

DataTH X Microsoft Azure

AZURE DATA FUNDAMENTALS TRAINING

EP. 1 Core Data Concepts



หากมีคำถาม ถามได้ที่ไหน

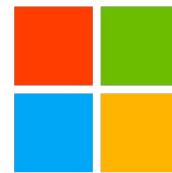


ถามคำถามไว้ใน chat ด้านขวา หัวข้อ Q&A

โดยจะมาตอบช่วงหลังเบรก หรือช่วงท้าย CLASS

หากพิมพ์เข้ามาในกล่องเขตทั่วไป คำถามอาจจะโดนดันจนไม่ได้ตอบ

Microsoft Azure คืออะไร

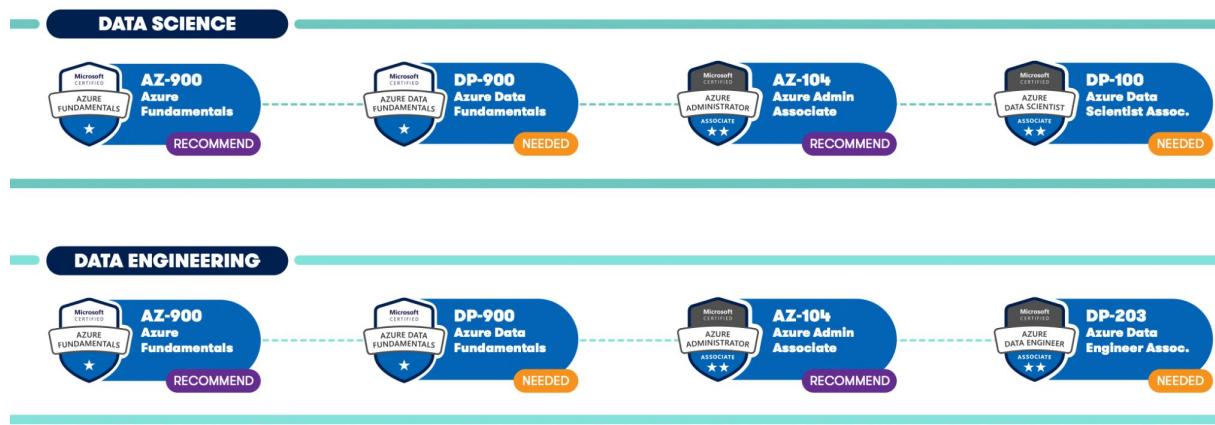


Microsoft
Azure

- ผู้ให้บริการ public cloud platform จาก Microsoft
- เปิดให้บริการมาตั้งแต่ปี 2008
- มี data center หลากหลาย location ทั่วโลก
- เป็น cloud ที่มีบริการให้ใช้หลากหลายด้าน เช่น จัดเก็บข้อมูล, ประมวลผลข้อมูล
- สามารถเริ่มต้นใช้งานได้ฟรี <https://azure.microsoft.com/en-us/free/>

Azure Certifications

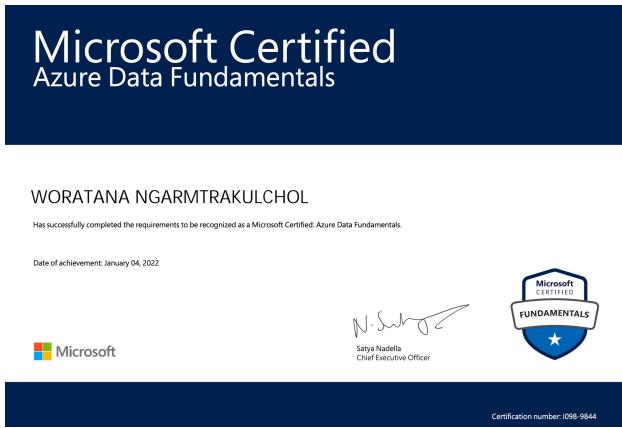
Azure มี Certificate (ประกาศนียบัตร) ที่จะต้องทำข้อสอบก่อน ให้เลือกสอบหลายระดับ หลายหัวข้อ สามารถเลือกด้านที่สนใจได้เลย เช่น



List of all Azure certificates by aCloudGuru:

<https://acloudguru.com/blog/engineering/which-azure-certification-is-right-for-me>
(Updated: June 2021)

Azure Data Fundamentals (DP-900) คืออะไร

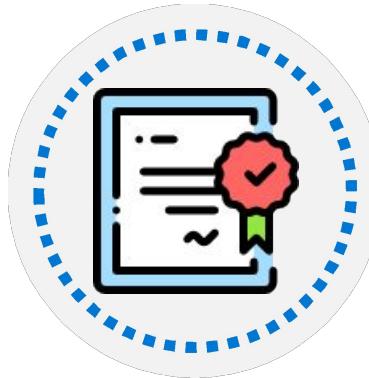


- Certification ที่ได้จากการสอบกับ Microsoft Azure โดยตรง (ค่าสอบปกติ 99 USD) แต่เดียว class 3 จะมีแนะนำวิธีสอบฟรี
 - วัดความรู้ด้าน Core Data Concepts และบริการด้าน Data บน Azure
 - ข้อสอบเป็น multiple choices ภาษาอังกฤษ 40-50 ข้อ
 - คะแนนผ่าน 700/1000
 - ใช้เวลาสอบ 1 ชั่วโมง
 - สอบ online หรือสอบที่ศูนย์สอบที่กำหนด

ทำไมต้องสอบ Azure Data Fundamentals



อัพสกิลด้าน Data และ Azure เพื่อ อัปความรู้ และ เงินเดือน



อัพเดท Certificate บน LinkedIn / Resume เพื่อดึงดูด Recruiters



ใช้ความรู้และ Certificate สมัครงานได้ ในไทยมีบริษัท จำนวนมากที่ใช้ Azure

FAQ: คอร์สนี้ต้องสมัครสมาชิก Azure เพื่อเรียนมั้ย ?

สมัครหรือไม่สมัครก็ได้

ในคอร์สนี้จะเป็นคอนเซปต์ความรู้ของ Azure ซึ่งเราจะมีกดใช้งานบริการให้ดู แต่จะดูเฉย ๆ หรือทำตามเพื่อเรียนรู้ไปด้วยกันก็ได้ (ขออนุญาตไม่หยุดรอ เพราะไม่ใช่ Workshop)

ถ้าทำตาม กดลงบริการที่สร้างมาทุกรุ่น หลังจากจบคลาส
ไม่อย่างนั้นถ้าลืมลงบริการไหน อาจจะโดนหักเงินไปเรื่อย ๆ

WARNING: THAI & ENGLISH CONTENT



เนื้อหาจะมีทั้งภาษาไทย เพื่อความเข้าใจง่าย

และบางคัพท์จะขอใช้ภาษาอังกฤษ เพื่อให้ทุกท่านคุ้นชินกับตอนสอบ Certificate ที่จะเจอคัพท์ภาษาอังกฤษ



ตารางเรียน Azure Data Fundamentals Training



We are here

CLASS 1

วันจันทร์ 17 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Core Data Concepts
- Explore relational data in Azure



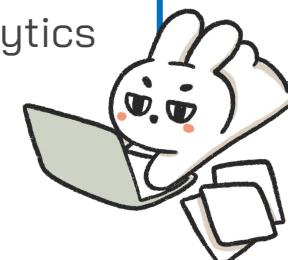
CLASS 2

วันพุธที่สุดวัน 20 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Explore non-relational data in Azure
- Explore modern data warehouse analytics



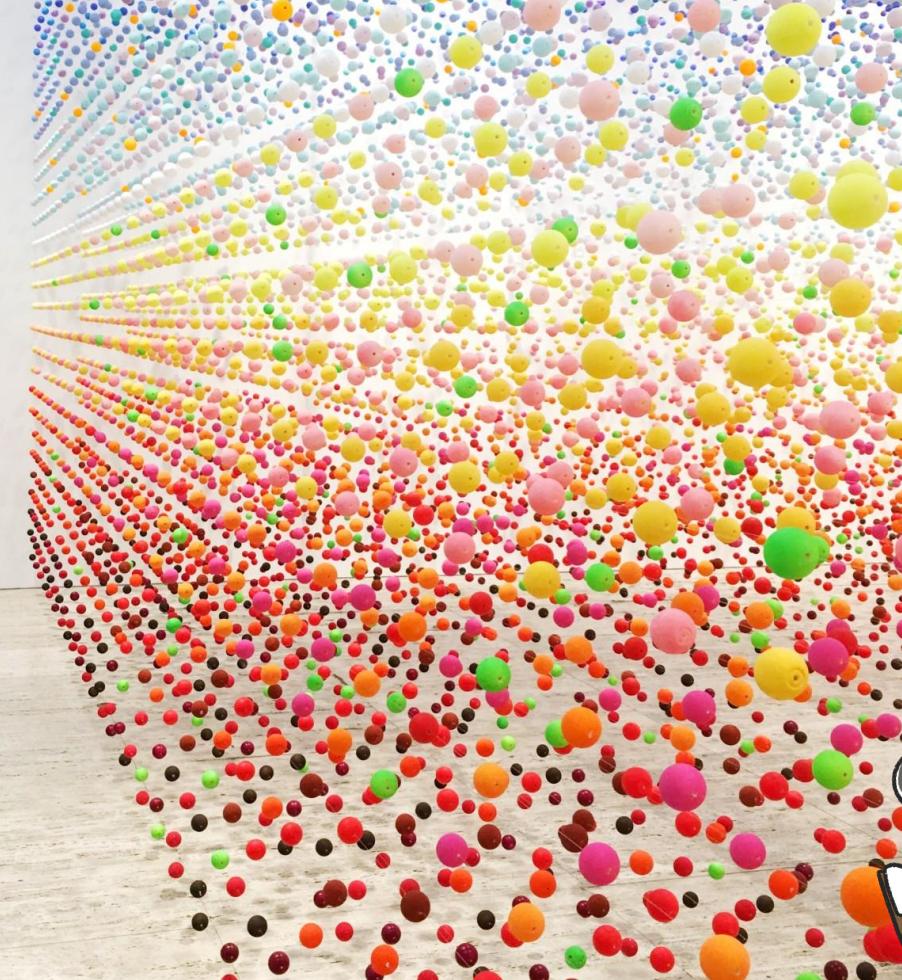
CLASS 3

วันจันทร์ 24 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Azure Data Architecture with Microsoft Thailand



CLASS 1

CH 1: CORE DATA CONCEPTS

1 HR

- BREAK 5 min -

5 MIN



CH 2: EXPLORE RELATIONAL DATA

1 HR

Q & A

15 MIN

AZURE DATA FUNDAMENTALS TRAINING

CH1: CORE DATA CONCEPTS



Chapter 1

ในบทนี้ เราจะเรียนกัน 5 หัวข้อ



Lesson 1: เรียนรู้ Core data concepts



Lesson 2: ตำแหน่งงานต่าง ๆ ในโลก Data



Lesson 3: ทำความรู้จักกับ Relational data

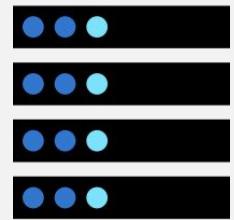


Lesson 4: ทำความรู้จักกับ non-relational data



Lesson 5: เรียนรู้คอนเซปต์ data analytics

Lesson 1: Explore core data concepts



Lesson 1 objectives



Types of Data: Relational vs Non-relational



Data Workloads: OLTP vs OLAP

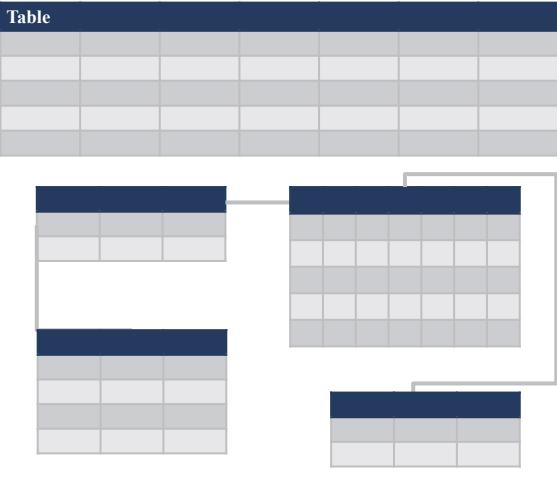


Data Processing: Batch vs Streaming

ประเภทของข้อมูล

ข้อมูลแต่ละประเภทต้องใช้ Database (ฐานข้อมูล) ที่แตกต่างกัน
แบ่งข้อมูลออกเป็น 3 ประเภทหลัก ๆ

Structured (Relational)

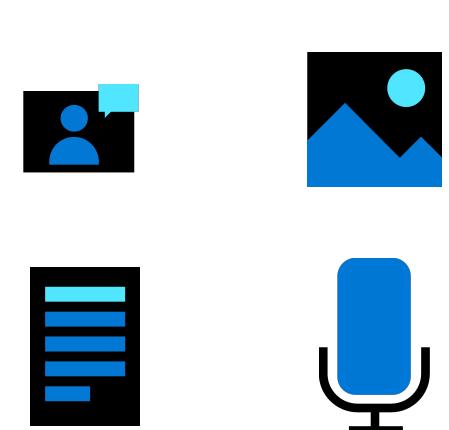


Semi-structured (Non-relational)

```
## Document 1 ## {
"customerID": "103248",
"name": { "first": "AAA",
"last": "BBB" }, "address": {
"street": "Main Street",
"number": "101", "city": "Acity",
"state": "NY" },
"ccOnFile": "yes",
"firstOrder": "02/28/2003" }
## Document 2 ## {
"customerID": "103249",
"name": { "title": "Mr",
"forename": "AAA",
"lastname": "BBB" },
"address": { "street": "Another Street",
"number": "202", "city": "Bcity",
"county": "Gloucestershire",
"country-region": "UK" },
"ccOnFile": "yes" }
```

ข้อมูลที่ความสัมพันธ์เป็นตาราง
(tabular data) เช่น Excel, SQL Server, MySQL

Unstructured



ข้อมูลที่ความสัมพันธ์ไม่เป็นตาราง
เช่น JSON, XML, MongoDB

ข้อมูลที่เป็นไฟล์ เช่น รูปภาพ เลียงไฟล์วีดิโอ

Data Workloads: OLTP vs OLAP

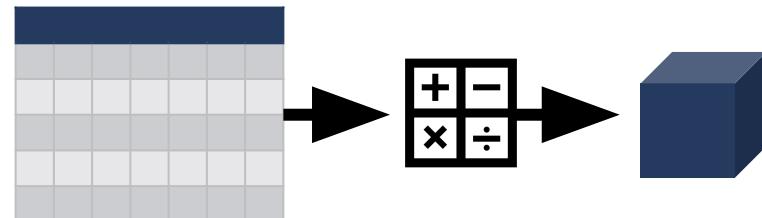
Online Transactional Processing (OLTP)

Customer		
CustomerID	CustomerName	CustomerPhone
Orders		
OrderID	CustomerID	OrderDate

ระบบประมวลผลข้อมูล Transaction Data (ข้อมูลธุรกรรม) เช่น ข้อมูลการซื้อขาย, ข้อมูลการฝากเงิน-ถอนเงิน

ตัวอย่างระบบ: ระบบเว็บไซต์, แอพพลิเคชันมือถือ

Online Analytical Processing (OLAP)



ระบบสำหรับวิเคราะห์ข้อมูล

มักจะดึงข้อมูลจากระบบ OLTP มาจัดเก็บไว้ในระบบที่แยกกัน เพื่อนำมาวิเคราะห์

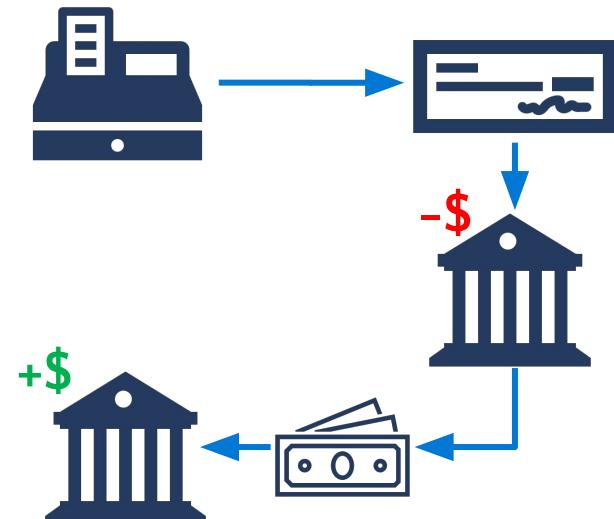
ตัวอย่างระบบ: ระบบคลังข้อมูล (Data Warehouse)

Transactional Processing (OLTP) ทำงานอย่างไร

ในธุรกิจ จะมีการเก็บข้อมูล Transaction Data ในทุกการกระทำ เช่น ซื้อ-ขายสินค้า, ฝาก-ถอนเงิน, ยืนยัน-ยกเลิกงาน ฯลฯ

งาน OLTP จะเน้นการจัดเก็บในฐานข้อมูลที่มี ACID

- **A - Atomicity** – ทุกธุกรรมจะมีแค่สำเร็จ หรือไม่สำเร็จ (ไม่มีแบบสำเร็จครึ่งธุกรรม ถ้าเป็นแบบนั้นต้องยกเลิกทั้งหมด)
- **C - Consistency** – ทุกธุกรรมจะได้ผลลัพธ์ตามกฎเกณฑ์ของฐานข้อมูล (เช่น ถ้าข้อมูลรับแต่ตัวเลข ก็ใส่ตัวหนังสือเข้าไปไม่ได้) และหากมีปัญหา จะต้องย้อนกลับไปได้เหมือนเดิม
- **I - Isolation** – ธุกรรมแต่ละตัว ไม่ส่งผลกระทบให้ธุกรรมอื่นเปลี่ยนแปลง
- **D - Durability** – ธุกรรมที่สำเร็จแล้วจะถูกจัดเก็บไว้อย่างปลอดภัย ผลลัพธ์ของธุกรรมจะไม่หายไปต่อให้ระบบมีปัญหา



Analytical Processing (OLAP) ทำงานอย่างไร

Analytical workloads ใช้สำหรับวิเคราะห์ข้อมูล เพื่อตัดสินใจทางธุรกิจ

ผลการวิเคราะห์

- สรุปข้อมูล เช่น ยอดขายปีนี้เป็นอย่างไรบ้าง, ลูกค้าซื้อของโดยเฉลี่ยกี่บาท
- เทรนด์ข้อมูล เช่น ยอดขายเติบโตขึ้นมั้ยในแต่ละ Quarter (3 เดือน)



Data Processing: Batch vs Streaming

Data processing เป็นขั้นตอนการแปลงข้อมูลดิบ (raw data) ให้กลายเป็นข้อมูลที่นำมาวิเคราะห์ได้ แบ่งออกเป็น 2 ประเภทหลัก ๆ

Batch Processing: ประมวลผลข้อมูลแบบรอทีละกลุ่ม ตามตารางเวลา เช่น ทุกชั่วโมง หรือตามเงื่อนไขอื่น ๆ เช่น ประมวลผลทีละ 100 แถว



Stream Processing: ประมวลผลข้อมูลทันทีที่เข้ามาในระบบ



Lesson 1: Knowledge check



Q1 How is data in a relational table organized?

Rows and Columns

Header and Footer

Pages and Paragraphs



Q2 Which of the following is an example of unstructured data?

An Employee table with columns Employee ID, Employee Name, and Employee Designation

Audio and Video files

A table within SQL Server database



Q3 What of the following is an example of a streaming dataset?

Data from sensor feeds

Sales data for the past month

List of employees working for a company

Chapter 1

ในบทนี้ เราจะเรียนกัน 5 หัวข้อ



Lesson 1: เรียนรู้ Core data concepts



Lesson 2: ตำแหน่งงานต่าง ๆ ในโลก Data



Lesson 3: ทำความรู้จักกับ Relational data



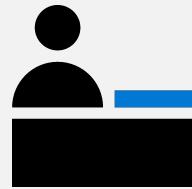
Lesson 4: ทำความรู้จักกับ non-relational data



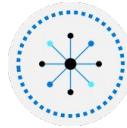
Lesson 5: เรียนรู้คอนเซปต์ data analytics

Lesson 2: Explore roles and responsibilities in the world of data

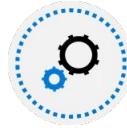
ตำแหน่งงานต่าง ๆ ในโลก Data



Lesson 2 objectives



ตำแหน่งงานในโลก Data: Database Admin, DA, DE



เครื่องมือต่าง ๆ ที่ใช้ในแต่ละตำแหน่งงาน

ตำแหน่งงานสาย Data

Database Administrator

ผู้ดูแลและจัดการระบบฐานข้อมูล
สำหรับจัดเก็บข้อมูล

หน้าที่ของ Database Admin:

- Database Management
- Implements Data Security
- Backups
- User Access
- Monitors performance



Data Engineer

ผู้จัดเตรียมข้อมูลให้พร้อมนำไป
วิเคราะห์

หน้าที่ของ Data Engineer:

- Data Pipelines and processes
- Data Ingestion storage
- Prepare data for Analytics
- Prepare data for analytical processing



Data Analyst

ผู้นำข้อมูลมาวิเคราะห์

หน้าที่ของ Data Analyst:

- Provides insights into the data
- Visual Reporting
- Modeling Data for Analysis
- Combines data for visualization and analysis



อีกตำแหน่ง คือ Data Scientist ที่จะดูแลเรื่องการนำข้อมูลมาสร้างโมเดล Machine Learning เพื่อทำนายผล ซึ่งเราอาจจะไม่ได้ลงลึกในคอร์สนี้

เครื่องมือของ Database administrator

Azure Data Studio

โปรแกรมสำหรับจัดการข้อมูลบนฐานข้อมูลหลากหลาย เชื่อมต่อได้ทั้งระบบฐานข้อมูลบนเซิร์ฟเวอร์บริษัท (on-premise) และบน Cloud

ใช้งานได้ทั้ง Windows, macOS, Linux

SQL Server Management Studio

โปรแกรมสำหรับจัดการข้อมูลบนฐานข้อมูลของ Microsoft

เชื่อมต่อได้ทั้งระบบฐานข้อมูลบนเซิร์ฟเวอร์บริษัท (on-premise) และบน Cloud

ใช้งานได้เฉพาะบน Windows

สามารถสั่งงานต่าง ๆ ที่ Azure Data Studio ทำไม่ได้ เช่น Import & Export Data

Azure Portal/CLI

ระบบจัดการ Azure

ใช้สั่งงานระบบ Azure เช่น สร้างฐานข้อมูล ลบฐานข้อมูล

Azure Portal = เว็บไซต์ portal.azure.com เปิดได้จาก Web Browser ไหนก็ได้

Azure CLI = เครื่องมือที่ติดตั้งได้บนเครื่อง หรือรันบน Azure Portal ก็ได้

เครื่องมือของ Data engineer

Azure Synapse Studio

บริการบน Azure Portal ที่จัดการ Data ได้ตั้งแต่ต้นจนจบ

ดึงข้อมูล (Data Ingestion) ด้วยเครื่องมือ Azure Data Factory

แปลงและวิเคราะห์ข้อมูล ด้วยบริการ Azure Synapse SQL Pools และ Spark Pools

SQL Server Management Studio

โปรแกรมสำหรับจัดการข้อมูลบนฐานข้อมูลของ Microsoft

เชื่อมต่อได้ทั้งระบบฐานข้อมูลบนเซิร์ฟเวอร์บริษัท (on-premise) และบน Cloud

ใช้งานได้เฉพาะบน Windows

Azure Portal/CLI

ระบบจัดการ Azure

ใช้ส่วนงานระบบ Azure เช่น สร้างฐานข้อมูล ลบฐานข้อมูล

เครื่องมือของ Data analyst

Power BI Desktop

โปรแกรมสำหรับนำข้อมูลมาทำ Data Visualization

ใช้งานได้เฉพาะบน Windows

Power BI Portal/ Power BI Service

ระบบเว็บไซต์ สำหรับจัดการ PowerBI

สามารถเปิดและแก้ไข Report (รายงาน) ที่สร้างโดย PowerBI ได้เล็กน้อย

สามารถทำ Dashboard ได้

แชร์ Report และ Dataset (ชุดข้อมูล) ให้กับคนอื่น ๆ ในองค์กรได้

Power BI Report Builder

โปรแกรมสำหรับสร้าง Paginated Report (รายงานที่มีการแบ่งเป็นหน้า ๆ)

ใช้งานได้เฉพาะบน Windows

Lesson 2: Knowledge check



Which one of the following tasks is a role of a database administrator?

- Backing up and restoring databases
 - Creating dashboards and reports
 - Identifying data quality issues
-



Which of the following tools is a visualization and reporting tool?

- SQL Server Management Studio
 - Power BI
 - SQL
-



Which one of the following roles is not a data job role?

- Systems Administrator
- Data Analyst
- Database Administrator

Chapter 1

ในบทนี้ เราจะเรียนกัน 5 หัวข้อ



Lesson 1: เรียนรู้ Core data concepts



Lesson 2: ตำแหน่งงานต่าง ๆ ในโลก Data



Lesson 3: ทำความรู้จักกับ Relational data

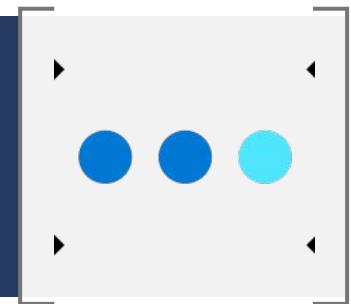


Lesson 4: ทำความรู้จักกับ non-relational data



Lesson 5: เรียนรู้คอนเซปต์ data analytics

Lesson 3: Describe concepts of relational data



Lesson 3 objectives



ทำความรู้จักกับข้อมูลแบบ Relational Data



รู้จักกับコンเซปต์ของ Table, index, view



Data workload สำหรับข้อมูลแบบ Relational Data

ตัวอย่างข้อมูลแบบ Relational data



IoT: (Internet of Things - อุปกรณ์อัจฉริยะ)

เช่น ข้อมูลจาก sensor ต่าง ๆ มักจะเข้ามาเป็นข้อมูลเล็ก ๆ แต่มีจำนวนมาก และเข้ามารวบเร็ว โดยเป็นได้ทั้งข้อมูลแบบ structured (relational data) หรือ semi-structured (non relational data)



Online transaction processing (OLTP)

ระบบประมวลผลธุกรรมมักจะอ่านเขียนข้อมูลเป็น structured data

For example order systems that perform many small transactional updates



Data warehousing: ระบบคลังข้อมูล

ระบบสำหรับเก็บข้อมูลจำนวนมาก จากหลายแหล่งข้อมูล (data source) และประมวลผลวิเคราะห์ด้วยความเร็วสูง

Tables (ตารางข้อมูล)

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX
107	Francis Ribeiro	XXX-XXX-XXXX

ข้อมูล Relational Data จะจัดเก็บในรูปแบบ
ตารางข้อมูล

ในตารางข้อมูล จะมี row (แนวอน) กับ
column (แนวตั้ง)

1 แถว = ข้อมูล 1 ก้อน

- ทุกก้อนจะมีคอลัมน์เหมือนกัน
- แต่ละคอลัมน์ก็จะมีข้อมูลประเภท
เดียวกัน เช่น ตัวเลข ตัวหนังสือ เบอร์
โทรศัพท์ วันที่ ฯลฯ

Entities

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX

ข้อมูล 1 แถว = 1 entity ซึ่งจะแทนของ 1
อย่าง เช่น ลูกค้า 1 คน, สินค้า 1 ชิ้น, การสั่งซื้อ
1 ครั้ง

ข้อมูลแต่ละ entity สามารถมี Relationship
(ความสัมพันธ์) กันได้

เช่น ลูกค้า 1 คนบนตาราง customers
สามารถเชื่อมกับการสั่งซื้อสินค้าได้หลายครั้ง
บน ตาราง orders

Normalization

Normalization คือ การแบ่งข้อมูลออกเป็นหลายตาราง

จากตัวอย่างนี้ จะเห็นว่า CustomerName กับ CustomerPhone มีการเก็บซ้ำทั้งตาราง Customers และ Orders

Customers		
CustomerID	CustomerName	CustomerPhone
100	Muisto Linna	XXX-XXX-XXXX
101	Noam Maoz	XXX-XXX-XXXX
102	Vanja Matkovic	XXX-XXX-XXXX
103	Qamar Mounir	XXX-XXX-XXXX
104	Zhenis Omar	XXX-XXX-XXXX
105	Claude Paulet	XXX-XXX-XXXX
106	Alex Pettersen	XXX-XXX-XXXX

Orders		
OrderID	CustomerName	CustomerPhone
AD100	Noam Maoz	XXX-XXX-XXXX
AD101	Noam Maoz	XXX-XXX-XXXX
AD102	Noam Maoz	XXX-XXX-XXXX
AX103	Qamar Mounir	XXX-XXX-XXXX
AS104	Qamar Mounir	XXX-XXX-XXXX
AR105	Claude Paulet	XXX-XXX-XXXX
MK106	Muisto Linna	XXX-XXX-XXXX

Normalization => Relational Data

ผลลัพธ์จากการทำ normalization เราจะเก็บ customer เป็น CustomerID ในตาราง Customers จะเห็นว่าเราไม่ต้องเก็บ CustomerName กับ CustomerPhone ซ้ำอีกรอบแล้ว

Customers			Orders		
CustomerID	CustomerName	CustomerPhone	OrderID	CustomerID	SalesPersonID
100	Muisto Linna	XXX-XXX-XXXX	AD100	101	200
101	Noam Maoz	XXX-XXX-XXXX	AD101	101	200
102	Vanja Matkovic	XXX-XXX-XXXX	AD102	101	200
103	Qamar Mounir	XXX-XXX-XXXX	AX103	103	201
104	Zhenis Omar	XXX-XXX-XXXX	AS104	103	201
105	Claude Paulet	XXX-XXX-XXXX	AR105	105	200
106	Alex Pettersen	XXX-XXX-XXXX	MK106	105	201

เวลาเราจะดูในตาราง Orders ว่าลูกค้าคนนี้เป็นใคร เราก็เอาไป join (เชื่อมต่อ) กับตาราง Customers ได้

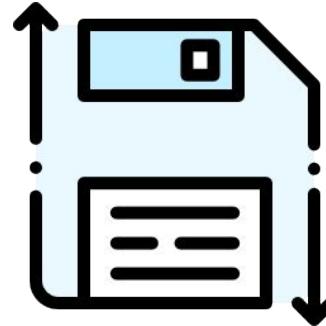
CustomerID ในตารางหลัก (ตาราง Customers) เรียกว่า primary key

CustomerID ในตารางอื่น ๆ ที่นำไปใช้ (เช่น ตาราง Orders) เรียกว่า foreign keys

สาเหตุของการทำ Normalization

ประโยชน์ของการทำ Normalization

1. ประยุกต์พื้นที่เก็บข้อมูล
2. ลด data duplication (ข้อมูลซ้ำ)
3. เพิ่ม data quality (คุณภาพข้อมูล)



กรณีศรีม:

ในงาน OLTP เราจะนิยมทำ Normalization หลายระดับ

ส่วนงาน OLAP จะไม่นิยมทำ Normalization หลายระดับ เนื่องจากถ้าข้อมูลแบ่งเป็นหลาย Table
มาก ต้องรอการ Join Table หลายชั้น จะทำให้ดึงข้อมูลซ้ำ

Index

Index เป็นความสามารถของฐานข้อมูล โดยเราบอกให้ index (จดจำ) เลขที่เดียวของค่า

ลั่มน้ำสำคัญไว้ ทำให้ในการค้นหาข้อมูล Database ที่เกี่ยวข้องกับคอลัมน์นี้ ไม่ต้องอ่านข้อมูล
ตั้งแต่แรก เพื่อให้ค้นหาข้อมูลได้เร็วขึ้น (faster data retrieval)

score_id	student_id	subject	score
1	1	Maths	70
2	1	Computer	85
3	1	Science	74
4	2	Maths	68
5	2	Computer	72
6	2	Science	83
7	3	Maths	56
8	3	Computer	92
9	3	Science	47

Index on **student_id**

แบ่ง Partition

Student_id <= 2

Student_id > 2

score_id	student_id	subject	score
1	1	Maths	70
2	1	Computer	85
3	1	Science	74
4	2	Maths	68
5	2	Computer	72
6	2	Science	83

score_id	student_id	subject	score
7	3	Maths	56
8	3	Computer	92
9	3	Science	47

เช่น ถ้าต้องการหาข้อมูลที่ student_id = 1 ก็อ่านเฉพาะ Partition แรกได้เลย

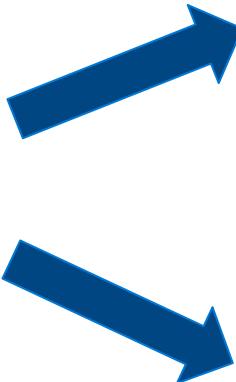
Table & View

View คือ การนำข้อมูลใน Table มาสร้างเป็น Virtual Table เพื่อคำนวณค่าต่าง ๆ และเก็บไว้ในชื่อที่เรียกใช้ช้าได้ง่าย หรือนำมาซ่อนบาง colum ที่เราไม่อยากแชร์ให้คนอื่นก็ได้

ตัวอย่างเช่น:

Table ข้อมูลการขายสินค้า

customer	product	product_category
Mark	corn	vegetable
Two	carrot	vegetable
John	pen	stationery
James	corn	vegetable



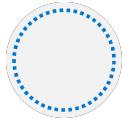
View: จำนวนครั้งการซื้อ ของลูกค้าแต่ละคน

customer_name	purchase_count
Mark	1
Two	1
John	1
James	1

View: ยอดขายสินค้าแต่ละชนิด

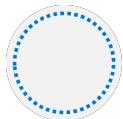
product_name	purchase_count
corn	2
carrot	1
pen	1

Lesson 3: Knowledge check



Which one of the following statements is a characteristic of a relational database?

- All data must be stored as character strings
 - A row in a table represents a single entity
 - Different rows in the same table can contain different columns
-



What is an index?

- A structure that enables you to locate rows in a table quickly, using an indexed value
- A virtual table based on the result set of a query
- A structure comprising rows and columns that you use for storing data

Chapter 1

ในบทนี้ เราจะเรียนกัน 5 หัวข้อ



Lesson 1: เรียนรู้ Core data concepts



Lesson 2: ตำแหน่งงานต่าง ๆ ในโลก Data



Lesson 3: ทำความรู้จักกับ Relational data



Lesson 4: ทำความรู้จักกับ non-relational data



Lesson 5: เรียนรู้คอนเซปต์ data analytics

Lesson 4: Explore concepts of non-relational data



Lesson 4 objectives



ทำความรู้จักกับ non-relational data



ประเภทของ non-relational data

Non-relational Data คืออะไร

Non-relational data คือ ข้อมูลที่ไม่ได้มีความสัมพันธ์แบบเป็นตาราง

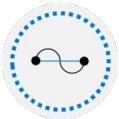
การเก็บข้อมูลจะมีความยืดหยุ่นมากขึ้น เพราะไม่มีการบังคับว่าข้อมูลทุกແவจะต้องมีค่าล้วนๆ เช่นกัน

ตัวอย่างข้อมูล Non-Relational Data

```
## Customer 1 ID: 1
Name: Mark Hanson
Telephone: [ Home: 1-999-9999999, Business: 1-888-8888888, Cell: 1-777- 7777777 ]
Address: [ Home: 121 Main Street, Some City, NY, 10110,
           Business: 87 Big Building, Some City, NY, 10111 ]

## Customer 2 ID: 2
Title: Mr
Name: Jeff Hay
Telephone: [ Home: 0044-1999-333333, Mobile: 0044-17545-444444 ]
Address: [ UK: 86 High Street, Some Town, A County, GL8888, UK,
           US: 777 7th Street, Another City, CA, 90111 ]
```

ตัวอย่างข้อมูลที่เป็น Non-relational



IoT: (Internet of Things - อุปกรณ์อัจฉริยะ)

เช่น ข้อมูลจาก sensor ต่าง ๆ มักจะเข้ามาเป็นข้อมูลเล็ก ๆ แต่มีจำนวนมาก และเข้ามาราดเร็ว โดยเป็นได้ทั้งข้อมูลแบบ structured (relational data) หรือ semi-structured (non-relational data)



Retail and Marketing: ข้อมูลการการซื้อขาย และการตลาด

ปัจจุบันมีการใช้ document database (จะอธิบายเพิ่มเติมหลังจากนี้) เข้ามาเก็บข้อมูลประเภทนี้ มากขึ้นเรื่อย ๆ เพราะรองรับข้อมูลขนาดใหญ่ทั่วโลก และใช้งานได้รวดเร็ว



Gaming: ข้อมูลเกม และโลก social

เช่น สติติในเกม, ตารางคะแนน, ข้อมูล social media



Web and Mobile: ข้อมูลเว็บไซต์หรือแอพมือถือ ที่ต้องการความเร็วสูง

เช่น ข้อมูลการคลิกโฆษณา, การเข้าชมเว็บไซต์, การใช้งานแอ�플ิเคชั่นมือถือ

2 ประเภทของ non-relational data (1)

1. Semi-structured data

ข้อมูลที่จัดเก็บแบบไม่ได้เป็นตาราง ยังมีรูปแบบ (Structure) ของตัวเองบ้าง แต่ไม่มีความยึดหยุ่น เพิ่ม-ลดคอลัมน์ได้

ตัวอย่าง semi-structured data:

JSON

AVRO

ORC

Parquet

ประเภทของ Semi-Structured data (NoSQL)

Key-value stores

เหมือน Structured Data ที่มีแค่ 2 คอลัมน์

คือ 1 แถวจะมีแค่ key (ชื่อเดา) กับ value (ข้อมูล)

Document based

เก็บข้อมูลในรูปแบบเอกสาร (Document)

เหมือน key-value store ที่ในคอลัมน์ value เก็บข้อมูลที่ซับซ้อนขึ้นได้

Column family databases

หน้าตาคล้ายตาราง Relational Data ที่มี key ชัดเจน (เรียกว่า row key) และไม่มีชื่อ จำกัดว่าแต่ละคอลัมน์จะต้องเป็นข้อมูลประเภทเดียวกัน

Graph Databases

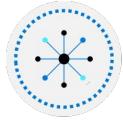
เก็บข้อมูลในรูปแบบของกราฟ

เช่น กราฟเพื่อนใน Social Media

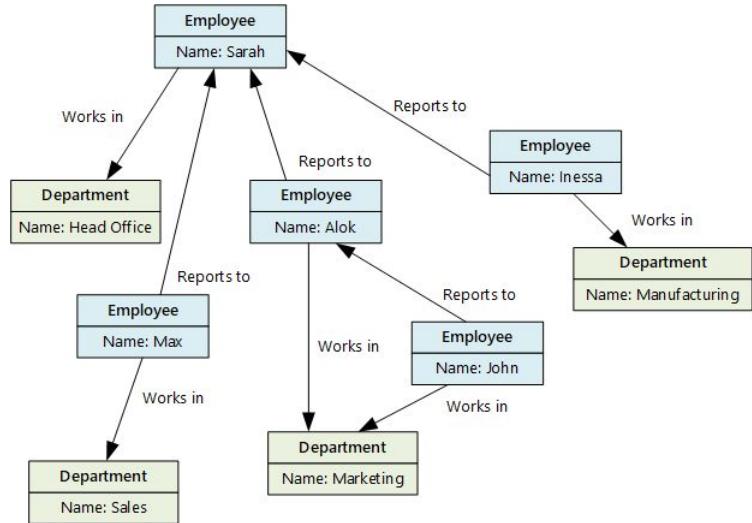
Graph database ทำงานอย่างไร



เก็บความล้มพันธ์ระหว่างแต่ละ entity



สามารถค้นหาข้อมูลจาก entity (เรียกว่า **node**)
และความสัมพันธ์ของ entity (เรียกว่า **edge**)

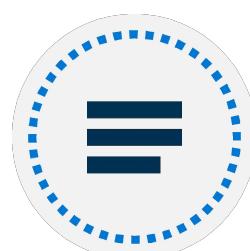
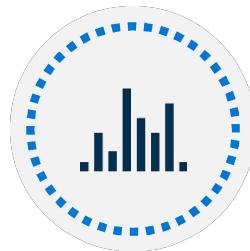


2 ประเภทของ non-relational data (2)

2. Unstructured Data

เป็นข้อมูลที่ไม่มีพิลด์ข้อมูลชัดเจน

ตัวอย่าง Unstructured Data เช่น ไฟล์วีดิโอ, ไฟล์เสียง, ไฟล์รูปภาพ, ไฟล์ตัวหนังสือ



มักจะใช้ร่วมกับ Machine Learning เพื่อวิเคราะห์ข้อมูล

เช่น ใช้บริการด้าน ML ของ Azure ซึ่ง Vision API ในการวิเคราะห์รูปภาพ

Lesson 4: Knowledge check



Which of the following services should you use to implement a non-relational database?

- Azure Cosmos DB
 - Azure SQL Database
 - The Gremlin API
-



Which of the following is a characteristic of non-relational databases?

- Non-relational databases contain tables with flat fixed-column records
 - Non-relational databases require you to use data normalization techniques to reduce data duplication
 - Non-relational databases are either schema free or have relaxed schemas
-



You are building a system that monitors the temperature throughout a set of office blocks, and sets the air conditioning in each room in each block to maintain a pleasant ambient temperature. Your system has to manage the air conditioning in several thousand buildings spread across the country or region, and each building typically contains at least 100 air-conditioned rooms.

What type of NoSQL data store is most appropriate for capturing the temperature data to enable it to be processed quickly?

- A key-value store
- A column family database
- Write the temperatures to a blob in Azure Blob storage

Chapter 1

ในบทนี้ เราจะเรียนกัน 5 หัวข้อ



Lesson 1: เรียนรู้ Core data concepts



Lesson 2: ตำแหน่งงานต่าง ๆ ในโลก Data



Lesson 3: ทำความรู้จักกับ Relational data

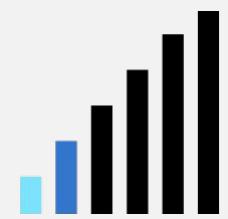


Lesson 4: ทำความรู้จักกับ non-relational data



Lesson 5: เรียนรู้คอนเซปต์ data analytics

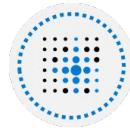
Lesson 5: Explore concepts of data analytics



Lesson 5 objectives



ทำความรู้จักระดับต่าง ๆ ของ data analytics (การวิเคราะห์ข้อมูล)



ทำความรู้จักรอบ data ingestion (การดึงข้อมูล) and processing (การประมวลผลข้อมูล)



ทำความรู้จักรอบ data visualization

ระดับต่าง ๆ ของ data analytics (การวิเคราะห์ข้อมูล)



Level 1
Descriptive

วิเคราะห์จากข้อมูล
ที่มีแบบเบื้องต้น

เพื่อตอบคำถามว่า
**WHAT IS
HAPPENING in
my business?**



Level 2
Diagnostic

วิเคราะห์จากข้อมูล
ที่มีแบบละเอียด

เพื่อตอบคำถามว่า
**WHY IS IT
HAPPENING in
my business**



Level 3
Predictive

นำข้อมูลที่มีอยู่ มา
ทำนายผลว่าจะเกิด
อะไรขึ้นต่อไป

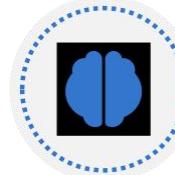
เพื่อตอบคำถามว่า
**WHAT IS LIKELY
TO HAPPEN in the
future based on
historical data?**



Level 4
Prescriptive

นำผลลัพธ์จากการ
ทำนายผลด้วยข้อมูล
มาใช้ประกอบการ
ตัดสินใจ

เพื่อตอบคำถามว่า
**WHAT IS THE
BEST course of
action?**



Level 5
Cognitive

สร้างระบบที่ทำนาย
อนาคต เรียนรู้จาก
ปัจจุบัน และตัดสินใจ
จากสิ่งที่เรียนรู้มาได้
เหมือนมนุษย์จริง

การเดินทางของข้อมูล ตั้งแต่ต้นจนจบ

Data Ingestion

การดึงข้อมลมาจากการแหล่งข้อมูล (Data Source) และจดเก็บในระบบเก็บข้อมูลของเรา



Data Processing

นำข้อมูลดิบ (Raw Data) มาแปลงเป็นข้อมูลที่มีคุณภาพสูง (Good Data Quality) เพื่อให้พร้อมใช้งาน

ETL: Extract and transform in one server and load to another datawarehouse
ETL

Extract, Transform, and Load



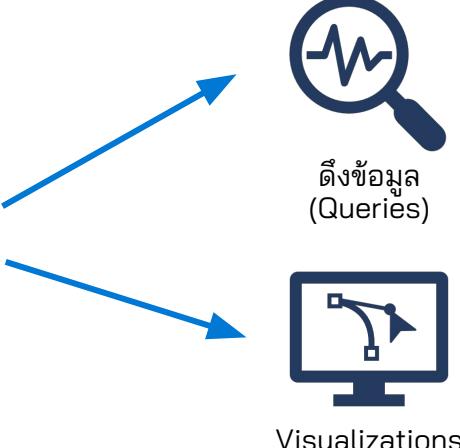
ELT
Extract, Load, and Transform



ELT: Extract then load and transform in same Datawarehouse

Data Visualization

นำข้อมูลมาทำเป็นกราฟฟิกเพื่อสื่อสารให้ผู้อื่นเข้าใจได้ง่าย และสะดวกต่อการนำไปวิเคราะห์



Data visualization

การนำข้อมูลจำนวนมากราบเปลี่ยนรูปแบบเป็นกราฟฟิก เพื่อให้เข้าใจข้อมูลได้ง่าย ซึ่งสามารถรวมกราฟฟิกมาเป็นรายงาน (Report) หรือ Dashboard ได้อีกด้วย

มีประโยชน์ทั้งในด้านการสื่อสารให้ผู้อื่นเข้าใจได้ง่าย (เช่น ระดับ C-level หรือทีมอื่น ๆ) และสะดวกต่อการนำไปวิเคราะห์

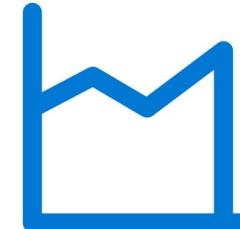
สามารถทำออกมาได้หลายรูปแบบ:



Report (รายงาน)



นำข้อมูลธุรกิจมาวิเคราะห์ หรือ
Business intelligence (BI)



แปลงข้อมูลเป็นกราฟฟิก
(Data visualization)

Lesson 5: Knowledge check



What is data ingestion?

- The process of transforming raw data into models containing meaningful information
 - Analyzing data for anomalies
 - Capturing raw data streaming from various sources and storing it
-



Which one of the following visuals displays the major contributors to a selected result or value?

- Key influencers
 - Column and bar chart
 - Matrix chart
-

在此处键入文本



Which type of analytics helps answer questions about what has happened in the past?

- Descriptive analytics
- Prescriptive analytics
- Predictive analytics

Reference - Key influencer Chart:

<https://docs.microsoft.com/en-us/power-bi/visuals/power-bi-visualization-influencers>

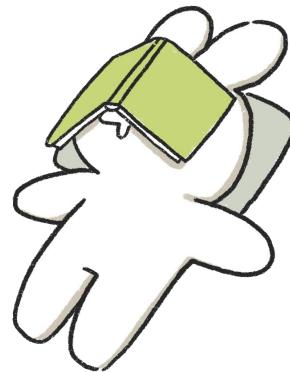
หากมีคำถาม ถามได้ที่ไหน



ถามคำถามไว้ใน chat ด้านขวา หัวข้อ Q&A

โดยจะมาตอบช่วงหลังเบรก หรือช่วงท้าย CLASS

หากพิมพ์เข้ามาในกล่องเขตทั่วไป คำถามอาจจะโดนดันจนไม่ได้ตอบ



BREAK TIME



หากมีคำถาม ถามได้ที่ไหน



ถามคำถามไว้ใน chat ด้านขวา หัวข้อ Q&A

โดยจะมาตอบช่วงหลังเบรก หรือช่วงท้าย CLASS

หากพิมพ์เข้ามาในกล่องเขตทั่วไป คำถามอาจจะโดนดันจนไม่ได้ตอบ

AZURE DATA FUNDAMENTALS TRAINING

CH2: RELATIONAL DATA



Chapter 2

ในบทนี้ เราจะเรียนกัน 3 หัวข้อ



Lesson 1: รู้จักกับ relational data services



Lesson 2: ทดลองเปิดใช้ (provision) บริการ relational database



Lesson 3: ดึงข้อมูล (Query) จากบริการ relational data

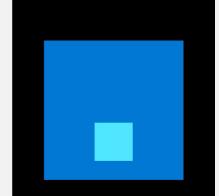
FAQ: คอร์สนี้ต้องสมัครสมาชิก Azure เพื่อเรียนมั้ย ?

สมัครหรือไม่สมัครก็ได้

ในคอร์สนี้จะเป็นคอนเซปต์ความรู้ของ Azure ซึ่งเราจะมีกดใช้งานบริการให้ดู แต่จะดูเฉย ๆ หรือทำตามเพื่อเรียนรู้ไปด้วยกันก็ได้ (ขออนุญาตไม่หยุดรอ เพราะไม่ใช่ Workshop)

ถ้าทำตาม กดลงบริการที่สร้างมาทุกรุ่น หลังจากจบคลาส
ไม่อย่างนั้นถ้าลืมลงบริการไหน อาจจะโดนหักเงินไปเรื่อย ๆ

Lesson 1: Explore relational data services in Azure



Lesson 1 objectives



บริการด้าน Data บน Azure



IaaS vs PaaS



SQL Server บนบริการ Azure virtual machines



บริการ Azure SQL DB



บริการ PostgreSQL, MySQL, MariaDB
บน Azure

บริการด้าน Data ของ Azure



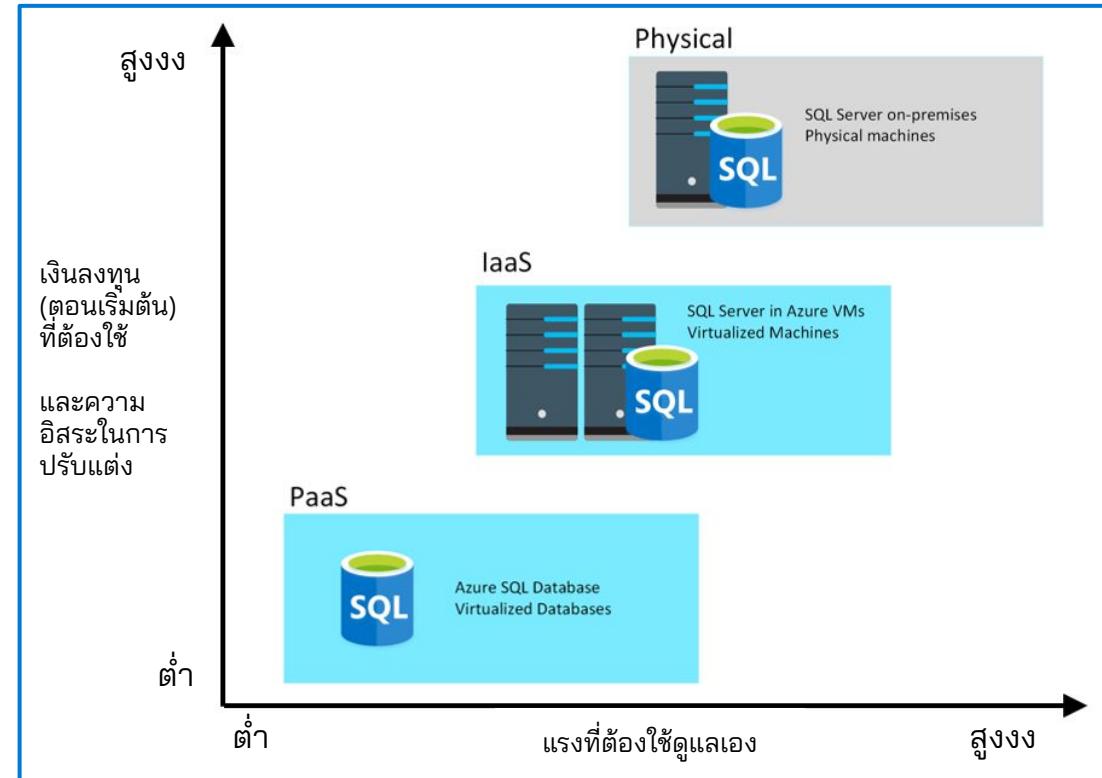
Azure มีบริการสำหรับเก็บข้อมูลบน Cloud ให้เลือกใช้ตามความเหมาะสมหลายบริการ ซึ่งแต่ละบริการก็จะมีจุดเด่นที่แตกต่างกันออกไป รวมถึงบประมาณที่ใช้ก็แตกต่างกันออกไป

บริการบน Cloud: IaaS (Infrastructure as a Service) vs PaaS (Platform as a Service)

Physical = ติดตั้งระบบฐานข้อมูลบนเครื่อง Server บริษัท (ไม่อยู่บน Cloud)
ผู้ใช้ต้องดูแลเองทั้งหมด และใช้เงินลงทุนเยอะ

IaaS = บริการที่ผู้ใช้เลือกเช่าแค่เครื่อง Server และมาติดตั้งระบบฐานข้อมูลเอง ต้องดูแลเอง

PaaS = บริการที่ Azure ดูแลระบบฐานข้อมูลให้แทนทั้งหมด ผู้ใช้แทนไม่ต้องออกแบบ แค่เข้ามายกdraftสร้าง แล้วใช้ได้เลย



Microsoft SQL Server (หรือ SQL Server) คืออะไร

ระบบฐานข้อมูล Relational Database อันดับ
ต้น ๆ ของโลก ที่พัฒนาโดย Microsoft

จุดเด่น คือ มีภาษาในการดึงข้อมูลที่พัฒนาขึ้นมา^{เพิ่มเติม} ชื่อ Transact SQL (หรือ T-SQL) ทำให้
เพิ่มความยืดหยุ่นในการใช้งาน

ปกติแล้ว ลูกค้าจะต้องซื้อใบอนุญาต (License)
ของ Microsoft SQL Server เพื่อนำมาใช้ใน
เซิร์ฟเวอร์ของบริษัท (On-premise)



ถ้าต้องการใช้ SQL Server บน Azure ทำได้อย่างไรบ้าง

— Infrastructure as a Service — | — Platform as a Service —



Option 1)
ติดตั้ง SQL Server บนบริการ Azure
Virtual Machines



Option 2)
Azure SQL Database



Option 3)
Azure SQL Managed Instance

Option 1) ติดตั้ง SQL Server บนบริการ Azure Virtual Machines

การติดตั้ง SQL Server บนบริการ Azure Virtual Machines (บริการเช่าเครื่องเซิร์ฟเวอร์บน Azure) เป็นบริการแบบ IaaS ที่ช่วยให้เรามีความยืดหยุ่นสูงสุด เพราะสามารถจัดการเซิร์ฟเวอร์ได้เองทั้งหมด โดยที่ไม่ต้องดูแลเซิร์ฟเวอร์ท่ากับการตั้งเซิร์ฟเวอร์เอง

- มั่นใจได้ว่าจะรองรับกับ SQL Server ในเซิร์ฟเวอร์บริษัท (On-premise) แน่นอน เพราะสามารถจัดการเซิร์ฟเวอร์ได้เองทั้งหมด
- ข้อเสีย คือ ต้องจัดการเองทั้งหมด เช่น การอัพเดท OS, การอัพเดท software, backups, การขยายเซิร์ฟเวอร์ (replication & scaling)
- ยังจำเป็นต้องจ่ายค่าใบอนุญาตของ SQL Server



Option 2) ใช้บริการ Azure SQL Database

บริการ Azure SQL Database เป็นบริการแบบ PaaS ที่ผู้ใช้ได้รับฐานข้อมูลพร้อมใช้งานที่ไม่ต้องติดตั้งอะไรเพิ่มเองเลย มีฐานข้อมูลให้เลือกหลากหลายยี่ห้อ ไม่รองรับทุกฟีเจอร์ของ SQL Server เต็มที่

- ค่าใช้จ่ายไม่แพง และ Azure ดูแลฐานข้อมูลให้แทนทั้งหมด
- เหมาะสมกับโครงการใหม่ที่ไม่ต้องย้ายฐานข้อมูลจาก On-premise ขึ้นมา เพราะอาจจะปรับแต่งไม่ได้ 100%
- สามารถเพิ่ม - ลดขนาด (scale up & down) เพื่อรับรองการใช้งานที่เหมาะสมได้รวดเร็ว ไม่ต้องรีสตาร์ทเซิร์ฟเวอร์

สามารถเลือกสร้างได้ 2 แบบ

Single Database



ฐานข้อมูล 1 ฐาน ใช้ทรัพยากร่อง
ทั้งหมด

Elastic Pool



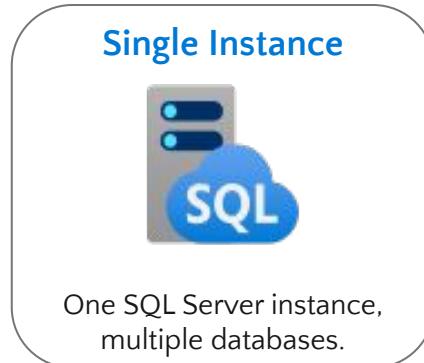
สร้างเป็น Pool รวมขึ้นมา ที่จะแชร์
ทรัพยากรกัน เช่น พื้นที่, RAM,
Processing Power ฯลฯ

Option 3) ใช้บริการ Azure SQL Managed Instance

บริการ Azure SQL Managed Instance เป็นบริการ PaaS เป็นบริการที่เกิดมาสำหรับการย้าย (Migration) จาก SQL Server บน on-premise ขึ้นมาอยู่บน Azure รองรับฟีเจอร์แทบทั้งหมดของ SQL Server และ Azure ดูแลระบบฐานข้อมูลให้

- ตัวอย่างสิ่งที่ Azure ดูแลให้: Automatic backups, software patching, database monitoring, and other administrative tasks
- รองรับฟีเจอร์ของ SQL Server แบบ on-premise แทบทั้งหมด

สามารถเลือกสร้างได้ 2 แบบ



บริการ PostgreSQL, MySQL, MariaDB บน Azure

Azure SQL DB รองรับฐานข้อมูลยี่ห้ออื่น นอกจาก Microsoft SQL Server



Azure Database for PostgreSQL



Azure Database for MySQL



Azure Database for MariaDB

MariaDB เป็นฐานข้อมูลที่เกิด (fork) มาจาก MySQL ที่ถูก Oracle ซื้อไปในปี 2008

ข้อดีของการใช้บริการ Azure Database for PostgreSQL, MySQL, MariaDB



Fully managed community database:

ใช้ระบบฐานข้อมูลฟรีที่คุณชื่นชอบแล้วบน Azure ได้เลย



Built-in high availability for lowest TCO (Total Cost of Ownership):

เซิร์ฟเวอร์ไม่ล่มง่าย ๆ โดยไม่ต้องจ่ายเงินซื้อเซิร์ฟเวอร์ราคาสูง



Intelligent performance and scale:

ขยายขนาดตามการใช้งานได้อัตโนมัติ พื้นที่เก็บสูงสุดถึง 16TB และพลังการประมวลผลสูงสุดถึง 20K IOPs



Industry-leading security and compliance:

ความปลอดภัยสูง และรองรับบริการด้าน Security กับ Compliance ของ Azure เช่น Advanced Threat Protection



Integration with the Azure ecosystem:

เชื่อมต่อกับบริการต่าง ๆ ของ Azure ได้ง่าย

Lesson 1: Knowledge check (continued on next slide)



Which deployment requires the fewest changes when migrating an existing SQL Server on-premises solution?

- Azure SQL Database Managed Instance
 - SQL Server running on a virtual machine
 - Azure SQL Database Single Database
-



Which of the following statements is true about SQL Server running on a virtual machine?

- You must install and maintain the software for the database management system yourself, but backups are automated
 - Software installation and maintenance are automated, but you must do your own backups
 - You're responsible for all software installation and maintenance, and performing back ups
-



Which of the following statement is true about Azure SQL Database?

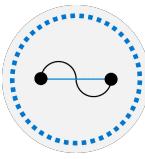
- Scaling up doesn't take effect until you restart the database
- Scaling out doesn't take effect until you restart the database
- Scaling up or out will take effect without restarting the SQL database

Lesson 1: Knowledge check (continued)



When using an Azure SQL Database managed instance, what is the simplest way to implement backups?

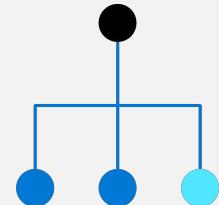
- Manual Configuration of the SQL server
 - Create a scheduled task to back up
 - Backups are automatically handled
-



What is the best way to transfer the data in a PostgreSQL database running on-premises into a database running Azure Database for PostgreSQL service?

- Export the data from the on-premises database and import it manually into the database running in Azure
- Upload a PostgreSQL database backup file to the database running in Azure
- Use the Azure Database Migration Services

Lesson 2: Explore provisioning and deploying relational database services in Azure



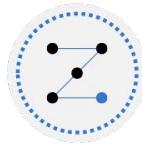
Lesson 2 objectives



การ provision บริการ relational data services



การตั้งค่า relational data services



จะจัดกับปัญหาการเชื่อมต่อ (connectivity issues) ที่พบบ่อย



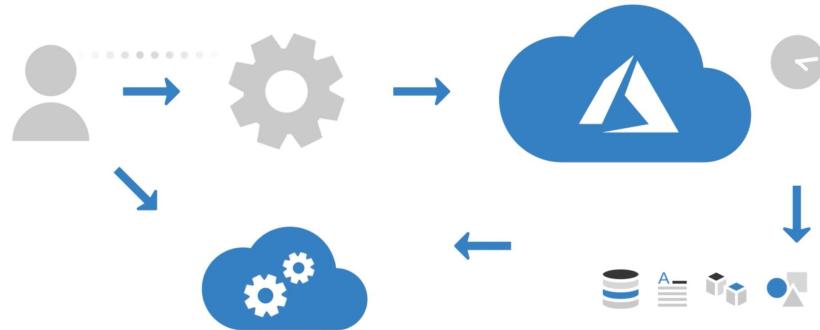
จะจัดกับ data security

การ Provision คืออะไร

Provision คือ การแจ้ง Azure ว่าต้องการสร้างบริการไหน เช่น ต้องการสร้าง Azure SQL DB ใหม่

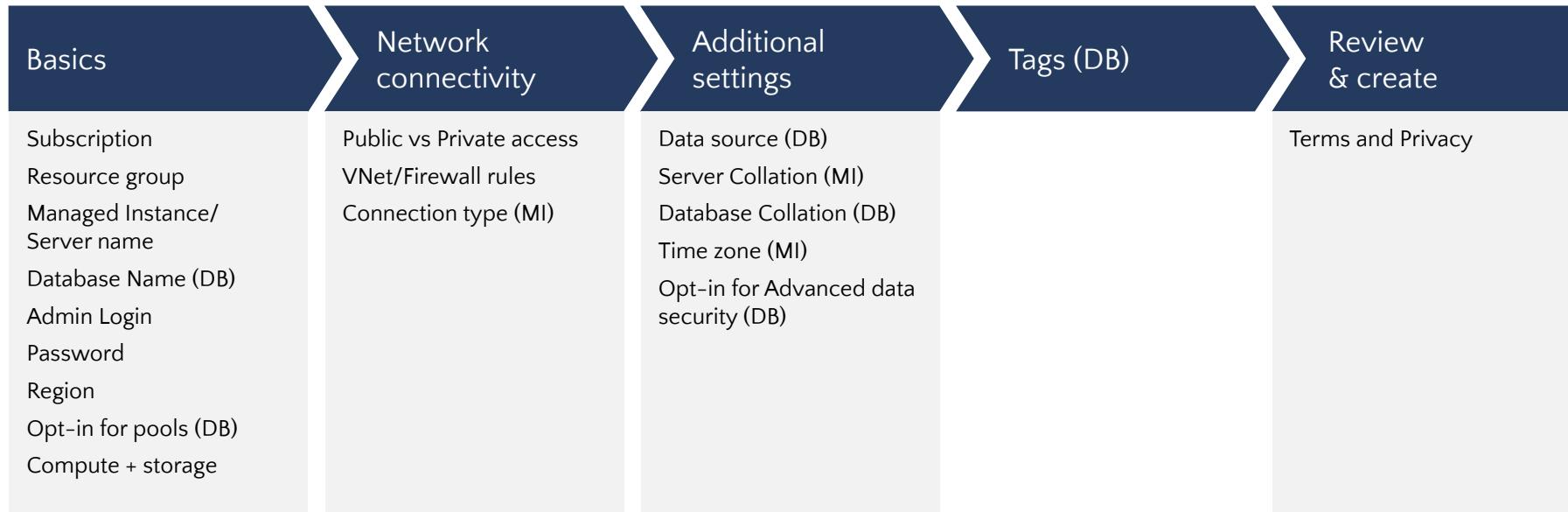
ในการ Provision เราต้องตั้งค่าบริการด้วย เช่น จะใช้เซิร์ฟเวอร์ขนาดไหน, username - password อะไร ฯลฯ

หลังจากนั้น Azure ก็จะสร้างบริการนั้นเพื่อให้เราใช้งาน ซึ่งบางบริการ หลังจากสั่ง Provision อาจจะต้องรอหลายนาที



การตั้งค่า relational data services

ตอนสร้างฐานข้อมูลใหม่ เราสามารถตั้งค่าต่าง ๆ ได้ดังนี้



Demo: พาไปลองสร้างฐานข้อมูล ด้วยบริการ Azure SQL Database

เข้าเว็บไซต์: <https://portal.azure.com> และสมัครสิทธิใหม่ หรือล็อกอิน เพื่อเริ่มต้นใช้งาน Azure

ค้นหาคำว่า “SQL Database” ในแท็บเลิร์จด้านบน

แล้วคลิก Create

สำคัญ:

ในการสร้างฐานข้อมูล จะต้องกำหนด username / password ของแอดมิน เพื่อเข้าใช้งานฐานข้อมูล

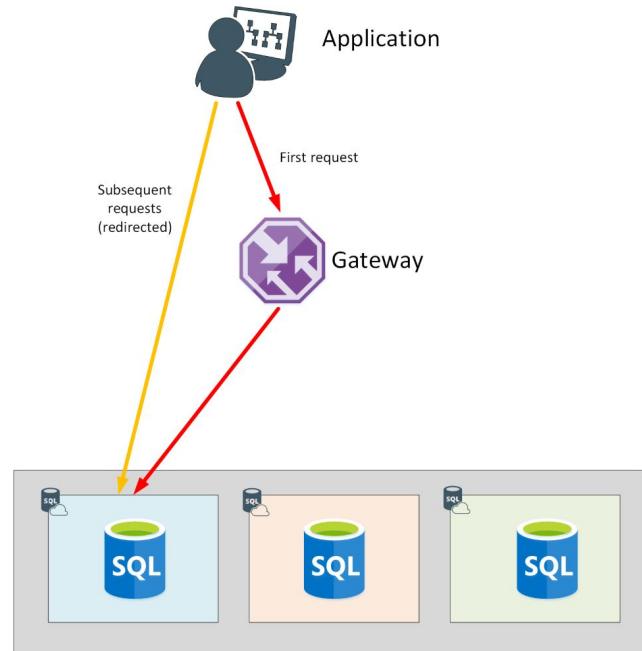
The screenshot shows the Microsoft Azure portal homepage. At the top is a search bar with the placeholder "Search resources, services, and docs (G+)" and a "Create a resource" button. Below the search bar is a row of service icons: SQL databases, Managed databases, SQL elastic pools, App Services, Cost Management ..., Data factories, Virtual machines, Storage accounts, and a "More services" link. A large blue arrow points down from the search bar area to the "Recent resources" section. This section lists recent items with columns for Name, Type, and Last Viewed. The items listed are: Visual Studio Enterprise FY16 (Subscription, 17 hours ago), test-vm-ip (Public IP address, 8 months ago), perth-rg (Resource group, 8 months ago), test-perth-azfunction (Function App, 10 months ago), perth-adf-test (Data factory (V2), 10 months ago), and testperth (Synapse workspace, a year ago). A "See all" link is at the bottom of the list.

The screenshot shows the Microsoft Azure portal with a blue header bar containing the Microsoft logo and the word "Azure". Below the header is a "Home" link and a navigation menu icon. The main content area is titled "SQL databases" and features a "Student Ambassadors" badge. At the top of this section are three buttons: "Create", "Reservations", and "Manage".

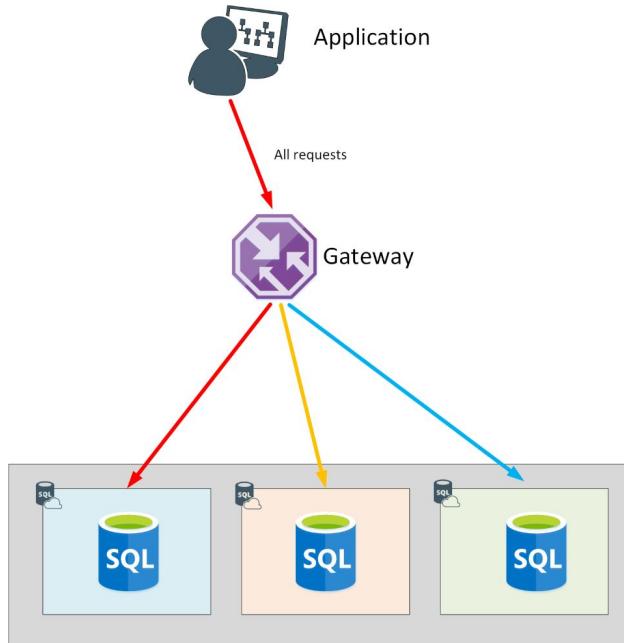
การเชื่อมต่อ (Connectivity) จากด้านใน Azure

Policy of Redirect

- ในการเชื่อมต่อครั้งแรก จะเชื่อมต่อผ่านตัวกลาง (gateway) ซึ่งจะทำตัวเป็นส่วนหนึ่งของฐานข้อมูล
- หลังจากการเชื่อมต่อครั้งแรก จะเชื่อมต่อ (redirect) เข้าสู่ฐานข้อมูลได้โดยตรงเลย



การเชื่อมต่อ (Connectivity) จากด้านนอก Azure



Policy of Proxy

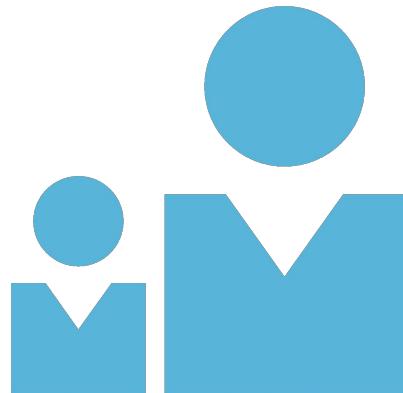
การเชื่อมต่อทั้งหมด ไม่ว่าจะครั้งแรกหรือครั้งต่อ ๆไป จะผ่านตัวกลาง (Gateway หรือ Proxy) ทุกครั้ง

Azure Role Based Access Control (RBAC)

Azure Role Based Access Control (RBAC) เป็นระบบที่ช่วยในการกำหนดสิทธิ์ การเข้าถึงบริการต่าง ๆ ของ Azure (รวมถึงฐานข้อมูล) ซึ่งทำให้กำหนดสิทธิ์ได้แบบ ปลอดภัยกว่าการใช้ username / password

การกำหนดสิทธิ์ (Role Assignment) มี 3 ส่วนประกอบหลัก

- **Security principal:** user หรือแอพพลิเคชัน ที่ต้องการ เชื่อมต่อ
- **Role:** ตำแหน่ง เช่น อ่านได้อย่างเดียว (Reader)
- **Scope:** กำหนดว่า user หรือบริการนี้ เชื่อมต่อบริการ อะไรได้บ้าง



Azure DB – Read replicas

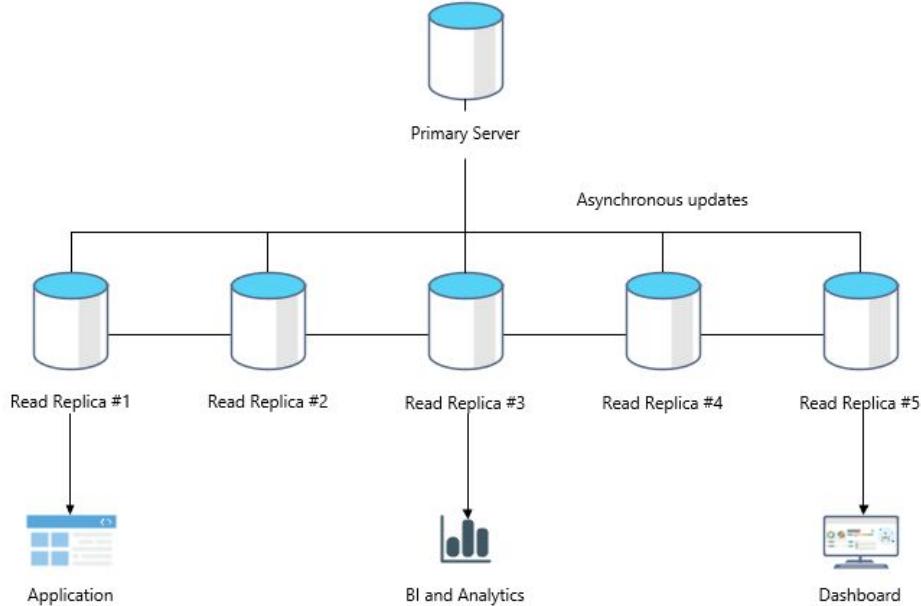
ฐานข้อมูลที่มีผู้ใช้อ่านข้อมูลเยอะ ๆ จะต้องใช้พลังในการประมวลผลเยอะ และอาจทำให้ฐานข้อมูลล่มได้ เช่น งานวิเคราะห์ข้อมูล (Data Analytics, Business Intelligence), ระบบเว็บไซต์ที่มีคนใช้เยอะ

Read Replicas เป็นการสร้าง Server ฐานข้อมูลขึ้นมาอีกตัว ก็อปปี้ข้อมูลจากฐานข้อมูลหลักมา แล้วให้คนที่อ่านเยอะ ๆ ก็ไปอ่านฐานข้อมูลอันใหม่แทน

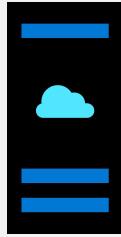
ประโยชน์ของ Read Replica

- เพิ่มประสิทธิภาพ และความสามารถในการรองรับผู้ใช้งานจำนวนมาก (scale) เฉพาะผู้ใช้ที่อ่านข้อมูล (read) เท่านั้น
- ใช้แทนฐานข้อมูลหลักหากเกิดภัยพิบัติ (Disaster Recovery) ได้ เช่น ไฟไหม้ Data Center ของฐานข้อมูลหลัก

Create up to five read-only replicas of the primary server



Lesson 3: Query relational data in Azure



Lesson 3 objectives



การดึงข้อมูลแบบ relational data



การดึงข้อมูลโดยใช้ภาษา SQL

ทำความรู้จักกับภาษา SQL



SQL เป็นภาษาพื้นฐานที่ดึงข้อมูลจาก relational databases



SQL มีมาตรฐานชื่อ ANSI and ISO



ระบบ RDBMS (Relational Database Management System หรือ ระบบฐานข้อมูล Relational Data) มักจะมีภาษา SQL ที่พัฒนาเพิ่มขึ้นของตัวเอง เช่น T-SQL, PL/SQL, pgSQL

กลุ่มของคำสั่ง ในภาษา SQL

DML

Data Manipulation Language

ใช้สำหรับดึงข้อมูล
ตัวอย่างคำสั่ง:

SELECT, INSERT, UPDATE,
DELETE

DDL

Data Definition Language

ใช้สำหรับสร้าง และจัดการฐาน
ข้อมูล รวมถึงลิงต่าง ๆ ด้านใน
ฐานข้อมูล เช่น Table

ตัวอย่างคำสั่ง:

CREATE, ALTER, DROP,
RENAME

DCL

Data Control Language

ใช้สำหรับจัดการด้านความ
ปลอดภัย เช่น ให้สิทธิอ่านข้อมูล
กับผู้ใช้

ตัวอย่างคำสั่ง:

GRANT, REVOKE, DENY

Use DML statements

Statement	Description
SELECT	เลือกข้อมูล / ค้นหาข้อมูลจาก Table
INSERT	เพิ่มข้อมูลใหม่ใส่ table
UPDATE	แก้ไขข้อมูลใน table
DELETE	ลบข้อมูลบางส่วนใน table

ส่วนประกอบของ SELECT Statement

Clause	Expression
SELECT	<select list>
FROM	<table or view>
WHERE	<search condition>
GROUP BY	<group by list>
ORDER BY	<order by list>

ตัวอย่างคำสั่ง SELECT statement

```
SELECT EmployeeId, YEAR(OrderDate) AS OrderYear  
FROM Sales.Orders  
WHERE CustomerId = 71  
  
GROUP BY EmployeeId, YEAR(OrderDate)  
HAVING COUNT(*) > 1  
ORDER BY EmployeeId, OrderYear;
```

ตัวอย่างคำสั่ง INSERT statement

การ INSERT ข้อมูลทีละเดียว

```
INSERT INTO Sales.OrderDetails  
    (orderid, productid, unitprice, qty, discount)  
VALUES  (10255,39,18,2,0.05);
```

การ INSERT ข้อมูลทีละหลายเดียว

```
INSERT INTO Sales.OrderDetails  
    (orderid, productid, unitprice, qty, discount)  
  
VALUES  
(10256,39,18,2,0.05),  
(10258,39,18,5,0.10);
```

การใช้ DDL statements

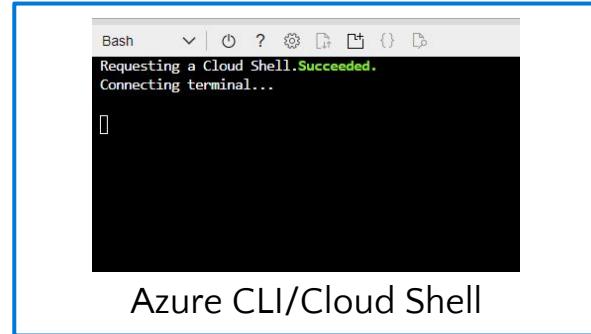
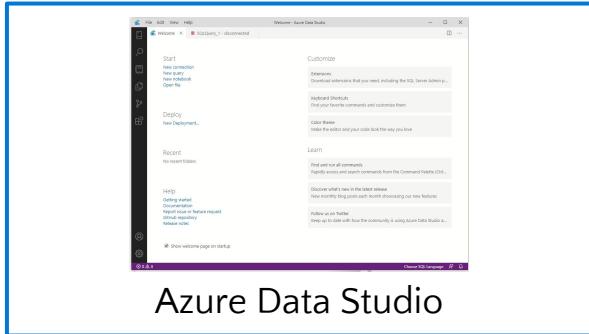
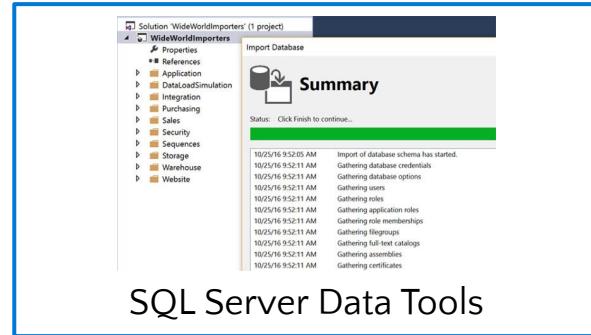
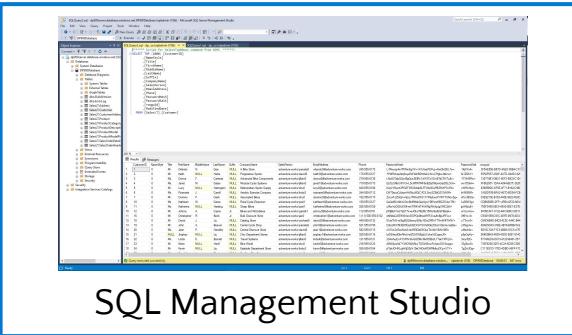
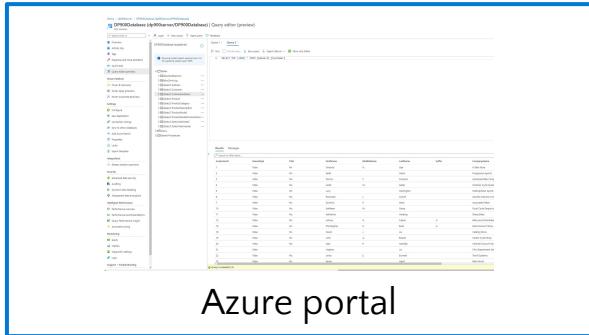
Statement	Description
CREATE	การสร้าง object ในฐานข้อมูลใหม่ เช่น table หรือ view
ALTER	แก้ไขโครงสร้าง (structure) ของ object เช่น ALTER ตัว table เพื่อเพิ่มคอลัมน์ใหม่
DROP	ลบ object ออกจากฐานข้อมูล เช่น ลบ table
RENAME	เปลี่ยนชื่อ object

ตัวอย่าง CREATE statement

```
CREATE TABLE Mytable
```

```
(Mycolumn1 int NOT NULL PRIMARY KEY, Mycolumn2 VARCHAR(50) NOT  
NULL , Mycolumn3 VARCHAR(10) NOT NULL
```

เครื่องมือที่ใช้ในการดึงข้อมูล (Query)



ตัวอย่างการดึงข้อมูลจาก Azure Database for PostgreSQL

ใช้ภาษา PSQL ในการดึงข้อมูล โดยใช้โปรแกรมต่าง ๆ ดังนี้ได้

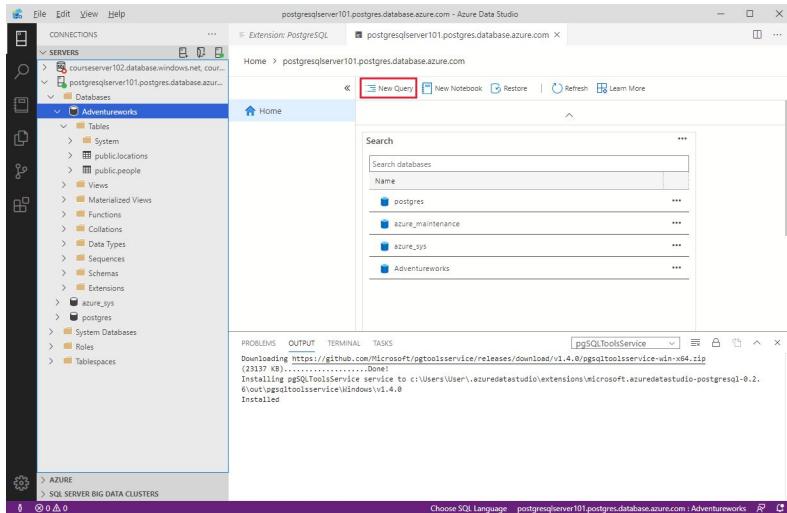
Azure Cloud Shell

```
psql --host=<server-name>.postgres.database.azure.com  
--username=<admin-user>@<server-name>  
--dbname=postgres
```

```
CREATE DATABASE "Adventureworks";
```

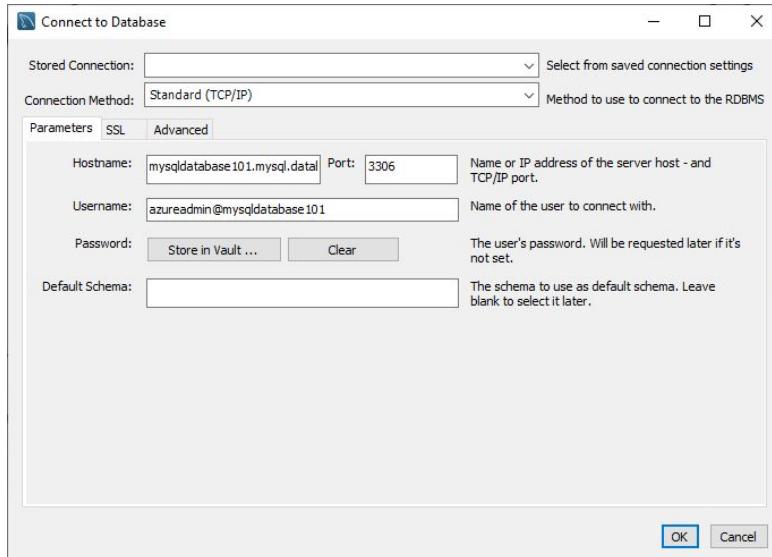
```
CREATE TABLE PEOPLE(NAME TEXT NOT NULL, AGE INT NOT  
NULL);  
INSERT INTO PEOPLE(NAME, AGE) VALUES ('Bob', 35);  
INSERT INTO PEOPLE(NAME, AGE) VALUES ('Sarah', 28);  
CREATE TABLE LOCATIONS(CITY TEXT NOT NULL, STATE  
TEXT NOT NULL);  
INSERT INTO LOCATIONS(CITY, STATE) VALUES ('New York',  
SELECT * FROM PEOPLE;  
SELECT * FROM LOCATIONS;
```

Azure Data Studio



ตัวอย่างการดึงข้อมูลจาก Azure Database for MySQL

สามารถใช้โปรแกรม MySQL Workbench เพื่อดึงข้อมูลได้



The interface shows the following:

- File menu: MySQL Workbench, File, Edit, View, Query, Database, Server, Tools, Scripting, Help
- Toolbar: Standard MySQL icons
- Navigator pane:
 - MANAGEMENT: Server Status, Client Connections, Users and Privileges, Status and System Variables, Data Export, Data Import/Restore
 - INSTANCE: Startup / Shutdown, Server Logs, Options File
 - PERFORMANCE: Dashboard, Performance Reports, Performance Schema Setup
- Query Editor pane:
 - Query 1: SELECT * FROM inventory;
 - Query 2: -- Delete
 - Query 3: DELETE FROM inventory WHERE id = 2;
 - Query 4: SELECT * FROM inventory;
- Result Grid pane:

ID	Name	Quantity
1	banana	200
3	apple	100
4	milk	200
- Output pane:

Action	Time	Action	Message	Duration / Fetch
7	18:44:19	INSERT INTO inventory (name, quantity) VALUES ('apple', 100)	1 row(s) affected	0.094 sec
8	18:44:19	SELECT * FROM inventory LIMIT 0, 5000	3 row(s) returned	0.094 sec
9	18:44:19	UPDATE inventory SET quantity = 200 WHERE id = 1	1 row(s) affected Rows matched: 1 Changed: 1 Warnings: 0	0.094 sec
10	18:44:19	SELECT * FROM inventory LIMIT 0, 5000	3 row(s) returned	0.093 sec / 0.000 sec
11	18:44:19	DELETE FROM inventory WHERE id = 2	1 row(s) affected	0.078 sec
12	18:44:19	SELECT * FROM inventory LIMIT 0, 5000	2 row(s) returned	0.094 sec / 0.000 sec

DEMO: การดึงข้อมูลด้วย Azure Data Studio

Azure Data Studio เป็นโปรแกรมฟรีจาก Microsoft ที่สามารถใช้เชื่อมต่อกับ Database เพื่อดึงข้อมูลได้ด้วยภาษา SQL

สามารถต่อ Database ได้หลากหลายตัว

The screenshot shows the Azure Data Studio interface. On the left is a sidebar with server groups: Development Server Group (selected), Test Server Group, and Production Server Group. The main area displays a query results grid titled 'QDS_Query_Perf.sql' with the following data:

query_id	last_execution_time	max_duration	plan_id
41996	2017-10-11 10:...	61931	525
42563	2017-10-12 03:...	221918	665
42561	2017-10-12 03:...	61144	663
42561	2017-10-12 03:...	51189	663

Below the grid is a chart titled 'QDS: Top 5 slowest queries' showing execution time over time. To the right are two charts: 'Space used by tables' and 'Disk Usage'.

ตารางเรียน Azure Data Fundamentals Training



Next Class

CLASS 1

วันจันทร์ 17 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Core Data Concepts
- Explore relational data in Azure



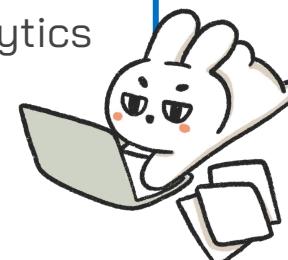
CLASS 2

วันพุธที่สุดวันเดือน 20 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Explore non-relational data in Azure
- Explore modern data warehouse analytics



CLASS 3

วันจันทร์ 24 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Azure Data Architecture with Microsoft Thailand



AZURE DATA FUNDAMENTALS TRAINING

Q&A TIME

THANK YOU FOR COMING TODAY



