From Data Analytics

22 May 2019 @ MCRHD

FC for MES Offrs

Sudhir Voleti

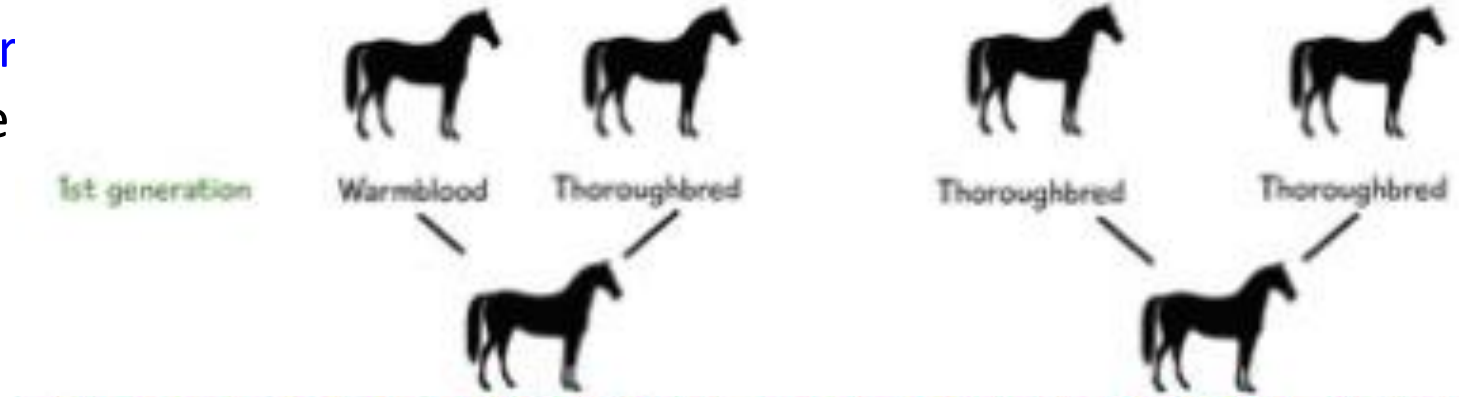
Associate Professor of Marketing, ISB

Sudhir_Voleti@isb.edu

ISB

ry

- Hor
the



BLOODSTOCK FOR THE 21ST CENTURY™

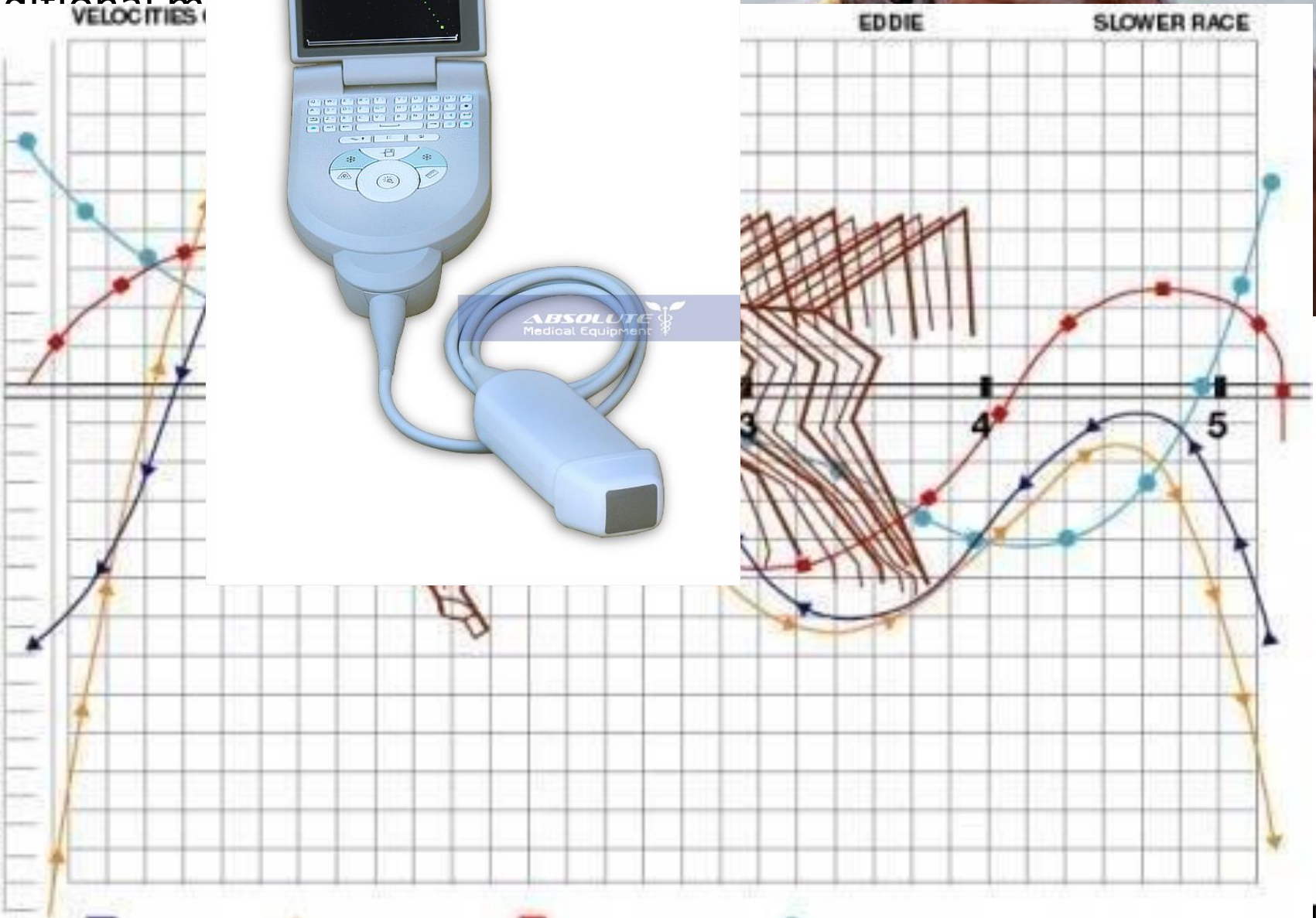
EQB®

Agents & Consultants

- Enter Jeff Seder of EQB, a boutique consulting firm.

ISB

- T
h
- S
c
- H
h
a
- T
b
- A
v



A Motivating Example

Grand Slam Tennis Tours

- raoh became the first horse in 37 years to

A Motivating Example: Concluded

- So, what is the example trying to motivate?
- [1] Importance of having a clear **Objective** to pursue or **Question** to answer.
- [2] **Data is paramount**, when studying, measuring, modeling or understanding any phenomenon of interest.
- [3] Good predictors of an outcome *can* show up in unexpected places - where nobody thought to look, **overtaking theories & explanations** - involves **trial-&-error**, guesswork & analytics.
- [4] Important to keep an eye out for **new tech**, which may enable new data to be collected & analyzed.
- [5] Data alone is NOT enough. **Analytics is required**, and an open mindset.

ISB

Session Outline

- Motivating Example for Data Analytics
- Preliminaries
- Introduction to Problem Formulation
- Determining Data Requirements
- Some Thoughts on Report Writing: Best Practices
- Session Wrap-up

ISB

Some Preliminaries

ISB

Preliminaries: About me…

- **Academic Credentials:**
  - PhD in Marketing – Univ of Rochester (2009)
  - MS in Applied Statistics – Univ of Rochester (2006)
  - PGDM – IIM Calcutta (2001)
  - B.E. – BIT Mesra (1998)

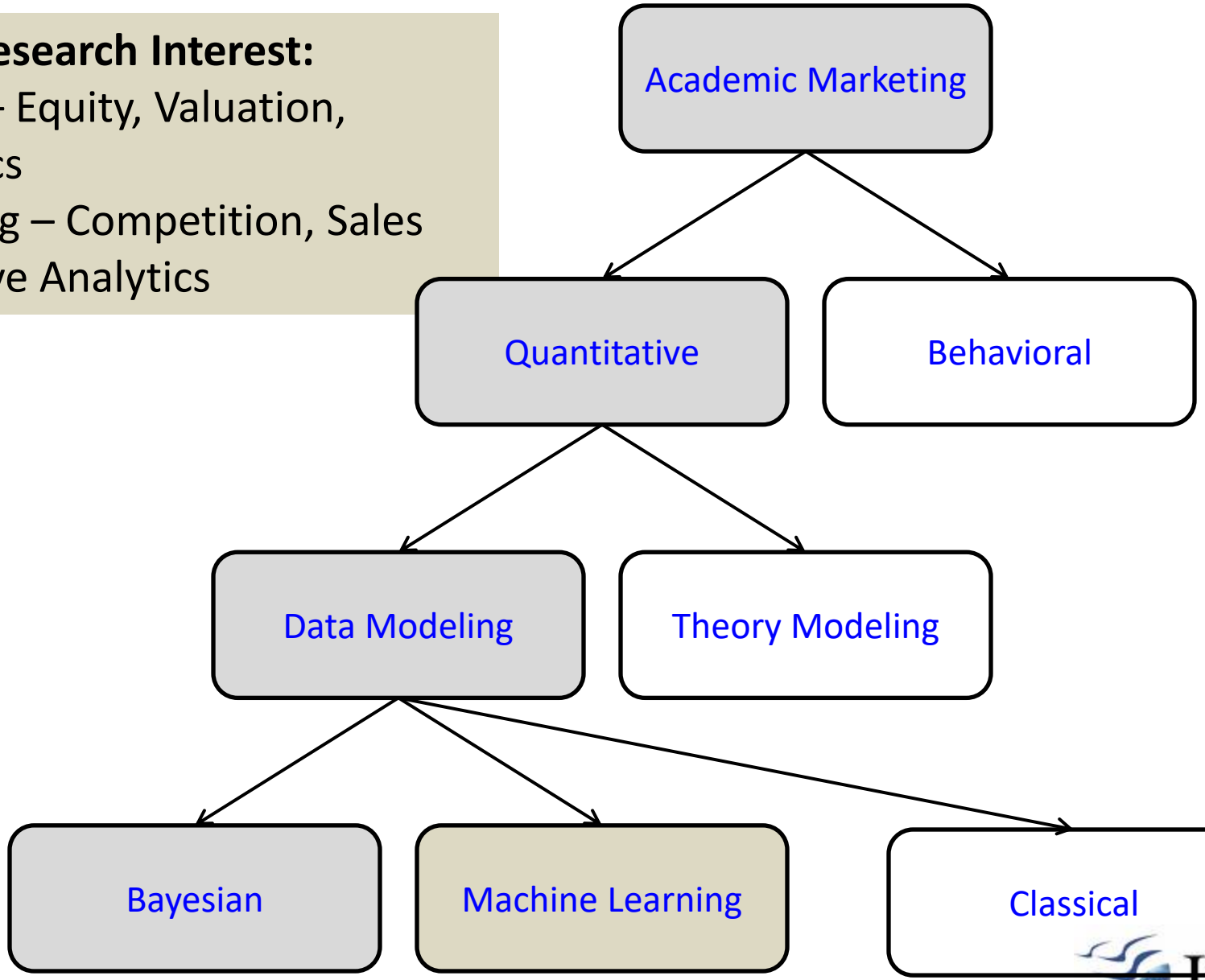
- **Industry Experience:**
  - Software Programmer with Cognizant 1998-99
  - Management Consultant with Accenture 2001-02
  - Data Analyst – Daymon Consumer Insights Division 2006-08
  - Academic Faculty with ISB – 2009 onwards
  - Been involved in a Tech Startup – Modak Analytics – 2012

ISB

Preliminaries: About my Research…

Topics of Research Interest:

1. Brands – Equity, Valuation, Dynamics
2. Modeling – Competition, Sales
3. Predictive Analytics



ISB

Motivating

Problem Formulation

ISB

Motivating Example

- What's the Mongolian landscape like?

Motivating Example

- First, they analyzed the most common diseases needing hospital access.
- Next, they developed DIY (Do-it-Yourself) medicine kits, which like first aid, could be self-medicated after self-diagnosis.
- The DIY kits were placed in each home and their use explained.
- Next, paramedical staff were assigned territories they'd cover once every 6-12 months.
- On each visit, they'd **audit the kit** and the family would **pay only for what medicine was consumed**.
- Simple model, eh? But was it effective? What was the result?

ISB

Motivating Example

- Hospital visits declined **45%** in many remote areas → pressure eased on hosp resources and budgets.

- House-call demand for doctors fell **17%** → precious doctor time freed up for other work.

- But more importantly, look at the seemingly simple business model…
- Medicine as *postpaid* rather than *prepaid.*

- Extensions? Implications? Further possibilities? Plentiful.

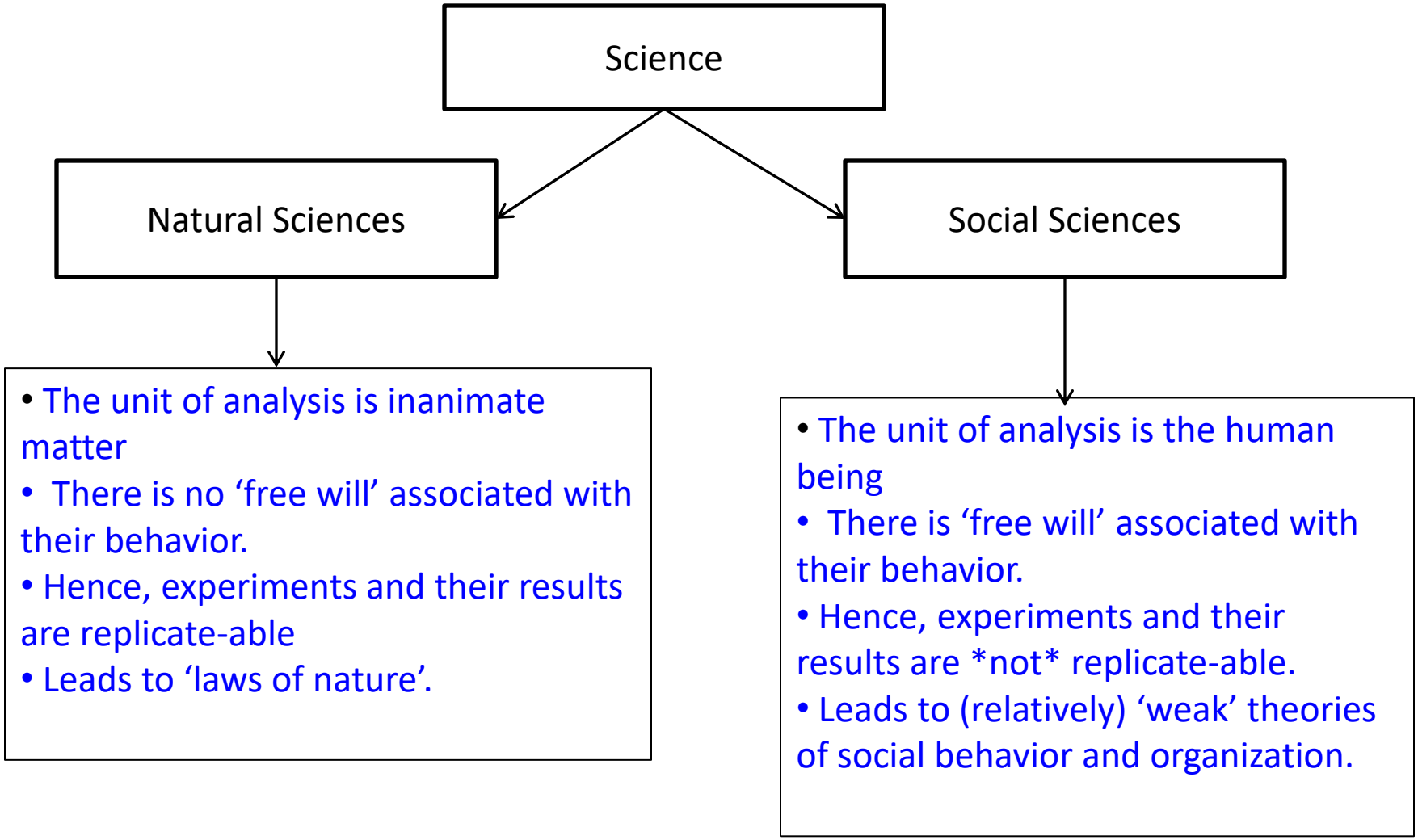
- But remember how it all began… at the problem formulation stage…
- By changing one Q with another, we transformed the problem from "increasing supply of healthcare" to "reducing healthcare demand"…

ISB

Conceptual Preliminaries

ISB

Preliminaries: Is 'Analytics' Scientific?



Bottomline: There's only so much **precision** in our **measurements** and our results that we can expect.

Why Identify the Units of Analysis

- Because without units of analysis, there is no Measurement.
- Without Measurement, there is no Data.
- Without Data, there is no Analysis.
- Without Analysis, there is no Modeling.
- Without Modeling, there is no Explanation and Prediction.
- Without Explanation, there is no Insight.
- Without Prediction, there can be no Optimization.
- Without Insight & Optimization, there is no Management.

ISB

The Data Story

and History

ISB

The Age of Data

"If *Land* was the primary raw material of the agricultural age,

and *Iron* that of the industrial age,

then *Data* is the primary raw material of the information age."

Nice quotation. But what's its practical significance?

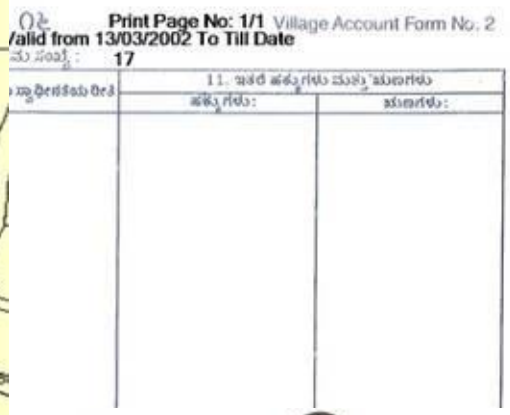
Consider this Q:

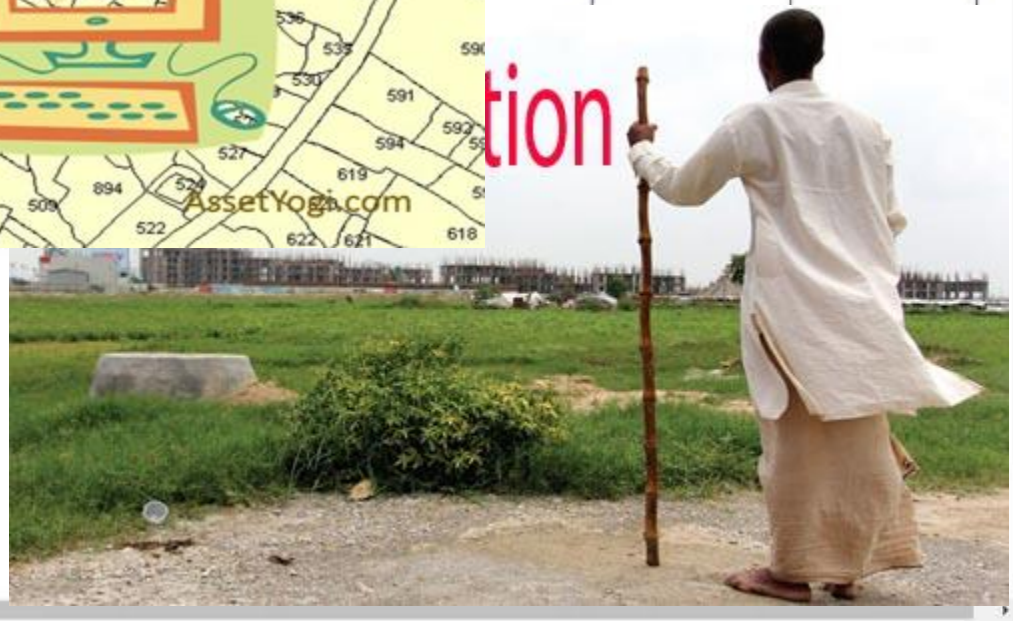
"How many of our present day laws, institutions, societal norms and governance structures actually derive from the agricultural age?"

ISB

The Agricultural Age, Data and Governance

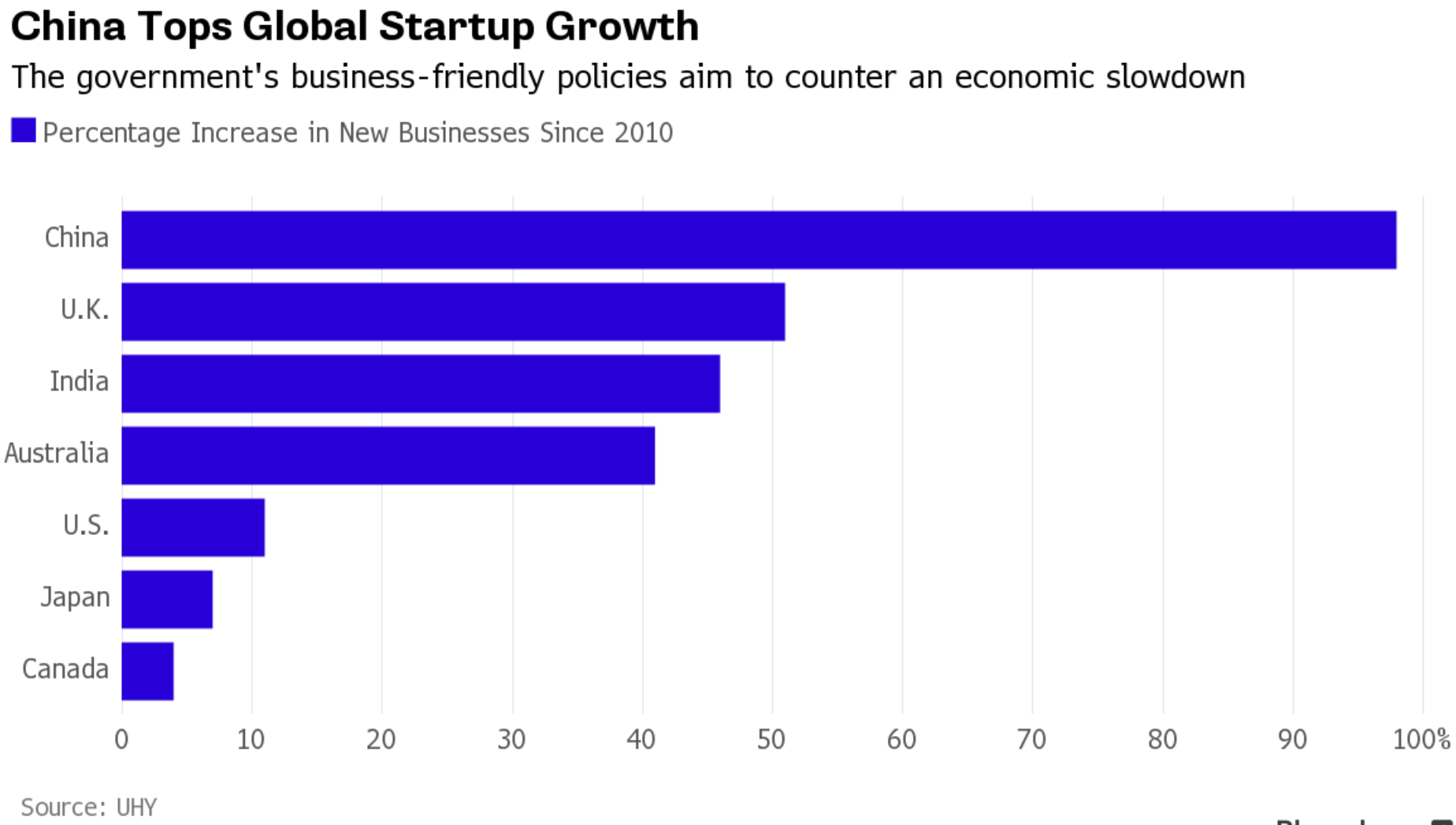
Q: How many of our present day laws, institutions, societal norms and governance structures actually derive from the agricultural age?


Print Page No: 1/1 Village Account Form No. 2
Valid from 13/03/2002 To Till Date
17

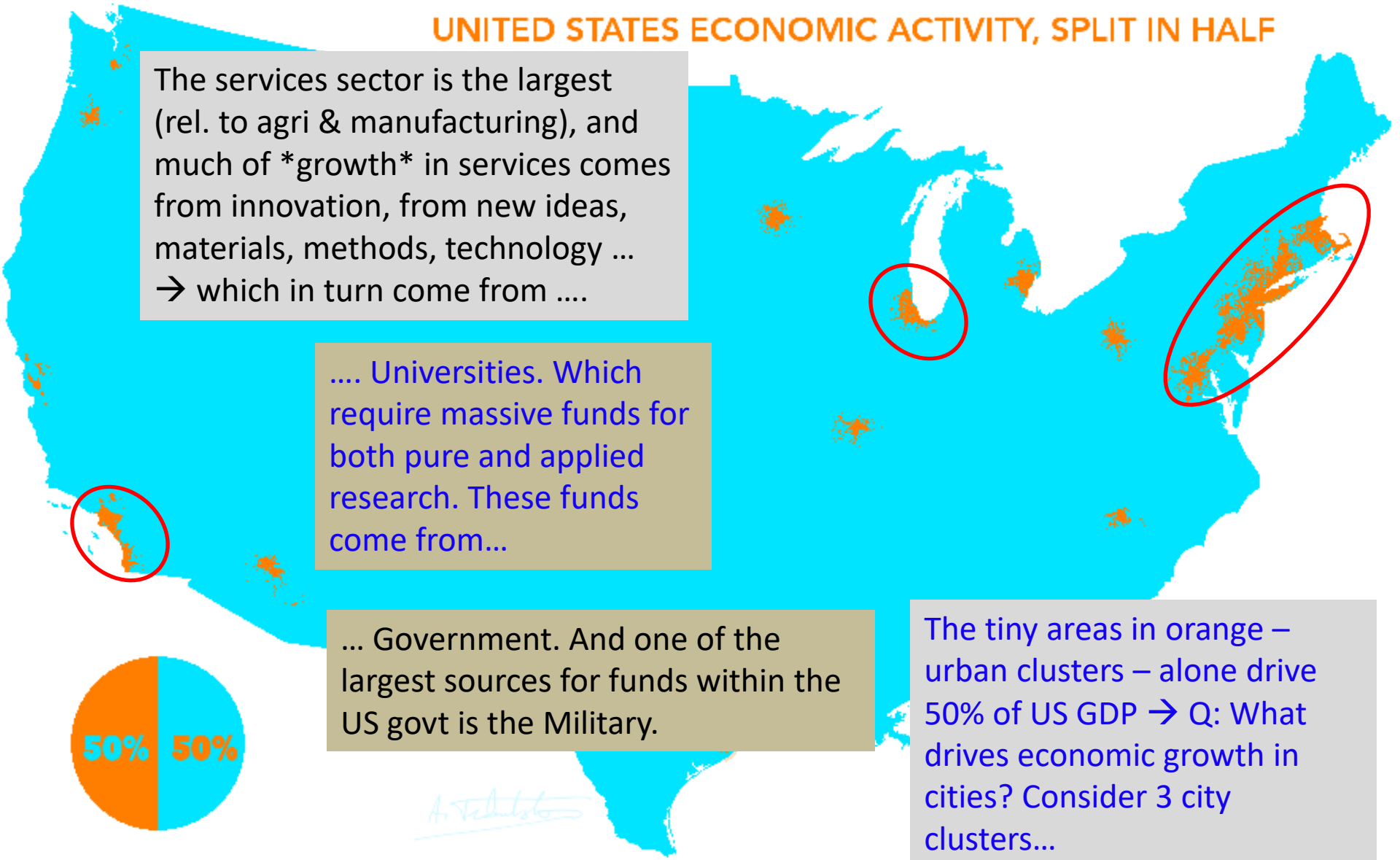


2017-2018 ಮುಂಗಾರು	ತಳವಾರ ಇನಾಂತಿ	1
2017-2018 ಹಿಂಗಾರು	ತಳವಾರ ಇನಾಂತಿ	1
2017-2018 ಬೇಸಿಗೆ	ತಳವಾರ ಇನಾಂತಿ	1

The Industrial Age, Data and Governance



Q: What Drives [US] Economic Growth?



The services sector is the largest (rel. to agri & manufacturing), and much of *growth* in services comes from innovation, from new ideas, materials, methods, technology …
→ which in turn come from ….

…. Universities. Which require massive funds for both pure and applied research. These funds come from…

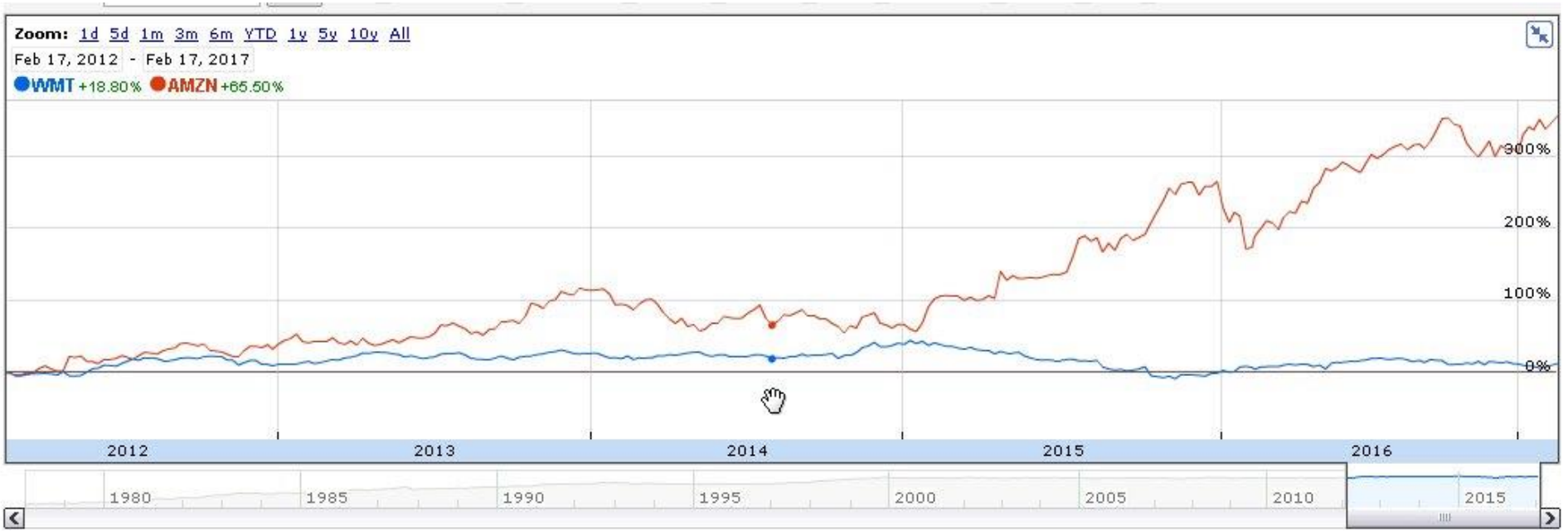
… Government. And one of the largest sources for funds within the US govt is the Military.

The tiny areas in orange – urban clusters – alone drive 50% of US GDP → Q: What drives economic growth in cities? Consider 3 city clusters…

ISB

The Information Age, Data and Governance: Example

- Consider the stock performance of Amazon (AMZN) vs Walmart (WMT)



- Valuation, February 2012:
- Walmart: $202 billion; Amazon: $82 billion

- Valuation, February 2017:
- Walmart: $210 billion; Amazon: $400 billion

ISB

Cost of Lost Opportunity: Quick Example

- 2000: Blockbuster had the opportunity to buy Netflix for $50M

n Disney's

- 2017: @Netflix worth $61 Bi

Disruption in Action …

- The world's largest taxi company owns no taxis (Uber)
- The largest accommodation provider owns no rooms (Airbnb)
- Largest phone co.s own no telco infra (Skype, WeChat)
- World's most valuable media firm creates no content (Facebook)
- The world's largest Movie house owns no theatres (Netflix)
- The world's largest software vendors don't write their own code (Apple, Google)
- Etc.

ISB

How does Disruption happen?

The next big thing will start out looking like a toy

One of the amazing things about the internet economy is how different the list of top internet properties today looks from the list ten years ago. It wasn't as if those former top companies were complacent – most of them acquired and built products like crazy to avoid being displaced.

The reason big new things sneak by incumbents is that **the next big thing always starts out being dismissed as a "toy."** This is one of the main insights of Clay Christensen's "disruptive technology" theory. This theory starts with the observation that technologies tend to get better at a faster rate than users' needs increase. From this simple insight follows all kinds of interesting conclusions about how markets and products change over time.

ISB

Data an

Technology News / Latest Technology News / Startups

Startups » | Data localisation | payments data | digital payment

How local compan

ISB

Data in the Information Age:

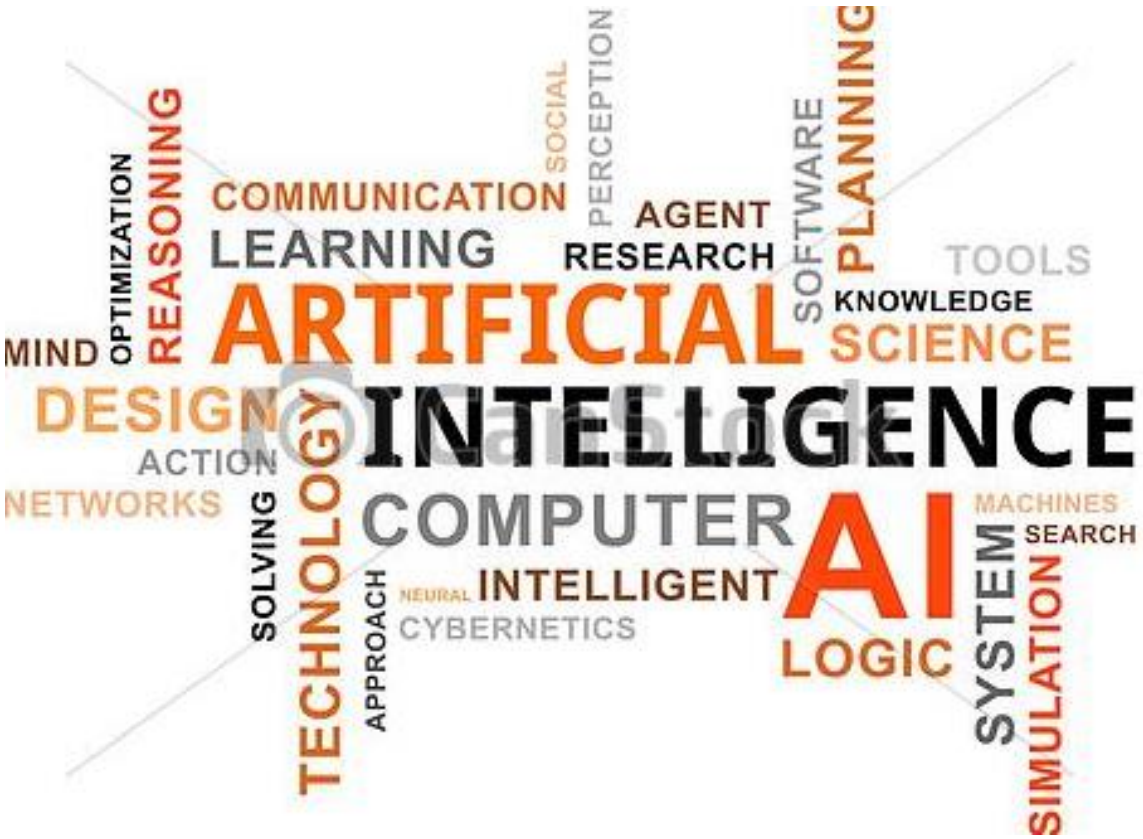
The Exponential Learning Curve

ISB

The Exponential Learning Curve

- ‘Data Analytics’ often leads to other terms such as ‘machine learning’, ‘artificial intelligence’, blockchain’, etc.

- So what do they mean anyway? How about an example to start figuring out what and how machines *learn* in this century?



The Exponential Learning Curve

- Till 1954, it was widely believed that human beings couldn't run 1 mile in 4 minutes of less. *Why?*

- In 1954, Roger Bannister broke that barrier.

- By 1

- … w *model* bre

- Wh ata driv

- Let'

ISB

The Exponential Learning Curve

- March 13, 2004. The Mojave desert, Calif., site of the DARPA Grand challenge. $1 million prize money.
- 150 mile race course, numerous [small] obstacles. 15 participants.
- What happened?

  None of the vehicles did even 10% of the course. CMU's modified Humvee did 7.5 miles before crashing into a ditch.

- October 8, 2005. Same venue. Re-match.
- Prize is now $2 million. Obstacles are now tougher – tunnels, narrow roads along cliff-edges.
- What happened?

  5 completed the race, 4 did so within 7.5 hours. Stanford's Sebastian Thurn's creation emerged winner by a 10 minute margin.

ISB

The Exponential Learning Curve

- November 10, 2007. Re-rema
- This time in an urban setting.
- Cars must now obey all of CA'
  merge into traffic, park by the

Its important to have an appreciation for growing processing, sensory and cognitive power of the machines.

Implications for Business and for managers? Plentiful.

ISB

Moving from Bits to Atoms …

TurtleBot 2

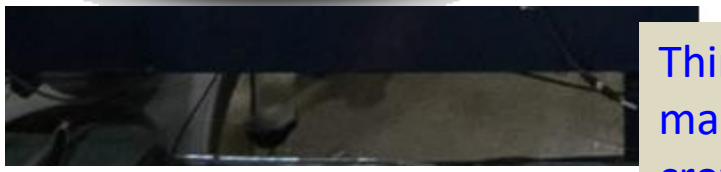
Open-source robot development kit for apps on wheels.

Quick Links

- I want to buy a TurtleBot!
- I want to build my own TurtleBot!
- Help me get started!

What is a TurtleBot?

TurtleBot is a low-cost, personal robot kit with open-source software. With TurtleBot, you'll be able to build a robot that can drive around your house, see in 3D, and have enough horsepower to create exciting applications.



- Implications?

Think of much of evolving tech as Platforms that enable mass collaboration, Co-creation, shareware, and the crowdsourcing of ideas + funding + design + programming + feedback + …
Welcome to the future.

Moving from Bits to Atoms …

2018 世界人工智

WORL

About Why ROS? Getting Started

VENTURE CAPITAL INVESTMENT INTO AI

The difference between China and the US in AI investment is big. The US has already invested heavily, while China is now catching up. Since the first US investment into AI in 1999, AI development has accelerated globally. In 18 years, total venture capital invested into AI has reached $28.9 billion.

progran

Low-Tech Analytics

The iCow story

ISB

- Ve
- **Pr**
  du
- (a
- (b
- (c
- **M**
  to
- **Si**
- Er

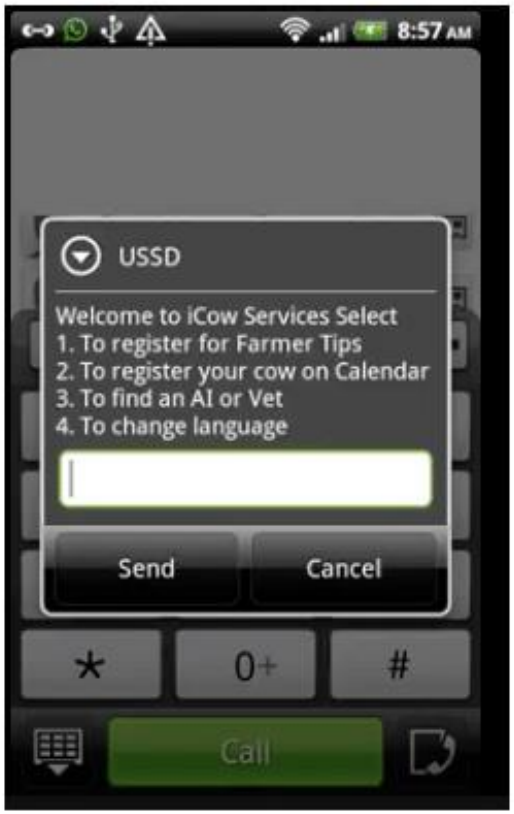
Motivating the iCow story

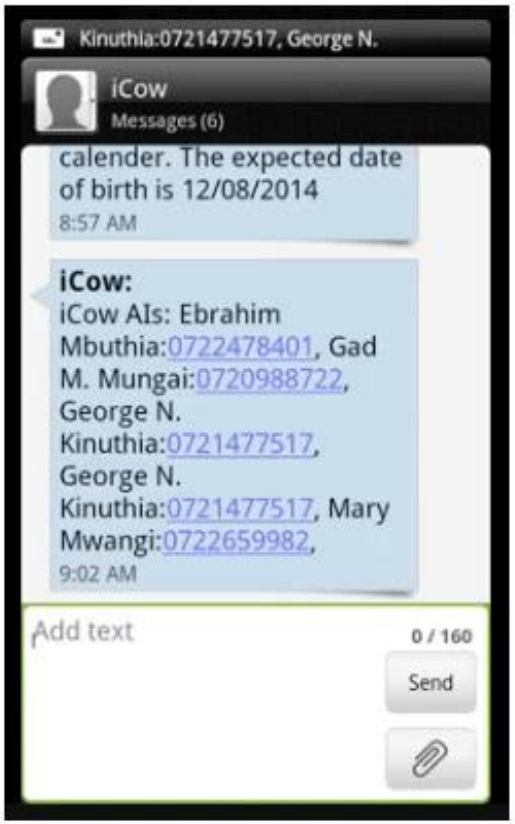
- iCow says "*SMS* me info on all 3 issues in standardized format. I'll SMS back instructions to maximize milk yield."

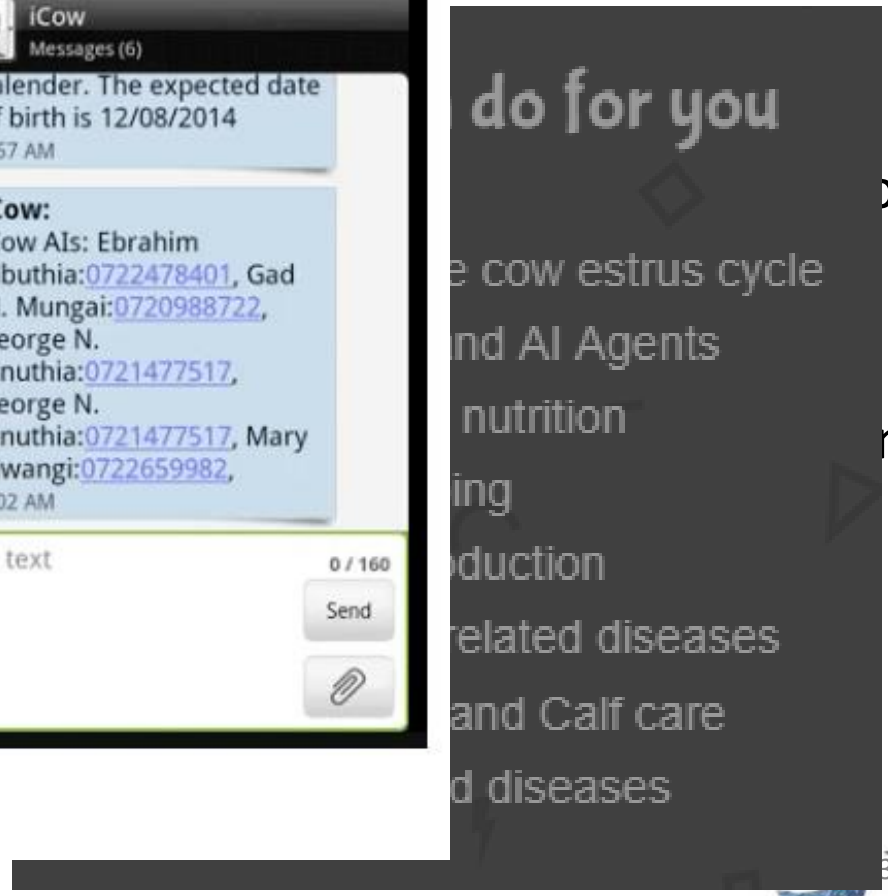
- Word s … start tuning in.

- Think … & yield d … on nation…

- Q: So h … ng only th…

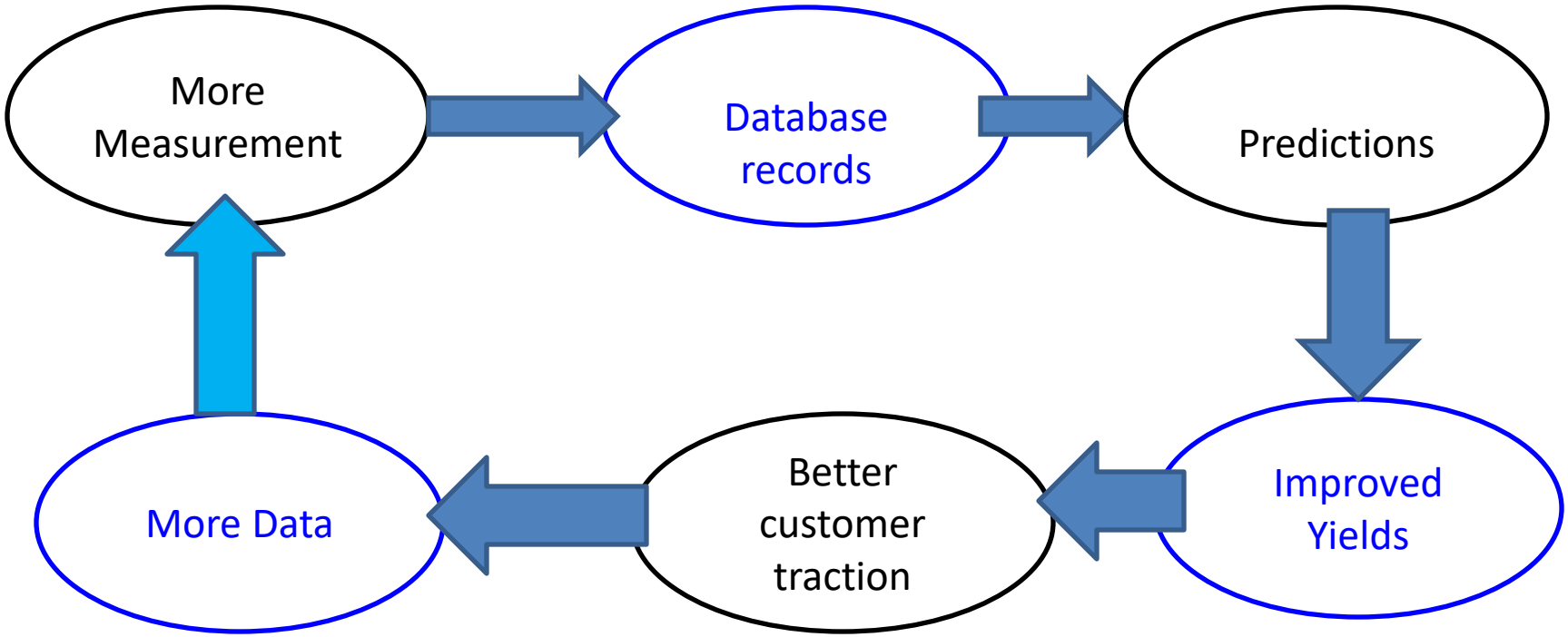






The iCow story: A Virtuous Cycle

- In the beginning, she starts with little or no data and relies primarily on theory and guesswork …. Later, when the data flow in, analytics is in.



Hence, more signups!

1. Bundled Service

We offer you the ability to bundle iCow with your service or product.

We are currently working with micro finance companies that offer financial products to smallholder farmers. products they are not only adding value to their customer, but also reducing their risk of non perfoming loans less of a risk than a farmer without adequate production knowledge.

- As a bus
- (a)
- (b)
- (c)
- (d)

2. Marketing

We offer marketing opportunities on the iCow platform as well as on the iCow website and social media sites

3. Surveys

- So i
- (i) i
- (ii)

We offer baseline surveys, longitudinal surveys, analytics, and reporting through our customer care and rese our partners to create in depth knowledge of their beneficiaries.

4. Sponsored content

- Q:

We offer organisations the ability to add value to their customers with pre-paid iCow products.

What is iCow?

iCow is an Agricultural Information Service with a variety of products available as a subscription service through *285# to help farmers enhance productivity

iCow Products

- Mashauri-Farmer Tips
- Upon subscription to Mashauri-Farmer Tips farmers receive 3 SMS tips per week at Ksh.3 per SMS
- Kalenda-Cow Calendar

Read More

Basic usage instructions:

Simply Dial *285# and follow the simple menu

iCow Forum

The iCow Story: Concluded

- So what was the example trying to motivate?

[1] Clear Prob Formulation → clarity in (Y, X); [2] Data Collection Op (low tech but sophisticated) → infused with domain knowledge; [3] ML engaged (connective function discovery); [4] Risk & uncert. Esp. in the early stages → necessitated common sense, fast feedback loops & risk taking; [5] Org issues simplified → e.g., "pilot traps", data silos etc. avoided; [6] Laser-like focus on end-customer need and value; [7] Appreciation of the core data asset; [8] Partnering with collaborators to co-build value; [9] Etc.

- What learnings can be generalized and carried over to large orgs? And importantly, what can't?

Larger, established orgs in mature mkts will have 2 main challenges: [1] Org Issues and [2] Mkt conditions. In Org issues think of (a) Org culture → priorities, status quo, tools access, data silos, talent acquisition, etc. In Mkt Envmt, think of (a) established competitors in mature markets; (b) opportunity identification…

Thank You

Q & A

ISB

A Framework for Problem Formulation

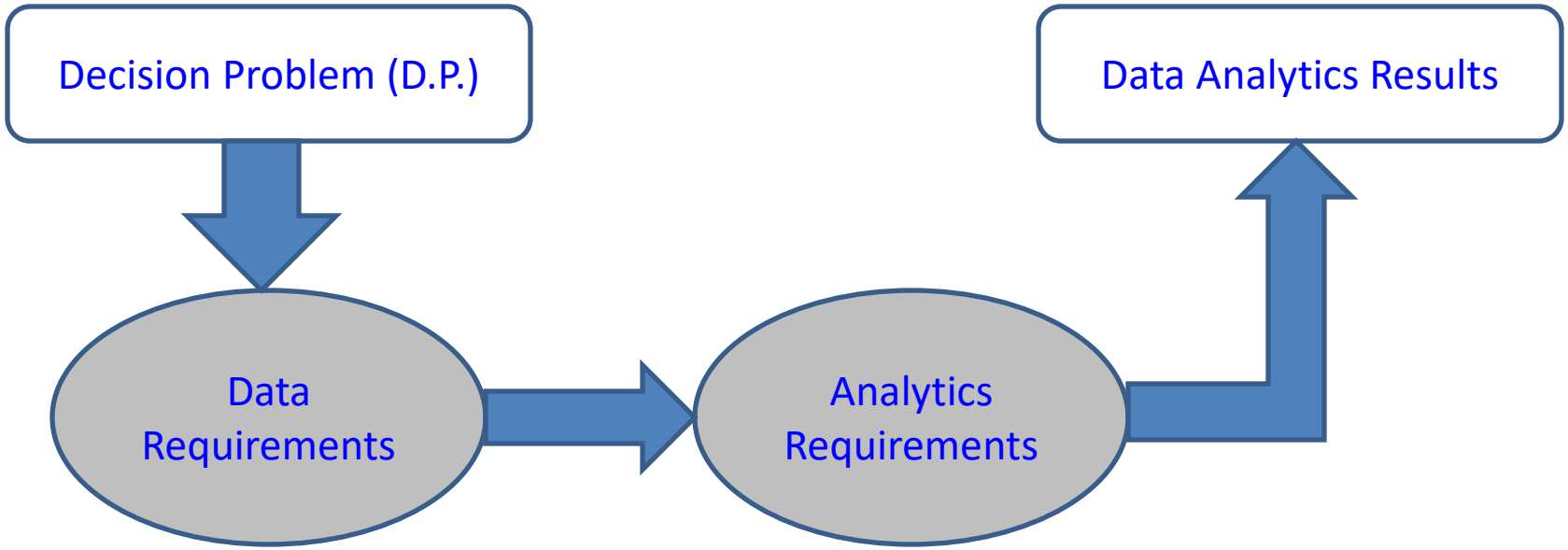
ISB

Problem Formulation Basics

- “Computers are useless. They can only give us **answers.”** ~Pablo Picasso (1881-1973)
- "A problem well formulated is half the job done."
- Problem formulation (P.F.) is critical because: (1) without P.F. we **wouldn't know what to look for**.
- (2) Hence, IF our P.F. goes wrong, our **data analytics will all be useless**.
- (3) P.F. impacts data side decisions - collection, cleaning, analysis - and thereby time and cost.
- Next, we'll see a P.F. framework that will help structure the P.F. process for us.

ISB

A Problem Formulation Framework



- D.P. is usually asked as a question. (E.g., “Can we raise supply?”)
- Data requirements are gaps in data needed to answer the question
- Analytics requirements are analytics tools and transformations needed on the data
- Data analytics results should ideally aid in solving the D.P.

ISB

P.F. Framework: From D.P. to Data Requirements

Decision Problem (D.P.)

Health Department:
"How healthy are T.S. people on average?"

Data Requirements

Problem Formulation: Recap

- Why is problem formulation critical? Challenging?

- How does problem formulation impact data side decisions – collection and analysis?

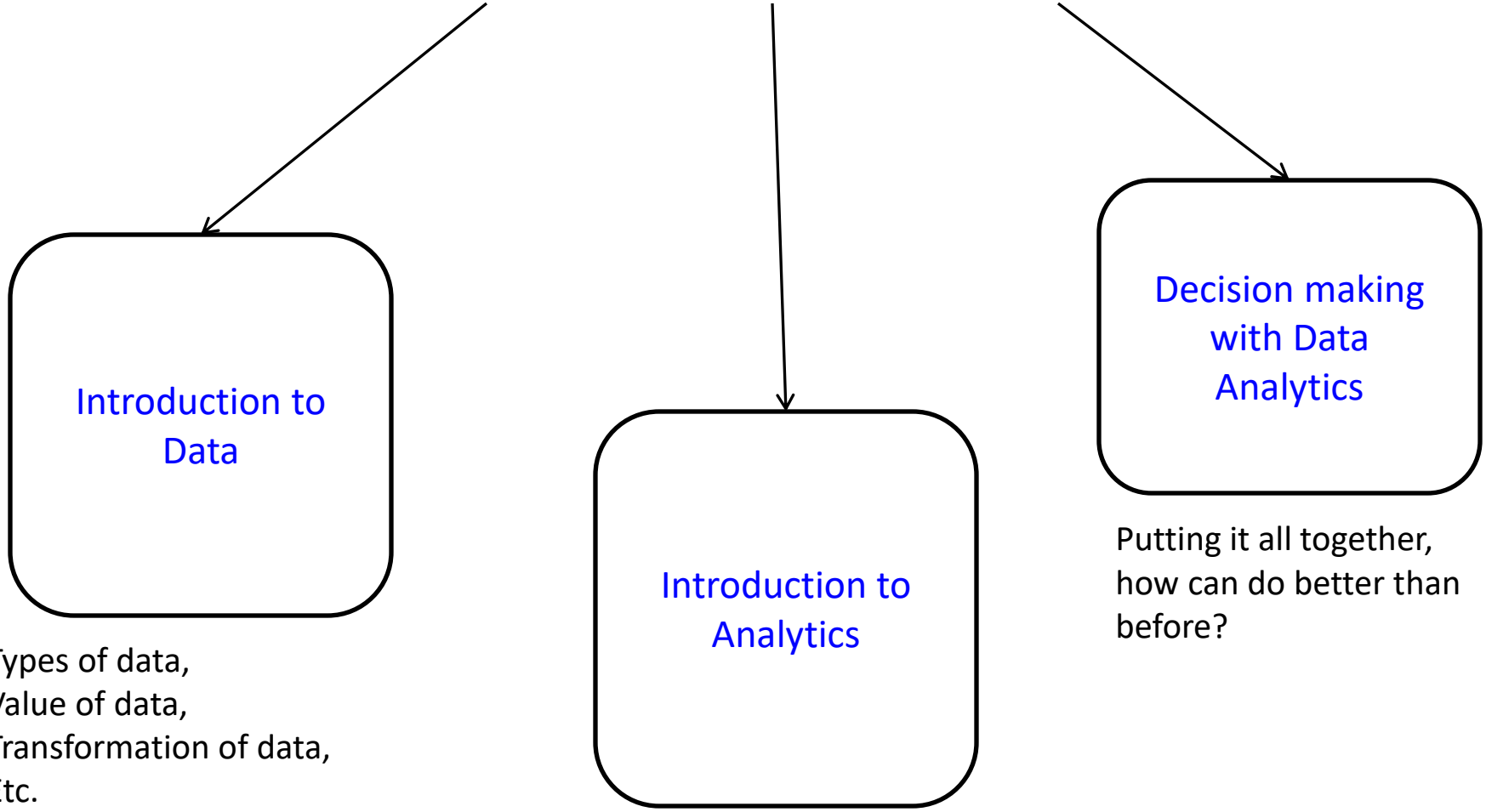
- Where does analytics come into the picture?

ISB

Blank Separator

ISB

Preliminaries: 3 Course Objectives



ISB

Analytics in Govt Action: Example

- Yesterday's news article has a nice example of Data Analytics in Govt Ac

NDTV

- Let

Over 11.44 Lakh PANs Deactivated, Says Junior Finance Minister

- "A

ide

"As on July 27, 11,44,211 PANs have been identified and deleted or de-activated in cases where multiple PANs were found allotted to one person," Santosh Kumar Gangwar said.

wh ase.

Be

Business | Press Trust of India | Updated: August 01, 2017 22:55 IST

r

ap

- Th

ac

ISB

Data and Measurement Basics

ISB

Background: The Data Story

- For millennia record keeping meant clay tablets, papyrus scrolls, parchments ...
- Modern paper was an enormous advance but what really set the revolution going was the Printing press.
- In 50 years, printing presses produced more books than had been produced in *all* of prior history.
- In subsequent centuries came the telegraph, telephone, radio, TV and computers.
- Digital storage first became cheaper than paper storage in the year 1996.
- In 2000, 25% of new data was stored digitally. By 2007, that figure rose to 94 %.

ISB

One perspective of the Digital Transformation

Vala Afshar @ValaAfshar · 14h

Both photos of the same place:

1957: 13 men deliver a computer

The Data Collection Story: Some Learnings

- Let's connect the last slide's facts with some from the 2007-2017 timeframe...

- If you consider the rate of content generation today:

  - 6 billion photos uploaded monthly to FB
  - Blogosphere doubles in content volume every 5 months
  - 72 hours of video uploaded onto YouTube every minute
  - 400 million daily tweets on twitter...

- 2 things stand out: (1) Evermore data is generated Year on year.

- (2) Evermore of that data is *native to digital* means of storage, processing, transformation.

ISB

Data Types and Data Dichotomies

ISB

Data Format: Simple Example

- Consider the following data with the SRTC. (Just for illustration)

			Departure			
Date	Route No.	Bus No.	Station	Time	Ticket Revenue	Occupancy
1/7/2017	83	AP 83QRTC	Nellore	1830	6400	80%
2/7/2017	84	AP 83QRTC	Vijaywada	830	6785	85%

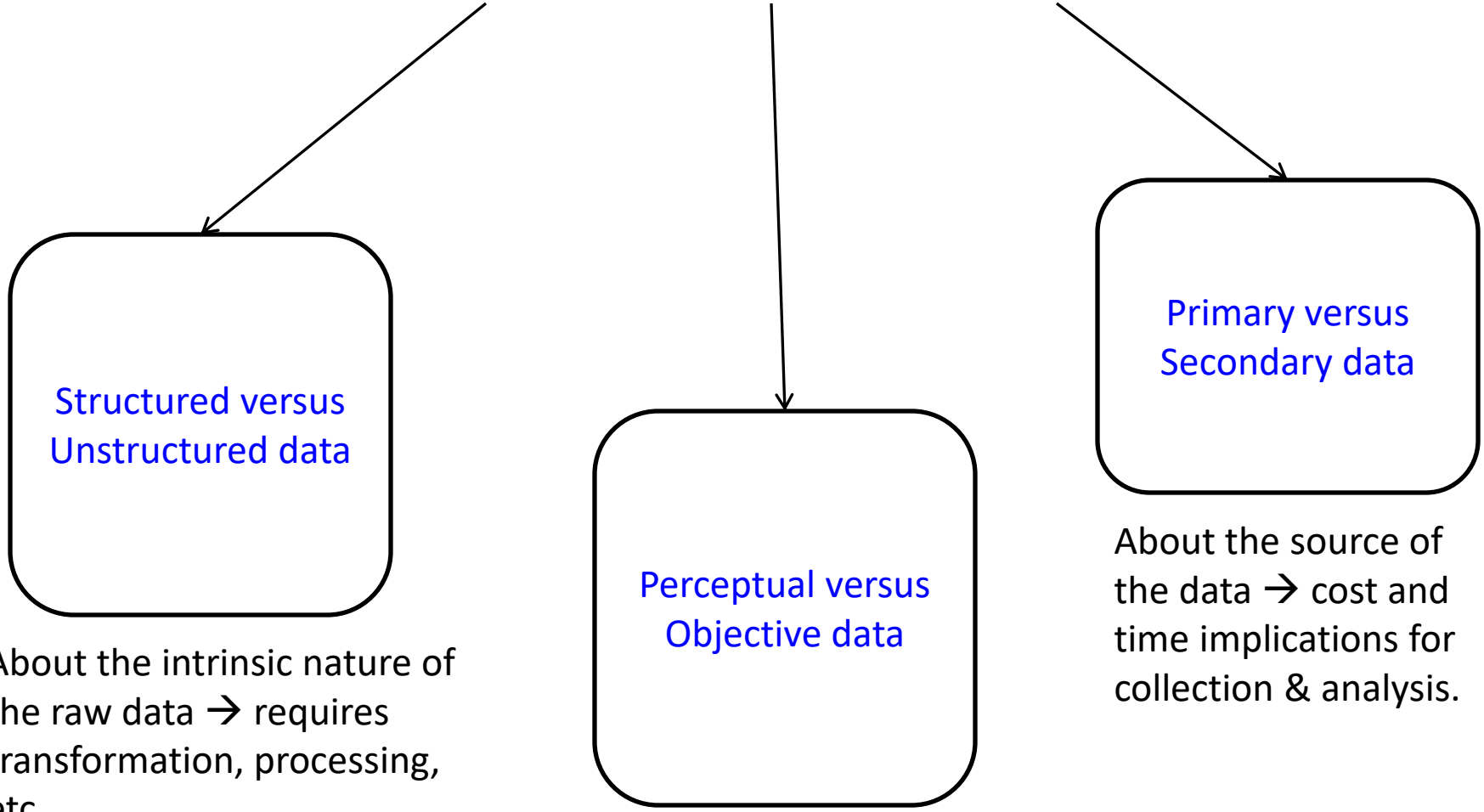
- This is only a small part of the full dataset, which is structured along rows and columns.

- Rows are also called observations, instances, cases etc. Columns are also called variables, attributes, features etc.

- Note the types of data we have present (date, time, names, numbers, percentages etc.).

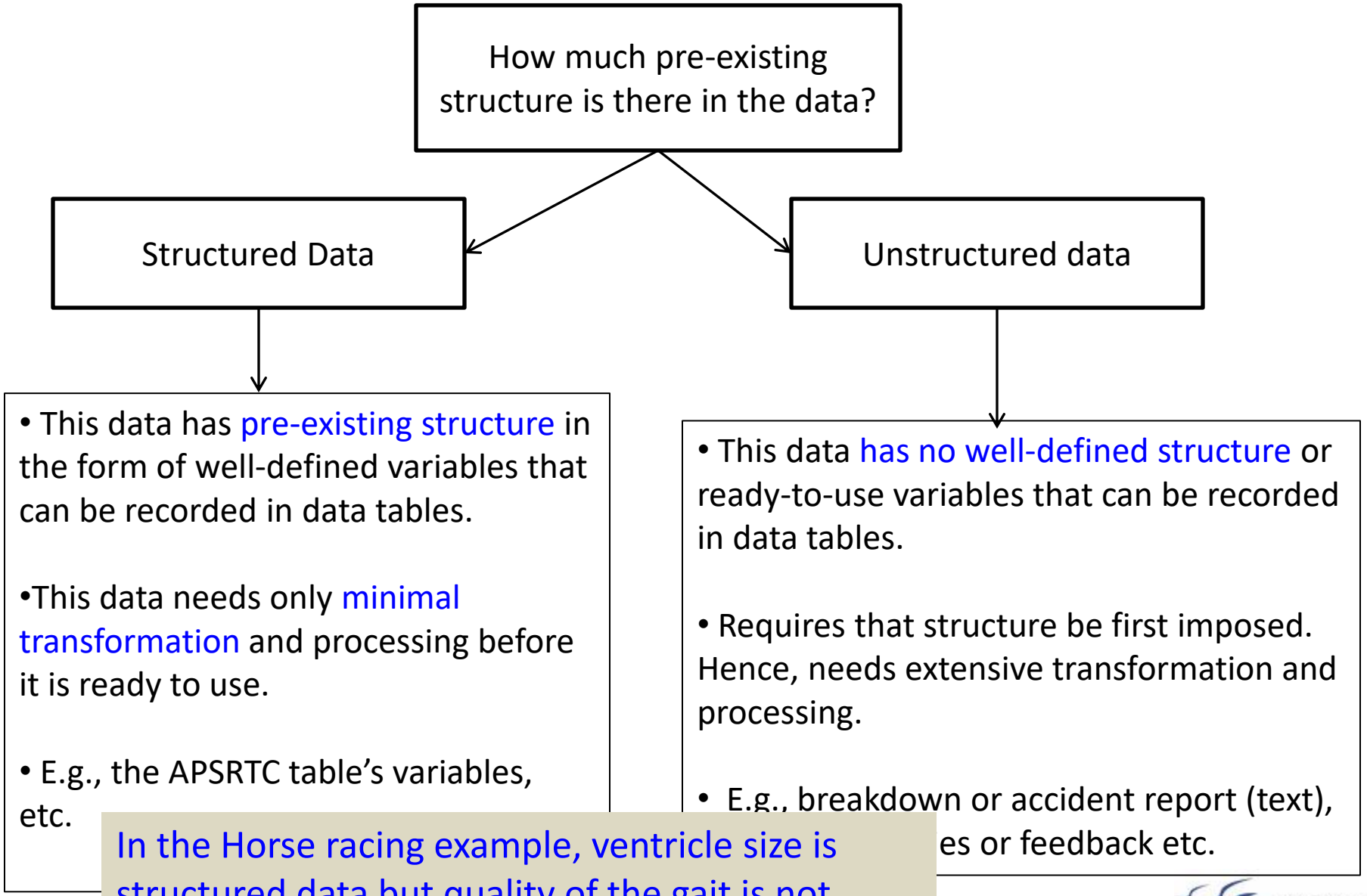
ISB

3 Basic Data Dichotomies



ISB

The Structured Vs Unstructured Data Dichotomy



In the Horse racing example, ventricle size is structured data but quality of the gait is not.

ISB

Quick Q on Structured vs Unstructured Data

- Which of the following data are Structured data - i.e., can directly be used as variables in a dataset? Why or why not?

- (a) Aadhaar fingerprints
- (b) PAN number
- (c) Address on the ration card
- (d) Jan dhan account number
- (e) Scheduled versus actual departure of APSRTC buses
- (f) availability of pulses in Srikakulam's PDS shops
- (g) date of birth on school certificate
- (h) photo on the passport

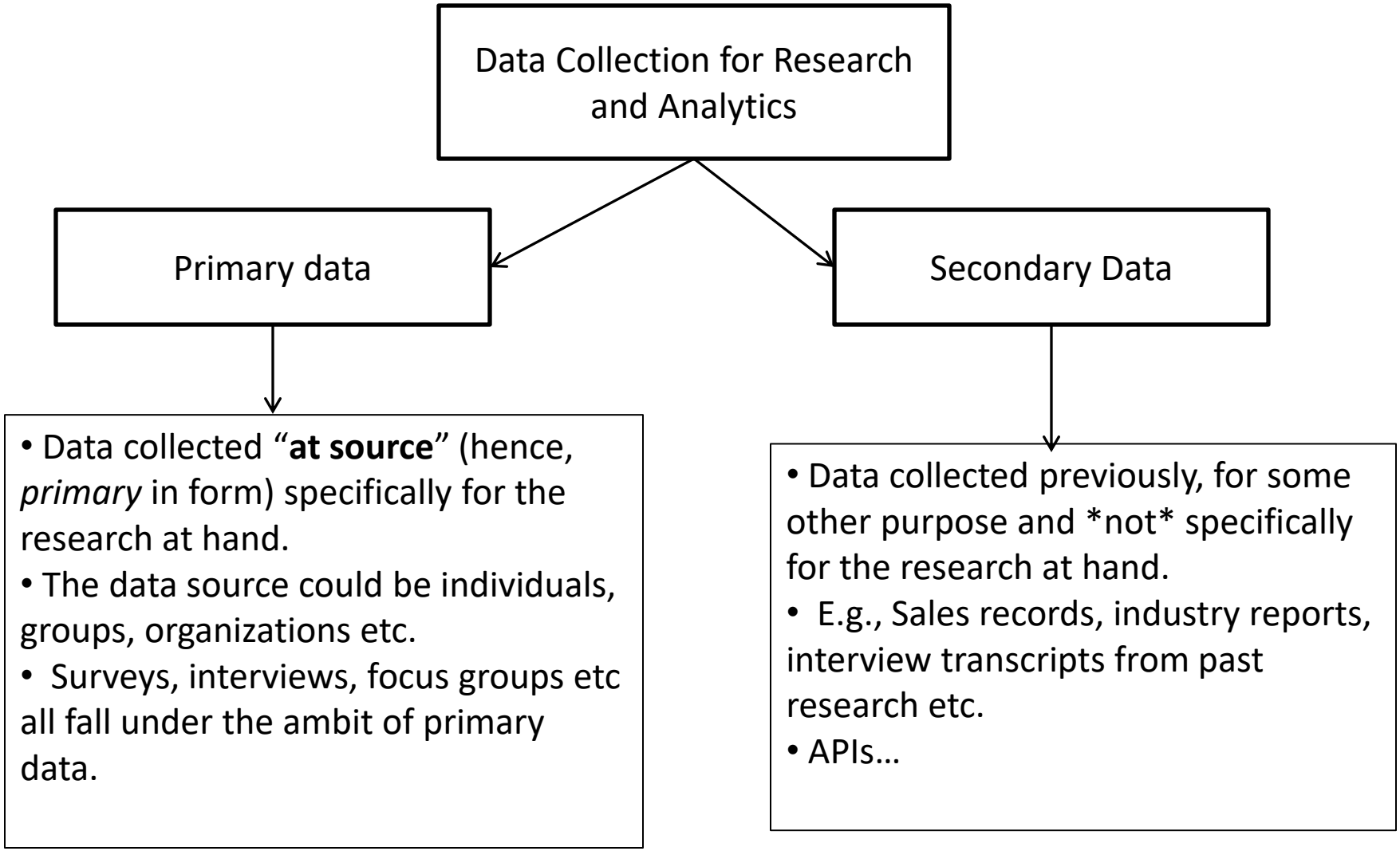
ISB

Perceptual versus Objective data

- **Perceptual Data:**
- Subjective data - about which two people can reasonably disagree.
- E.g., I give Virat Kolhli a 8/10, you give him a 7/10.
- Usually about people's perceptions of quality, service, performance, etc.
- Usually compared to some reference or prior expectations.

- **Objective data:**
- Facts that are independent of subjective perception.
- E.g., Virat's strike rate is 83.3.
- Usually about events measured in physical attributes, space, mass, time etc.

The Primary Vs Secondary Data Dichotomy



ISB

Group Exercise on Data Types & Dichotomies

- Meet as a group and brainstorm on the following: (10 minutes)

- 1. Examples of variables you usually work with - 1-2 for Structured data and 1-2 for Unstructured data.
- 2. What % of your dept's data (rough estimate) is Unstructured data?

- 3. Examples of variables you usually work with - 1-2 for Perceptual data and 1-2 for Objective data.
- 4. What % of your dept's data (rough estimate) is Perceptual data?

- 5. Examples of variables you usually work with - 1-2 from Primary sources and 1-2 from Secondary.
- 6. What % of your dept's data (rough estimate) is Primary data?

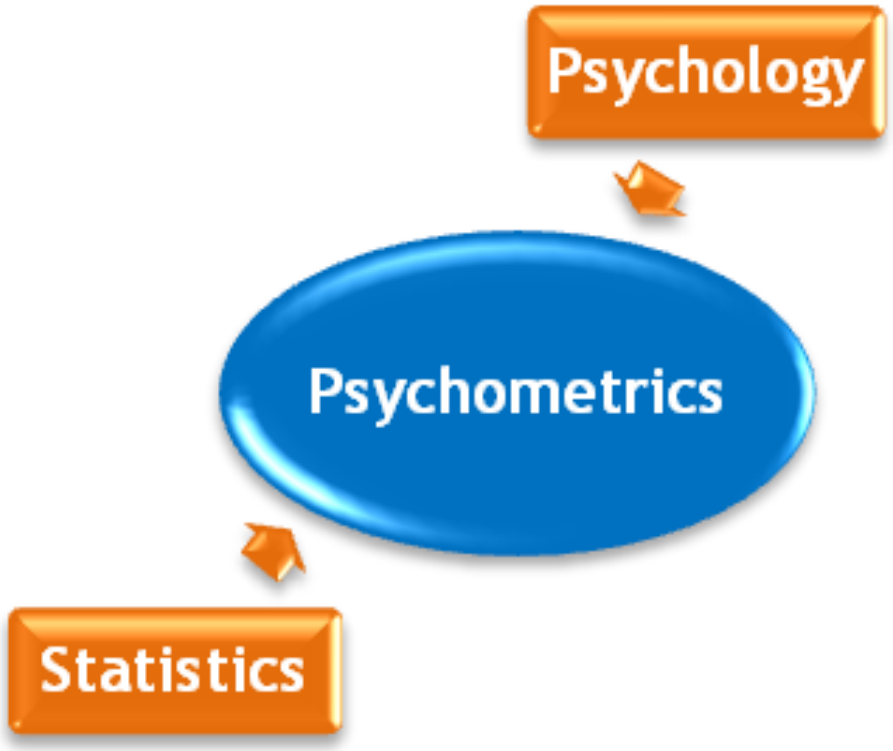
ISB

Thank You

Q & A

ISB

Basics of Psychometric Scaling



ISB

PsyScaling: Four Data Types

- There are 4 types of Data based on the quality of information contained and corresponding to these are 4 primary scales.
- **Nominal**
  - Merely labels. No further information can be gleaned.
  - Example: “Coke” and “Pepsi”.
- **Ordinal**
  - Conveys only upto preference information. Direction alone.
  - Example: “I prefer Coke to Pepsi”.
- **Interval**
  - Conveys relative magnitude information, in addition to preference.
  - Example: “I rate Coke a 7 and Pepsi a 4 on a scale of 10”.
- **Ratio**
  - Conveys information on an absolute scale.
  - Example: “I paid Rs 11 for Coke and Rs 12 for Pepsi”.

PsyScaling: Primary Scales of Measurement

Scale					
Nominal	Numbers Assigned to Runners				**Finish**
Ordinal	Rank Order of Winners	 **Third place**	 **Second place**	 **First place**	**Finish**
Interval	Performance Rating on a 0 to 10 Scale	**8.2**	**9.1**	**9.6**	
Ratio	Time to Finish, in Seconds	**15.2**	**14.1**	**13.4**	

ISB

PsyScaling: Examples of Common Analysis

NOMINAL	ORDINAL	INTERVAL	RATIO
Mode	Mode	Mode	Mode
Frequencies	Median	Median	Median
Percentages	Frequencies	Mean	Mean
	Percentages	Frequencies	Frequencies
	Some Statistical Analysis	Percentages	Percentages
		Variance	Variance
		Standard Deviation	Standard Deviation
		Most Statistical Analysis	Ratio of numbers
			All Statistical Analysis

4 MCQs on the primary Data types.

ISB

PsyScaling: Q1 – On Data scales

- What is the most informative measure possible if you are trying to measure the following *constructs*?
- Choose ONE from (A) Nominal, (B) Ordinal, (C) Interval, (D) Ratio for each of the items below.

  - (i) General Intelligence
  - (ii) Brand image
  - (iii) Consumer attitudes
  - (iv) Social impact of NGOs
  - (v) *Efficiency* of Govt policy in the Shipping sector
  - (vi) *Effectiveness* of Govt Policy.

ISB

PsyScaling: Q2

- Mr Fernando measures favorability of the Airtel brand on a 1-5 scale (higher means more favorable). Jai gives Airtel a 2 whereas Aditi gives it a 4.

- Which of the following statements hold true.

- (A) Airtel is twice as much favored by Aditi as Jai.
- (B) The difference between Jai's and Aditi's ratings is 2 points.
- (C) Jai is not favorably inclined towards Airtel. Aditi is.
- (D) On a 1-9 scale, Jai would have given 4 & Aditi would have given 6.
- (E) Can't say. It depends.

ISB

PsyScaling: Q3

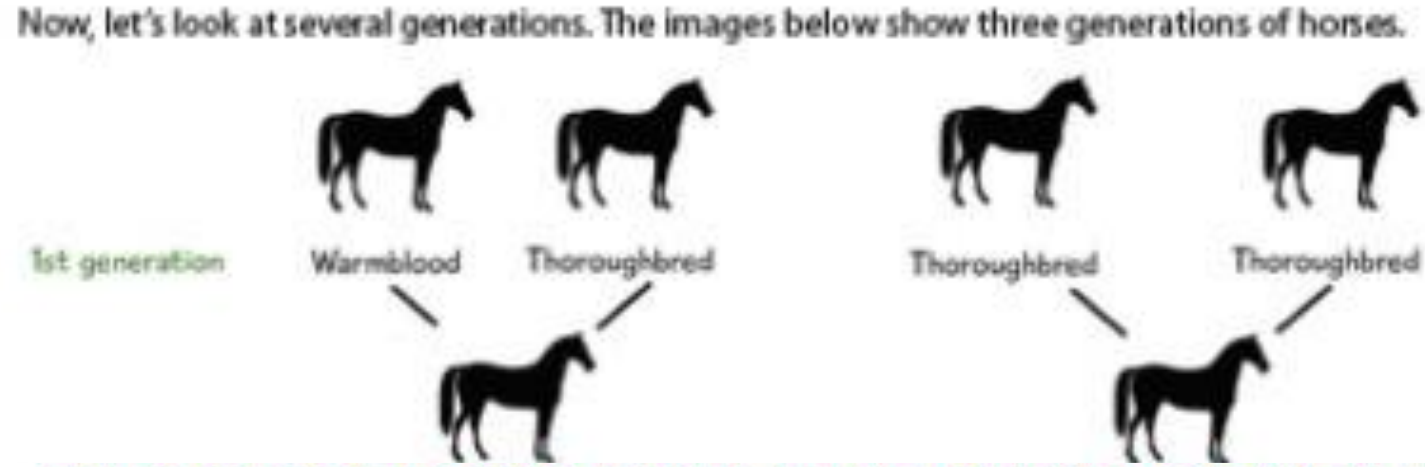
- Mr Fernando measures Airtel usage time in minutes/day. Jai reports an average of 20 minutes whereas Aditi reports an average of 40 minutes.

- Which of the following statements hold true.

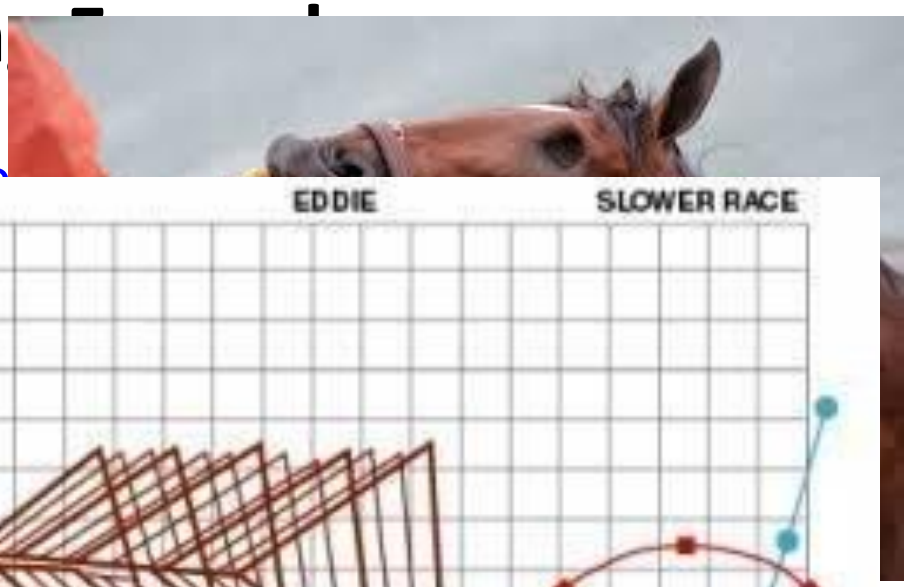
- (A) Airtel is used twice as much by Aditi as by Jai.
- (B) The difference between Jai's and Aditi's avg usage is 20 minutes.
- (C) Aditi uses Airtel more than Jai on any given day.
- (D) Aditi's Airtel bill is higher than Jai's.
- (E) Can't say. It depends.

ISB

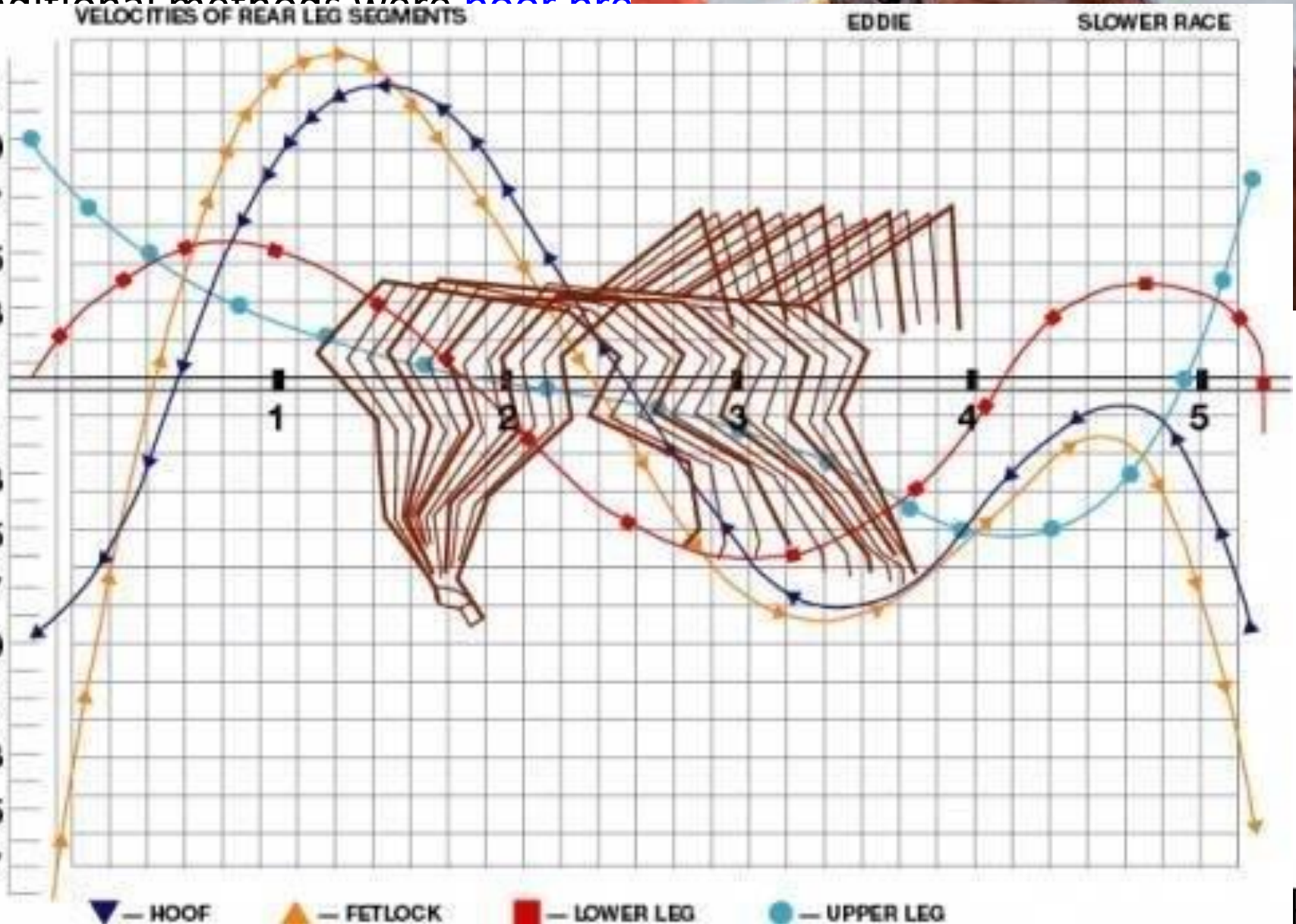
- H
- Enter Jeff Seder of EQB, a boutique consulting firm.



A Motivatin



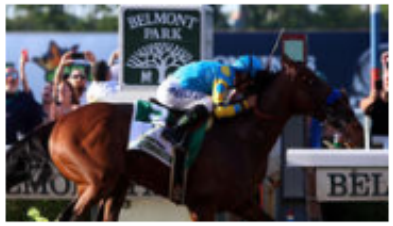
- T
- S
- H
- T
- A



A Motivating Example

American Pharoah wins the Triple Crown

JUNE 5, 2015, 7:08 AM



SPORTS NOW JUNE 6, 2015

American Pharoah wins the Belmont Stakes and Triple Crown

By John Cherwa

The wait is over. History has been made. The sporting world has its newest hero. On a near-perfect Saturday at Belmont Park in Elmont, N.Y., American Pharoah became the first winner of horse racing's Triple Crown in 37 years. He now takes his place as racing royalty by becoming the 12th horse to...

SPORTS NOW JUNE 5, 2015

Why American Pharoah will win the Triple Crown

By John Cherwa

We've stood in this exact position before. Thirteen times, in fact, since Affirmed won the Triple Crown in 1978. But there is something different about this year. Something that makes people believe that this could be the year. OK, it's fair to bring up Einstein's definition of insanity at this...

- haraoh became the first horse in 37 years to win the *Triple Crown*.

ISB

A Motivating Example: Concluded

- So, what is the example trying to motivate?
- [1] **Data is paramount**, when studying, measuring, modeling or understanding any phenomenon of interest.
- [2] Good predictors of an outcome *can* show up in unexpected places - where nobody thought to look.
- [3] Important to keep an eye out for **new tech**, which may enable new data to be collected & analyzed.
- [4] Finding the right set of predictors is challenging - involves **trial-&-error**, guesswork & analytics.
- [5] Data alone is NOT enough. **Analytics is required**, and an open mindset.
- Welcome to an exploration of the fascinating Data + Analytics world.

ISB

Session 1 Recap and Reconnect

- In Session 1, we started with Govt's objectives:

$$\begin{pmatrix}\text{Net Societal}\\ \text{Welfare}\end{pmatrix} = \begin{pmatrix}\text{Consumer}\\ \text{Surplus}\end{pmatrix} + \begin{pmatrix}\text{Producer}\\ \text{Surplus}\end{pmatrix}$$

- ➔ which entailed defining producers and consumers in a Govt dept context

  And importantly, what is the 'good' or 'product' that is being produced

- ➔ which in turn entailed examining data types, forms and dichotomies.

  Structured Vs Unstructured; Perceptual vs Objective; Primary vs Secondary

- Q arises, what if my dept.'s services are such that there maybe no clear 'product'? Hence, no clear producers?

ISB

Session 1 Recap and Reconnect: Concluded

- [1] Definitions are critical: Determine what gets considered vs not. What data types & forms are valid vs not.

- [2] Measurements are critical: Both for outcome variables (net welfare level, good production) and for inputs (all other variables)

- [3] Data collection is critical: Followed by collation, cleaning + processing, Analysis.

- Step 4: How to measure impact of Govt actions on producer & Consumer surplus?

- Given data, analytics tools & algorithms will connect inputs to outcomes → which inputs are relevant vs not in producing outcomes.

ISB

Problem Formulation: Group Exercise

- As a group, pls brainstorm and write down:
- [1] A D.P. for any one of your department's projects or programs
- [2] Map this D.P. to data requirements
- [3] Classify the data required into: (a) structured or unstructured, (b) perceived or objective, and (b) primary or secondary data types.
- 10 Minutes.

ISB

Preliminaries: Essential Equipment

- We'll need one of each for the journey ahead …

Day 1: Primarily about the WHAT, the WHY and the WHEN.

Day 2: Primarily about the HOW and the WHERE.

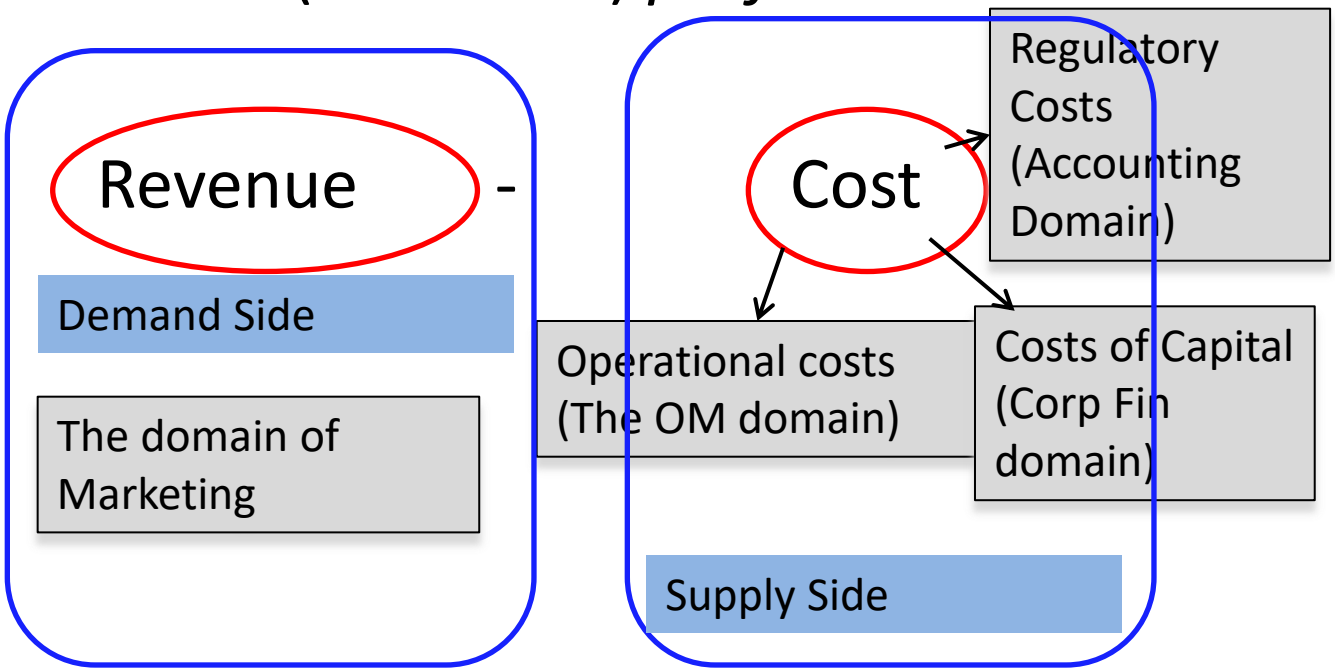
ISB

Preliminaries: Basic Concepts

- This is a Session on Data Analytics for Government officers.
- Q1. How is Business different from Government?
- Q2. What is a ‘business’? What does it do?
- Q3: What is Government? What does it do?

Preliminaries: The Objective of a Business

- Firms exist to maximize (economic) *profits*
- Profit = Revenue - Cost



- Business functions represent a logical way to deconstruct the enterprise → yield analytics that is function-specific.

- Market power derives from competencies on either the demand or the supply side.

ISB

2018 世界人工智

there
mple,

WORI

VENTURE CAPITAL INVESTMENT INTO AI

The difference between China and the US in AI investment is big. The US has already invested heavily, while China is now catching up. Since the first US investment into AI in 1999, AI development has accelerated globally. In 18 years, total venture capital invested into AI has reached $28.9 billion.

– We'r

On Data today

- The volume, variety and velocity (the famous three Vs of big data) of the data currently being captured is unprecedented.

- In the time it takes you to read this sentence (~ 6 seconds for the average reader), Google receives half a million queries from around the world.

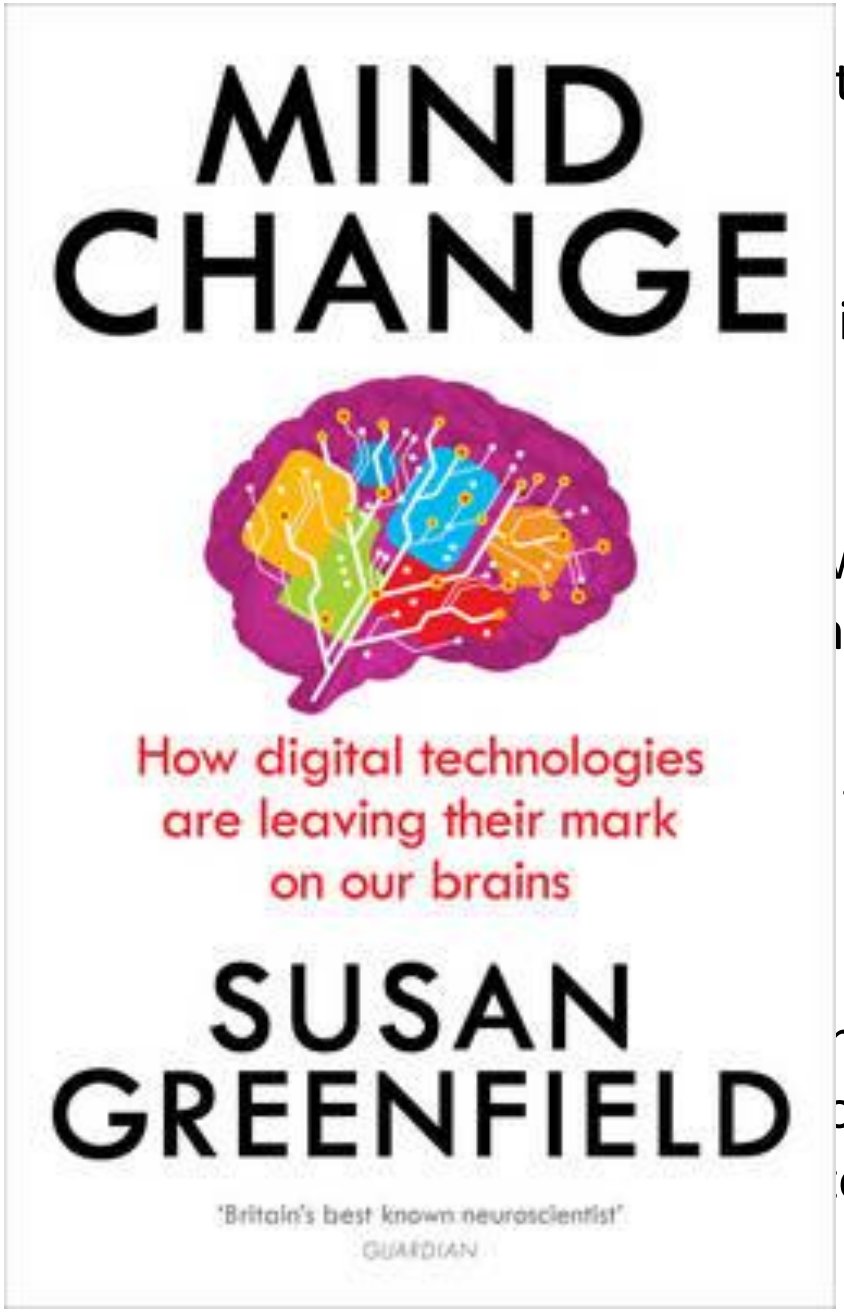
- In 2000, digitally stored data was a mere 25% of all data generated. By 2007, it jumped to 94% (and hasn't fallen since).

- Traditionally, Data analysis (say, D.A.) would adapt to whatever data form was available --> D.A. adapted to D.C. (Data Collection) --> In turn, D.C. adapted to Data Generation (say, D.G.).

- But the jump from Y2K to 2007 suggests something way more profound.... that perhaps D.G. is adapting to D.C. is adapting to D.A.?

ISB

Data and the Human Mind



- te down anything - pen to paper - to
- il, g de
- we k hat
  line
- func our
- now
  on, time horizon perceptions, value
  o the web remains to be seen.

Consider the effect of always-on social network access, binge-consumption of video games, audiovisual entertainment etc. and technologies to come…

Q is – are they changing children's brains? Rewiring circuits, coping and reward mechanisms? How about adult brains?

ISB

Some Report Writing

Best Practices

ISB

Report Writing: Typical Structure

- All reports will have 3 broad parts: Beginning, Middle and End.
- A best practice is to include a fourth part at the very beginning: The Executive Summary.
- The Executive summary is less than a page long and addresses the following :
- [1] Who is the audience for the report?
- [2] What are the objectives of the report?
- [3] A preview of main findings and conclusions.

ISB

Report Writing: Tying it all in Together

- What we discussed in the session today:

- [1] Appreciating the value of Data
- [2] Appreciating the value of Questions and Problem Formulation
- [3] Appreciating the process of Analysis

- All come together to form a complete report.

- Reports should ideally (and perhaps counter-intuitively) be:
  - Short (drop all non-relevant parts)
  - Simple (e.g., by being Factual , using simple words)
  - Complete (have a references section, data sources named in footnotes etc)
  - Actionable (e.g., set of recommendations, cost estimates etc.)

ISB

Preliminaries: The Objectives of Government

- What should government aim for?
- $$\begin{pmatrix}\text{Net Societal}\\ \text{Welfare}\end{pmatrix} = \begin{pmatrix}\text{Consumer}\\ \text{Surplus}\end{pmatrix} + \begin{pmatrix}\text{Producer}\\ \text{Surplus}\end{pmatrix}$$

  Ease of citizenry to improve consumption → living standards, at a given price level.

  Ease of business to improve production, productivity → profit, at a given price level.

- There is a *tradeoff* between consumer and producer surpluses. If social welfare is constant then raising one means lowering the other.

- Extent of control by government gives us different systems.

ISB

Preliminaries: Examples of Social Welfare Maximization

- To attain Govt's objectives, Govt actors must first identify 3 things:

- (1) What is the '*product*' produced by our department?
- (2) Who are the *producers* related to our department?
- (3) Who are the *consumers* related to our dept?

- Take an example of the Urban Traffic management department. Or the education dept. Or the *Home affairs* department.

- Who are the producers in this dept.? Consumers?

- How can we evaluate Govt policies and programs from a social welfare maximization perspective?

ISB

Class Exercise: The Police Department Example

- Consider (say) the **Police dept**.
- Step 1: What is the 'good' or product the dept. works with?

  e.g., Assurance of security, order and rule of law
- Step 2: Who are the producers? What is their surplus?

  e.g., Police of course + *all* law-abiding citizens. Form of surplus could be psychological, monetary, reputational etc.
- Step 3: Who are the consumers? What is their surplus?

  e.g., All residents incl. businesses, non-citizens, etc. Form of surplus could be investments, wealth generation, lower insur. premiums etc.
- Step 4: Govt actions that impact producer surplus? Consumer surplus?

  Incl. both incentives and disincentives. Examples?
- Once we have defined the above quantities, net social welfare can be measured --> modeled --> maximized (in principle).

ISB

Class Exercise: Measuring a Dept's Inputs & Outputs

- Take the Police Dept. example.
- Step 1: How to measure the 'good' or product the dept. works with?

  'feeling of security' is perceptional. Periodic surveys? [Social] Media chatter? etc.

- Step 2: Who are the producers? → How to measure their surplus?

  Form of surplus could be psychological (perceptual through surveys etc?), monetary (objective), reputational (perceptual again) etc.

- Step 3: Who are the consumers? →How to measure their surplus?

  Form of surplus could be investments, wealth generation, lower insur. premiums ((objective) etc.

- Leads us to think about data manifestations of even abstract, intangible quantities.
- Step 4: How to measure impact of Govt actions on producer & Consumer surplus? *

ISB

Learnings from the Group Exercise

- Some Qs we can now look back upon and ponder.
- Q: How easy or difficult is it to identify the producers and consumers?
- Q: How easy or difficult is it to identify the Govt policies and regulations that affect the above?
- Q: What data would help make it even more easier to *systematically* answer the above Qs?
- Q: Do we have that data with us already? Or must it be collected? What form is it in?
- Q: How can we analyze the data to easily, rapidly, systematically answer the Qs we put?

ISB