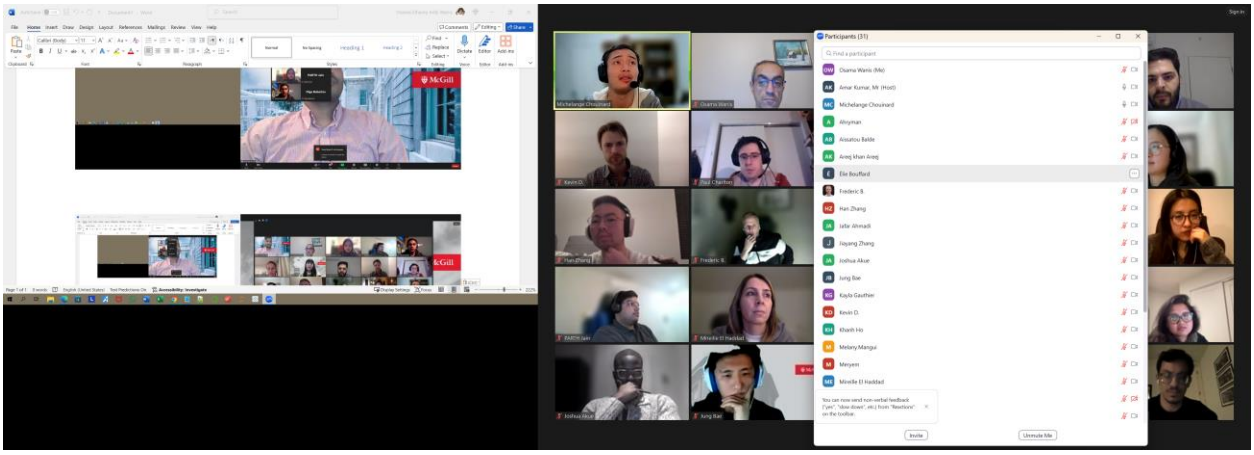
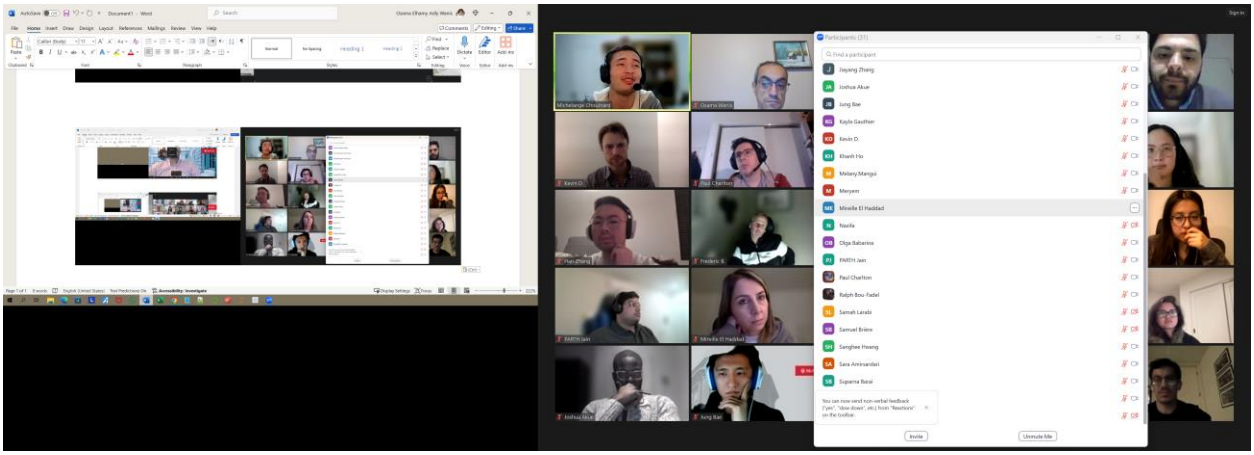
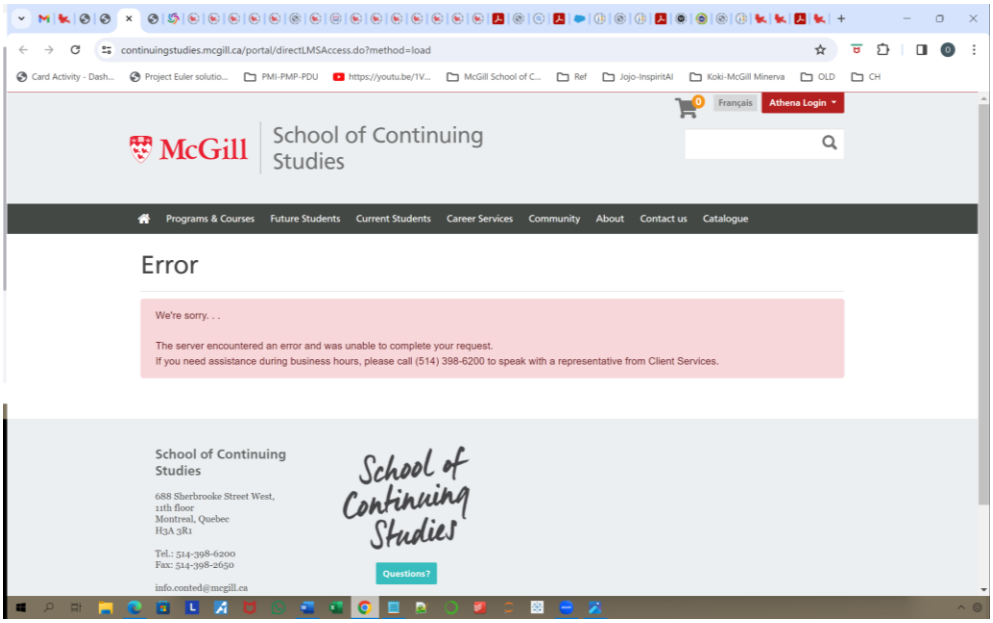


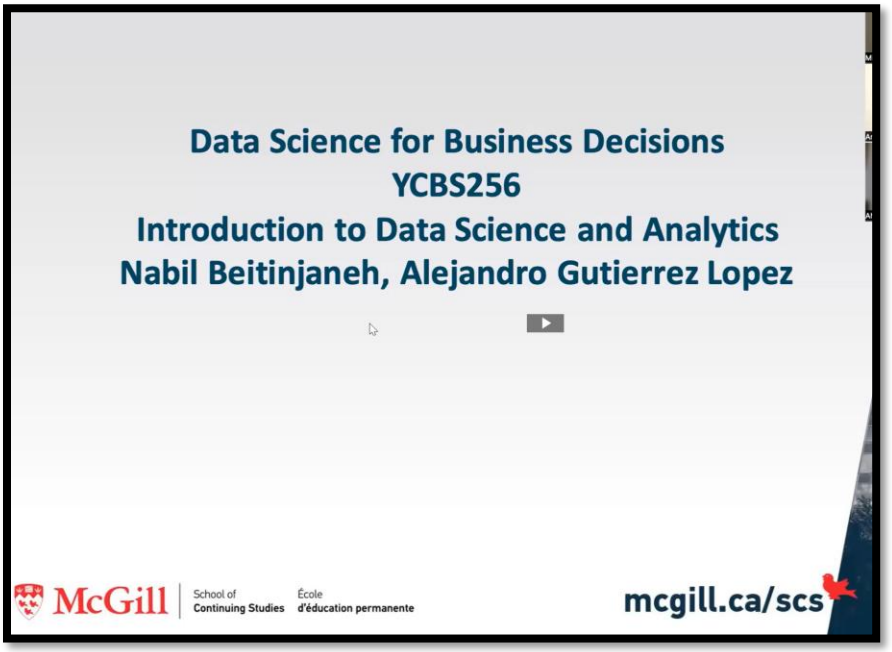


what is phone # and email address for IT

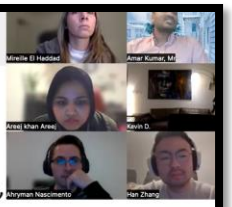
Michelange Chouinard 7:29 PM
514-398-6200

extension 1 > 4 > 2 if I remember
askd.scs@mcgill.ca



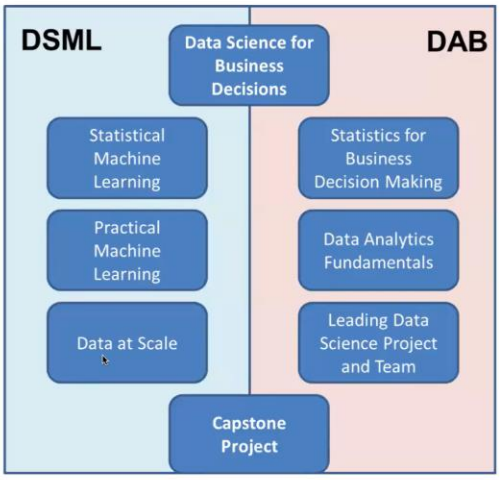


Session Overview



- Class Introductions
- Course fit within the overall program
- Course logistics (preparation, synchronous & synchronous sessions,
  - myCourses
  - The Course Outline
  - Assessments, assignments, participation
  - Tools (Alteryx and Tableau)
  - DataCamp access
  - Computer Lab 1285
  - How to reach us
- Why Data Science and Analytics?

Recommended Path(s) for McGill Certifica



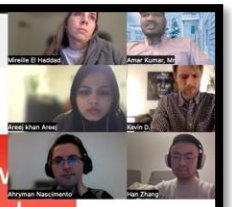
Skills

Please refer to the Study Plan document shared in the McGill One Drive

Study Plan

Data Science & Machine Learning | Data Analytics for Business

Advantages



Q If you were a top long distance runner, how much are you willing to pay to buy a special pair of shoes that will shave off a couple of minutes of a marathon?

Advantages

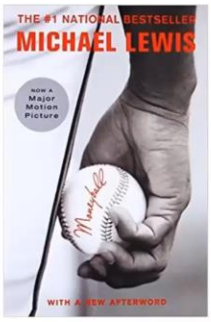
Q If you were a top long distance runner, how much are you willing to pay to buy a special pair of shoes that will shave off a couple of minutes of a marathon?

A Such a shoe exists…it's called the Vaporfly.

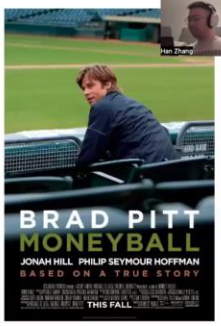
Do Nike's New Shoes Give Runners an Unfair Advantage?
Do Nike's New Marathon Shoes Actually Make you Run Faster?

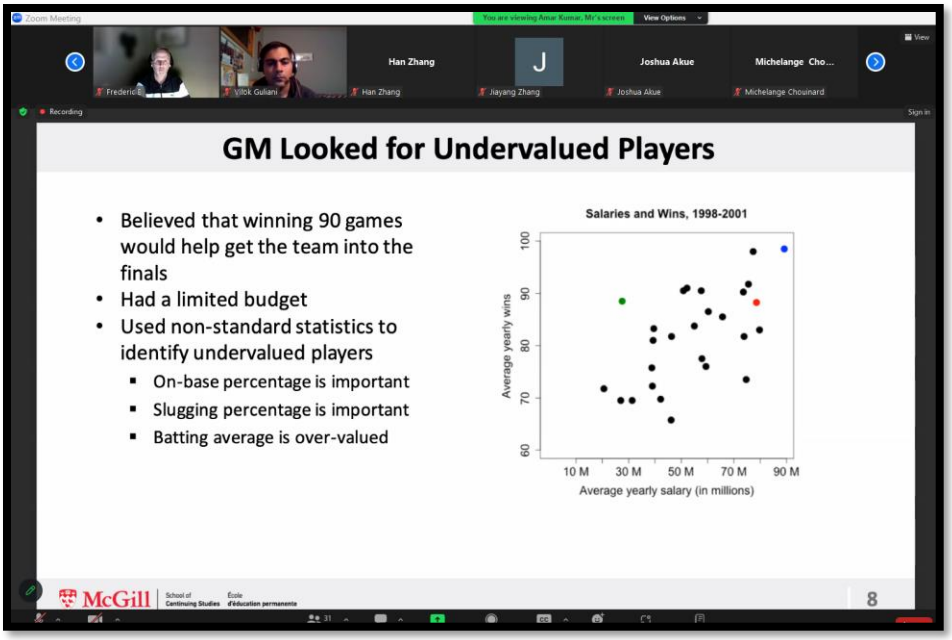
Why Data Science?

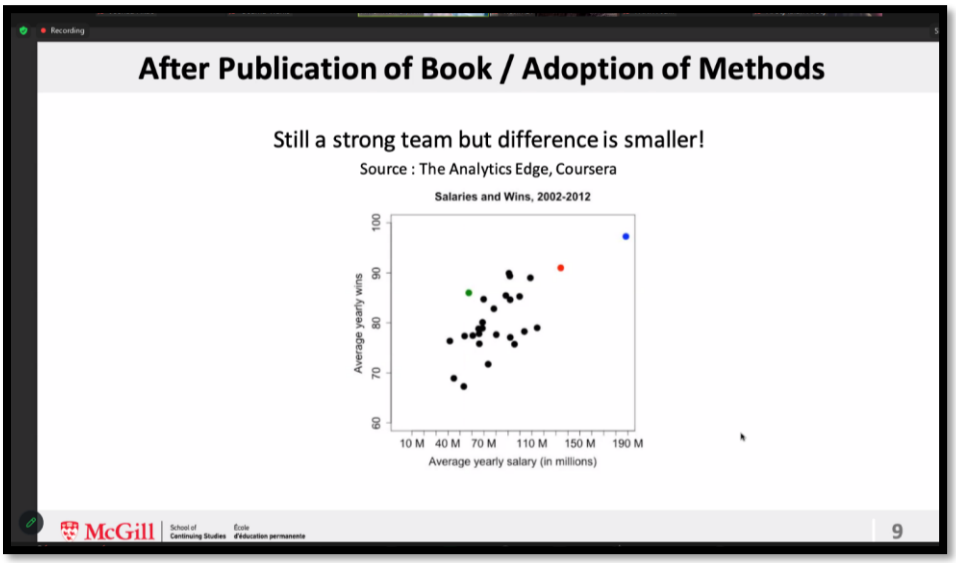
2003

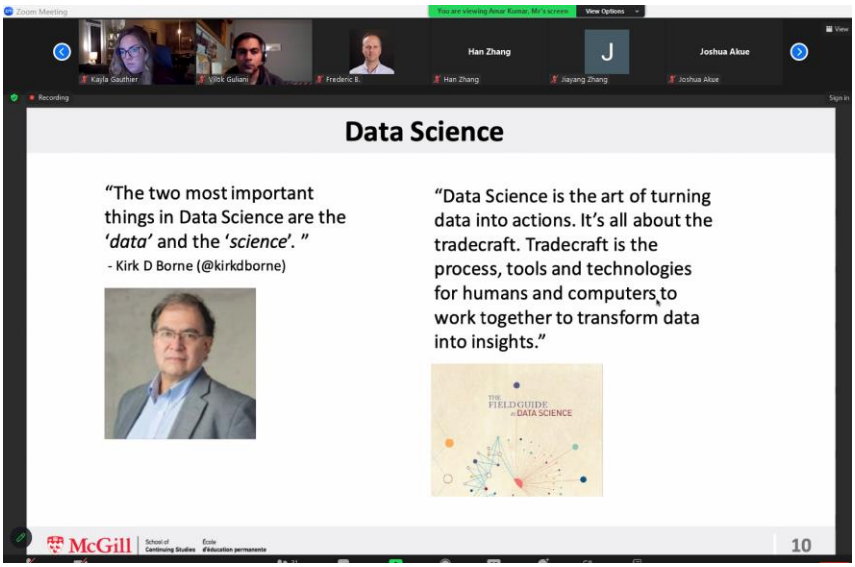


2011

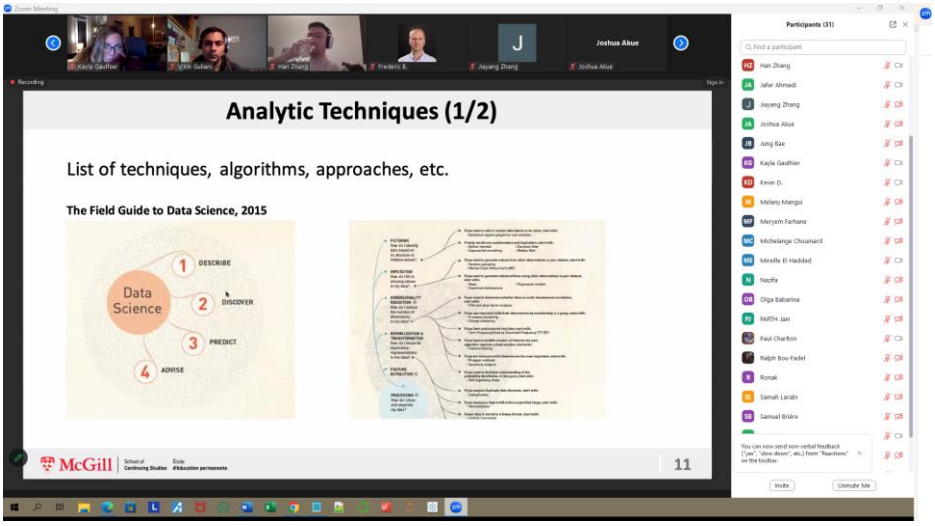


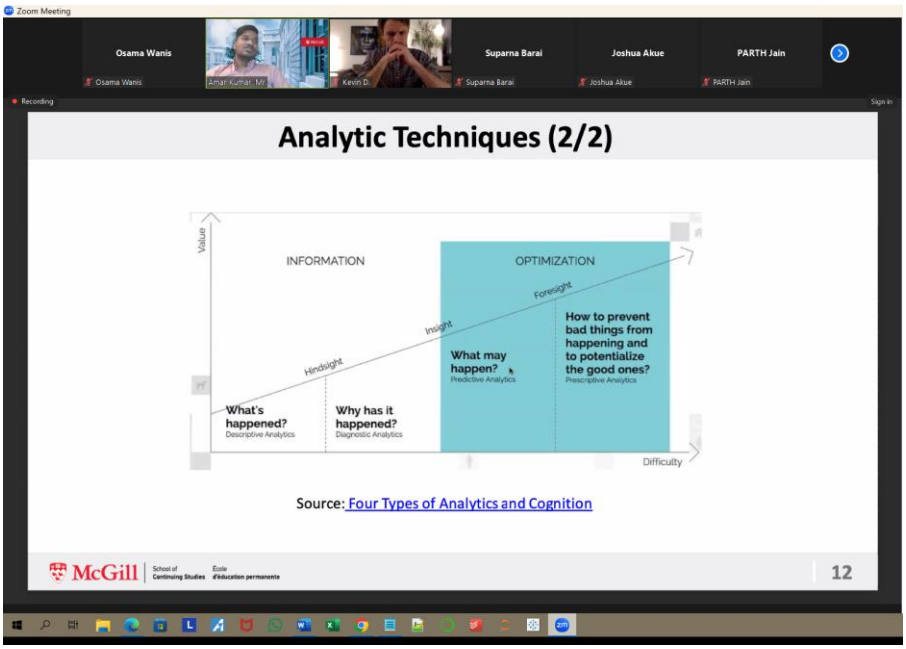

Zoom Meeting
You are viewing Amar Kumar, Mr's screen
View Options
Han Zhang
Joshua Akue
Michelange Cho...
Recording
Sign in
GM Looked for Undervalued Players
• Believed that winning 90 games would help get the team into the finals
• Had a limited budget
• Used non-standard statistics to identify undervalued players
▪ On-base percentage is important
▪ Slugging percentage is important
▪ Batting average is over-valued
Salaries and Wins, 1998-2001
Average yearly wins
60
70
80
90
100
10 M
30 M
50 M
70 M
90 M
Average yearly salary (in millions)
McGill
School of Continuing Studies
École d'éducation permanente
8

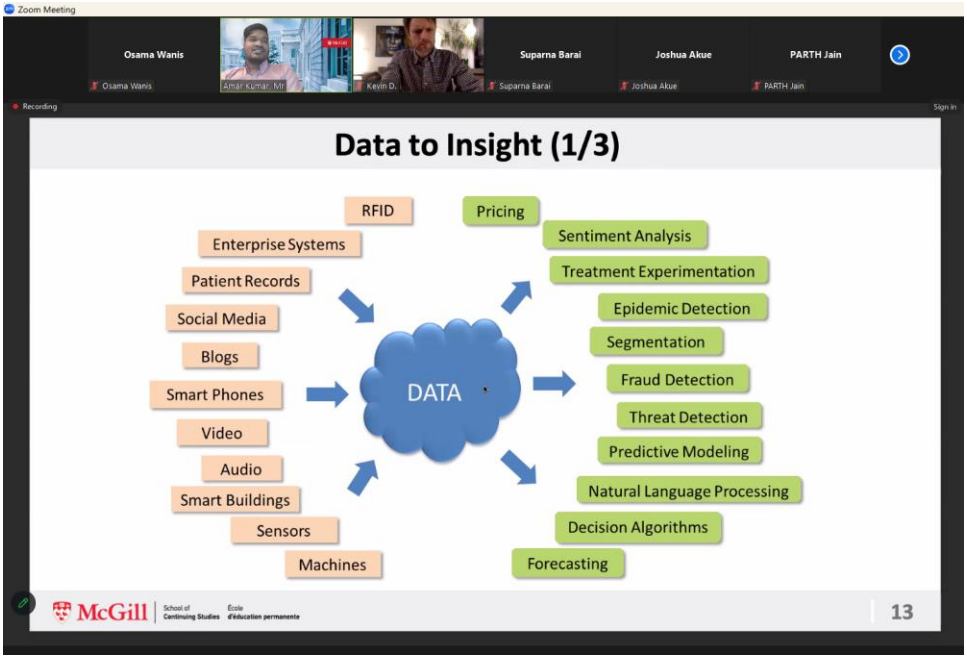

Recording
After Publication of Book / Adoption of Methods
Still a strong team but difference is smaller!
Source : The Analytics Edge, Coursera
Salaries and Wins, 2002-2012
Average yearly wins
60
70
80
90
100
10 M
40 M
70 M
110 M
150 M
190 M
Average yearly salary (in millions)
McGill
School of Continuing Studies
École d'éducation permanente
9

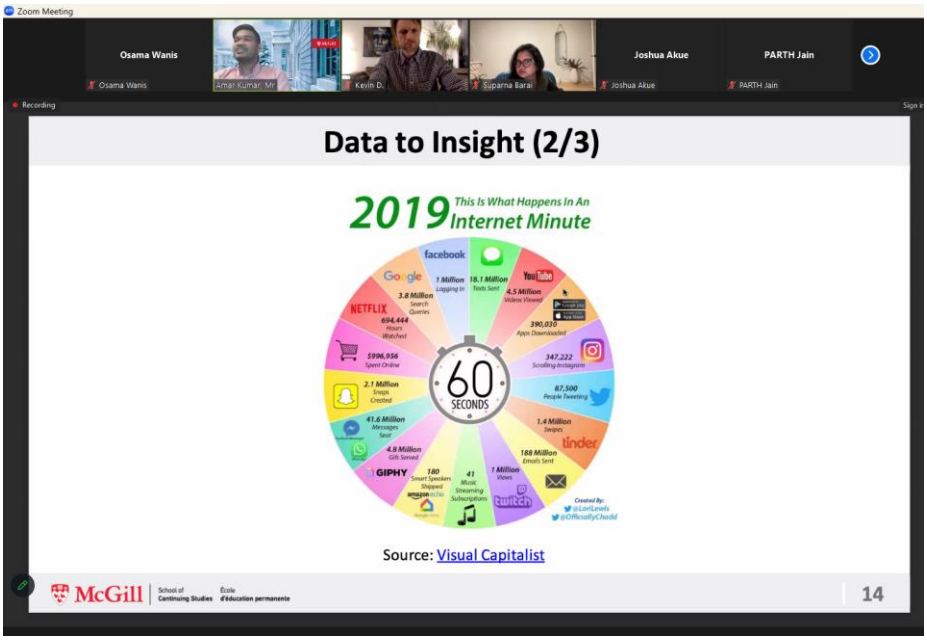


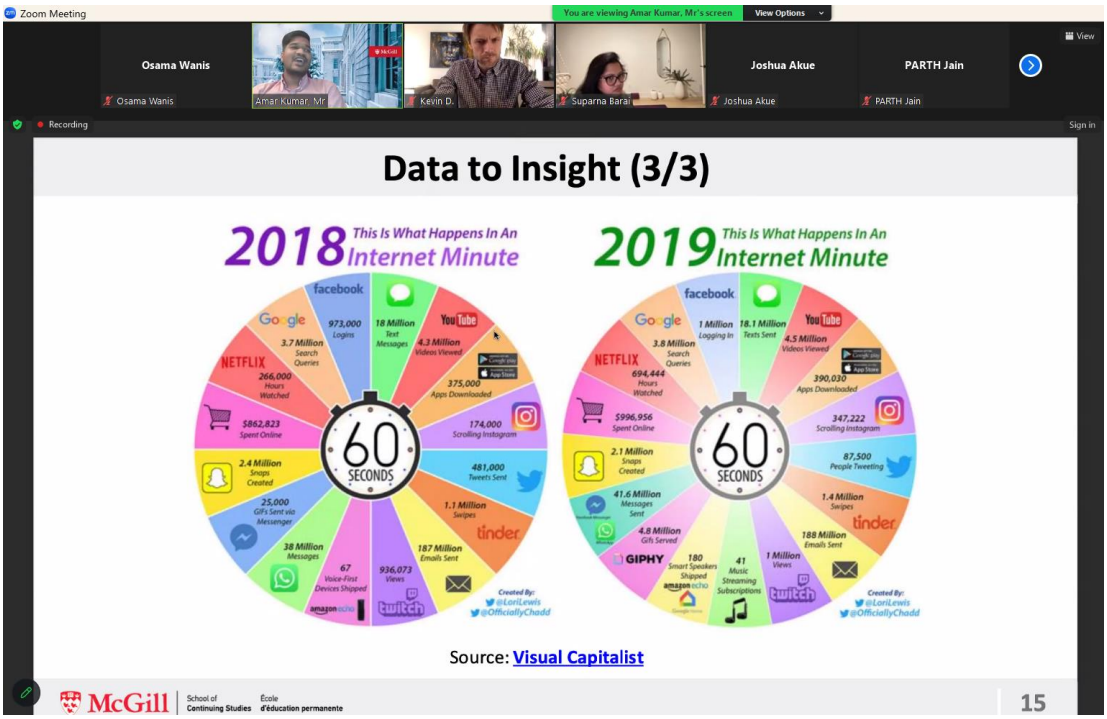
The book in the mcgill library I can access with mcgill ID

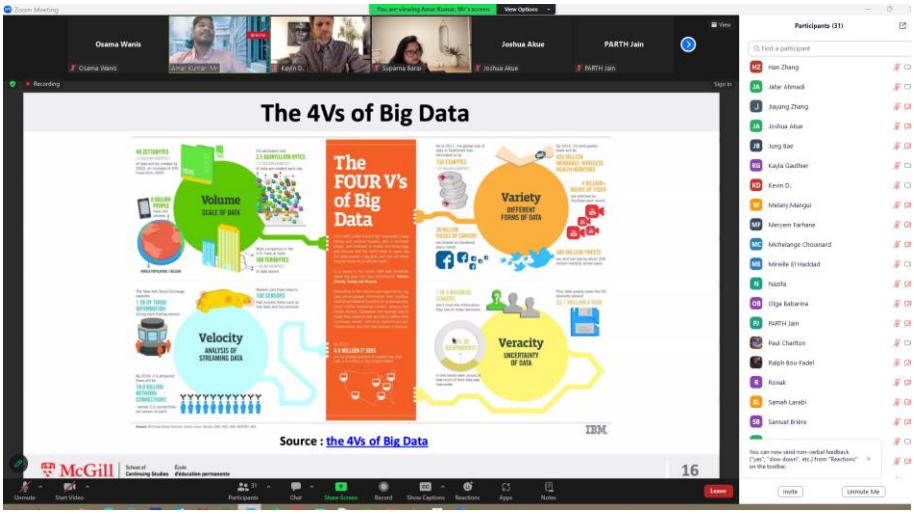


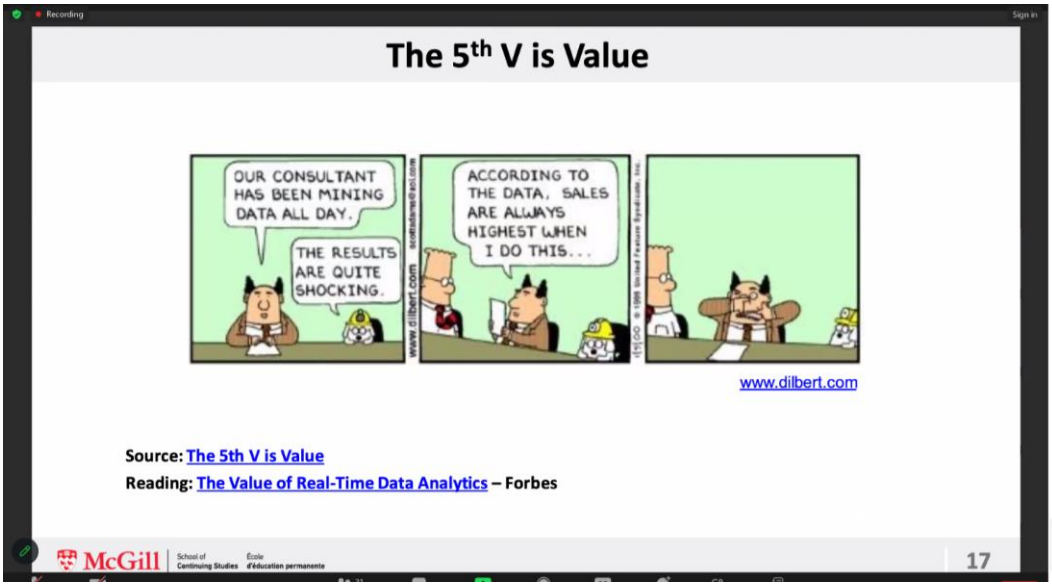

Zoom Meeting
Osama Wanis
Suparna Barai
Joshua Akue
PARTH Jain
Recording
Analytic Techniques (2/2)
Value
INFORMATION
OPTIMIZATION
Foresight
Insight
Hindsight
What may happen?
Predictive Analytics
How to prevent bad things from happening and to potentialize the good ones?
Prescriptive Analytics
What's happened?
Descriptive Analytics
Why has it happened?
Diagnostic Analytics
Difficulty
Source: Four Types of Analytics and Cognition
McGill
12

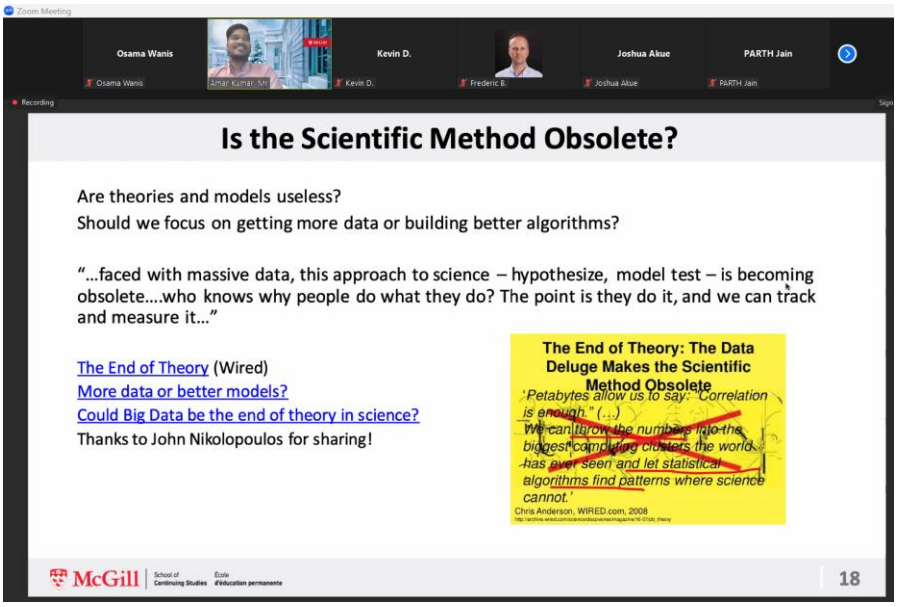

Zoom Meeting
Osama Wanis
Suparna Barai
Joshua Akue
PARTH Jain
Recording
Data to Insight (1/3)
RFID
Enterprise Systems
Patient Records
Social Media
Blogs
Smart Phones
Video
Audio
Smart Buildings
Sensors
Machines
DATA
Pricing
Sentiment Analysis
Treatment Experimentation
Epidemic Detection
Segmentation
Fraud Detection
Threat Detection
Predictive Modeling
Natural Language Processing
Decision Algorithms
Forecasting
McGill
13


Zoom Meeting
Osama Wanis
Joshua Akue
PARTH Jain
Recording
Data to Insight (2/3)
2019 This Is What Happens In An Internet Minute
60 SECONDS
Source: Visual Capitalist
McGill
School of Continuing Studies
École d'éducation permanente
14

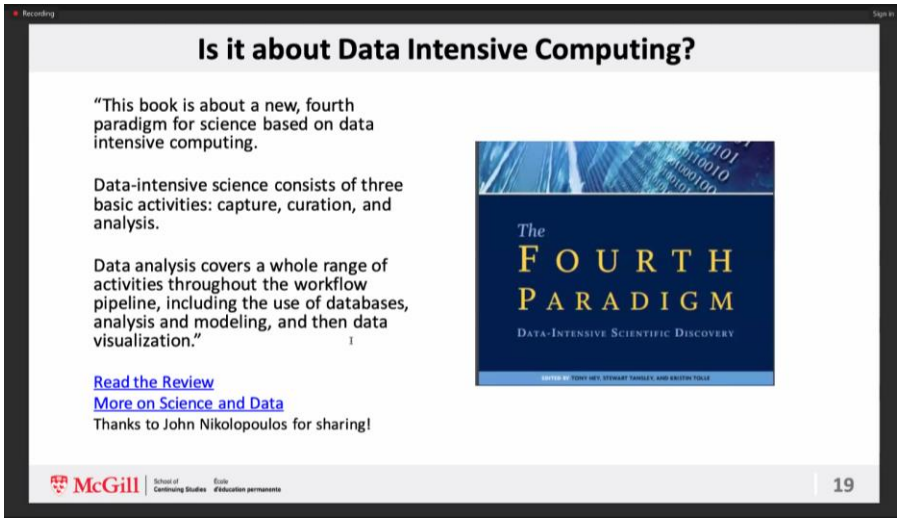

Zoom Meeting
You are viewing Amar Kumar, Mr's screen
View Options
Osama Wanis
Joshua Akue
PARTH Jain
Recording
Data to Insight (3/3)
2018 This Is What Happens In An Internet Minute
2019 This Is What Happens In An Internet Minute
60 SECONDS
Source: Visual Capitalist
McGill
School of Continuing Studies
École d'éducation permanente
15


Osama Wanis
Joshua Akue
PARTH Jain
Participants (31)
The 4Vs of Big Data
The FOUR V's of Big Data
Volume
SCALE OF DATA
Variety
DIFFERENT FORMS OF DATA
Velocity
ANALYSIS OF STREAMING DATA
Veracity
UNCERTAINTY OF DATA
IBM
Source : the 4Vs of Big Data
McGill
16


Recording
Sign in
The 5th V is Value
OUR CONSULTANT HAS BEEN MINING DATA ALL DAY.
THE RESULTS ARE QUITE SHOCKING.
ACCORDING TO THE DATA, SALES ARE ALWAYS HIGHEST WHEN I DO THIS...
www.dilbert.com
Source: The 5th V is Value
Reading: The Value of Real-Time Data Analytics – Forbes
McGill
17



Check the post/sharing by John Nikolopoulos



Check also the above

Is there a Unifying Theory?

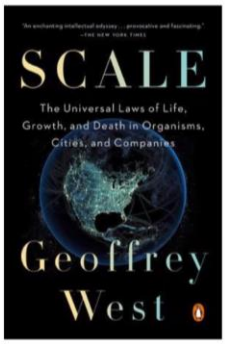
"The basic mathematical laws of physics governing growth in the physical world apply equally to biological, political and corporate organisms. On its face, his book's objective is to contribute to an overarching behavioral science of what it calls *highly complex systems*."

~ Jonathan A. Knee

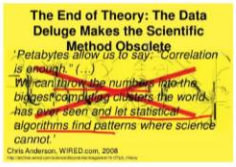
From the New York Times

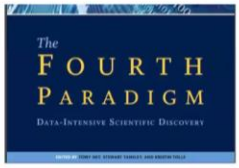
The hidden maths of organisms, cities and companies (The Economist)

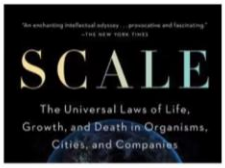
Thanks to John Nikolopoulos for sharing his reading list!



What 'Science' are we talking about?







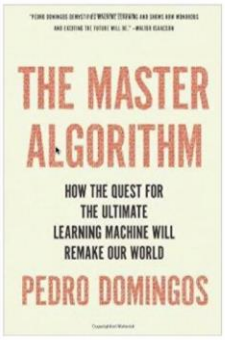
Where does Data Science Fit on this spectrum?

Now for Machine Learning

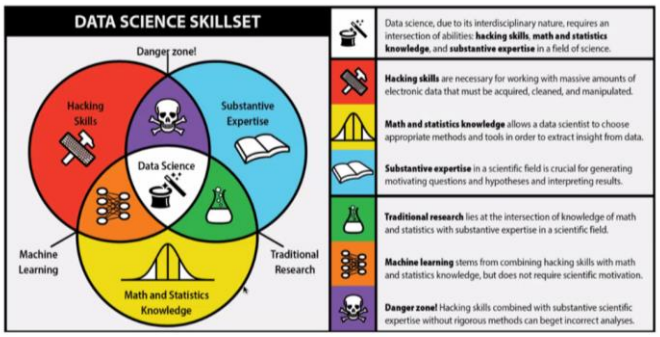
The Master Algorithm "*is an introduction to machine learning for a non-technical audience. It describes the five major paradigms in machine learning - symbolism, connectionism, evolutionism, Bayesianism and analogism - including the main ideas, the history, the leading researchers, applications, strengths and weaknesses, etc. It's framed as a quest for the eponymous master algorithm - a general-purpose learner capable of learning anything from data. In that search, we visit each of the five schools of thought to see what we can learn from it, and then we try to put all the pieces together into one overarching algorithm that is not the master algorithm yet, but hopefully brings us a step closer.*"

~ Pedro Domingos

The Master Algorithm

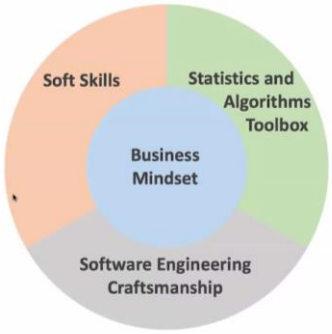


Data Scientists



Source: How to Become a Data Scientist Before you Graduate

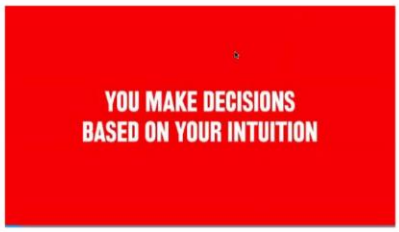
The Third Wave Data Scientist



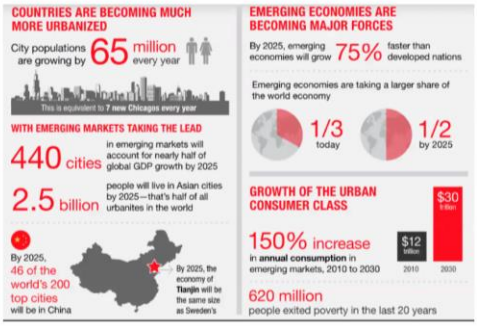
Source: The Third Wave Data Scientist

No Ordinary Disruption: Class Discussion

The 4 Global Forces breaking All the Trends

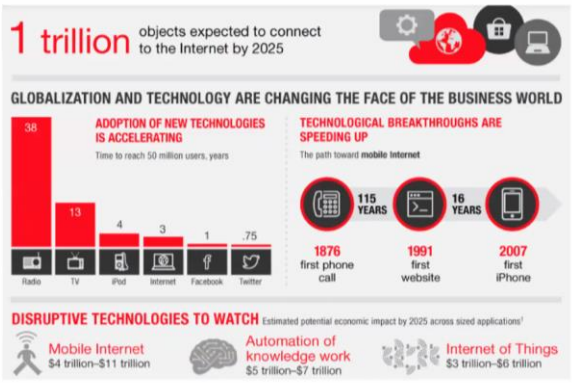


Disruption 1: The Age of Urbanization



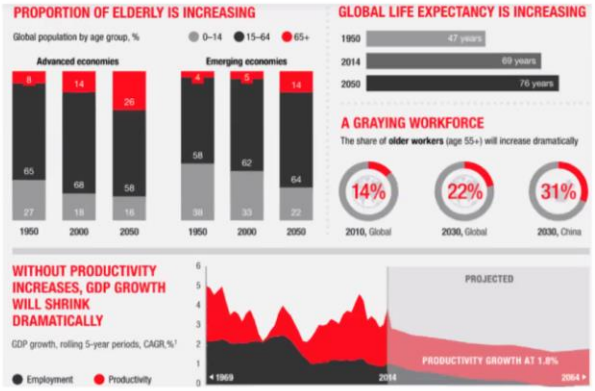
McKinsey web site

Disruption 2: Accelerating Technological Change



McKinsey web site

Disruption 3: Challenges of an Aging World



McKinsey web site

What Strategy Is: Discussion

Michael Porter : What Strategy Is?

Traditional vs. Big Bang Strategy

Big Bang Disruption Topples Traditional Business Strategy

Recording Sign in

Data Strategy

The Data-Strategy Spectrum

A company's industry, competitive and regulatory environment, and overall strategy will inform its data strategy.

FROM "WHAT'S YOUR DATA STRATEGY?," BY LEANDRO DALLEMULE AND THOMAS H. DAVENPORT, MAY–JUNE 2017 © HBR.ORG

Source:"What's your data strategy?"
by Leandro Dallemule and Thomas H. Davenport (2017)
Harvard Business Review, 95(3), 112-121

The Elements of Data Strategy

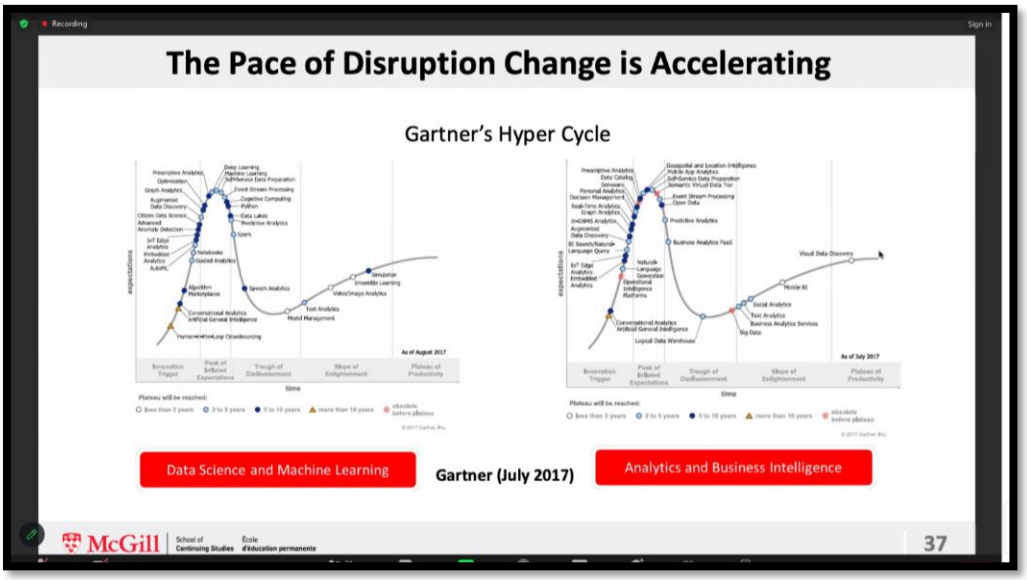
	DEFENSE	OFFENSE
KEY OBJECTIVES	Ensure data security, privacy, integrity, quality, regulatory compliance, and governance	Improve competitive position and profitability
CORE ACTIVITIES	Optimize data extraction, standardization, storage, and access	Optimize data analytics, modeling, visualization, transformation, and enrichment
DATA-MANAGEMENT ORIENTATION	Control	Flexibility
ENABLING ARCHITECTURE	SSOT (Single source of truth)	MVOTs (Multiple versions of the truth)

Source:"What's your data strategy?"
by Leandro Dallemule and Thomas H. Davenport (2017)
Harvard Business Review, 95(3), 112-121

Various Models Used by Business

- **Freemium model :** Basic products/services are provided free but users are charged for advance features
- **Pay as you go or Subscription Model :** Pay only for services which are used
- **Customer experience model :** provide the customer experience never before
- **On-Demand model** : provide customer service on demand with speed
- **Marketplace model :** provide a platform for buyer and seller interact with each other directly
- **Free model :** provide the typical services to users free and sell their behavior data to different businesses
- **Crowd-sourcing model :** receive money for engaging crowd for common goal, innovation, problem solving
- **Bundling model :** selling similar products or services together
- **Gamification model :** use of game like feature to simplify the interaction

Source: All You Need To Know About Business Models In Digital Transformation – Sandeep Raut



The Unicorn Club – Top 20 Private Tech Companies

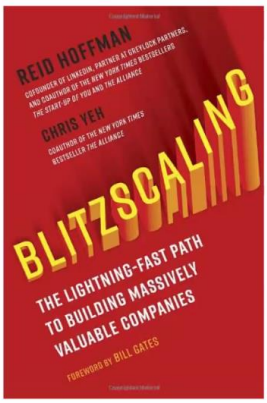
Company	Valuation ($B)	Date Joined	Country	Industry
Toutiao (Bytedance)	$75	04/07/2017	China	Digital Media/ AI
Didi Chuxing	$56	12/31/2014	China	On-Demand
WeWork	$47	02/03/2014	United States	Facilities
JUUL Labs	$38	12/20/2017	United States	Consumer Electronics
Airbnb	$29.30	7/26/2011	United States	eCommerce/Marketplace
Stripe	$22.50	1/23/2014	United States	Fintech
SpaceX	$18.50	12/01/2012	United States	Other Transportation
Epic Games	$15	10/26/2018	United States	Gaming
GrabTaxi	$14	12/04/2014	Singapore	On-Demand
Bitmain Technologies	$12	07/06/2018	China	Blockchain
Samumed	$12	08/06/2018	United States	Biotechnology
Palantir Technologies	$11	05/05/2011	United States	Big Data
Global Switch	$11.08	12/22/2016	United Kingdom	Computer Hardware & Services
Infor	$10	11/16/2016	United States	Internet Software & Services
DJI Innovations	$10	05/06/2015	China	Hardware
One97 Communications (operates Paytm)	$10	05/12/2015	India	Fintech
Go-Jek	$10	08/04/2016	Indonesia	On-Demand
Coupang	$9	5/28/2014	South Korea	eCommerce/Marketplace
Guazi (Chehaoduo)	$9	03/12/2016	China	eCommerce/Marketplace
Coinbase	$8	08/10/2017	United States	Fintech

Note that companies such as Uber were recently floated via IPO

Source : Research Unicorn Companies

How to get there?

- Presents a set of processes and frameworks in order to scale
- Prioritize growth over efficiency
- Needed in industries with high demand-side network effects (winner takes all markets)

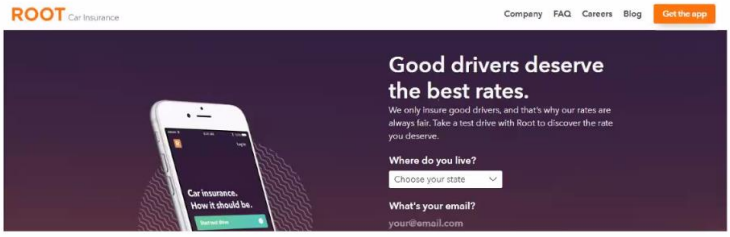


Osama Wanis | Amar Kumar, Mr | Frederic B. | Joshua Akue | Aissatou Balde | PARTH Jain

Recording | Sign

Real-Time, Context-Based Insurance

- For Tesla vehicle owners: Reduces price of policy the longer the vehicle runs in autonomous mode
- Rationale: autonomous mode is safer than manual drivers (Risk Assessment and Management)
- Root accesses GPS, accelerometers and gyroscopes data on your smartphone to evaluate risk
- How much time the vehicle is in use, frequent destinations, whether drivers change lane excessively, their driving speeds, to what extent they respect traffic rules, or if they use their smartphone while driving
- 30% of applicants rejected
- Accessing vehicle data would provide additional information
- Tesla is working with industry players and is considering becoming an insurer



Recording | Sign in

Data Product for Farmers – John Deere (1/2)

John Deere uses sensors added to their equipment to help farmers manage their fleet and to decrease downtime of their tractors as well as to save on fuel. The information is combined with historical and real-time weather data, soil conditions, crop features and many other data sets.

Variable Rate Application

When you put down seed and chemical where you need it most, you'll save seed and chemical (and money).

Field and Water Management

Whether it's water or soil, your resources are precious. Make the most of them with automated water and field management. Whether it's reducing irrigation costs or being able to have one person create ditches and levees in less time, you'll save on labor and fuel costs, as well.

Leave

Before we leave…

'Would you tell me please, which way I ought to go from here,' *(said Alice).*
'That depends a good deal on where you want to get to,' *said the Cat (Cheshire).*
'I don't much care where,' *said Alice.*
'Then it doesn't matter which way you go,' *said the Cat.*
—Lewis Carroll, *Alice in Wonderland*

For Next Week

- Do the readings!
- Get your McGill e-mail
- Install Alteryx
- Install Tableau
- Login to DataCamp