



Penang International
Education Forum 2017:
Future of Learning

Capturing the Attention of Millennials through Big Data

*J. Joshua Thomas,
Senior Lecturer,(Computing) @ KDU
Penang University College
jjoshua@kdupg.edu.my*

Overview



- ❖ Big Data for All
- ❖ Baby Boomers and Millennials
- ❖ Attractions through Data
 - A) AI helps to decipher millennials
 - B) Machine learning Attracts millennials
- ❖ Technology
- ❖ Examples

Big Data for All

Capturing the Attention of Millennials
through Big Data



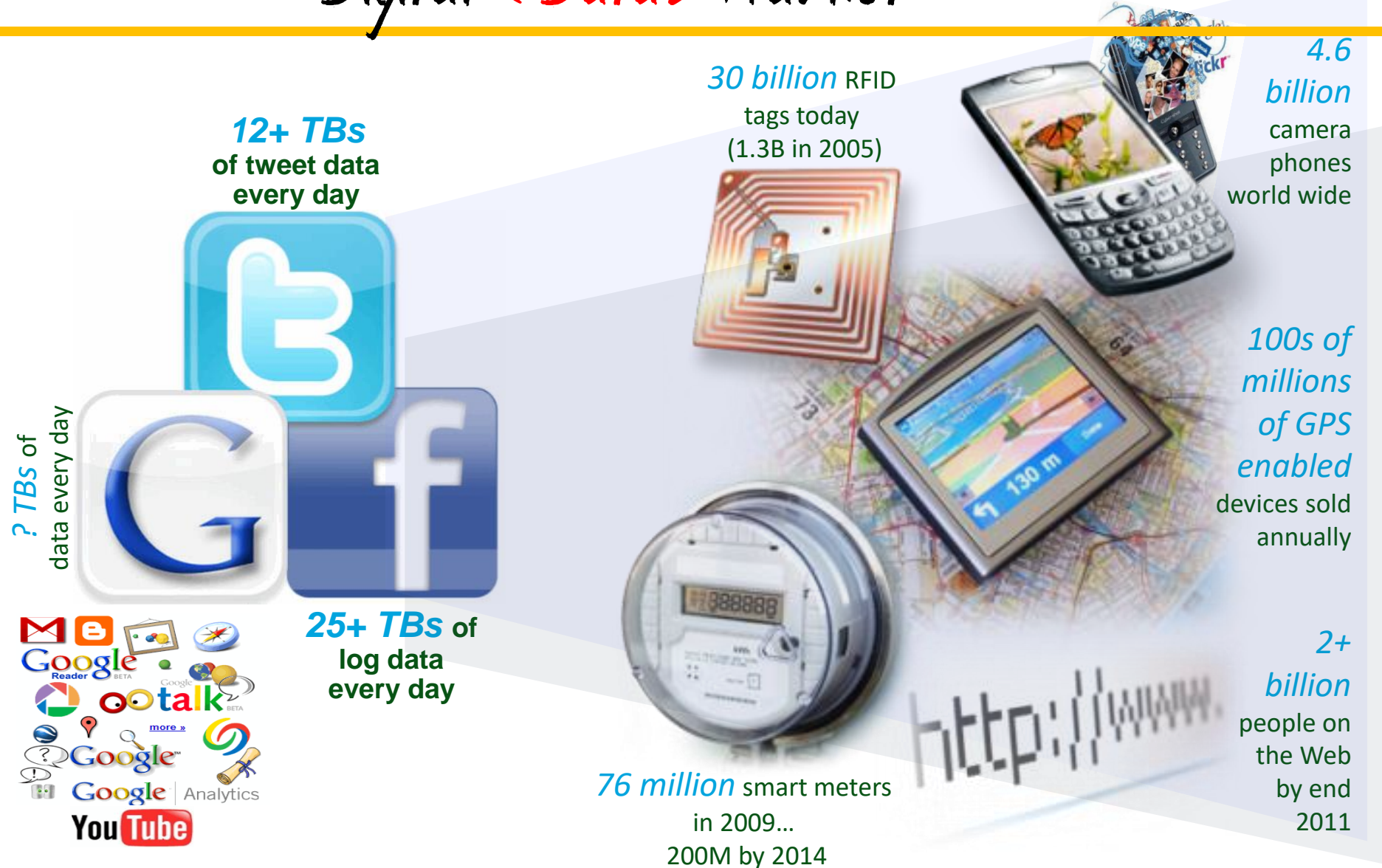
Big Data Is *High Performance Computing (HPC)*
Big Data Is About *Massive Data Volume*
Big Data Means *Hadoop (Big Data Framework)*
Big Data Need A *Very Large Data Warehouse*
Big Data Means *Collection of Unstructured Data*
Big Data Is for *Social Media & Sentiment Analysis*

Asian (Vegetable) Market

Capturing the Attention of Millennials
through Big Data



Digital (Data) Market



Asian (Big Data) Market

Capturing the Attention of Millennials
through Big Data

BIG DATA & ANALYTICS

Core to Nex-Gen Manufacturing

With more Asia Pacific manufacturers looking at making better and faster business decisions based on evidence, Big Data technologies can help them analyze larger volumes of data from a variety of sources and deliver the analysis at greater velocity. IDC observes that regional manufacturers are taking a cautious approach and the use of Analytics tools today is basic. Here is the 2013 landscape in Asia Pacific.

Asia Pacific manufacturers

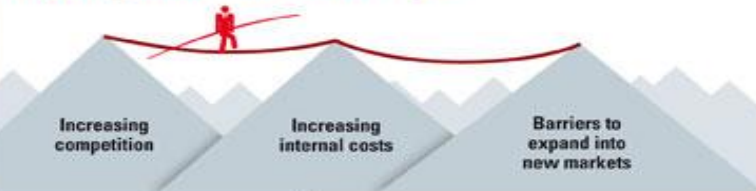
3 key highlights

4 of the top 10 automotive producing countries are from Asia Pacific

+370%
China's share of exports has nearly **quadrupled** over the past 15 years, rising from 3% in 1995 to 11% in 2012

x2
The average **labor cost in China** has nearly doubled in the past five years

Top 3 regional business challenges



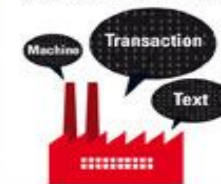
What Big Data means

Awareness

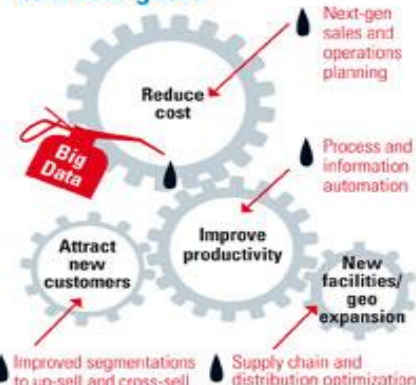
2/5 are still not aware of Big Data



Data that manufacturers are analyzing: Transactions (47%), text (45%) and machine (32%)



Big Data helps meet business goals



Common use cases

Today

37%
The % of Asia Pacific manufacturers using Big Data and Analytics technologies to improve **production quality management**

29%
The % of Asia Pacific manufacturers using Big Data and Analytics technologies for **inventory management**

Enabling technologies

Production reporting tools (46%)

Supply Chain analytics (35%)

Hadoop (8%)

Industry leader use cases



Top 3 IT challenges

37%
Not having the right people and skills

20%
Cost of Technology infrastructure

17%
Deciding what data is relevant

Moving forward

In the next 2 years

Asset Oriented Value Chain (AOVC) and Engineering Oriented Value Chain (EOVC) manufacturers will use Big Data and Analytics technologies to gain **Real time Asset Management**

Brand Oriented Value Chain (BOVC) manufacturers will use Big Data and Analytics technologies to improve **real time sales and operations planning**

Technology Oriented Value Chain (TOVC) manufacturers will use Big Data and Analytics technologies to improve **Supply Chain Management**

Enabling technologies



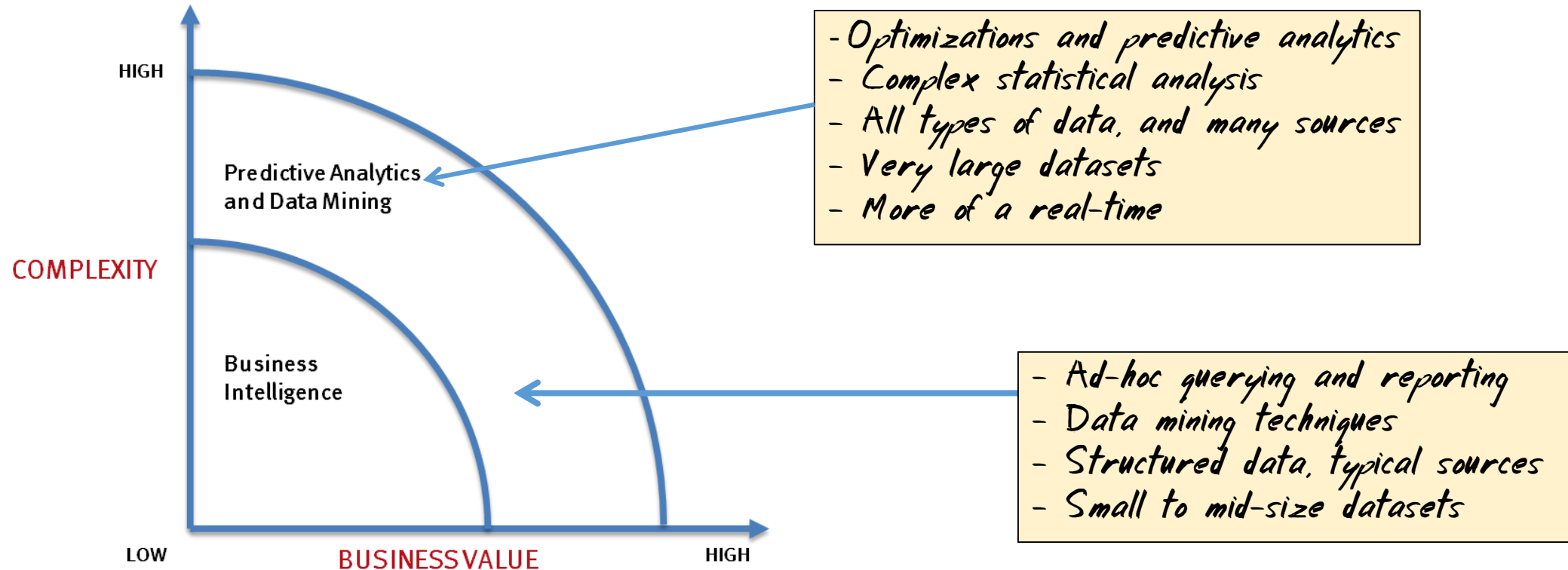
Source:
• IDC Big Data Pulse Survey
• IDC Manufacturing Insights Industry Survey
• IDC Big Data and Analytics QIS
• IDC Manufacturing Insights
• IMF
• China Ministry of Commerce

Research team:
• Chris Holmes
• Daniel-Zee Jimenez
• William Lee
• Craig Stiles
• Art direction
• ST Ling

This IDC infographic was produced by the IDC Asia-Pacific marketing department. The opinion, analysis, and research results presented herein are drawn from more detailed research and analysis independently conducted and published by IDC Asia-Pacific. Any information or reference to IDC Asia-Pacific that is to be used in advertising, press releases, or promotional materials requires prior written approval from IDC. For more information, visit www.ap.idc.asia-pacific.com

IDC
Analyze the Future

What's driving Big Data



Baby Boomers

- ❑ Baby boomers were born between 1946 and 1964.
- ❑ They call us that because after the World War II there was a boom of babies being born.
- ❑ People were like, we lived through the war, lets make whoopee (excitement)

Millennials

- ❑ Born between 1982 and 2004.
- ❑ Making whoopee is like **NetFlix** and **Chill**. But they were married.
- ❑ These are people who will **never know** the joy of using the end of a pencil to dial the phone.



Millennials

Capturing the Attention of Millennials
through Big Data



👍 Like 438K

I want to play a game

I am gonna test the knowledge between baby boomers and millennials and see how well they know each other.

(3 Minutes)

Big Data and Data Science

5 Aspects

Volume:

Speed:

Variability:

Value:

Visualization:



Attractions through Data

A) *AI helps to decipher millennials*



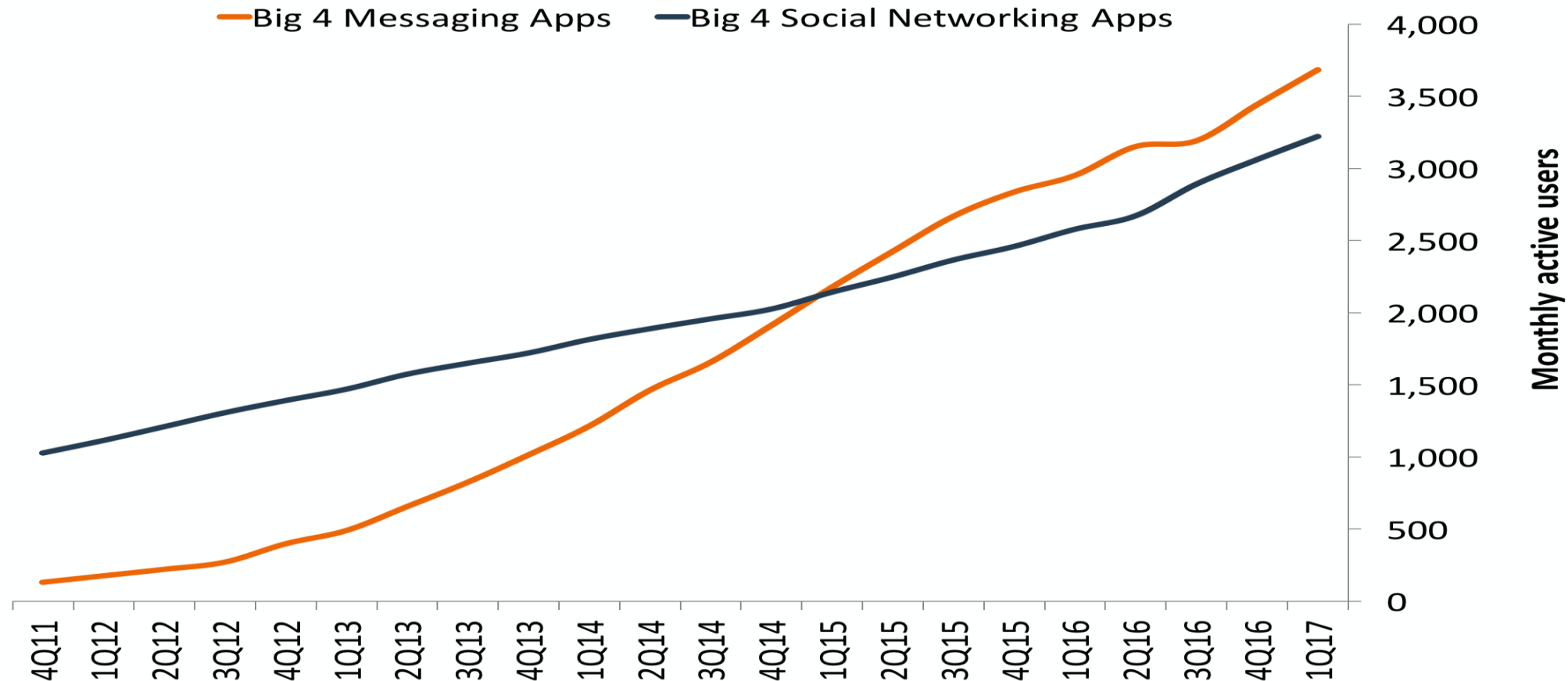
Image credit: Shutterstock / Anchiy

*Messaging apps are now
bigger than social networks*

Messaging apps are now bigger than social networks

Messaging Apps Have Surpassed Social Networks

Global monthly active users for the top 4 messaging apps and social networks,
In millions



Note: Big 4 messaging apps are WhatsApp, Messenger, WeChat, Viber.
Big 4 social networks are Facebook, Instagram, Twitter, LinkedIn
Source: Companies, Apptopia, TechCrunch, BI Intelligence estimates, 2017

Make a chatbot.



Image credit: Shutterstock / Anchiy

- ☐ *connect with millennials?*
- ☐ *Chatbots offer brands the powerful ability to fish where the fish are*
- ☐ *Brands spend millions, even billions, of dollars developing apps and buying up mobile ad space to capture the attention of millennials.*



- ❑ Google Home smart speakers will be able to make voice calls **free of charge** to contacts and millions of businesses in the United States and Canada

echo dot

Add Alexa to any room



- ❑ The added feature begins rolling out today and strikes a major distinction between **Google** and **Amazon**, two of the biggest names in the market for smart speakers with an **AI assistant** inside.





B) Machine learning Attracts millennials

- ❑ Machine learning is a crystal ball in the world of AI
- ❑ It analyses existing data and — through complex algorithms — predicts what will happen in similar cases in the future. *(millennials wants to know the future in a easy way!!)*
- ❑ Machine learning use cases that resonate the most with current business requirements and the availability of data.



Personal security



Data security



Healthcare



Online search



NLP



Financial trade



Fraud prevention



Recommendations

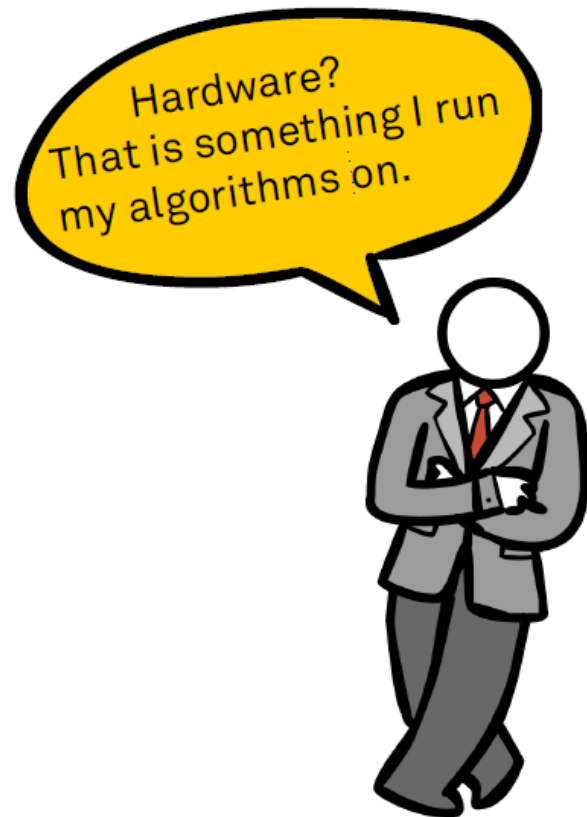


Smart cars



Marketing

Machine learning and hardware



MATLAB/Hortonworks takes care:

- distributes operations over cores,
- executes for-loops in parallel,
- executes on Hadoop

❑ Machine learning algorithms are independent of their execution platform.

Machine learning and new setting

Von Neumann bottleneck:

- instruction fetch and data operation sharing a bus.

New coprocessor:
shared memory GPU!

In 2013, with deep neural networks the computation demands on Google's data centres doubled.

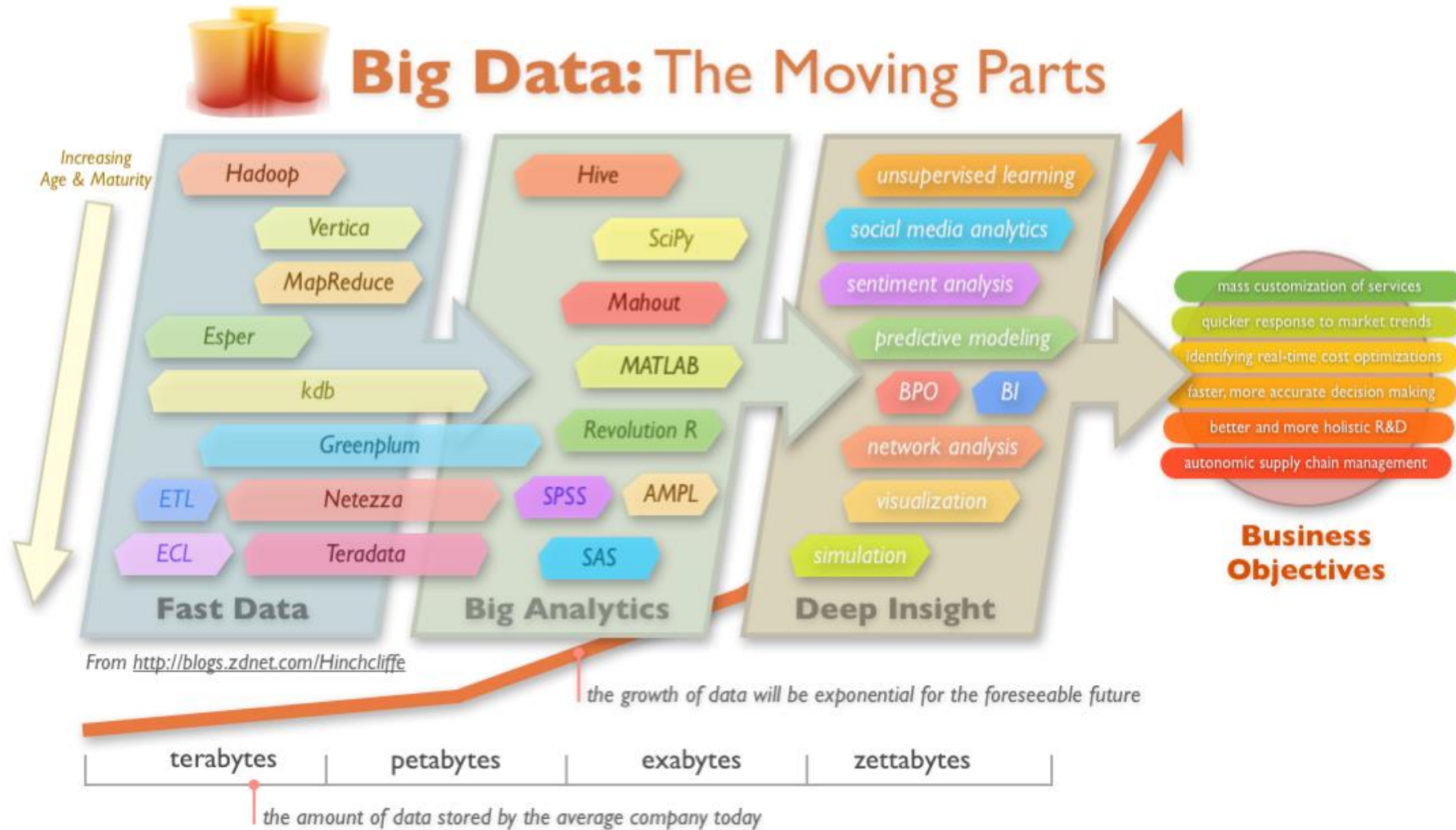
Even newer coprocessor:
inference by customized chip TPU!

- Intel: Lake Crest chip for learning
- Quantum computing (D-Wave): processor with more than 1000 Qubits for fast optimization.

Google Cooling, Georgia



Big Data Technology



Programming Paradigms



Cloud computing

Hadoop

- BigTable (Google Chrome)
- HBase (Apache Cassandra)

Lambda/Kappa paradigm

- Map reduce
- Stream processing

GPU

- Parallel computing
- Multiple instruction, multiple data



Big data Analytics and data Visualization

KDU Penang Millennials Projects

Capturing the Attention of Millennials
through Big Data

To visualize the air pollution data through data visualization technology
and give suggestions.



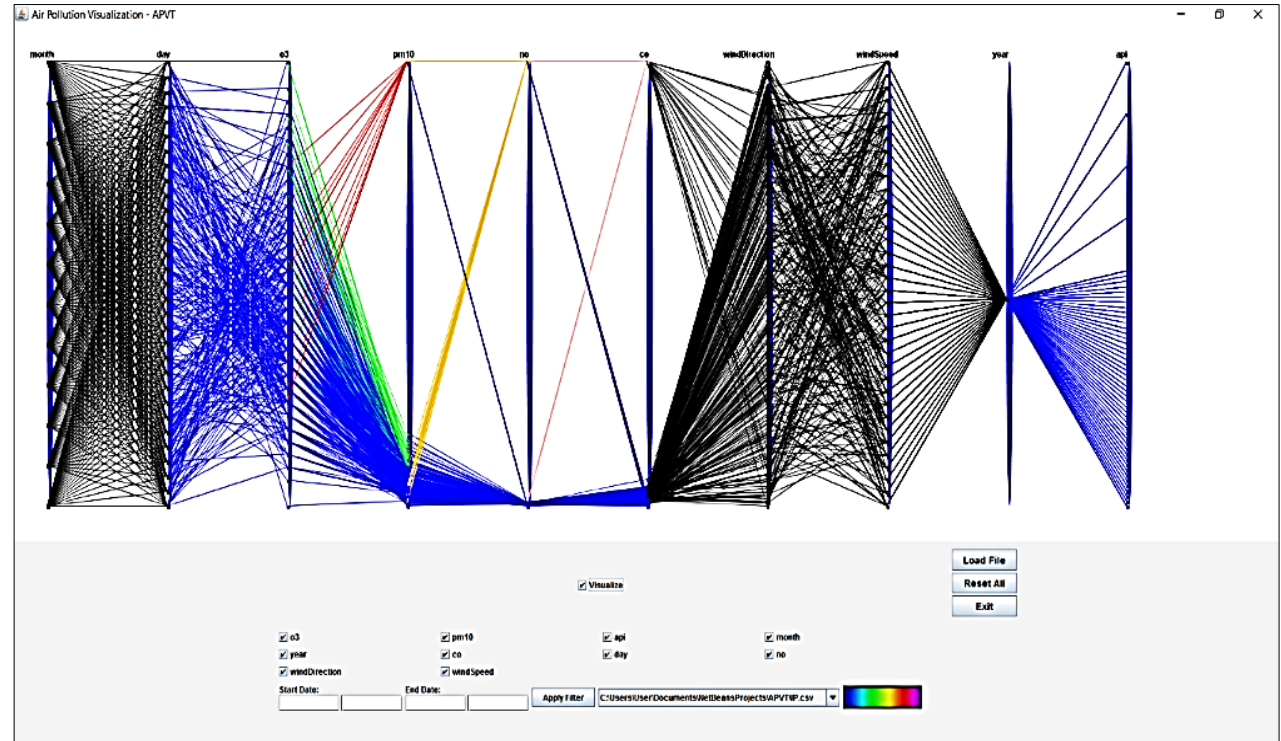
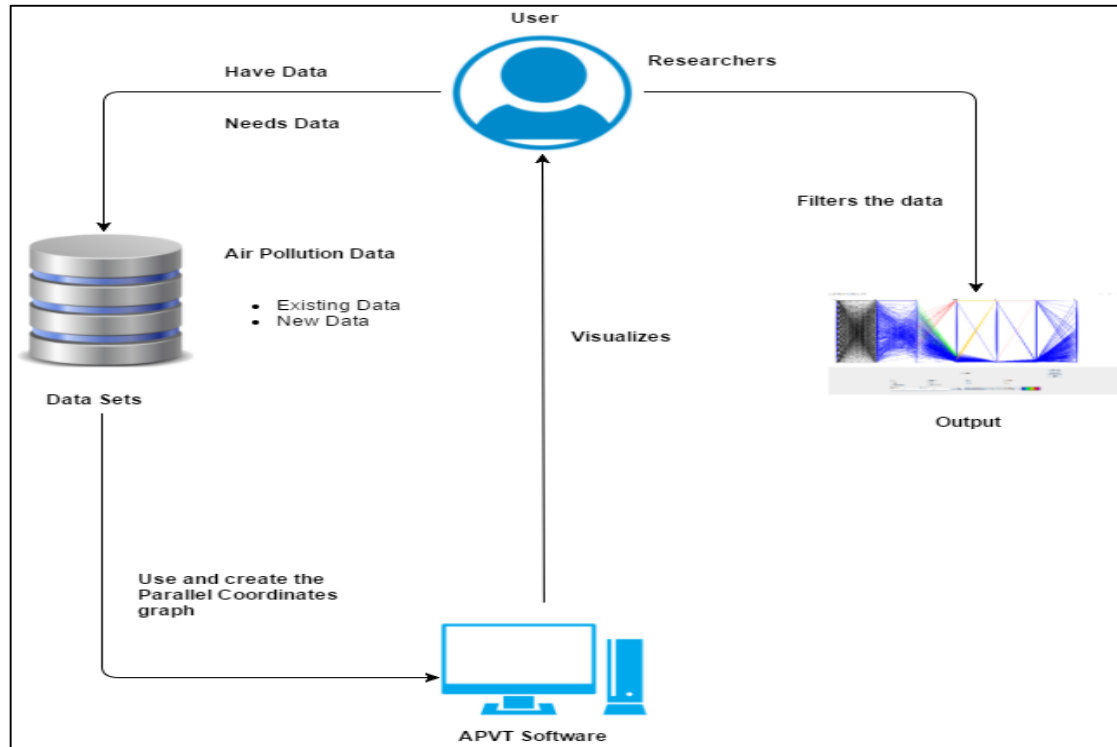
90%



50%



0%



Data Pipeline

Data plots and visual suggestions

Demo - #1

Air Pollution Visualization - APVT



☐ Visualize

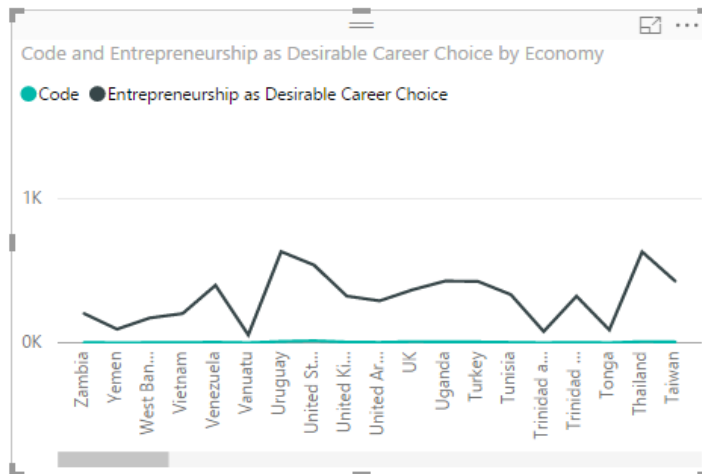
Load File	Reset All
Capture	Exit

Start Date: End Date:

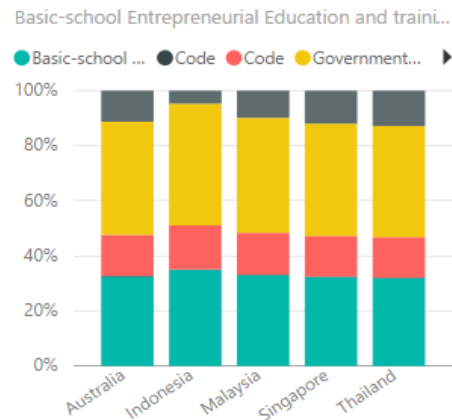
GDP indicators and immigrant data *visualization*

Global Entrepreneurship Landscape : (Asian Countries Analysis on Entrepreneurship)

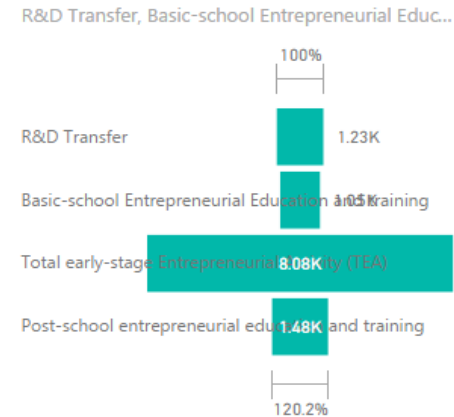
What is the trend for the level of confidence in any selected country to become a Entrepreneur



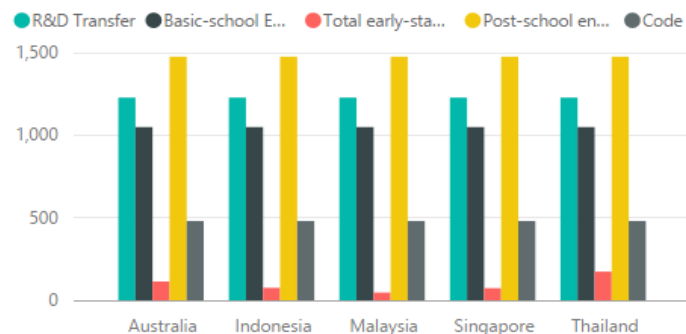
3 Key areas that the government need to focused on in order to drive entrepreneurship in the country



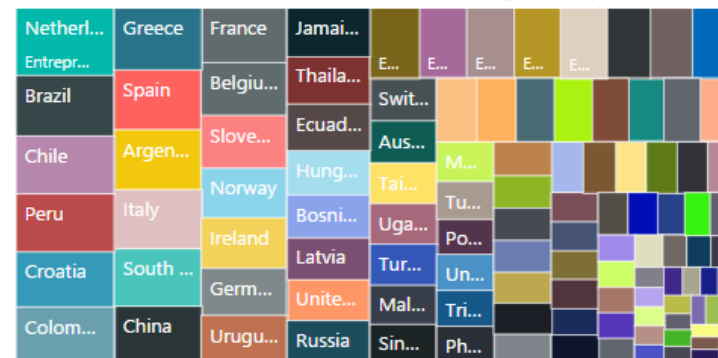
Perception of entrepreneurship in the eye of general public VS the chosen experts

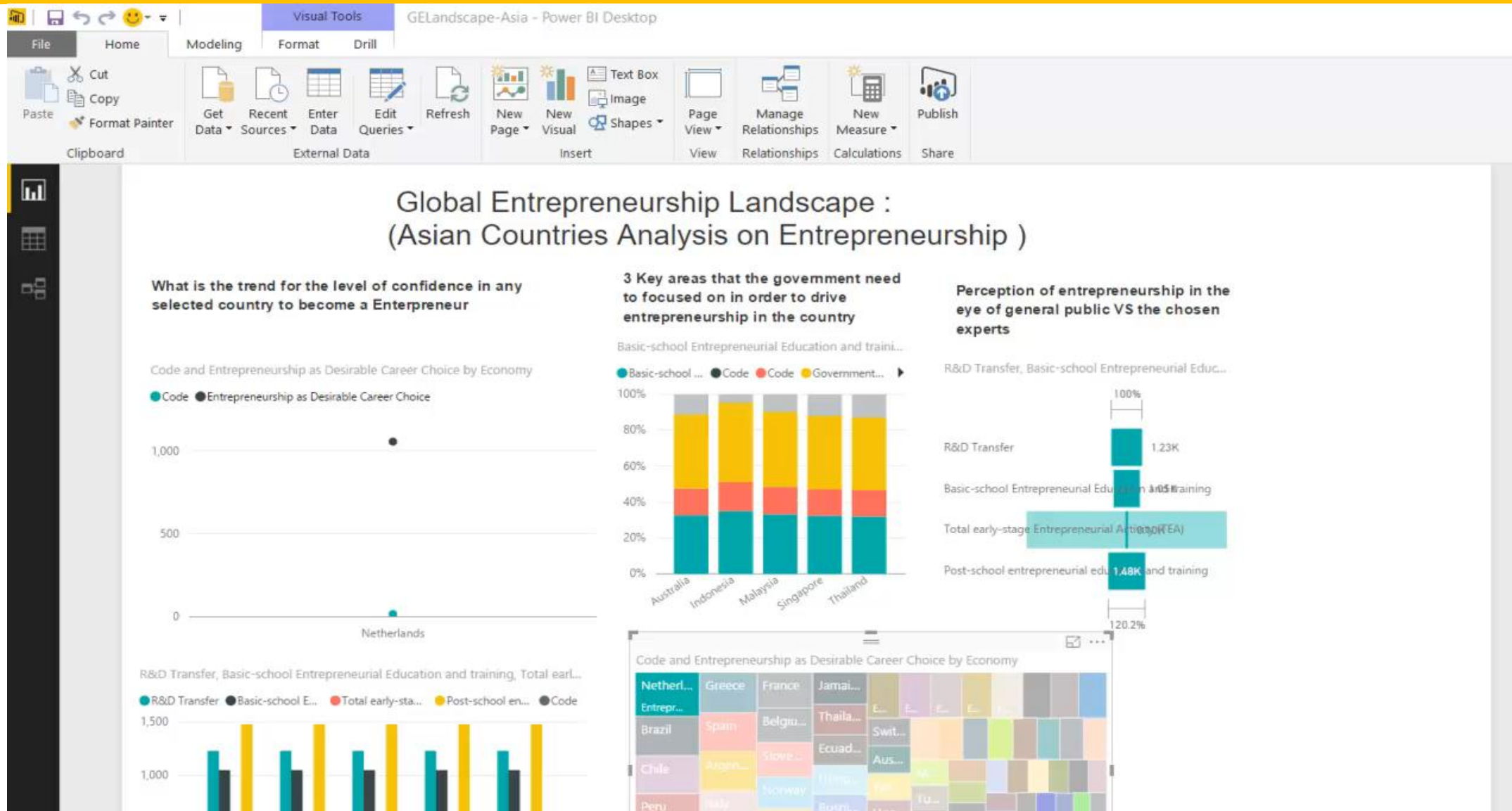


R&D Transfer, Basic-school Entrepreneurial Education and training, Total earl...



Code and Entrepreneurship as Desirable Career Choice by Economy





Demo - #3

Typhoon Haiyan Flood Data

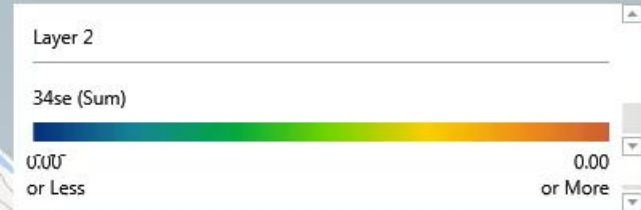
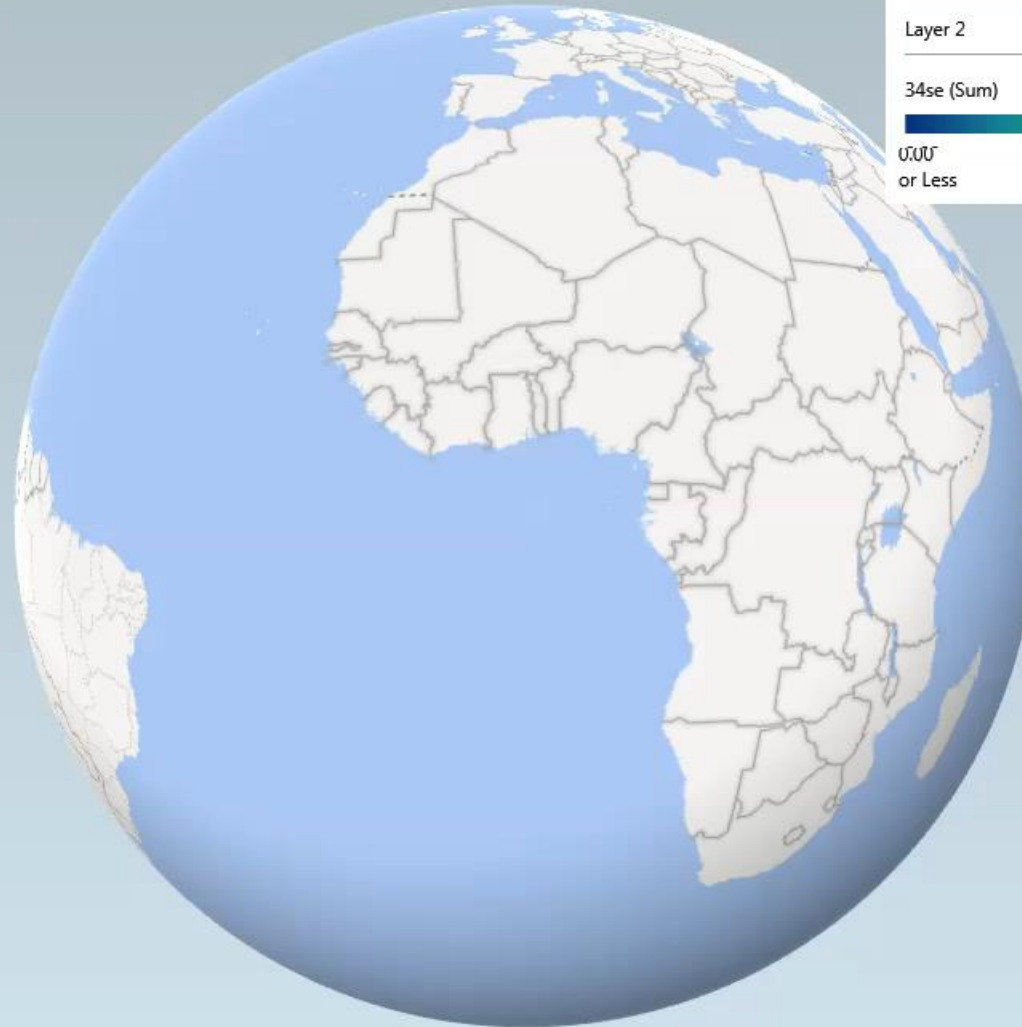
Wind Speed : 230 Km/h

Damage: \$2.86 Billion

Credits: Lukas Steindl's BI (Data)

Typhoon Haiyan Path

1/1/0001 12:00 AM



Penang Govt. initiatives over Big Data

- ❑ In Penang, ICT strategy and ongoing projects in the areas of big data analytics, open data, mobile government and geospatial information system.
- ❑ Malaysian Administrative Modernisation and Management Planning Unit (MAMPU) and other central government agencies and building public-private partnerships for delivering fast and efficient digital transformation.

Big Data Analytics Projects in Penang

Transportation & Traffic, Socio-economy, Environment and Integrity which will be implemented in phases. Three BDA projects will be implemented this year i.e. Traffic Management and Solid Waste Management under both local councils in Penang and Allocation of Affordable Housing for the low and middle income group.



5th INTERNATIONAL VISUAL INFORMATICS CONFERENCE

28-30th November 2017
Hotel Bangi-Putrajaya
Bangi, Selangor, Malaysia

Visual Informatics: Changing Landscapes in 4IR through Data Driven Decisions

Co-organised By:



Co-Sponsored By:



3RD CALL FOR PAPERS

Workshops

1. EXCAVATE DATA FOR DECISION, HORTON SANDBOX: HDP & HDF

Hortonworks Sandbox for Hadoop Data Platform (HDP) and Horton works Data Flow (HDF) is quick and easy personal desktop environment to get started learning, exploring, developing, testing and trying new features. This three hours workshop provides an overview of the Apache Hadoop Eco System. The HDF Sandbox is also a single node cluster for data flow and stream processing based on Apache NiFi, Apache Kafka and Apache Storm. Parton's gain hands on experience through HDP sandbox basics, real-world examples. The workshop covers the topics from visualization framework with end user capable of displaying analytics data at different geographic levels. The data sources will be from CSV or Spark.

Details of the workshop can be downloaded [here](#).

Speakers & Instructors :

- **J. Joshua Thomas, Ph.D.**
Department of Computing, KDU Penang University College.
- **Bahari Belaton, Ph.D., (Assoc. Prof. Dr.)**
School of Computer Sciences, Universiti Sains Malaysia.



Thank you for your attention

Credits

- <http://www.opengovasia.com/articles/7493-exclusive---ict-strategy-and-projects-for-enhancing-state-government-service-delivery-in-penang>
- Joshua Thomas, J., Raaj, Justtina (2017) Parallel Coordinates Visualization Tool on the Air Pollution Data for Northern Malaysia.: In Zelinka. Vasant & I. P (Eds.), Innovative Computing and Its Applications. Springer (2017 to be published)
- Joshua Thomas, J., Jodene Ooi (2017). DengueViz: A Knowledge-Based Expert System Integrated with Parallel Coordinates Visualization in the Dengue Diagnosis IVIC2017, November 28-30, 2017.