Aston Business School

Accredited by Association of MBAs | EFMD EQUIS ACCREDITED | AACSB ACCREDITED

Big Data Performance Management

Ali Emrouznejad

Professor and Chair in Business Analytics
Aston Business School
United Kingdom

BIG Data

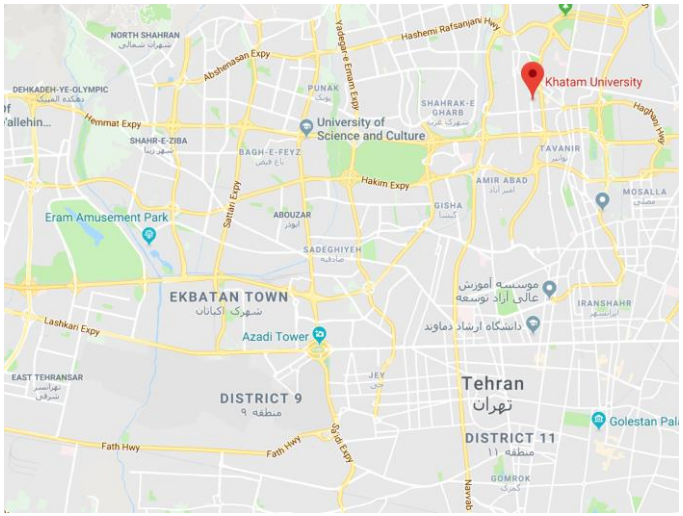
Performance Management

Efficiency & Productivity

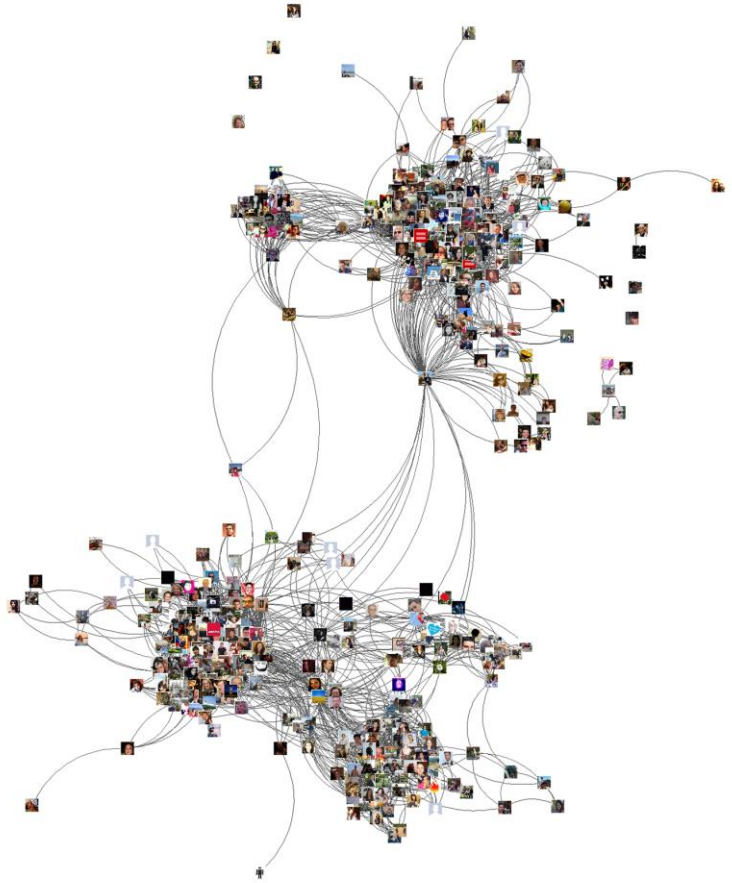
BIG Data

Why big data?

Who collect data?



Who collect data?



http://mbostock.github.io/d3/talk/20111116/force-collapsible.html

Social Network Analysis (SNA)

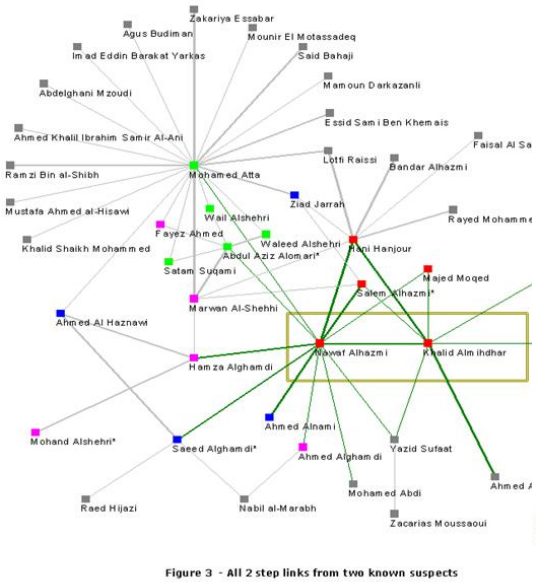
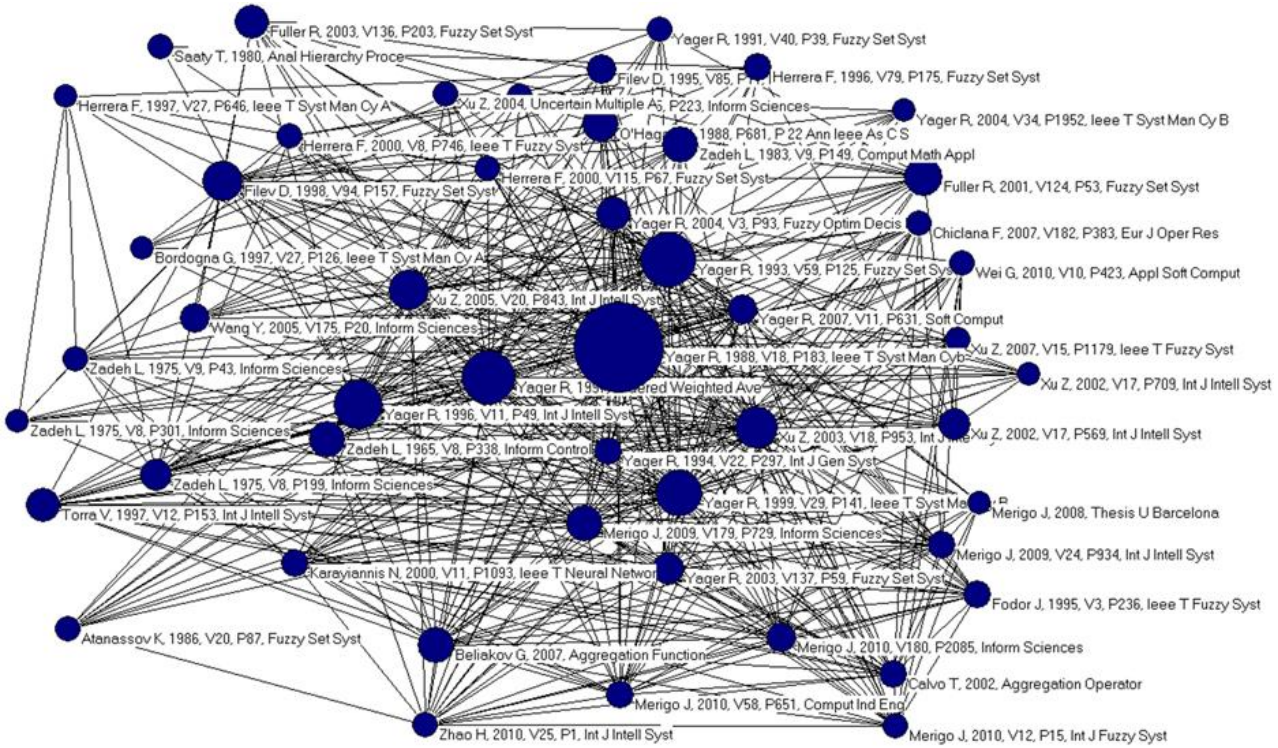
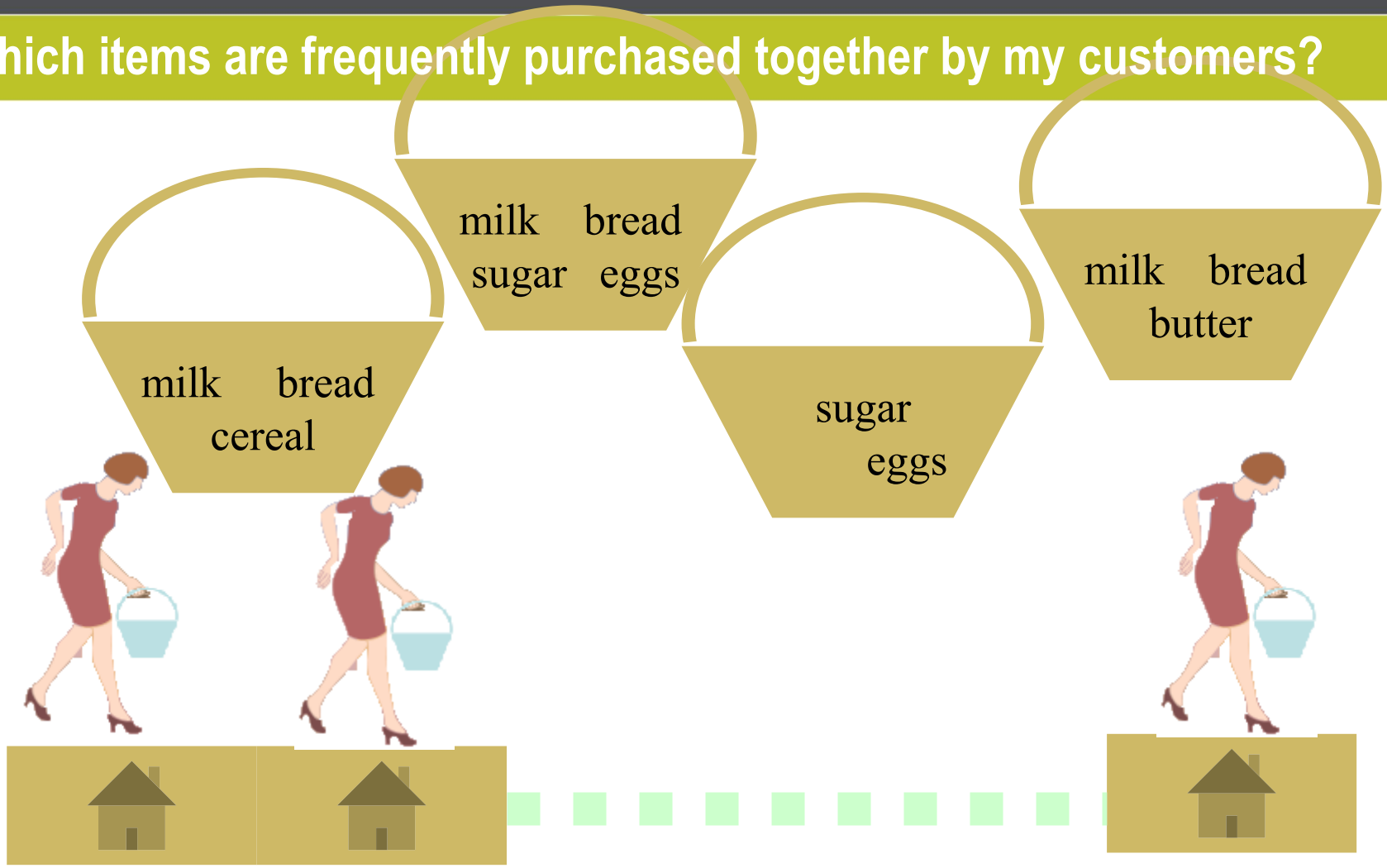


Figure 3 - All 2 step links from two known suspects



Marra M., **A. Emrouznejad**. W. Ho and J.S. Edwards (2015), The value of indirect ties in citation networks: SNA analysis with OWA operator weights, *Information Sciences*, **314** (2015) 135

Which items are frequently purchased together by my customers?



$$1000 * 2000 * 10 * 360 \approx 3{,}600{,}0000{,}000$$

Branches

Customer

Items

Supermarket deals: how the price promises match up

Which items are frequently purchased together by my customers?

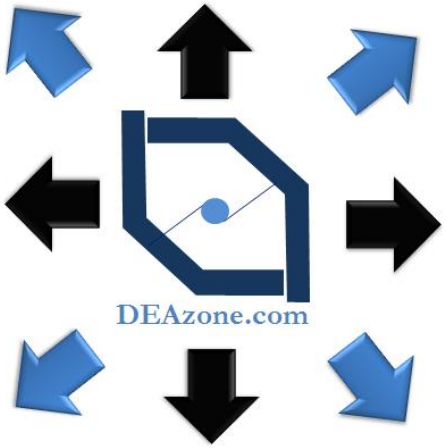
amazon web services™

Numbers online each minute

10000 * 20 * 60 * 24 * 365 ≈ 105,120,000,000

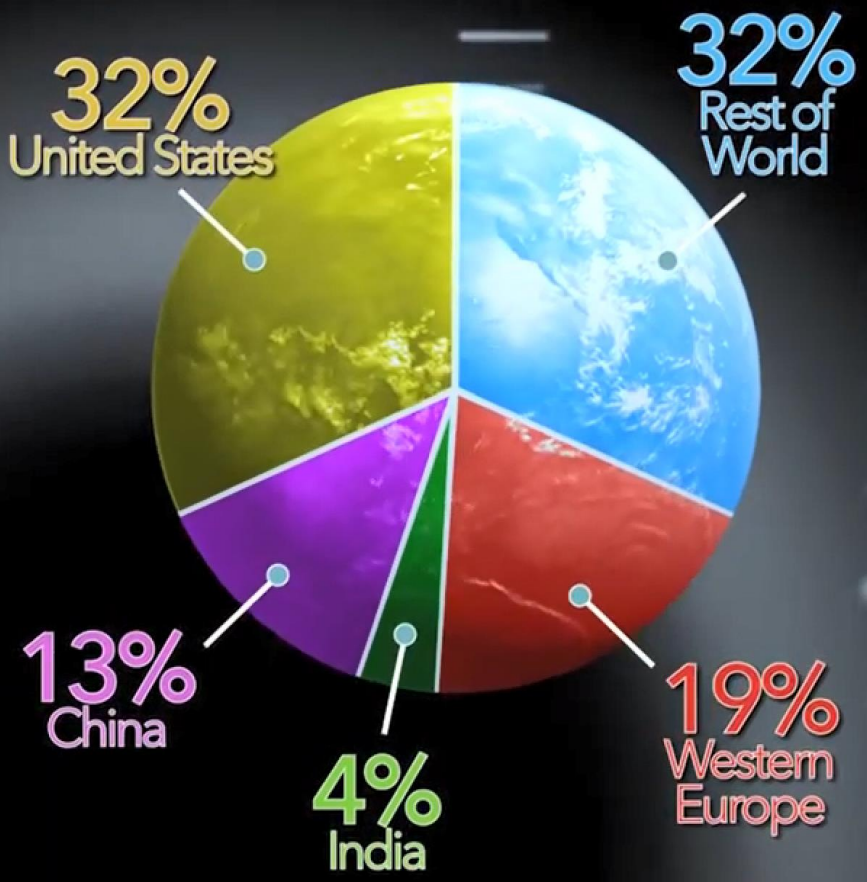
20 hits/ searches / links per person

Who collect data?



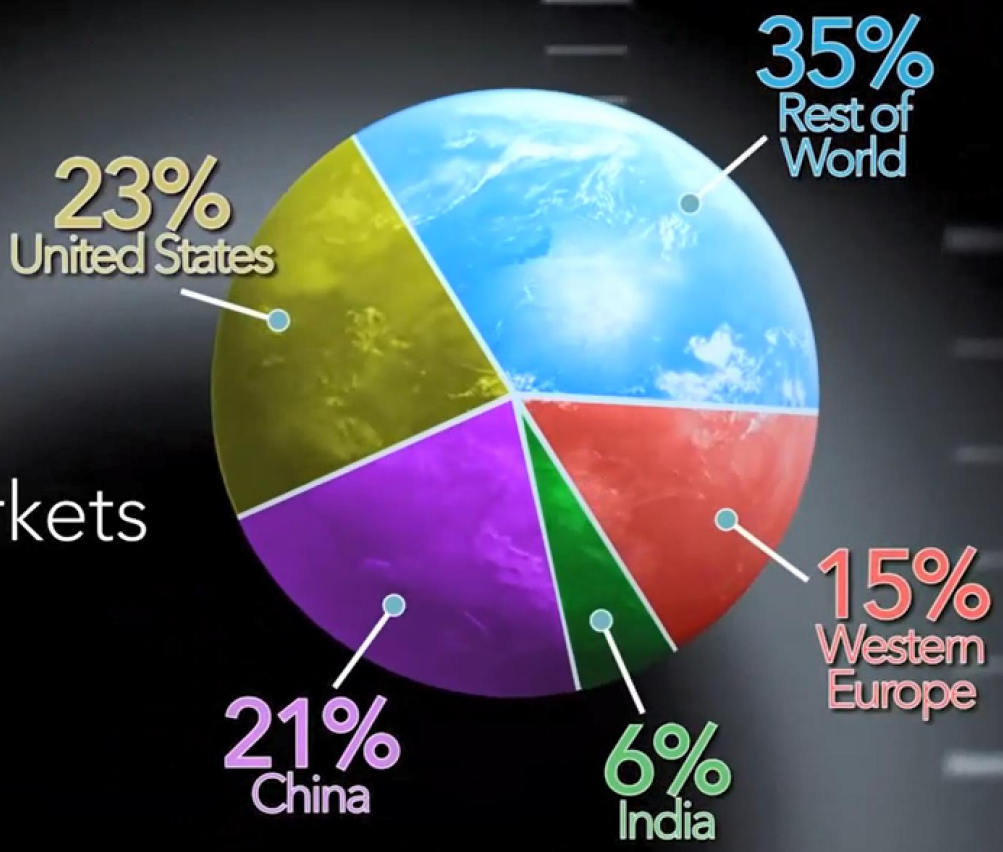
Prior to 2012 the US was the largest single contributor to global data

Where in 2012



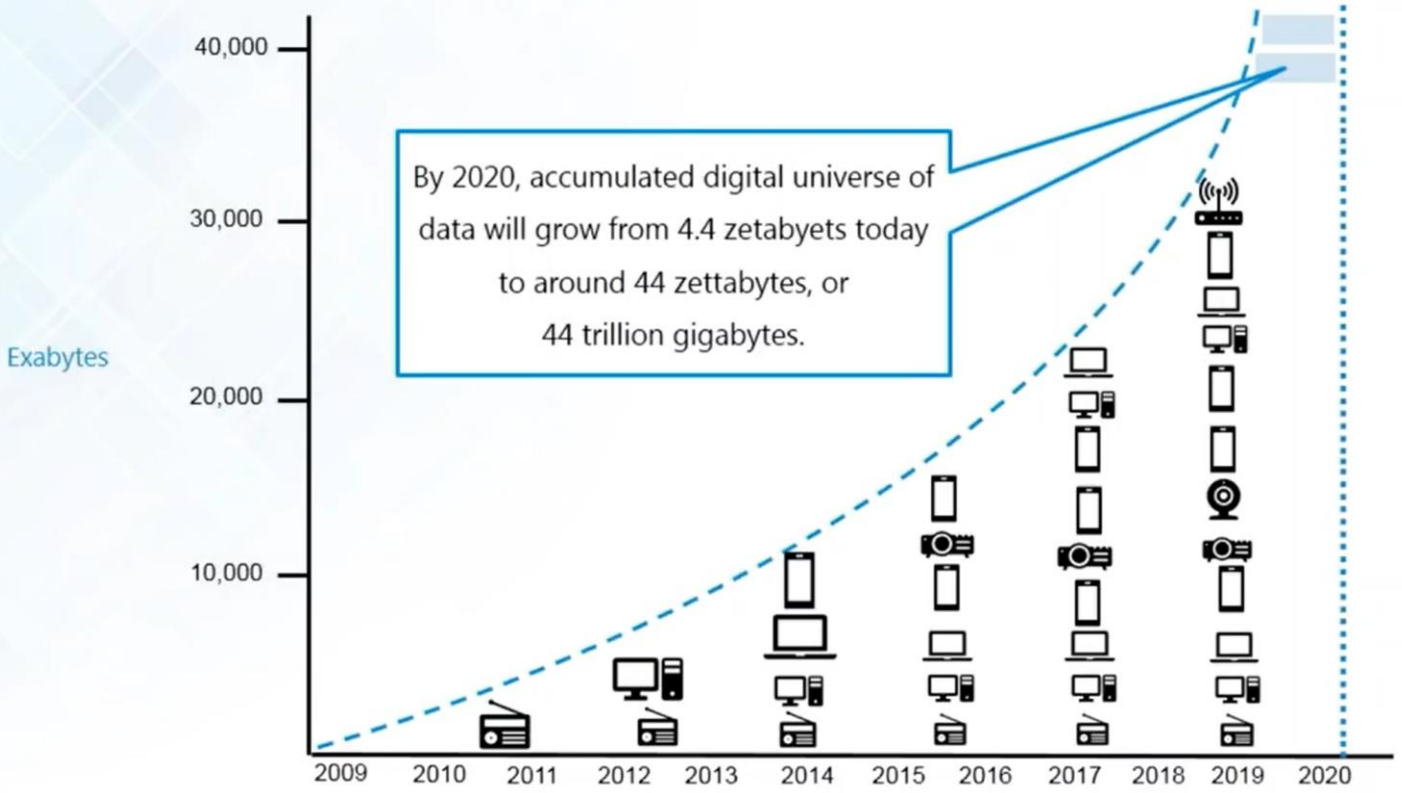
Where in 2020

The emerging markets are showing the largest increases in data growth

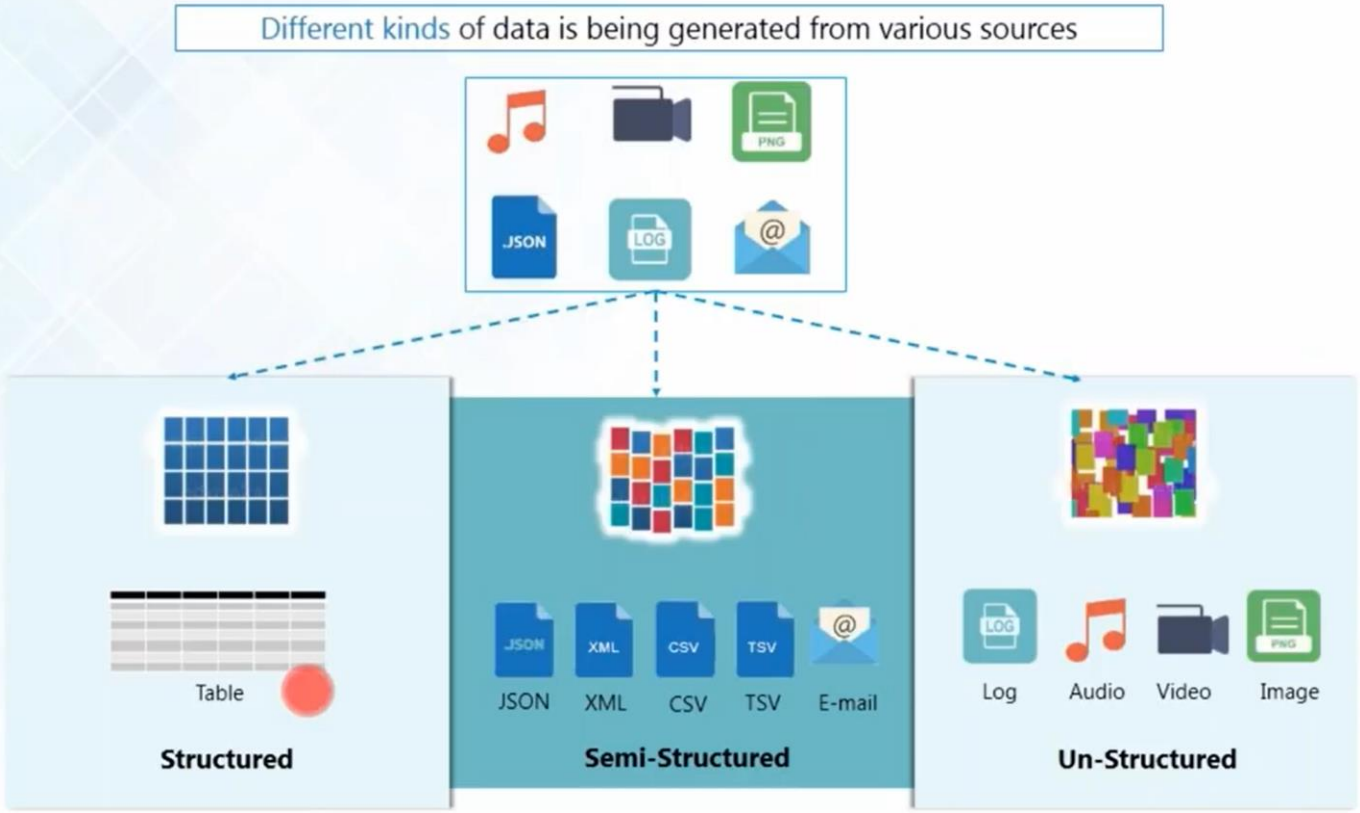


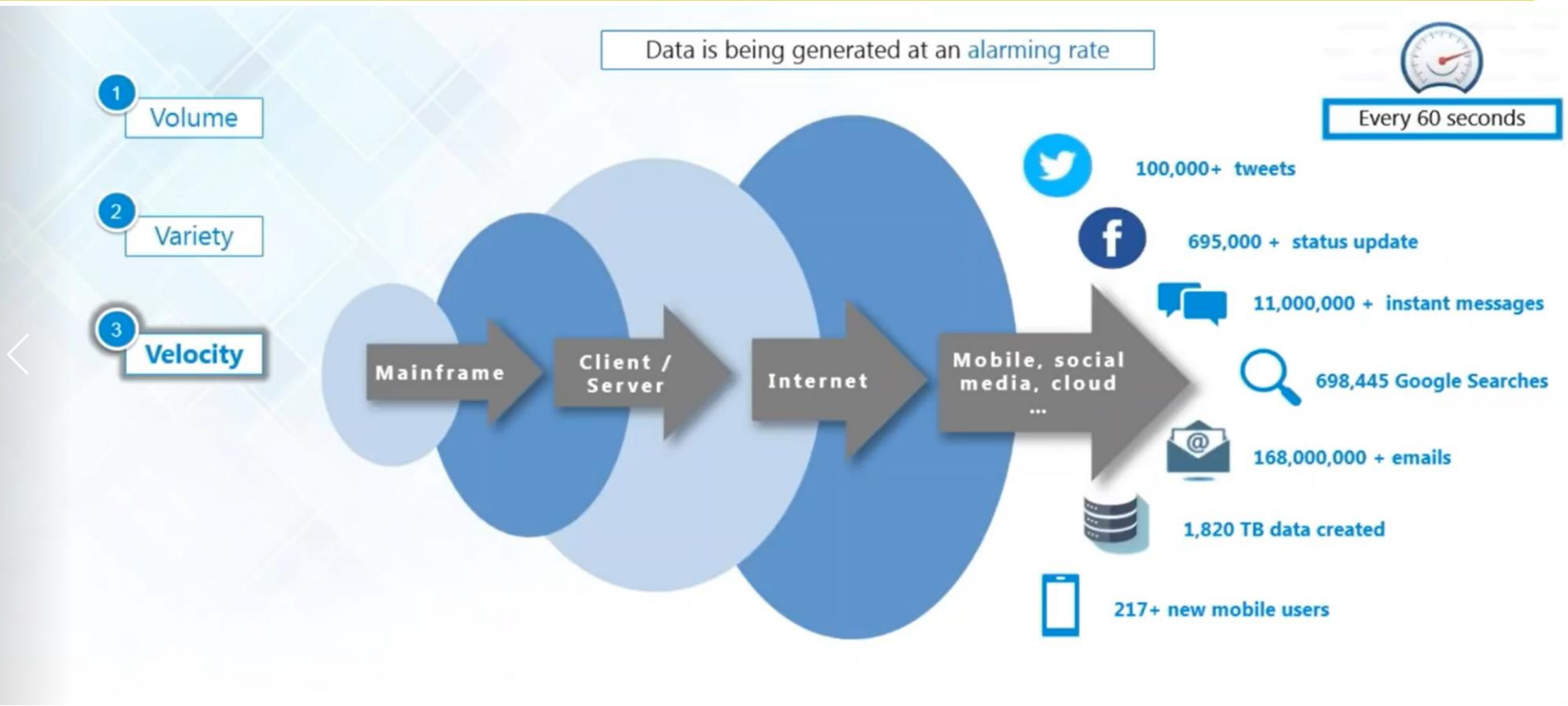

IN 60 SECONDS...
1 NEW DEFINITION IS ADDED ON urban
1,600+ READS ON Scribd.
13,000+ HOURS MUSIC STREAMING ON PANDORA
12,000+ NEW ADS POSTED ON craigslist
370,000+ MINUTES VOICE CALLS ON skype
98,000+ TWEETS
20,000+ NEW POSTS ON tumblr.
THE LARGEST SOCIAL READING PUBLISHING COMPANY!!
New Craigslist Ads
320+ NEW twitter ACCOUNTS
100+ NEW Linked in ACCOUNTS
13,000+ iPhone APPLICATIONS DOWNLOADED
1 associatedcontent NEW ARTICLE IS PUBLISHED
THE WORLD'S LARGEST COMMUNITY CREATED CONTENT!!
QUESTIONS ASKED ON THE INTERNET...
100+ Answers.com
40+ YAHOO! ANSWERS
6,600+ NEW PICTURES ARE UPLOADED ON flickr
600+ NEW VIDEOS
50+ WORDPRESS DOWNLOADS
25+ HOURS TOTAL DURATION
70+ DOMAINS REGISTERED
695,000+ facebook STATUS UPDATES
125+ PLUGIN DOWNLOADS
60+ NEW BLOGS
168 MILLION EMAILS ARE SENT
694,445 SEARCH QUERIES
1,700+ Firefox DOWNLOADS
1,500+ BLOG POSTS
79,364 WALL POSTS
510,040 COMMENTS
Google
Google Search
GO-Globe.com

1 **Volume**



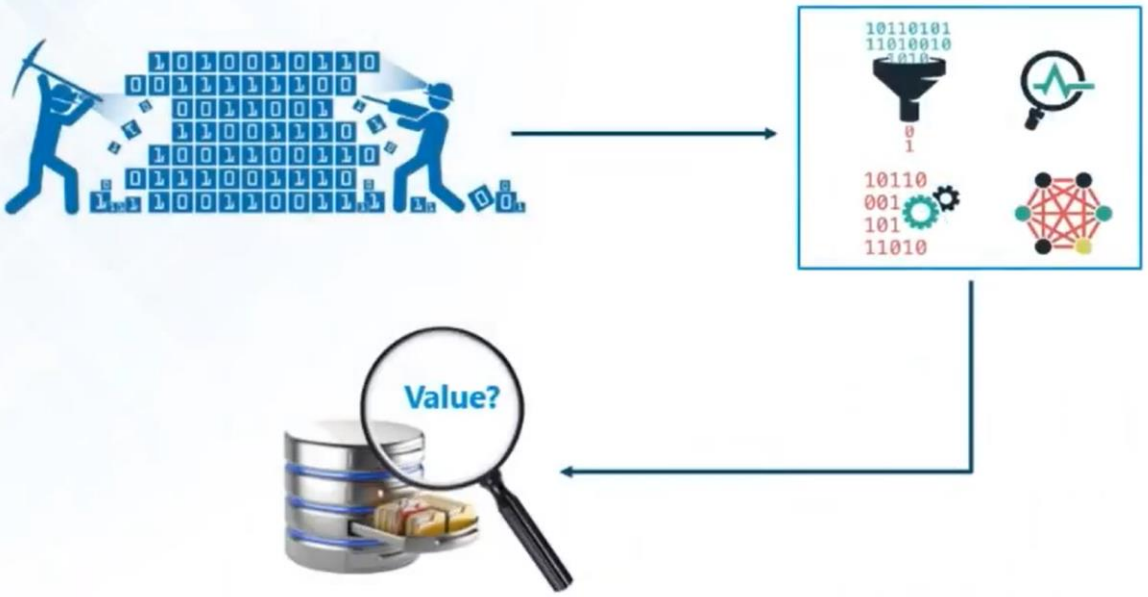
1. Volume
2. Variety




Data is being generated at an alarming rate
1 Volume
2 Variety
3 Velocity
Mainframe
Client / Server
Internet
Mobile, social media, cloud ...
Every 60 seconds
100,000+ tweets
695,000 + status update
11,000,000 + instant messages
698,445 Google Searches
168,000,000 + emails
1,820 TB data created
217+ new mobile users

1. Volume
2. Variety
3. Velocity
4. **Value**

Mechanism to bring the correct meaning out of the data

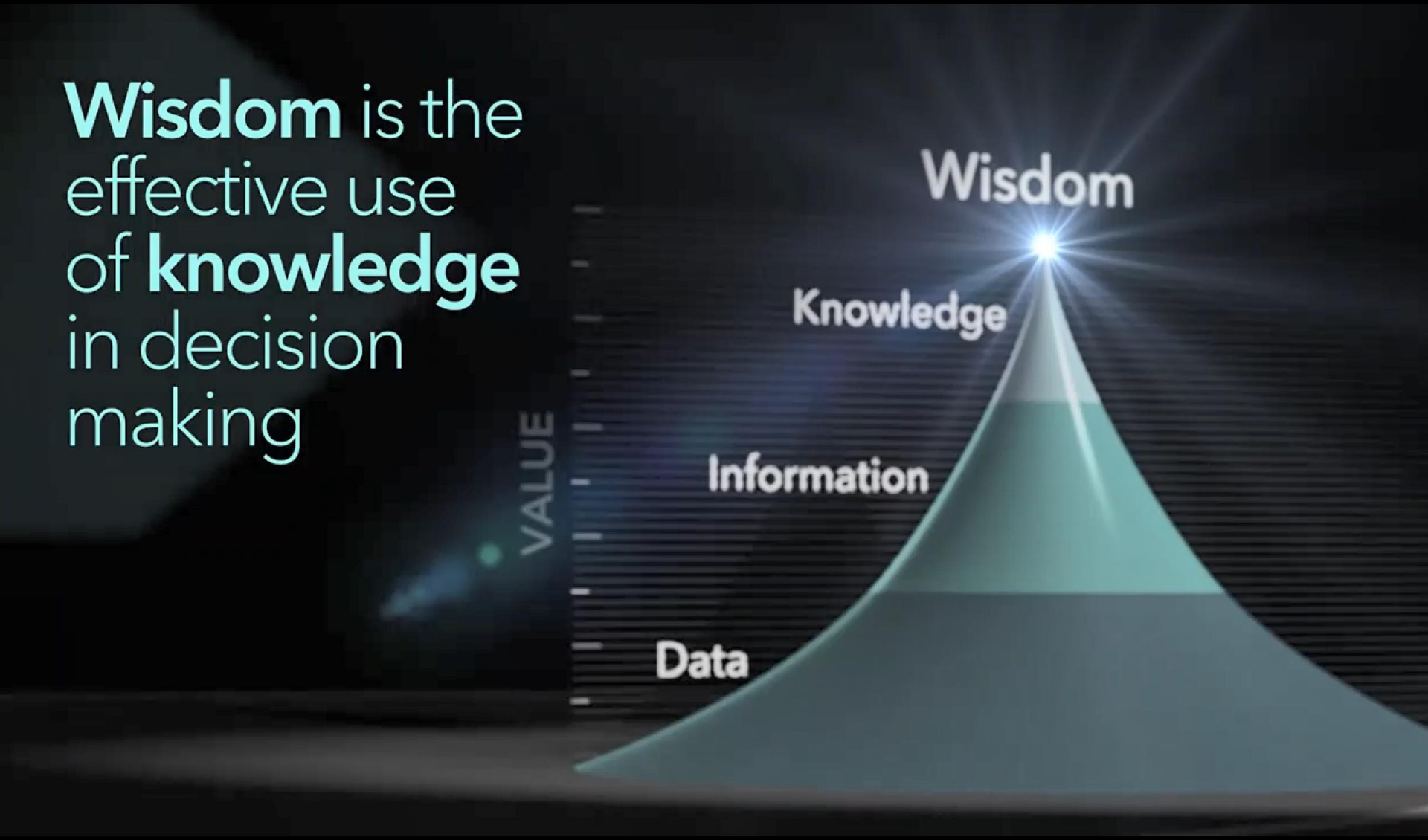


1. Volume
2. Variety
3. Velocity
4. Value
5. **Veracity**

Min	Max	Mean	SD
4.3	?	5.84	0.83
2.0	4.4	3.05	50000000
15000	7.9	1.20	0.43
0.1	2.5	?	0.76

Uncertainty and inconsistencies in the data

Wisdom is the effective use of **knowledge** in decision making



Smart Cities

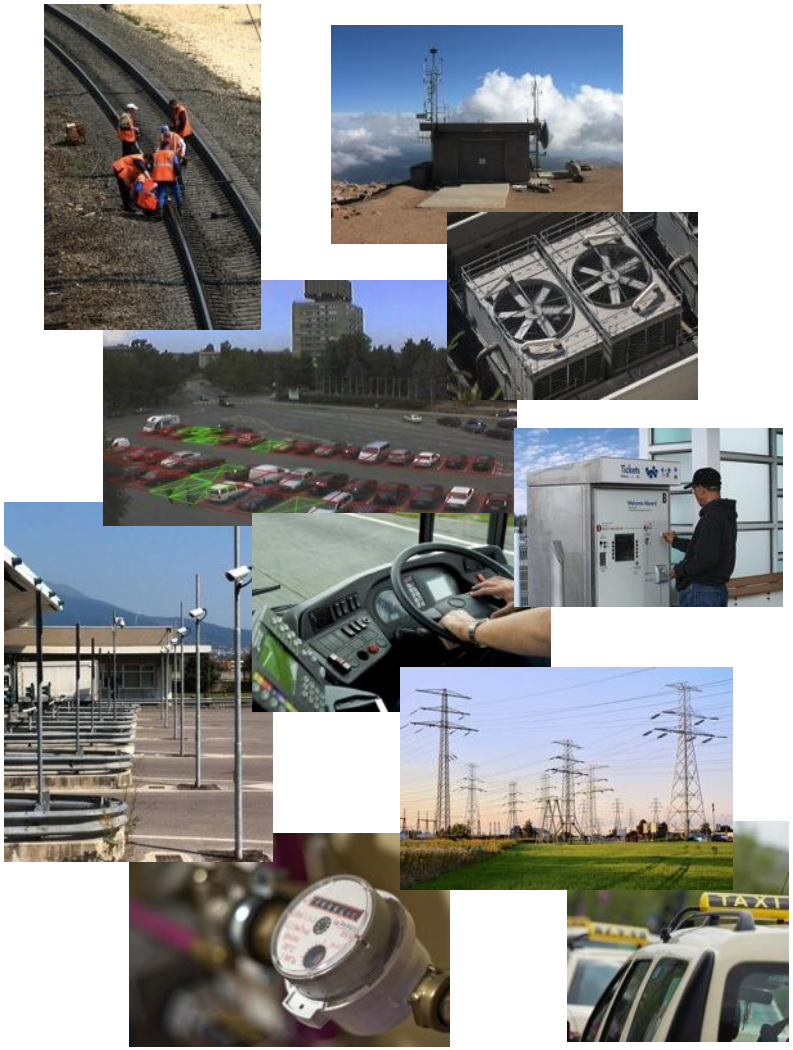

10:34:55:20
10:34:55:20
10:34:55:20
10:34:55:20
10:34:55:20
10:34:55:20
10:34:55:20
10:34:55:20
10:34:55:20

What's so smart about Smart Cities?

- A Smart City provides effective integration of physical, digital and human systems in the built environment to deliver a sustainable, prosperous and inclusive future for its citizens.
- How are "smarts" applied in cities
- To improve emergency response and resource management.
  - Centralized command and control system
- To enhance citizen access to the city information to inform individual and collective decision making

Smart city: A Combination of many applications

- Smart Transportation & Public Transport
- Smart Ticketing
  - Signage
  - Geo-Services
  - Communication Gateways
  - …
- Public Safety & Security
  - Surveillance & Security
  - Emergency Services
  - Public Infrastructure
  - …
- Smart Well-being
  - Healthcare
  - Elderly living
- Smart Energy / Smart Grid
- Smart Building
- Smart Water Management
- Smart Retail
- …



Smart City Enterprise Components

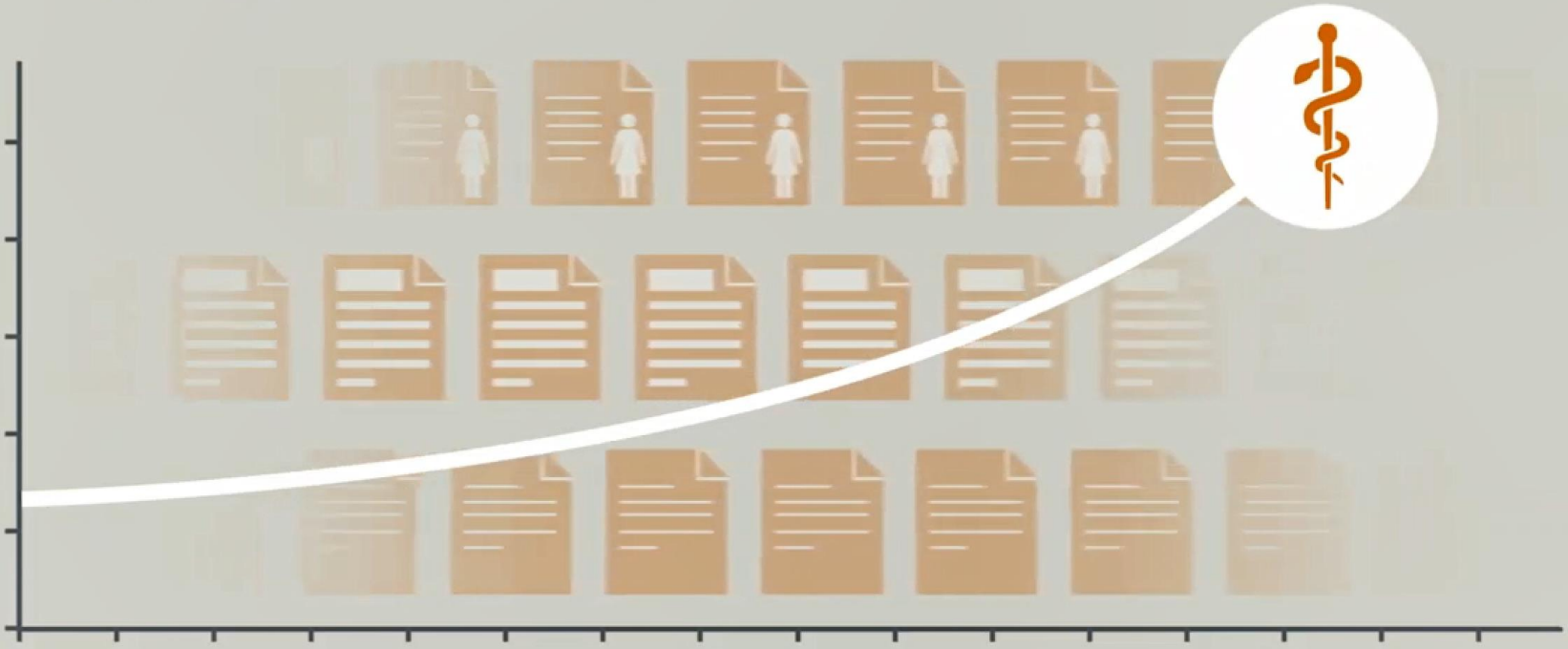
Application Layer

Health	Intelligent buildings	Intelligent transportation	Open data
Public safety and security	Environmental Protection	Urban planning	
Emergency Services	Utilities	Sanitation	Education

Cloud hosted resources

Security System

Tapping Efficiencies



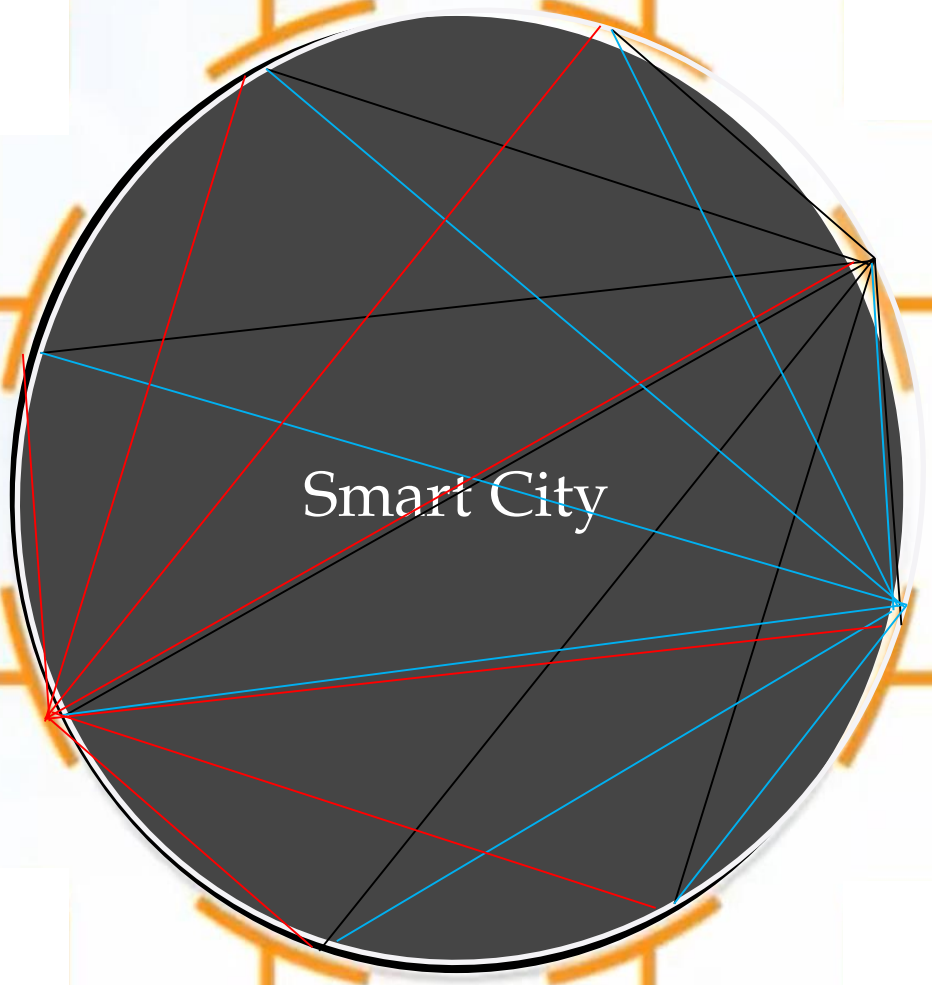
Define Key Performance Indicators

	Value to Customer	Value to Firm
Market		
Brand		
Customer		

http://square.github.io/crossfilter/

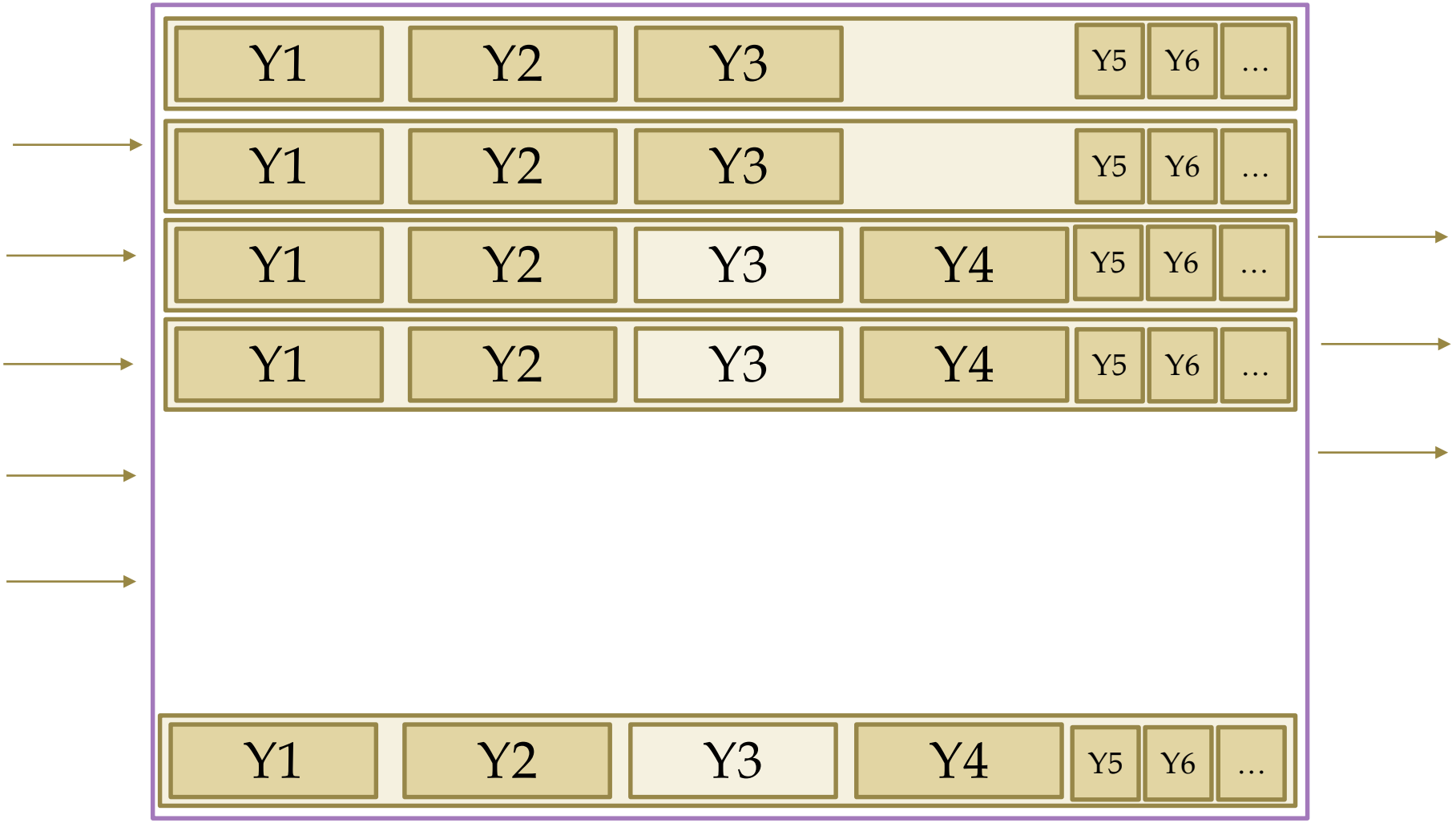
http://bost.ocks.org/mike/nations/

http://www.nytimes.com/interactive/2012/09/06/us/politics/convention-word-counts.html


Smart City

Education

Efficiency in Education

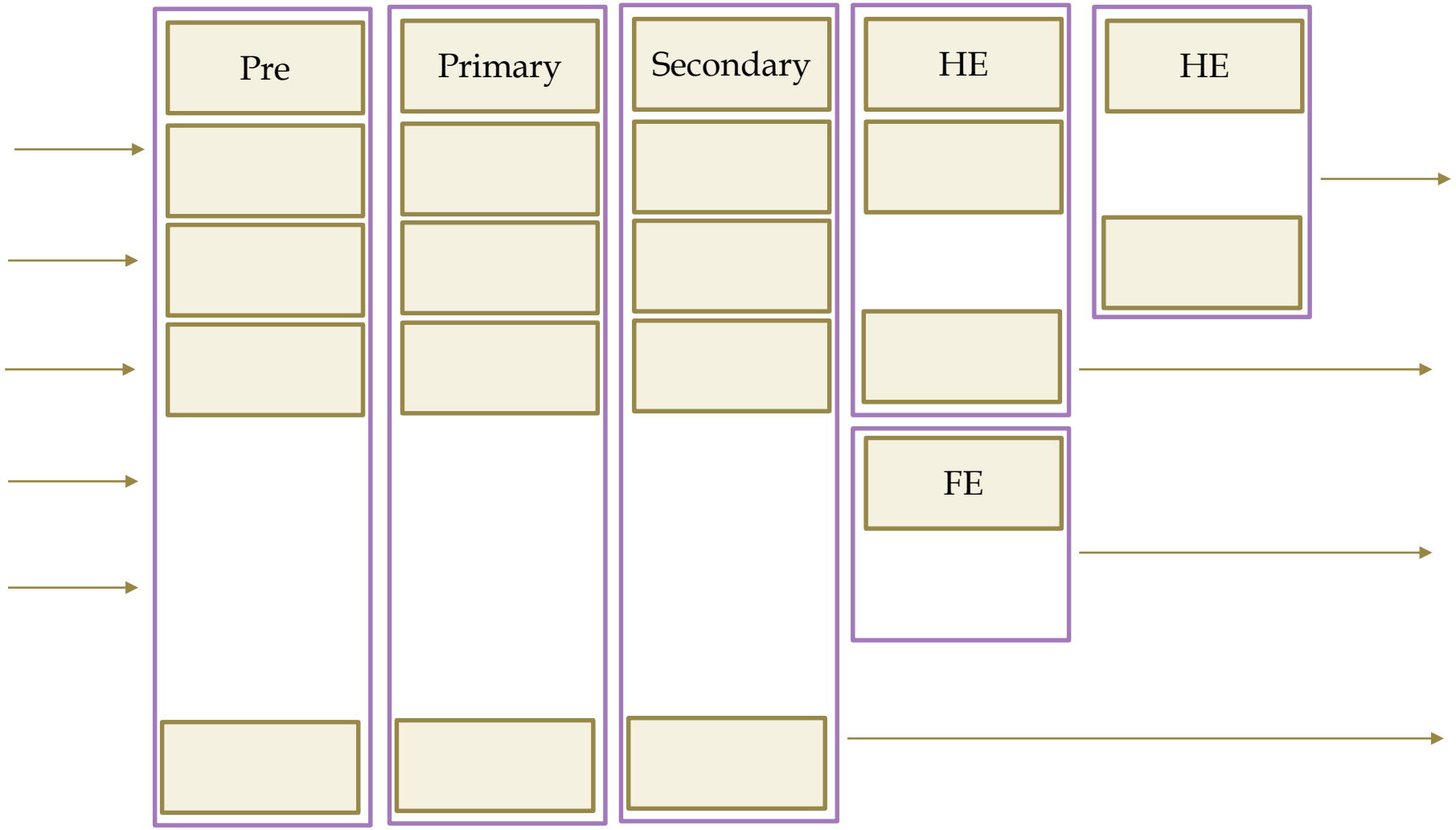


http://www.hefce.ac.uk/data/ukpi/

http://www.hefce.ac.uk/data/ukpi/

			Projected outcomes															
			Degree		Neither award nor transfer				Other award	Transfer	Not known	Projected average time to transfer or complete (years)		Efficiency				
Institution	Number of starters	Starters who are mature	Projected	Bench-mark	Projected	Bench-mark	Std dev		Projected values			Actual	Efficient	Projected	Bench-mark	Std dev		
		%	%	%	%	%			%	%	%			%	%			
UK																		
All UK Institutions	*267,823*	*27*	**77**		**16**				1	6	0	3.7	3.1	**85**				
England																		
All English institutions	*218,227*	*28*	**77**		**16**				1	6	0	3.6	3.1	**85**				
Anglia Polytechnic University	*2,238*	*41*	**67**	**72**	**24**	**20**	0.8	*	4	5	0	3.7	2.9	**78**	**82**	0.8	*	
Aston University	*1,217*	*11*	**86**	**83**	**9**	**10**	0.9		0	5	0	4.0	3.6	**91**	**89**	0.6		
University of Bath	*1,346*	*10*	**90**	**82**	**6**	**12**	0.9	*	0	4	0	3.9	3.6	**94**	**88**	0.4	*	
Bath Spa University College	*677*	*42*	**83**	**74**	**12**	**19**	1.5	*	1	4	0	3.3	3.0	**90**	**84**	1.0	*	
University of Birmingham	*3,685*	*6*	**90**	**86**	**6**	**8**	0.4		1	4	0	3.6	3.3	**92**	**90**	0.3		
Bishop Grosseteste College	*223*	*14*	**91**	**80**	**7**	**13**	2.2		0	3	0	3.9	3.7	**94**	**89**	1.4	*	
Bolton Institute of HE	*1,327*	*49*	**61**	**68**	**32**	**24**	1.2	*	2	4	1	3.7	2.6	**69**	**79**	1.2	*	
Bournemouth University	*2,076*	*21*	**73**	**73**	**20**	**19**	0.9		1	6	0	3.9	3.4	**85**	**83**	0.7		
University of Bradford	*1,800*	*24*	**79**	**75**	**16**	**17**	0.9		0	6	0	3.9	3.4	**87**	**83**	0.7	*	
Bretton Hall	*643*	*28*	**76**	**79**	**19**	**15**	1.4		0	5	0	3.5	3.1	**87**	**87**	1.0		
University of Brighton	*2,204*	*41*	**74**	**74**	**18**	**19**	0.8		3	5	0	3.6	3.1	**85**	**83**	0.7		
University of Bristol	*2,602*	*10*	**92**	**89**	**4**	**6**	0.5		0	4	0	3.6	3.4	**95**	**92**	0.3		
Brunel University	*2,697*	*28*	**75**	**72**	**17**	**20**	0.8		2	6	0	3.8	3.2	**83**	**82**	0.6		
Buckinghamshire Chilterns University College	*1,361*	*34*	**78**	**71**	**16**	**21**	1.1	*	1	5	0	2.9	2.4	**83**	**81**	0.9		
University of Cambridge	*2,946*	*10*	**96**	**89**	**1**	**6**	0.5	*	0	2	0	3.2	3.2	**98**	**92**	0.2	*	
Canterbury Christ Church University College	*1,130*	*38*	**84**	**74**	**13**	**18**	1.2	*	0	3	0	3.4	3.1	**90**	**83**	0.7	*	

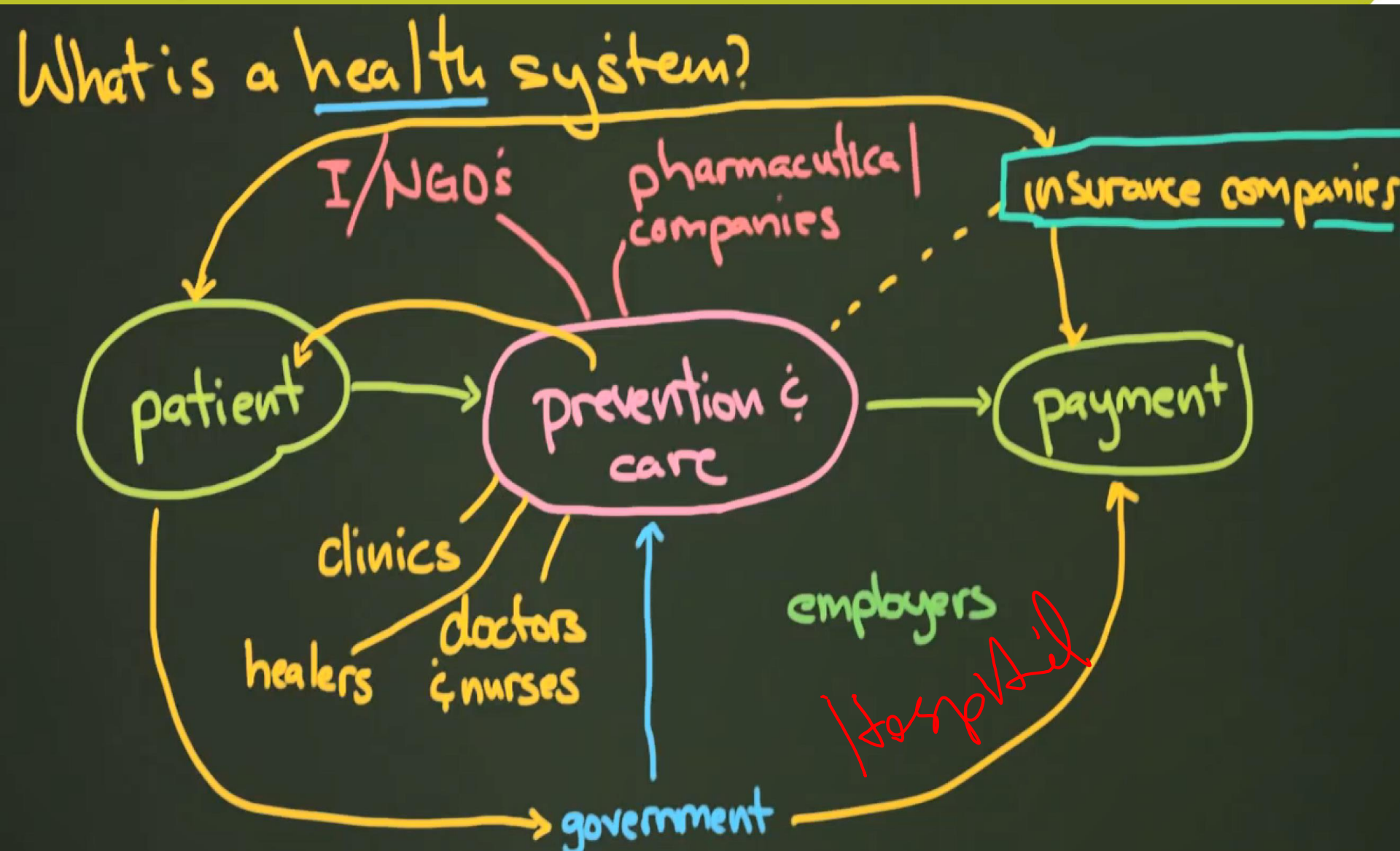
Efficiency in Education



Parents Students / Pupils LEA/ Gov Schools Uni /Colleges Teachers / Academics Admin VC/ Head teachers

Health

Health System



Banking

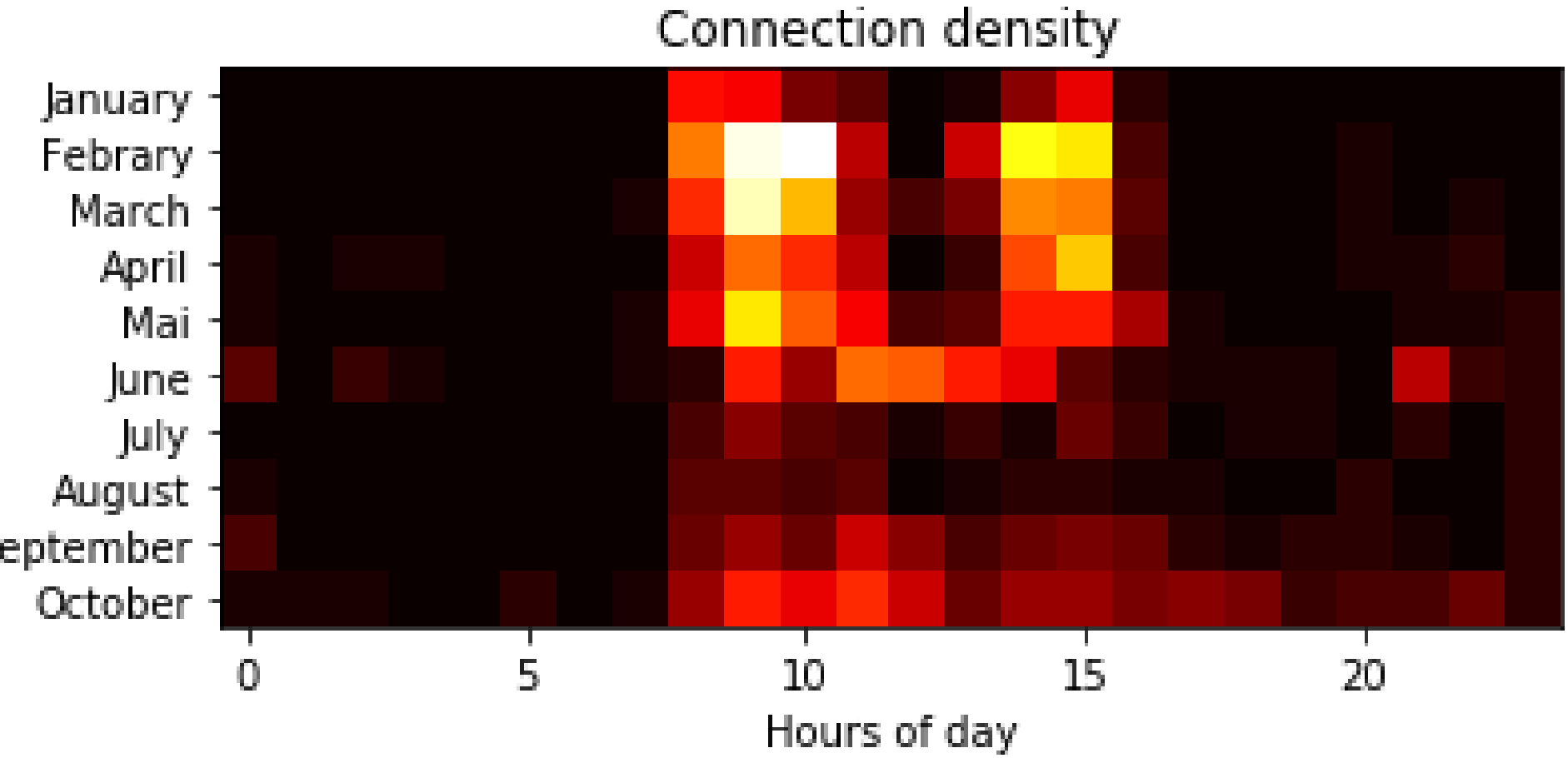
Customer with highest propensity to contract a consumer loan

- A group of 50,000 customers among the oldest users of mobile banking have been selected.

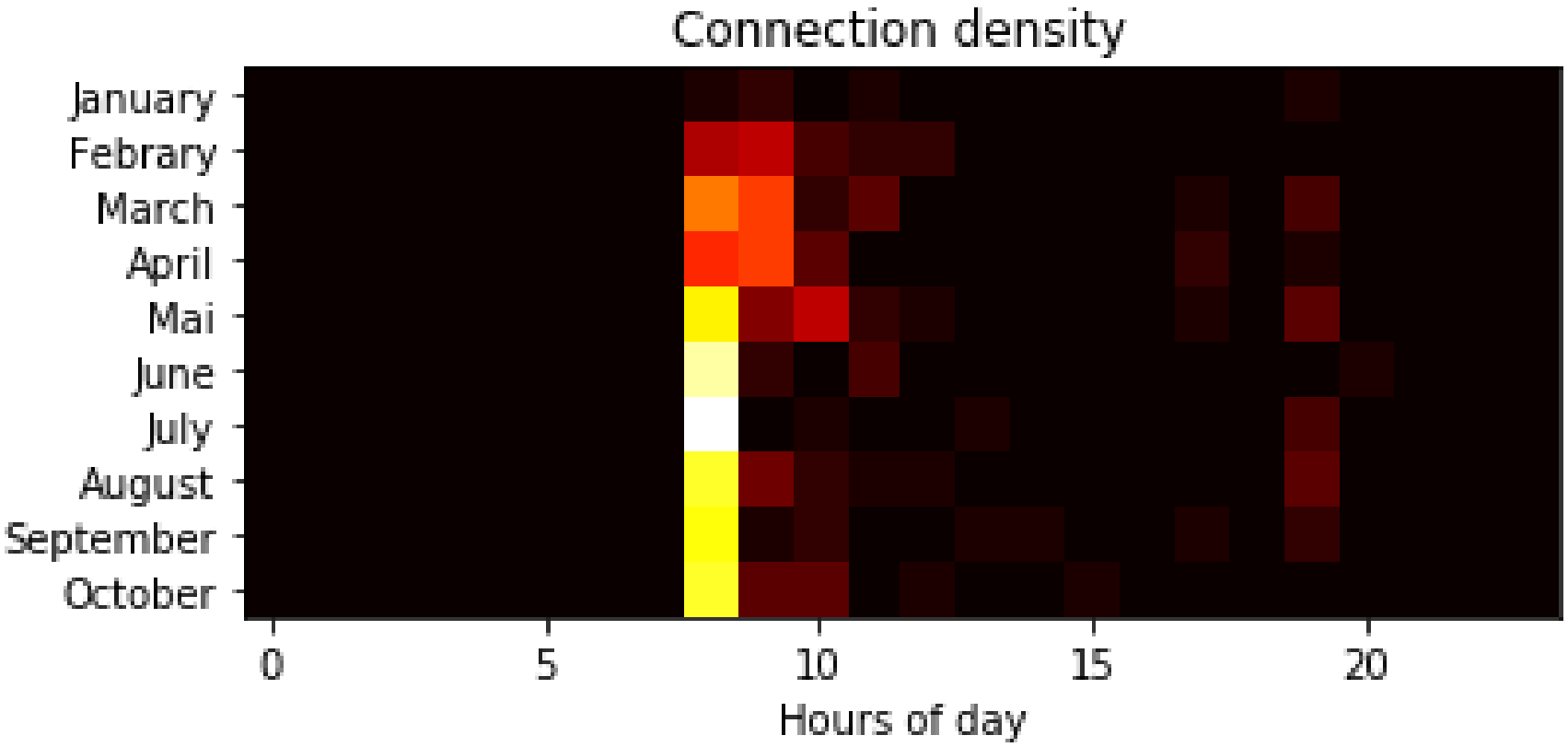
- Their mobile banking connections during the period January to October 2017 have been extracted and put in a local database resulting in a collection of about 8,900,000 records. A connection record is made of the customer ID and date and time of connection. The time of connection contains the hour and minute of connection.

- For the group of customers under study, the purchasing history has been extracted and put in a local database containing about 400,000 records. Each record contains the
  - Customer ID,
  - The product type
  - purchased and
  - the date of purchase.

Customer with highest propensity to contract a consumer loan



Customer with highest propensity to contract a consumer loan

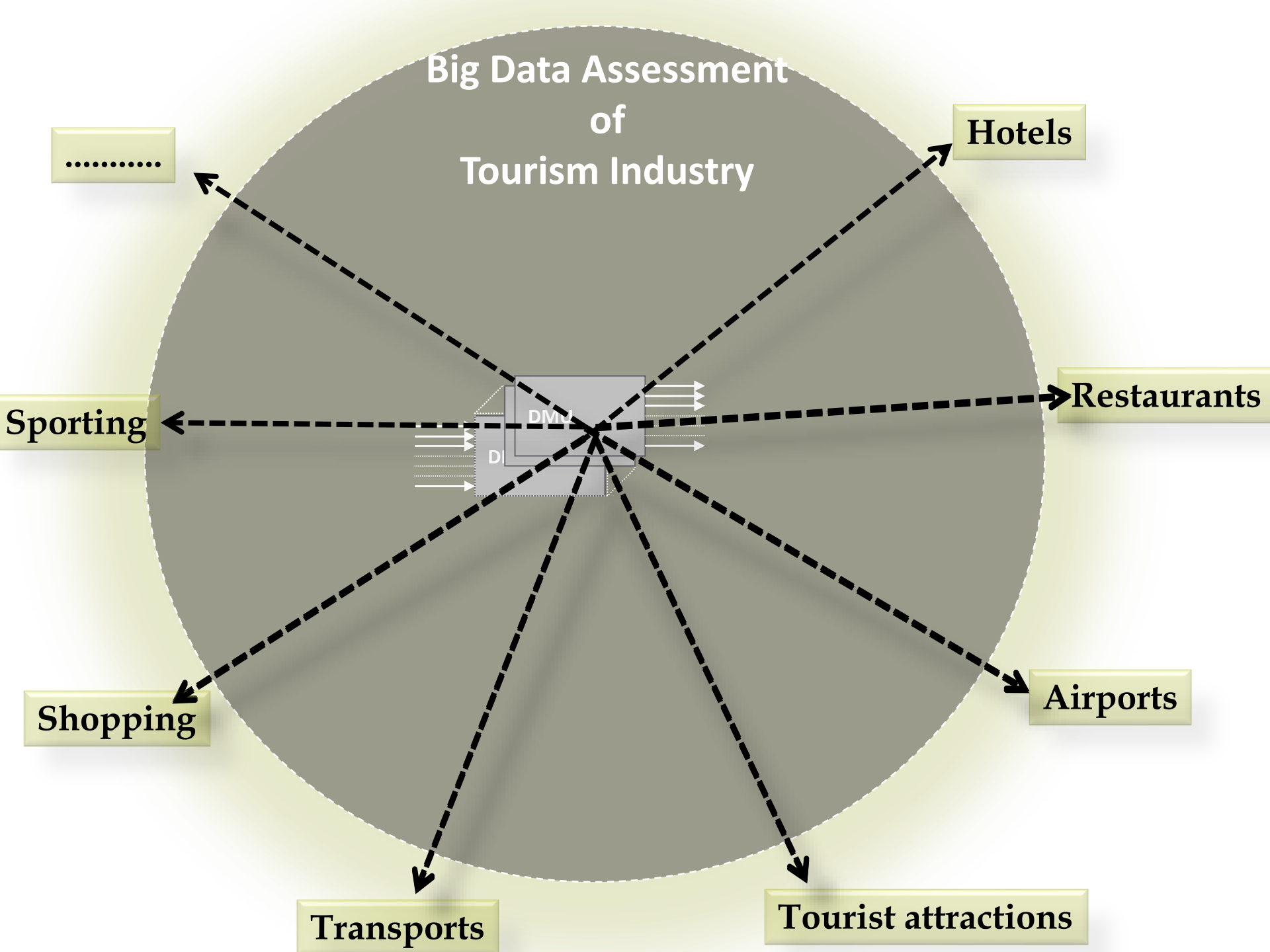


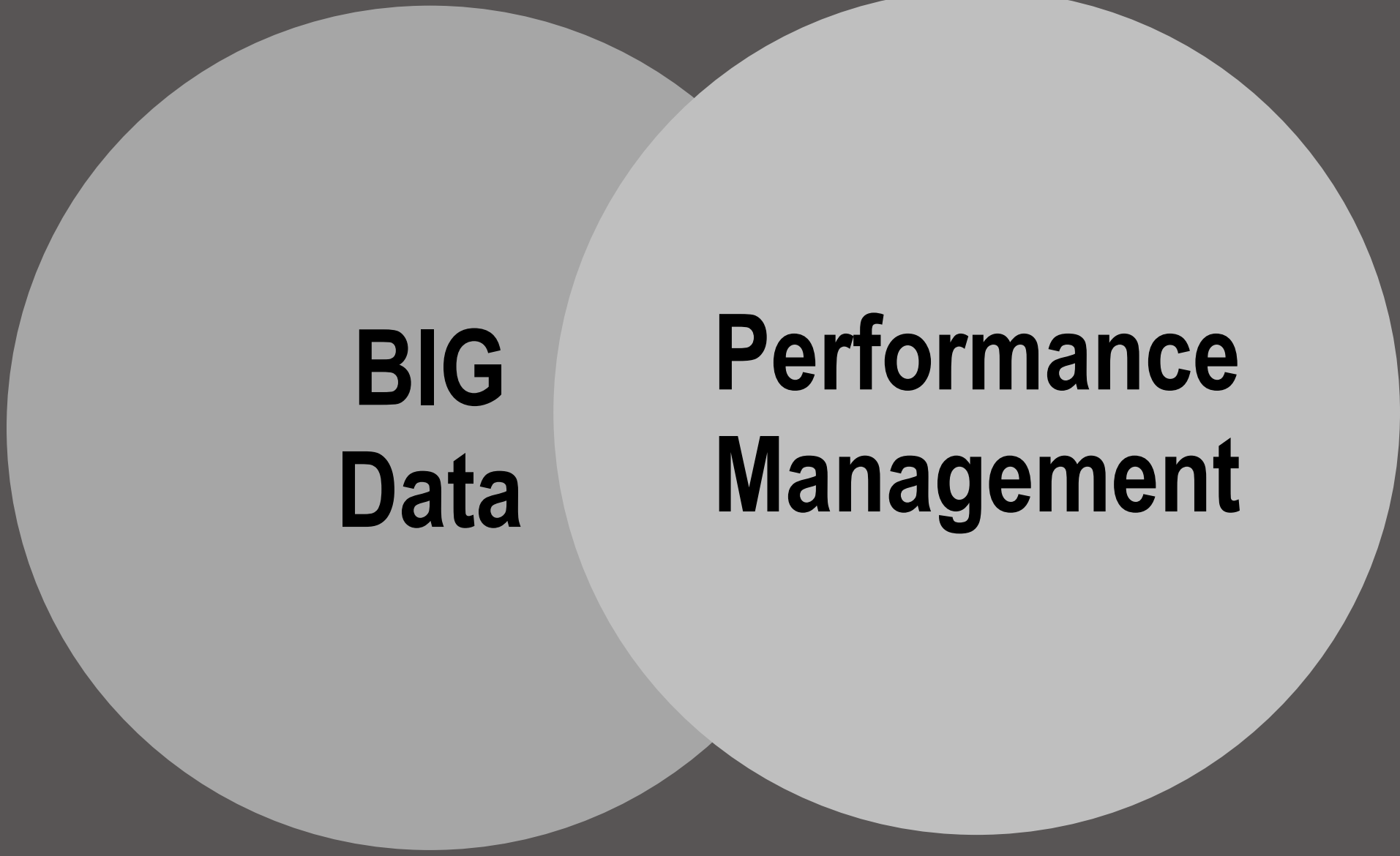
Energy

Modelling Residential Energy Consumption

- Residential energy consumption is a significant part of the overall energy consumption in UK.
- The objective of this study is to gain a clear understanding of the variation and the underlying determinants of electricity consumption in the residential sector by building data mining and statistical models.

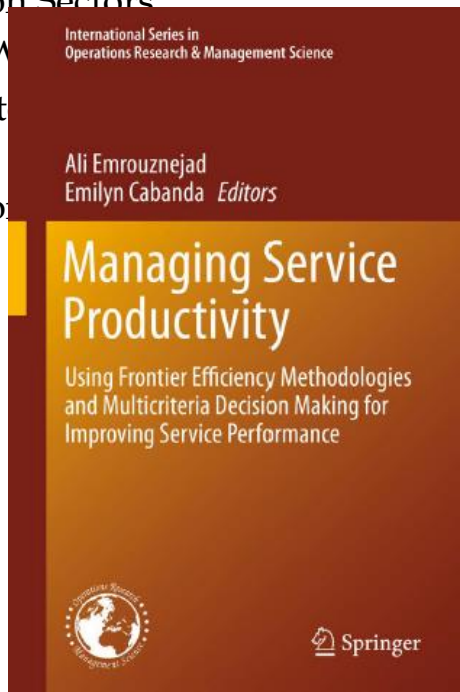
Tourism


Big Data Assessment
of
Tourism Industry
.....
Hotels
Sporting
Restaurants
Shopping
Airports
Transports
Tourist attractions


BIG
Data
Performance
Management

Managing service productivity

- **In this introductory chapter we present a classification scheme with seven main primary categories in service industry, namely, education, hospital and healthcare, tourism, banking and finance, information technology and media services, transportations, and utilities.**
- Development of Assessment Model for Research Efficiency of Universities
  - Estimating the Efficiency of Healthcare Facilities
  - Using Data Envelopment Analysis to Measure Good Governance
  - Measuring Efficiency of Courts: An Assessment of Brazilian Courts Productivity
  - Application of DEA in the Electricity Sector: The Case of Meralco Distribution Sectors
  - Improving Energy Efficiency Using Data Envelopment Analysis: A Case of W
  - Service Productivity in IT: A Network Efficiency Measure with Application t System
  - Efficiency of Software Development Projects: A Case Study on an Informatio Company in India
  - Protocol for Comprehensive Efficiency Analysis of Multi-Service
  - Metropolitan Transit Agency Operators
  - Measuring the Sustainability of Air Navigation Services
  - Measuring and Managing the Productivity of U.S.
  - Public Transit Systems: An Un-oriented Network DEA
  - Using DEA to Improve the Efficiency of Pupil Transportation



Aston Business School

Accredited by Association of MBAs · EFMD EQUIS ACCREDITED · AACSB ACCREDITED

Big data optimization

Big data optimization

- **There are many issues arising from the emergence of big data, from computational capacity to data manipulation techniques, all of which present challenging opportunities. These large amounts of data present various challenges, one of the most intriguing of which deals with knowledge discovery and large-scale data-mining.**
- Although these vast amounts of digital data are extremely informative, and their enormous possibilities have been highlighted on several occasions, **issues related optimization remains to be addressed,** For example formulation of optimization problems with unprecedented sizes (millions or billions of variables) is inevitable.

- ✓ summary of the current research in big data and large-scale optimization
- ✓ comprehensive discussion on setting up a big data project
- ✓ existing techniques for data reduction, at scale to facilitate big data processing optimization
- ✓ performance tools for big data optimization
- ✓ very good application of big data optimization that use for analysing big images, Real-life graphs and other images
- ✓ explains interlinking big data to web of data
- ✓ several techniques for optimizing big data processing including computational efficient implementations like parallel and distributed architectures
- ✓ Large-scale and big data optimization based on Hadoop
- ✓ discussion about metaheuristics for continuous optimization of high-dimensional problems

PIM