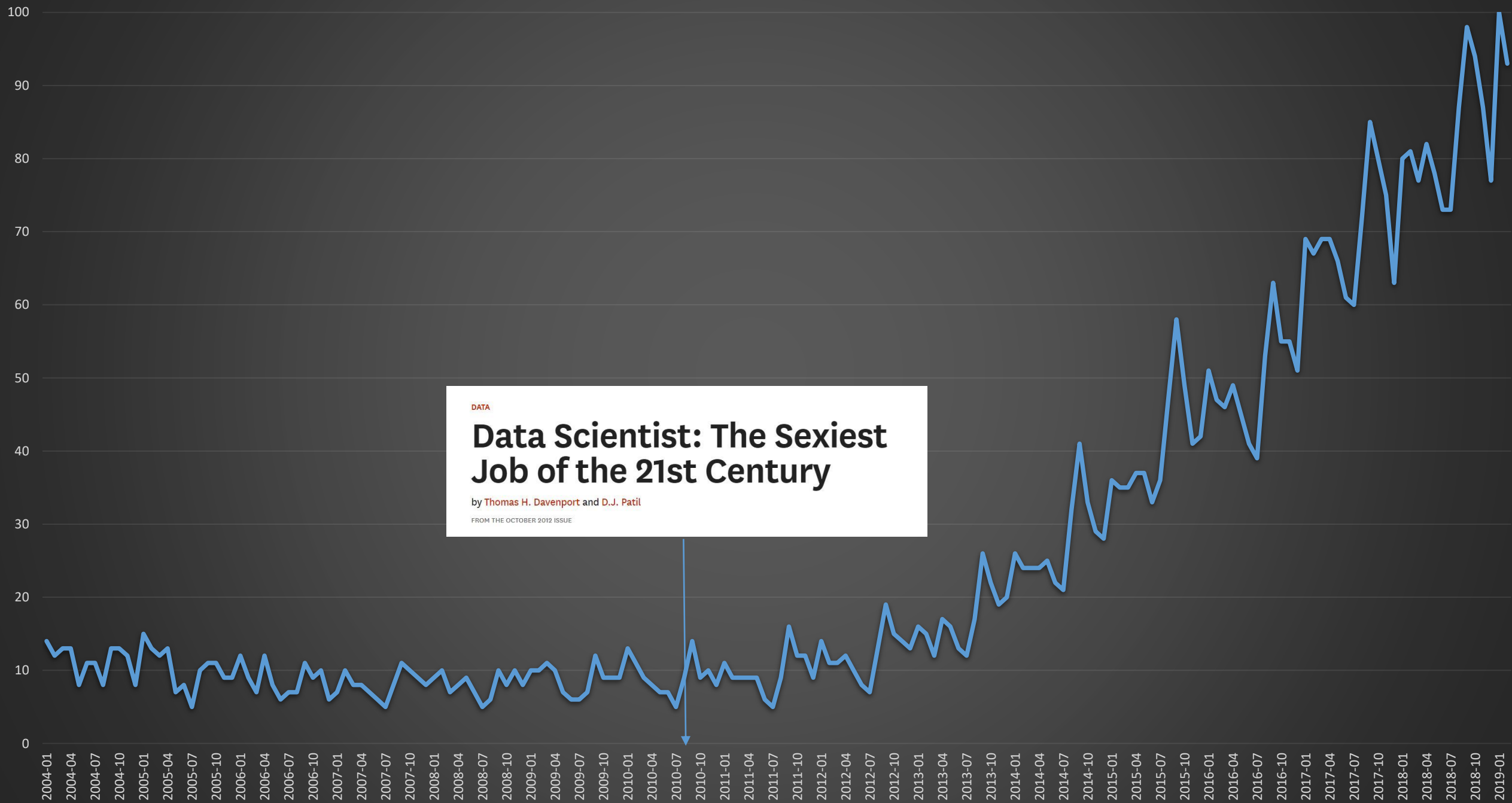
2018-07

2018-10

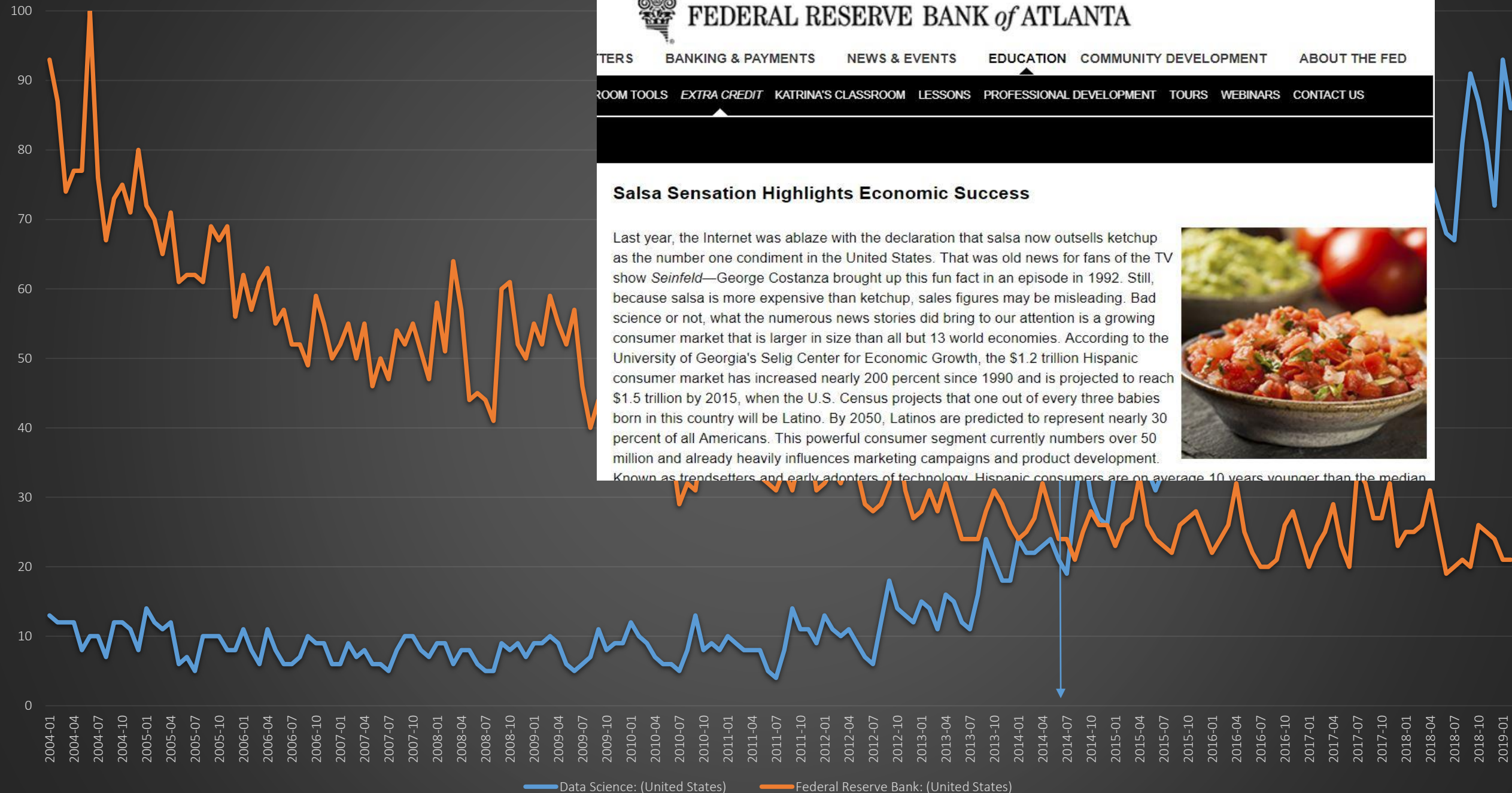
2019-01



Data Science: (United States)



Searches on "Data Science" vs "Federal Reserve Bank"



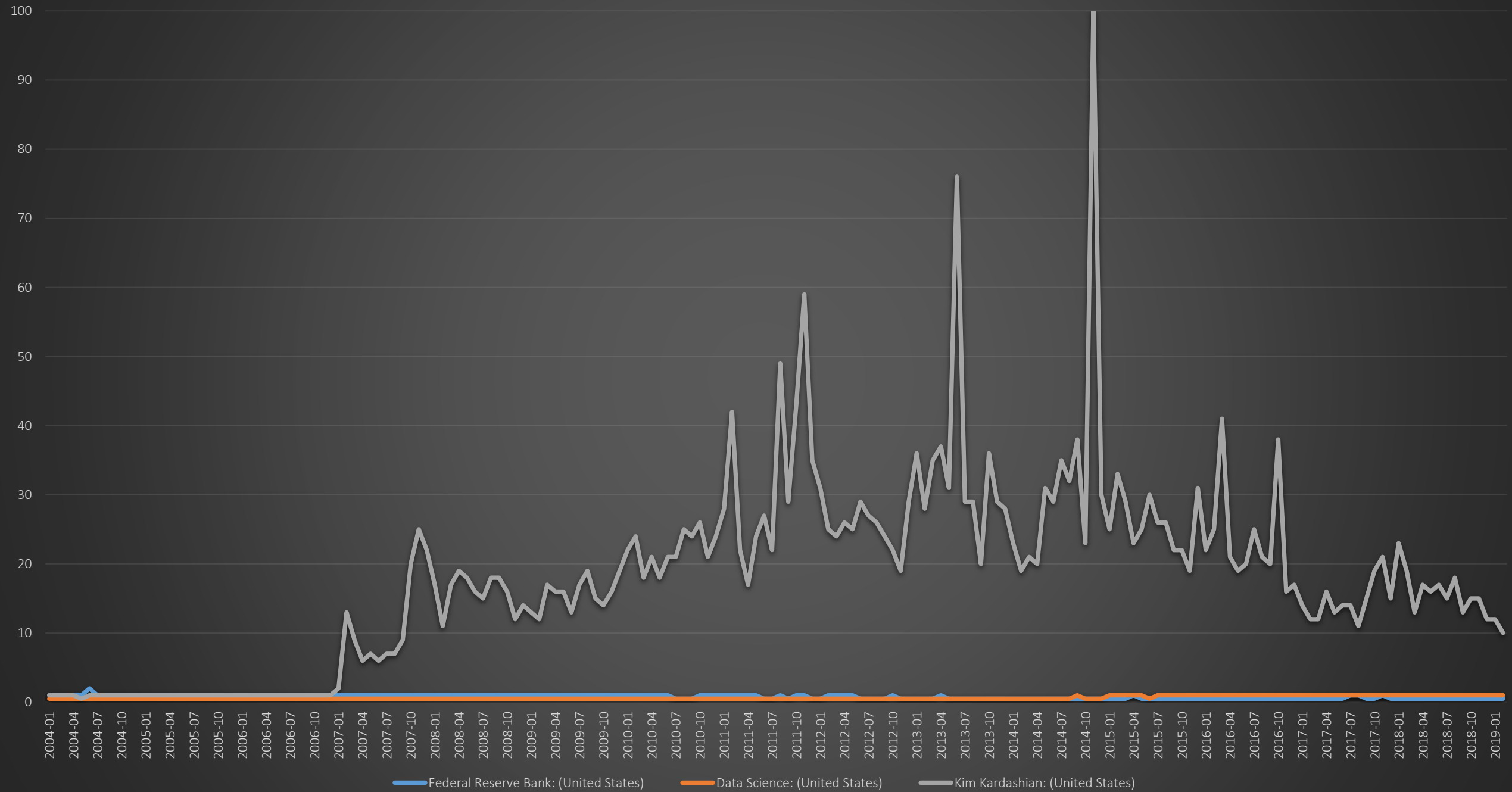
Salsa Sensation Highlights Economic Success

Last year, the Internet was ablaze with the declaration that salsa now outsells ketchup as the number one condiment in the United States. That was old news for fans of the TV show *Seinfeld*—George Costanza brought up this fun fact in an episode in 1992. Still, because salsa is more expensive than ketchup, sales figures may be misleading. Bad science or not, what the numerous news stories did bring to our attention is a growing consumer market that is larger in size than all but 13 world economies. According to the University of Georgia's Selig Center for Economic Growth, the \$1.2 trillion Hispanic consumer market has increased nearly 200 percent since 1990 and is projected to reach \$1.5 trillion by 2015, when the U.S. Census projects that one out of every three babies born in this country will be Latino. By 2050, Latinos are predicted to represent nearly 30 percent of all Americans. This powerful consumer segment currently numbers over 50 million and already heavily influences marketing campaigns and product development.

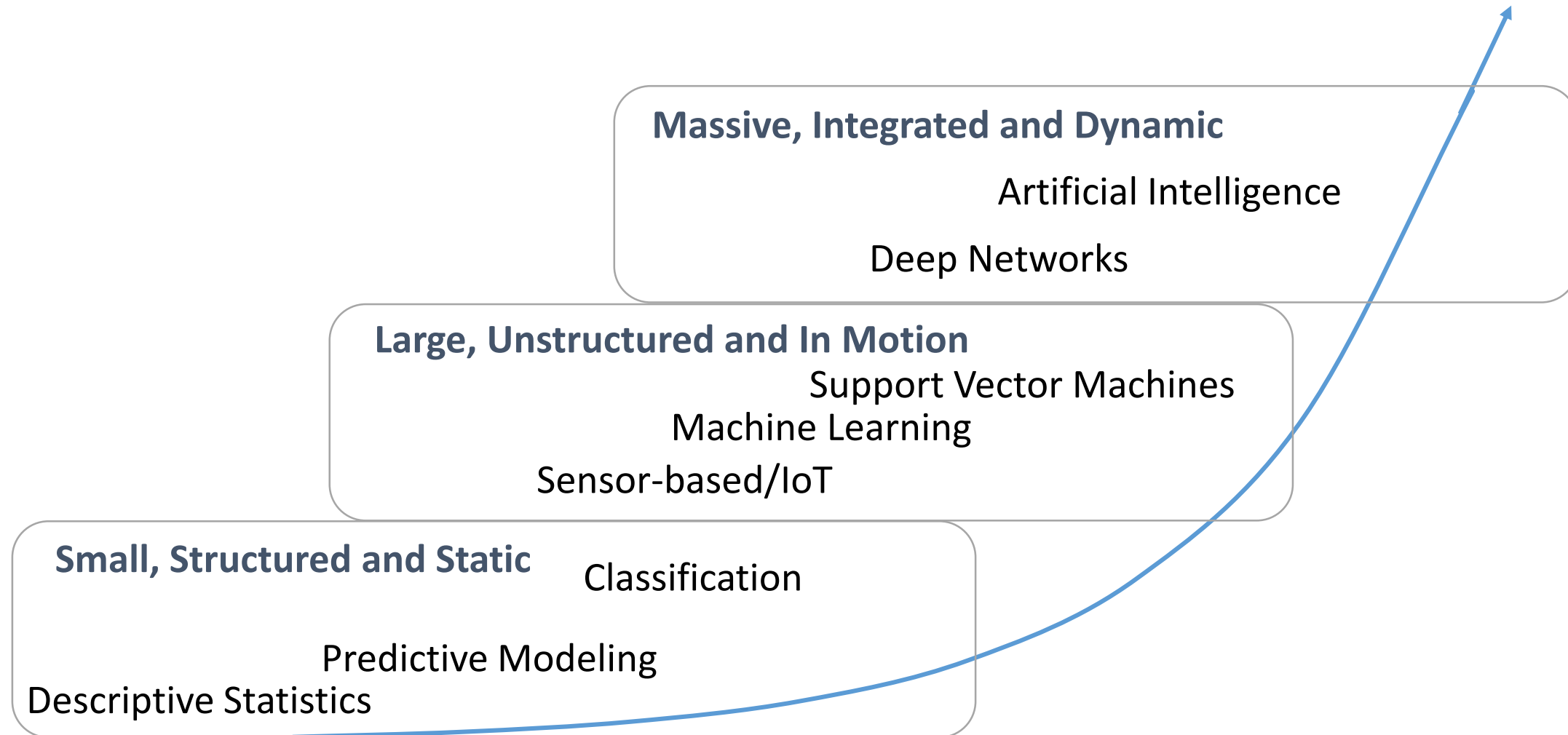


Known as trendsetters and early adopters of technology, Hispanic consumers are on average 10 years younger than the median

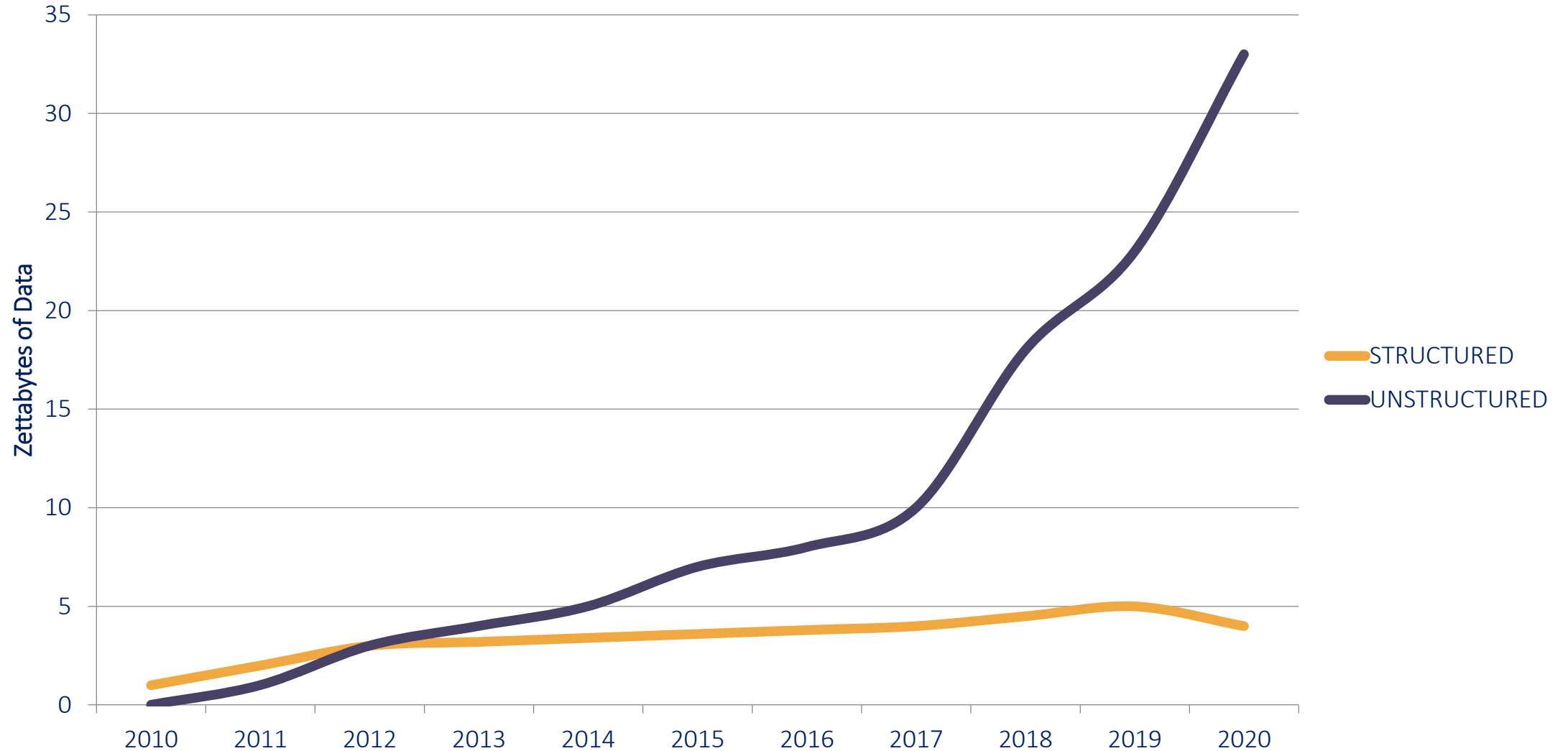
Searches on “Kim Kardashian” vs “Federal Reserve Bank” vs “Data Science”



How the Data Ecosystem is Evolving...



Structured Versus Unstructured Data Generated by Year



Everyone is chasing the same talent...

Finance
Retail
Political Science
Consulting
Healthcare
Economics
Manufacturing



Who ARE these People?

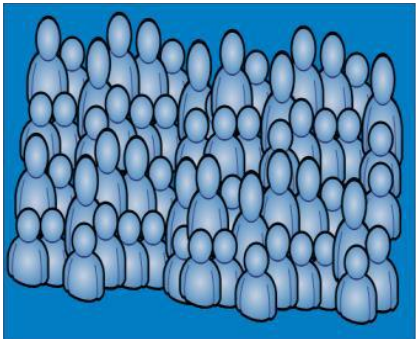


“Person who is better at statistics than any software engineer and better at software engineering than any statistician” – Josh Wills, Director of Data Engineering at Slack

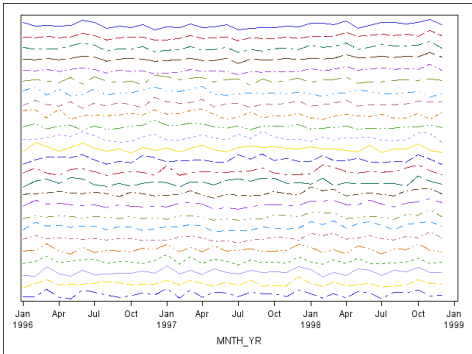


“Person who is better at explaining the business implications of the results than any scientist and better at science than any business school graduate”
Jennifer Priestley, Ph.D. Data Nerd

Lets be more specific about the data...



Cross Sectional



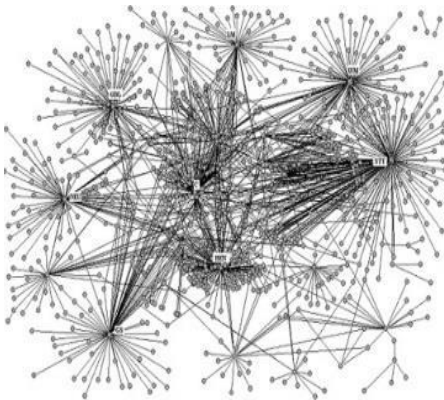
Streaming



Image/Video



Geo Spatial



Networks

A grid of data points with movie posters (Mickey Blue Eyes, The Matrix, etc.) above it, representing links or relationships between entities.

2			4	5	2.94*
5		4			1
		5		2	2.48*
	1		5		4
		4			2
4	5		1		1.12*

Link



Text

Dark Data/Alternative: Data used in non-traditional or unintended ways.



Could Yelp indicate whether your business deserves a loan?



TIMOTHY A. CLARY/AFP/GETTY IMAGES/FILE

New York-based JP Morgan Chase is incorporating data compiled online to offer quick small-business loans of up to \$250,000.





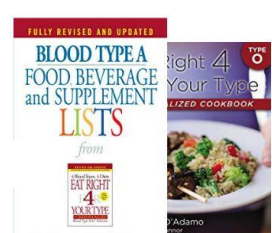
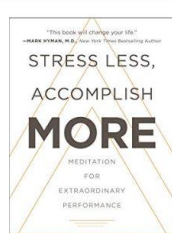
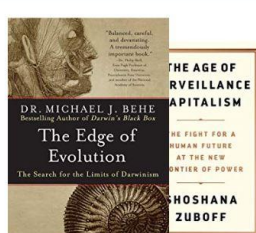
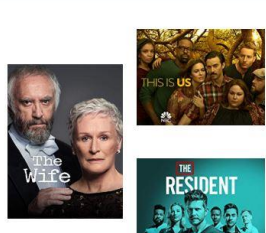






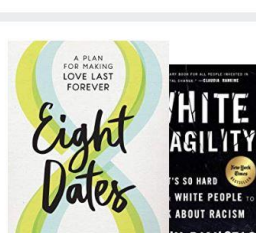
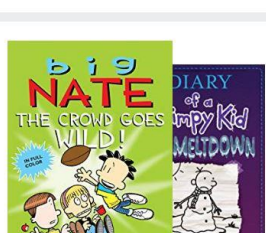

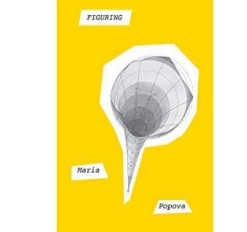

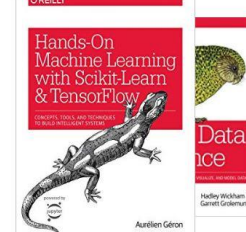




Social Media and Credit Risk



WORKING PAPER SERIES

On the Rise of the FinTechs—Credit Scoring using Digital Footprints

What does Amazon know about me?

 <p>Electronics for Kids 7 ITEMS</p>	 <p>Home Appliances & Accessories 40 ITEMS</p>	 <p>Girls' Shoes 21 ITEMS</p>	 <p>Cooking & Baking 11 ITEMS</p>	 <p>Health, Fitness & Dieting Books 99 ITEMS</p>	 <p>Religion & Spirituality Books 98 ITEMS</p>	 <p>Engineering Books 62 ITEMS</p>	 <p>Movies & TV 100 ITEMS</p>
 <p>Electrical Equipment 11 ITEMS</p>	 <p>Industrial Electrical 13 ITEMS</p>	 <p>Amazon Video - TV 100 ITEMS</p>	 <p>Boys' Shoes 29 ITEMS</p>				
 <p>Office & School Supplies 100 ITEMS</p>	 <p>Politics & Social Science Books 99 ITEMS</p>	 <p>Tableware 100 ITEMS</p>	 <p>Computer & Technology Books 100 ITEMS</p>	 <p>Recently Added TV Shows in Prime Video 75 ITEMS</p>	 <p>Recently Added Movies in Prime Video 99 ITEMS</p>	 <p>Kindle Unlimited 100 ITEMS</p>	 <p>Prime Pantry 42 ITEMS</p>

How does Kabbage work?



How should data be valued?

MISSION: AHEAD

Companies use your data to make money. California thinks you should get paid



By [Heather Kelly, CNN Business](#)

Updated 1:10 PM ET, Wed February 13, 2019

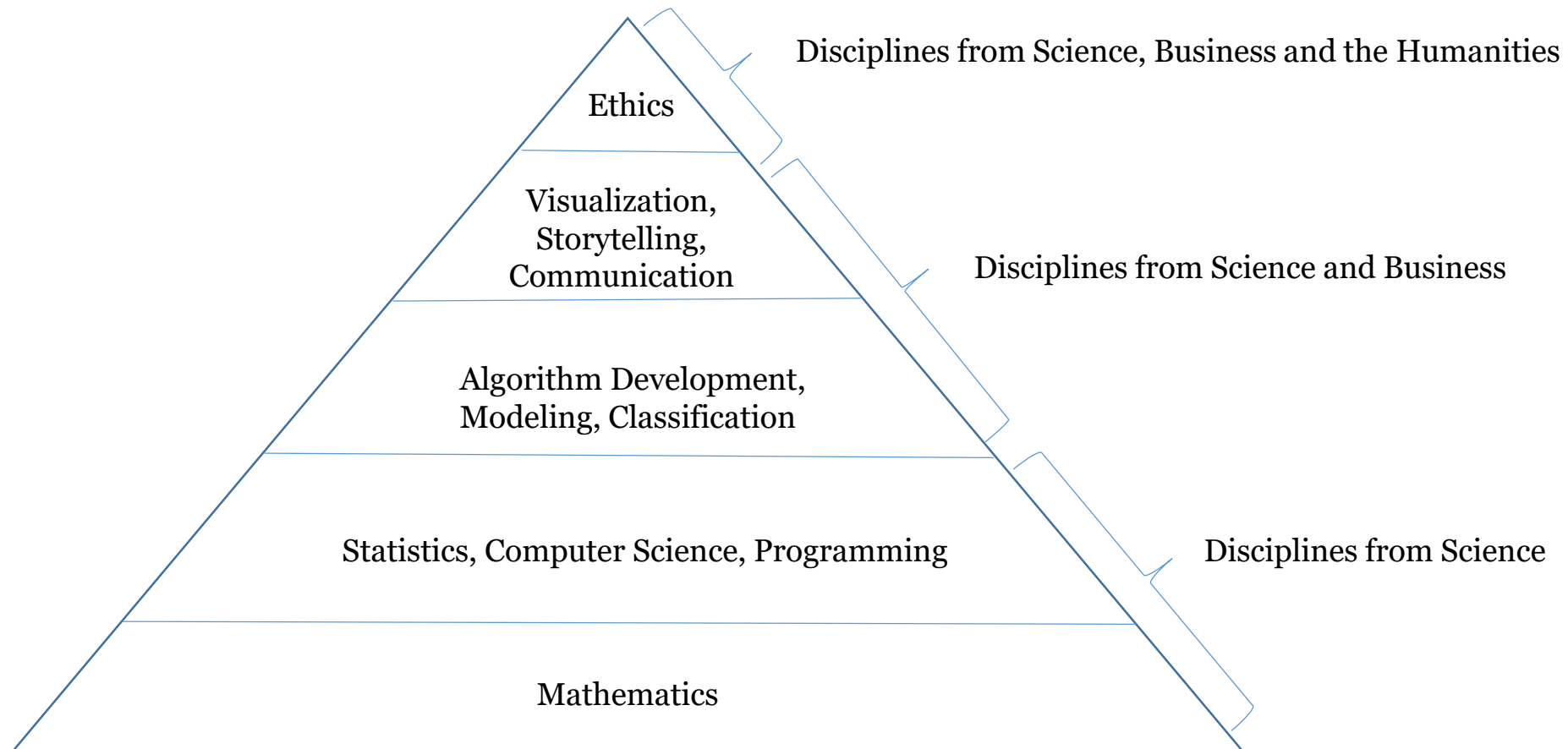


Ethics and Privacy

- ✓ Maslow's Hierarchy of Data Science
- ✓ The Paradox of Thrift
- ✓ An FDA for Data
- ✓ GDPR in the US

Ethics and Privacy – Maslow’s Hierarchy of Data Science

Maslow’s Hierarchy of Data Science



Ethics and Privacy – Paradox of Thrift

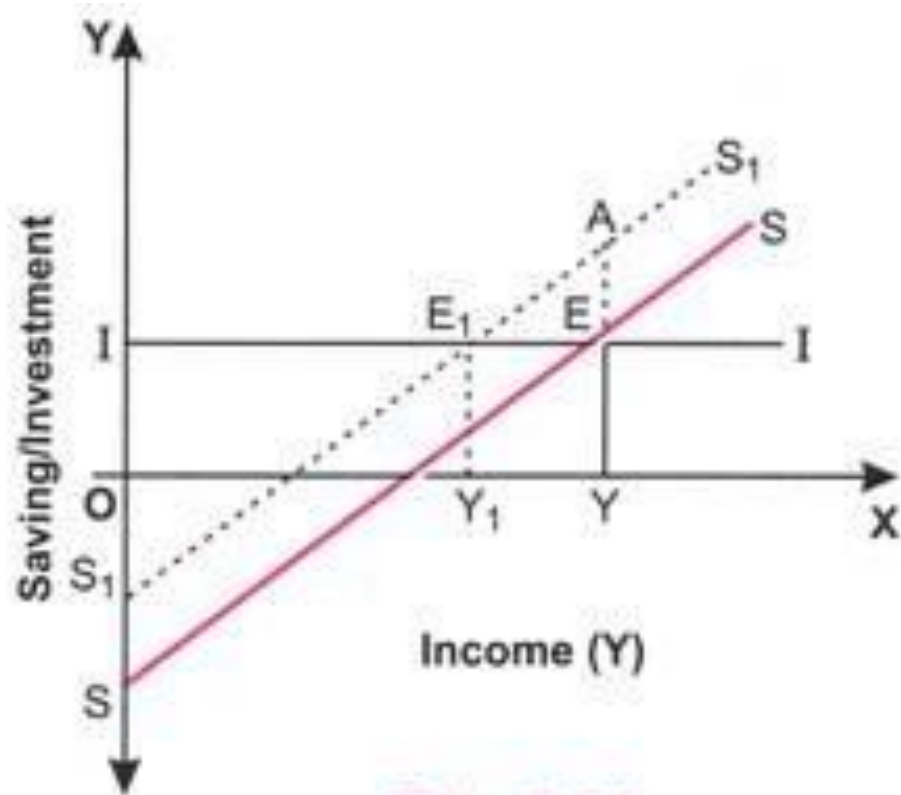


Fig. 8.14

From Floridi and Taddeo – “...overemphasizing the protection of individual rights in the wrong contexts may lead to regulations that are too rigid and thus in turn can cripple the chances to harness the social value of data science”.

Ethics and Privacy – FDA for Data

Mission of the FDA:

The Food and Drug Administration is responsible for protecting the public health by ensuring the safety, efficacy, and security of human and veterinary drugs, biological products, and medical devices; and by ensuring the safety of our nation's food supply, cosmetics, and products that emit radiation.



Ethics and Privacy – GDPR

The background of the slide is a photograph of the European Parliament building in Brussels. The building is a large, modern, circular structure with a glass facade, reflecting the sky and surrounding environment. It is situated on a grassy area with a river in the foreground. To the right of the building, there is a row of flags on tall poles. The sky is overcast.

The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years.

The regulation will fundamentally reshape the way in which data is handled across every sector, from healthcare to banking and beyond.

Implications for Universities



- ✓ Undergraduate/Graduate Curricula
- ✓ Research/Publication/Scholarship
- ✓ Ethics/Policy
- ✓ Engagement with the Private and Public Sectors



KENNESAW STATE
UNIVERSITY

Jennifer Lewis Priestley, Ph.D.

jpriestl@kennesaw.edu

datascience.kennesaw.edu

EU General Data Protection Regulation (GDPR) Scope

Data may not be processed unless there is at least one lawful basis to do so:

- ✓ The data subject has given consent to the processing of personal data for one or more specific purposes.
- ✓ Processing is necessary for the performance of a contract to which the data subject is party or to take steps at the request of the data subject prior to entering into a contract.
- ✓ Processing is necessary for compliance with a legal obligation to which the controller is subject.
- ✓ Processing is necessary to protect the vital interests of the data subject or of another natural person.
- ✓ Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller.

EU General Data Protection Regulation (GDPR) Key Changes

- ✓ **Increased Territorial Scope (extra-territorial applicability)** Arguably the biggest change to the regulatory landscape of data privacy comes with the extended jurisdiction of the GDPR, as it applies to all companies processing the personal data of data subjects residing in the Union, regardless of the company's location.
- ✓ **Penalties** Under GDPR organizations in breach of GDPR can be fined up to 4% of annual global turnover or €20 Million (whichever is greater). It is important to note that these rules apply to both controllers and processors -- meaning 'clouds' will not be exempt from GDPR enforcement.
- ✓ **Consent** The conditions for consent have been strengthened, and companies will no longer be able to use long illegible terms and conditions full of legalese, as the request for consent must be given in an intelligible and easily accessible form, with the purpose for data processing attached to that consent. Consent must be clear and distinguishable from other matters and provided in an intelligible and easily accessible form, using clear and plain language. It must be as easy to withdraw consent as it is to give it.
- ✓ **Breach Notification** Under the GDPR, breach notification will become mandatory in all member states where a data breach is likely to “result in a risk for the rights and freedoms of individuals”. This must be done within 72 hours of first having become aware of the breach. Data processors will also be required to notify their customers, the controllers, “without undue delay” after first becoming aware of a data breach.
- ✓ **Right to Access** Part of the expanded rights of data subjects outlined by the GDPR is the right for data subjects to obtain from the data controller confirmation as to whether or not personal data concerning them is being processed, where and for what purpose. Further, the controller shall provide a copy of the personal data, free of charge, in an electronic format. This change is a dramatic shift to data transparency and empowerment of data subjects.
- ✓ **Right to be Forgotten** Also known as Data Erasure, the right to be forgotten entitles the data subject to have the data controller erase his/her personal data, cease further dissemination of the data, and potentially have third parties halt processing of the data.
- ✓ **Data Portability** GDPR introduces data portability - the right for a data subject to receive the personal data concerning them, which they have previously provided in a '*commonly use and machine readable format*' and have the right to transmit that data to another controller.
- ✓ **Privacy by Design** Privacy by design as a concept has existed for years now, but it is only just becoming part of a legal requirement with the GDPR. At its core, privacy by design calls for the inclusion of data protection from the onset of the designing of systems, rather than an addition. More specifically - '*The controller shall..implement appropriate technical and organisational measures..in an effective way.. in order to meet the requirements of this Regulation and protect the rights of data subjects*'.
- ✓ **Data Protection Officers** Currently, controllers are required to notify their data processing activities with local DPAs, which, for multinationals, can be a bureaucratic nightmare with most Member States having different notification requirements. Under GDPR it will not be necessary to submit notifications / registrations to each local DPA of data processing activities, nor will it be a requirement to notify / obtain approval for transfers based on the Model Contract Clauses (MCCs). Instead, there will be internal record keeping requirements, as further explained below, and DPO appointment will be mandatory only for those controllers and processors whose core activities consist of processing operations which require regular and systematic monitoring of data subjects on a large scale or of special categories of data or data relating to criminal convictions and offences.