Mahidol University
Wisdom of the Land

BIG DATA PROCESSING

EGCI 466

Week 1

Mahidol University
Wisdom of the Land

Google classroom

- rdhi44x5

Course Outline

- Schedule

Evaluation	Percentage		
Weekly Lab Report/ Cloud Skill	10%		
Participation	10%		
Homework/Assign ent	20%		
Exam/ Quizzes	30%		TBC
Midterm Project	15%		TBC
Final Project/ Presentation	25%		
	110%		

Mahidol University
Wisdom of the Land

Learning Objectives (Today)

- Understand what "Big Data" means
- Explore real-world Big Data challenges
- Overview of key Big Data tools and architecture
- Set expectations for labs and projects

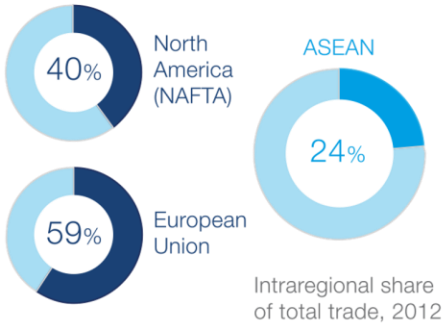
Why big data

Three global trends create opportunities to transform Southeast Asia by 2030.

1 Capturing a greater share of global flows

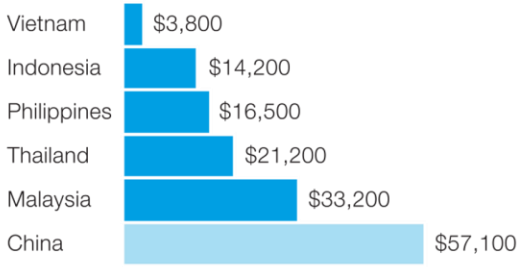
Up to $615 billion in annual economic value

The ASEAN[1] Economic Community (AEC) sets the stage for greater intraregional trade



Intraregional share of total trade, 2012

To attract more global production, Southeast Asia must raise labor productivity

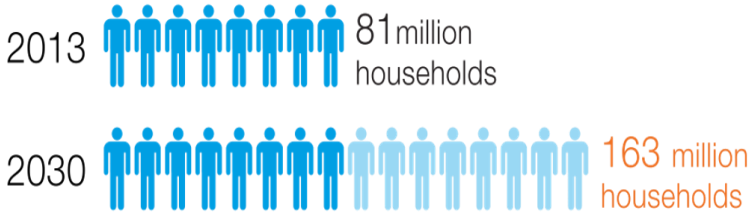


Annual manufacturing output per worker, 2012

2 Riding the urbanization wave

Up to $930 billion in annual economic value

An expanding consumer class



$7 trillion in investment needed for infrastructure, housing, and commercial space

3 Deploying disruptive technologies

Up to $625 billion in annual economic value

[1] Association of Southeast Asian Nations.

Source: McKinsey Global Institute analysis

Source:https://www.mckinsey.com/featured-insights/asia-pacific/three-paths-to-sustained-economic-growth-in-southeast-asia

What is big data?

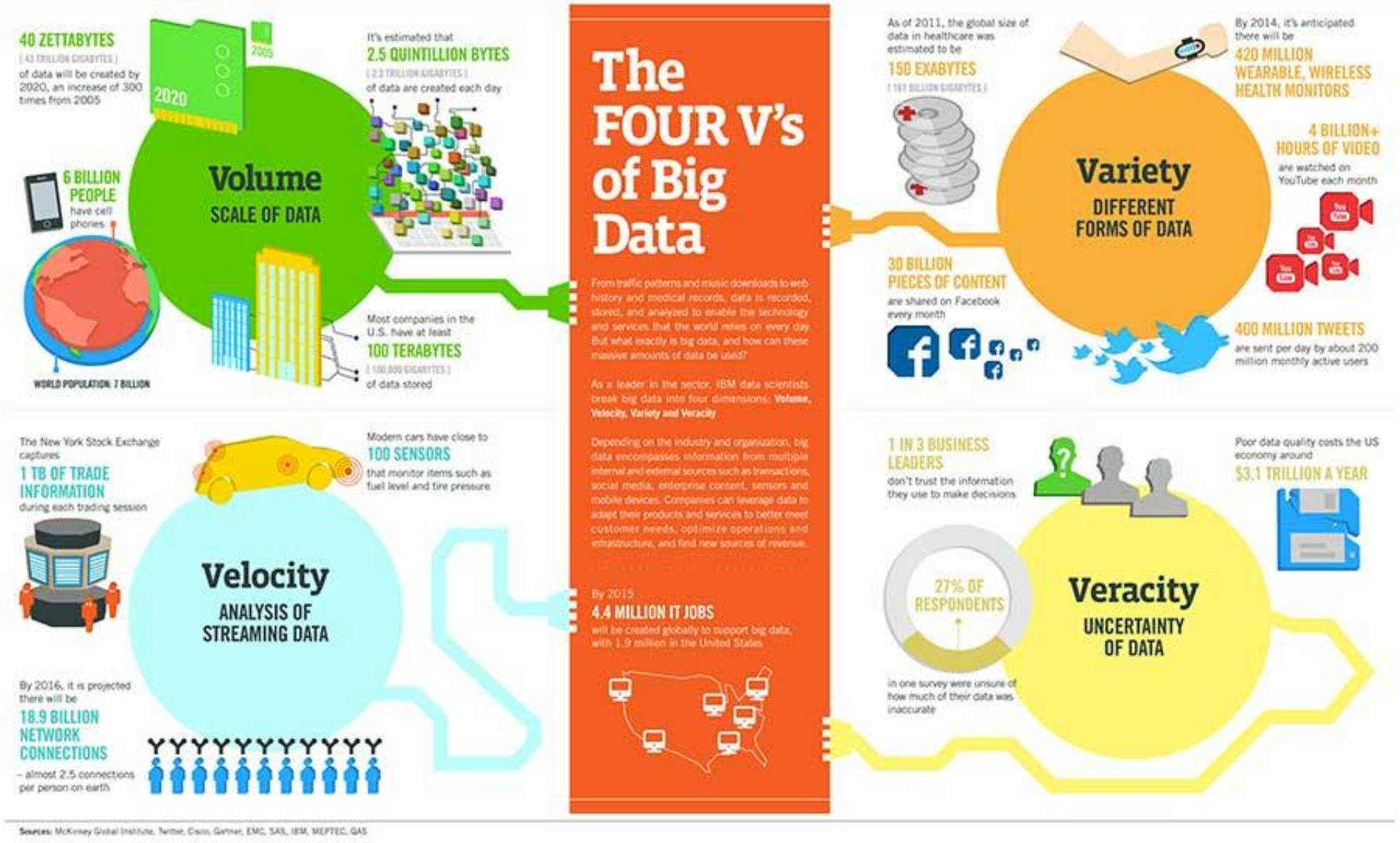
- 4 Vs. → 5 Vs.
- Volume
- Variety
- Velocity
- Veracity

  - Value
  - Variability
  - Valence

Mahidol University
Wisdom of the Land

What is big data?

- 4 Vs. → 5 Vs.
- Volume
- Variety
- Velocity
- Veracity
  - Value
  - Variability



Mahidol University
Wisdom of the Land

Applications

Big Data → Better Model → Better Precision

Personalized marketing

Mahidol University
Wisdom of the Land

Personalized marketing

- Recommendations Engines
  - Netflix, Amazon, Shopee
- Sentiment Analysis
- Real-time Fraud Detection
- Smart Cities with IoT sensors

Mobile Advertising

- Location based advertising

- Geolocation data

Is this a big data? If so why?

Mahidol University
Wisdom of the Land

Smart city

Mahidol University
Wisdom of the Land

Fun Fact about big data

https://www.slideshare.net/BernardMarr/big-data-25-facts/2-Every_2_dayswe_create_as

Where do data are from??

- Machine generated
- Human Generated
- Organization generated

Machine generated data

- Big plane →big data
  - 0.5 terabytes / flight

- Smart phones/ Smart Watch
  - GPS, HR, O2,

Mahidol University
Wisdom of the Land

Why it is useful?

- Detect mal-function
- Notify of unhealthful behaviour

Business

- Fraud detection
- Business Planing
- System moniting/control
- Customer statisfication

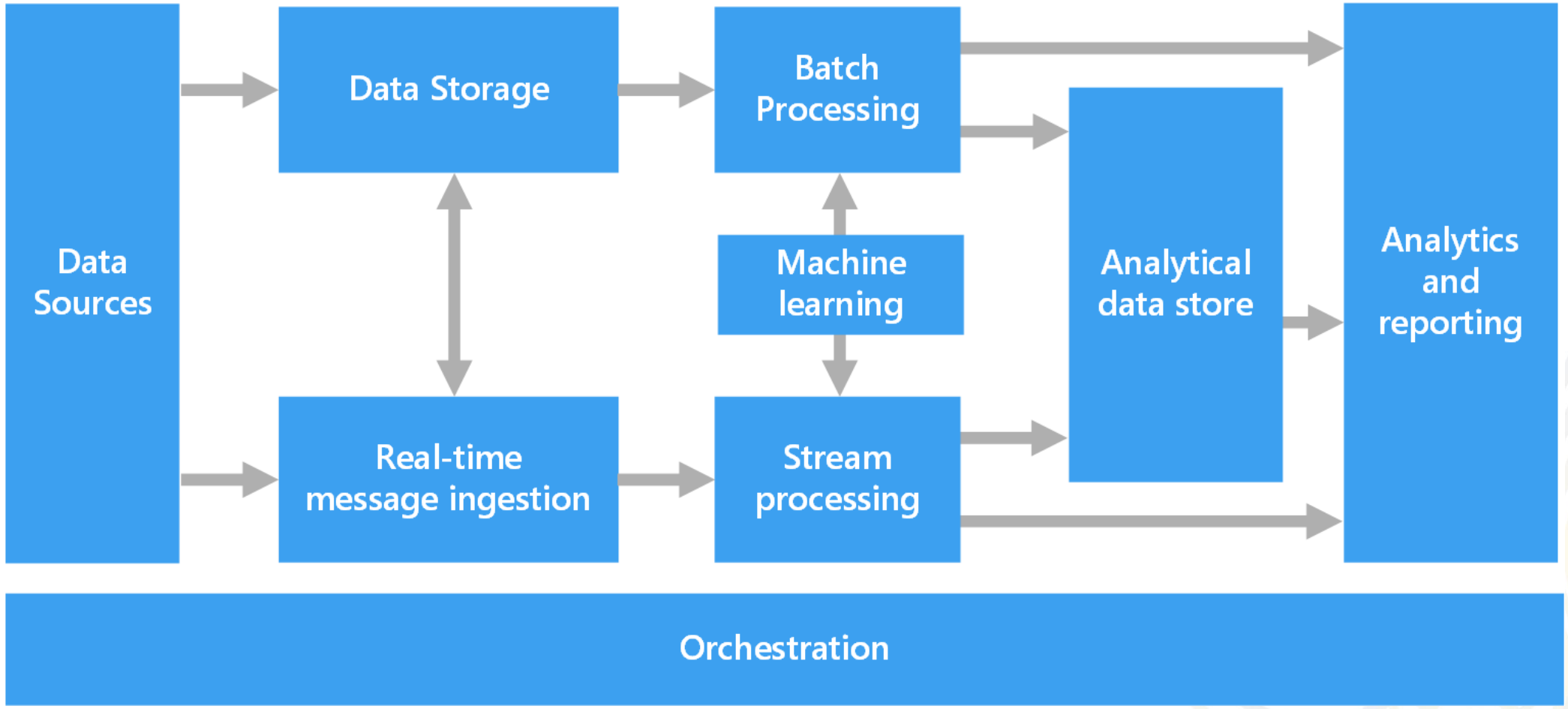
Mahidol University
Wisdom of the Land

Why traditional System fail?

- Single-node databases can't handle petabytes
- Slow processing times
- Limited scalability and fault tolerance
- Need for distributed computing

Mahidol University
Wisdom of the Land

Topic covered



Mahidol University
Wisdom of the Land

Course covered

- Hadoop Basics: HDFS, MapReduce
- Cloud Simulation
- Spark Processing: RDDs, DataFrames, SQL
- NoSQL Databases: MongoDB, Cassandra, Neo4J
- Real-Time Streaming: Spark Streaming

- Midterm Project:
  - Build a mini Big Data pipeline with no data processing and literature review
- Final Project:
  - Build a mini Big Data pipeline with **data processing**

Mahidol University
Wisdom of the Land

Setting Up for Success

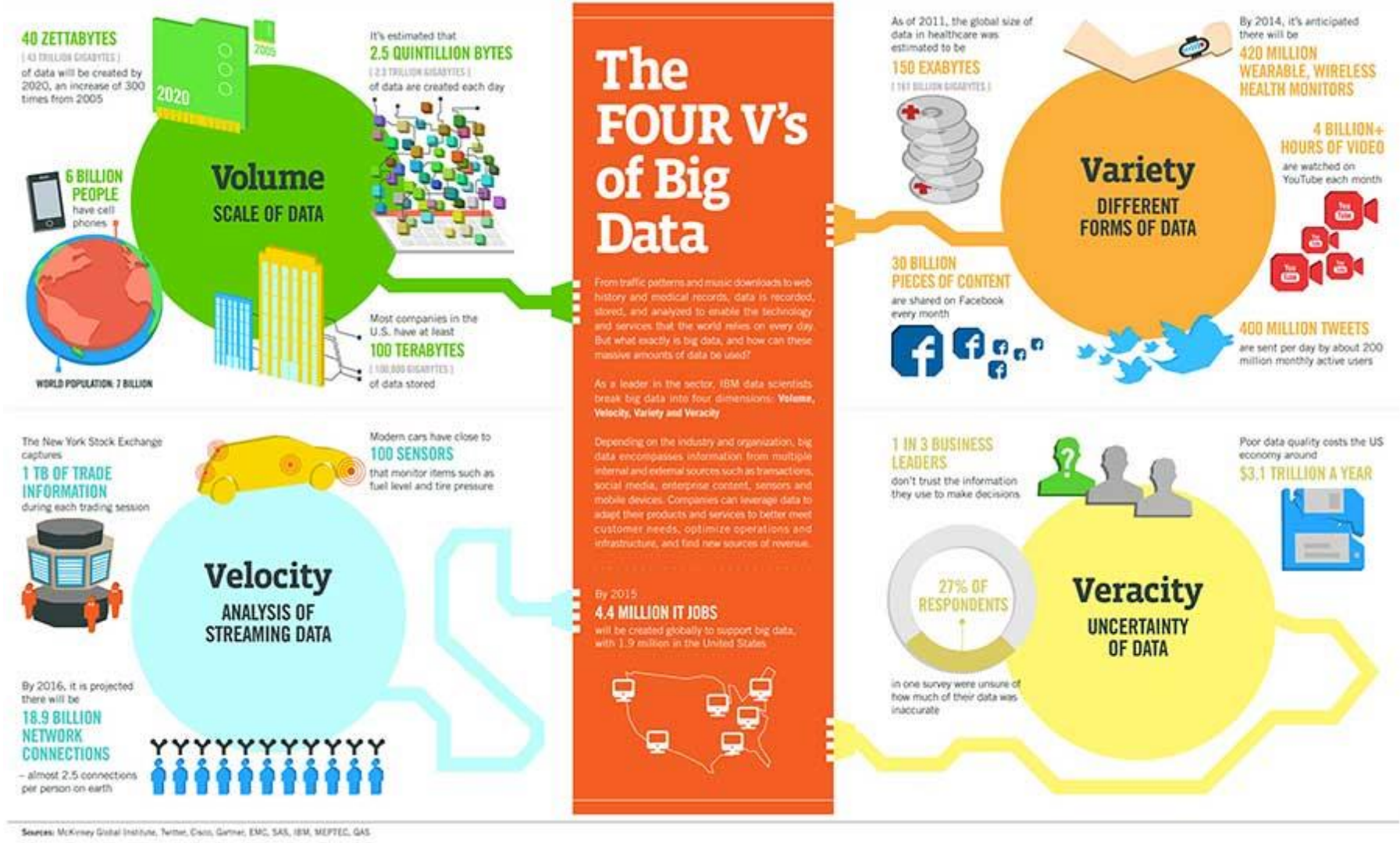
- Installations: Spark local, Python, Docker (optional)
- Join Discord group (optional)

  https://discord.gg/VyWp8pSw

- Join Cloudskill (I will add your email)
- Apply for Google cloud credit (I will provide education credit later)

Mahidol University
Wisdom of the Land

Group Discussion



Mahidol University
Wisdom of the Land

Discuss 4 Vs

- What is it?
- Where do you think the data come from?
- What is the big challenge?
- What solutions have been introduced?
- What tools have been used?

Mahidol University
Wisdom of the Land