

AZURE DATA FUNDAMENTALS TRAINING

CH4:

MODERN DATA WAREHOUSE

ANALYTICS



Agenda



เรียนรู้ components ในระบบ data warehouse สมัยใหม่



เรียนรู้การนำเข้าข้อมูล (Data ingestion)



เรียนรู้การจัดเก็บข้อมูล (data storage) และการประมวลผล (data processing)



เริ่มต้นการสร้าง Data Visualization ด้วย Power BI

Lesson 1: เรียนรู้ components ในระบบ data warehouse สมัยใหม่



Lesson 1 objectiv es



เรียนรู้คุณลักษณะของ Data warehouse



มาดูบริการต่าง ๆ ของ Azure สำหรับ data warehouse สมัยใหม่

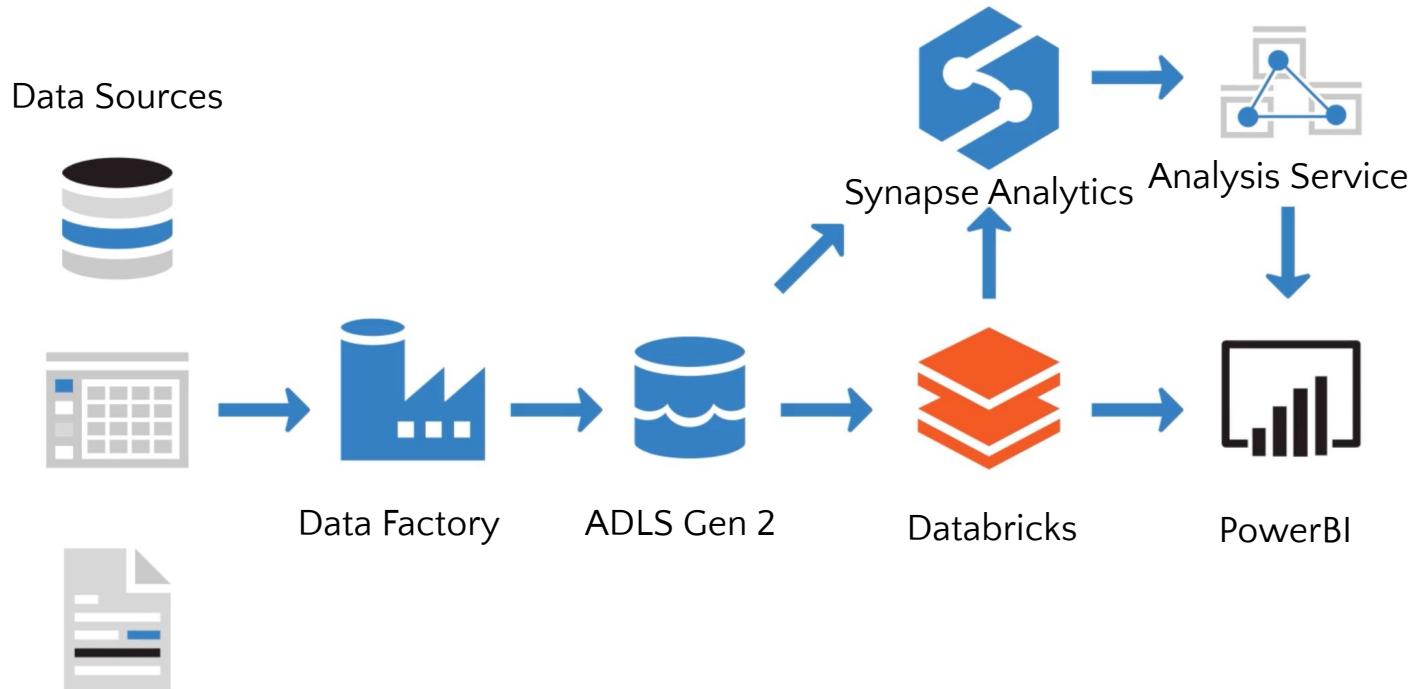


เรียนรู้การออกแบบ และ workload ของ data warehouse สมัยใหม่

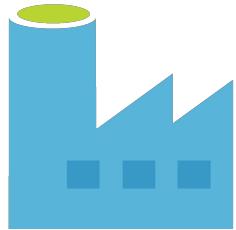


มาลองใช้บริการด้าน Data ของ Azure ด้วย Azure portal

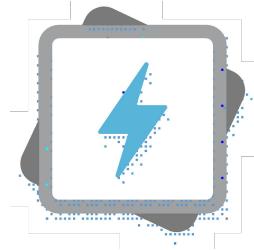
ຕົວຢ່າງ: Components ຂອງຮະບບ Data warehouse ສມໍຍໃຫ້



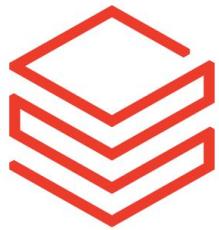
Azure services สำหรับ the modern data warehouse



Azure Data
Factory



Azure Data
Lake



Azure
Databricks



Azure HDInsight

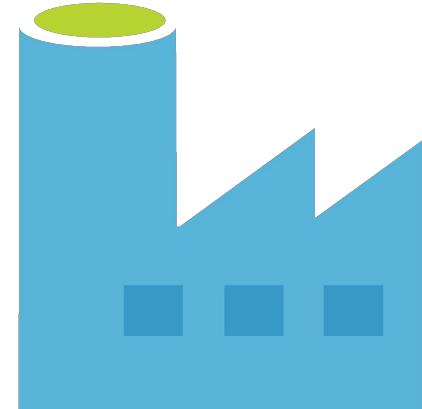


Azure Synapse
Analytics

Azure Data Factory คืออะไร?

Azure Data Factory เป็น data integration service

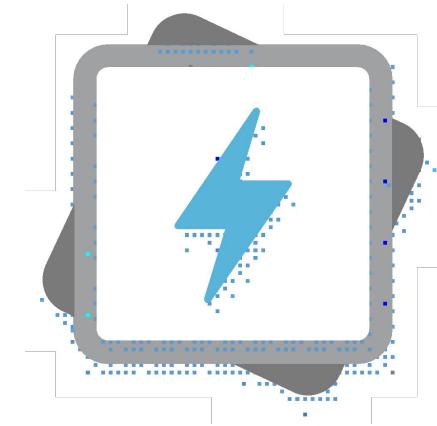
- ดึงข้อมูลมาจากแหล่งข้อมูลหลายแห่ง และส่งต่อไปยังจุดที่ต้องการได้
- กรองข้อมูลที่เป็น noise เพื่อเก็บข้อมูลส่วนที่สำคัญไว้
- สร้างเป็น Data Pipeline ที่ทำงานได้อย่างต่อเนื่อง เช่น ตามเวลาที่กำหนด



Azure Data Lake Storage คืออะไร?

Azure Data Lake Storage คือ ที่เก็บข้อมูล ใช้กับสำหรับระบบคลังข้อมูล เช่น Synapse Analytics ได้

- จัดระเบียบข้อมูลลงในไฟล์เดอร์เพื่อการเข้าถึงไฟล์ที่ดีขึ้น
- รองรับการอนุญาตแบบ POSIX และ RBAC
- ใช้งานได้กับ Hadoop Distributed File System



Azure Databricks คืออะไร?

Azure databricks คือแพลตฟอร์ม Apache Spark ที่ช่วยสำหรับการจัดเก็บข้อมูล Big data ทั้งส่วนของการประมวลผลข้อมูล และการรับข้อมูล Stream Data

- ลดความยุ่งยากในการจัด kuteriym และการทำงานร่วมกันของระบบบุคลากรที่ใช้ Apache Spark
- รองรับบริการแพลตฟอร์มข้อมูล Azure data และ Power BI
- ใช้ความสามารถด้านความปลอดภัยของ Azure



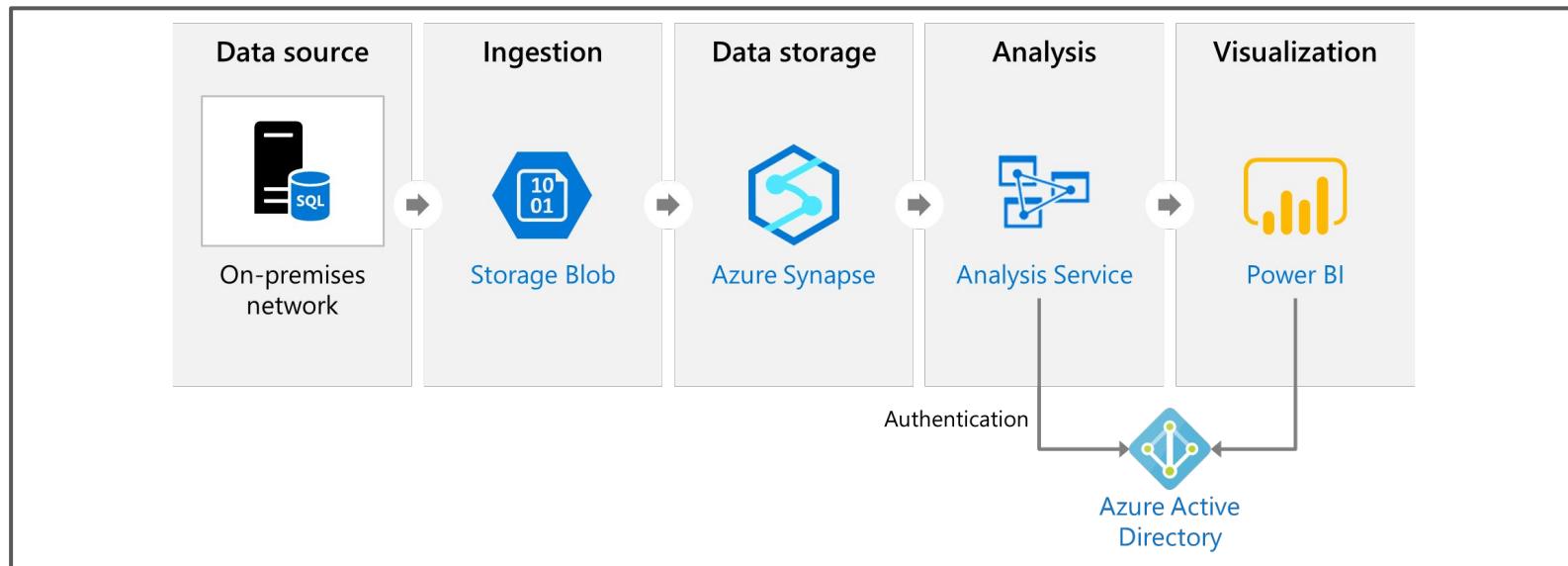
databricks

Azure Analysis Services คืออะไร?

Azure Analysis Services สร้าง tabular models เพื่อรองรับการดึงข้อมูลแบบ online analytical processing (OLAP) ที่รวดเร็ว

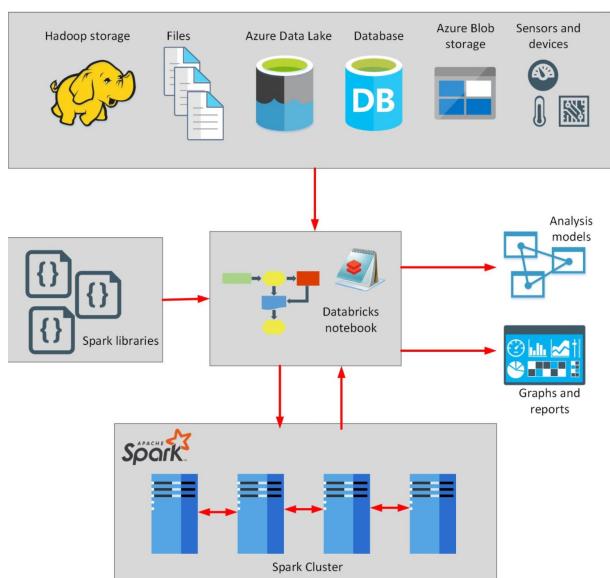
X

โดยสามารถรวมข้อมูลได้จากหลายแหล่ง ไม่ว่าจะเป็น Azure SQL Database, Azure Synapse Analytics, Azure Data Lake Store, Azure Cosmos DB และอื่นๆ



Azure HDInsight គីវូវ៉ាវ់

Azure HDInsight เป็นบริการ Hadoop บน Azure



Lesson 1: Knowledge check



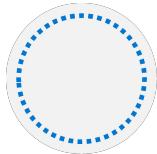
When should you use Azure Synapse Analytics?

- To perform very complex queries and aggregations
 - To create dashboards from tabular data
 - To enable large number of users to query analytics data
-



What is the purpose of data ingestion?

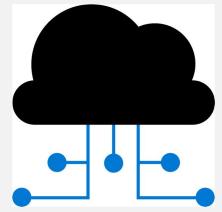
- To perform complex data transformations over data received from external sources
 - To capture data flowing into a data warehouse system as quickly as possible
 - To visualize the results of data analysis
-



What is the primary difference between a data lake and a data warehouse?

- A data lake contains structured information, but a data warehouse holds raw business data
- A data lake holds raw data, but a data warehouse holds structured information
- Data stored in a data lake is dynamic, but information stored in a data warehouse is static

Lesson 2: เรียนรู้การนำเข้าข้อมูลในระบบ Azure



Lesson 2 objectiv es



อธิบายการนำเข้าข้อมูลใน Azure



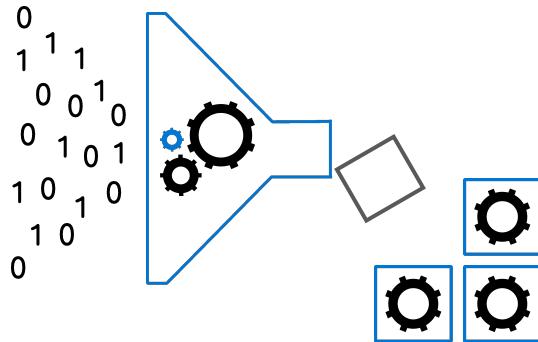
อธิบาย components ของ Azure Data Factory



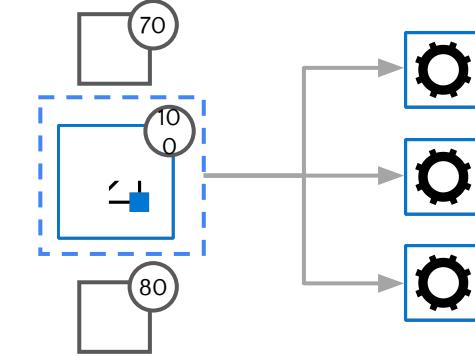
มาดูวิธีการใช้ Azure Data Factory ในการนำเข้าข้อมูลไปยัง data warehouse

วิธีต่าง ๆ ในการนำเข้าข้อมูลสู่ Azure

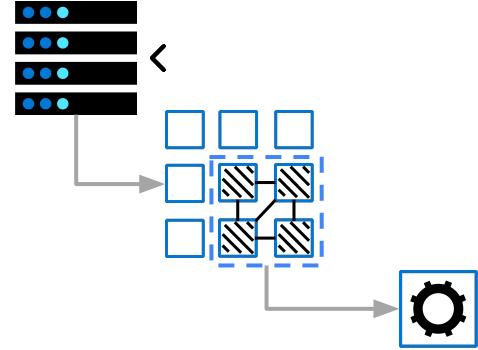
Azure Data Factory



PolyBase



SQL Server Integration Services (SSIS)

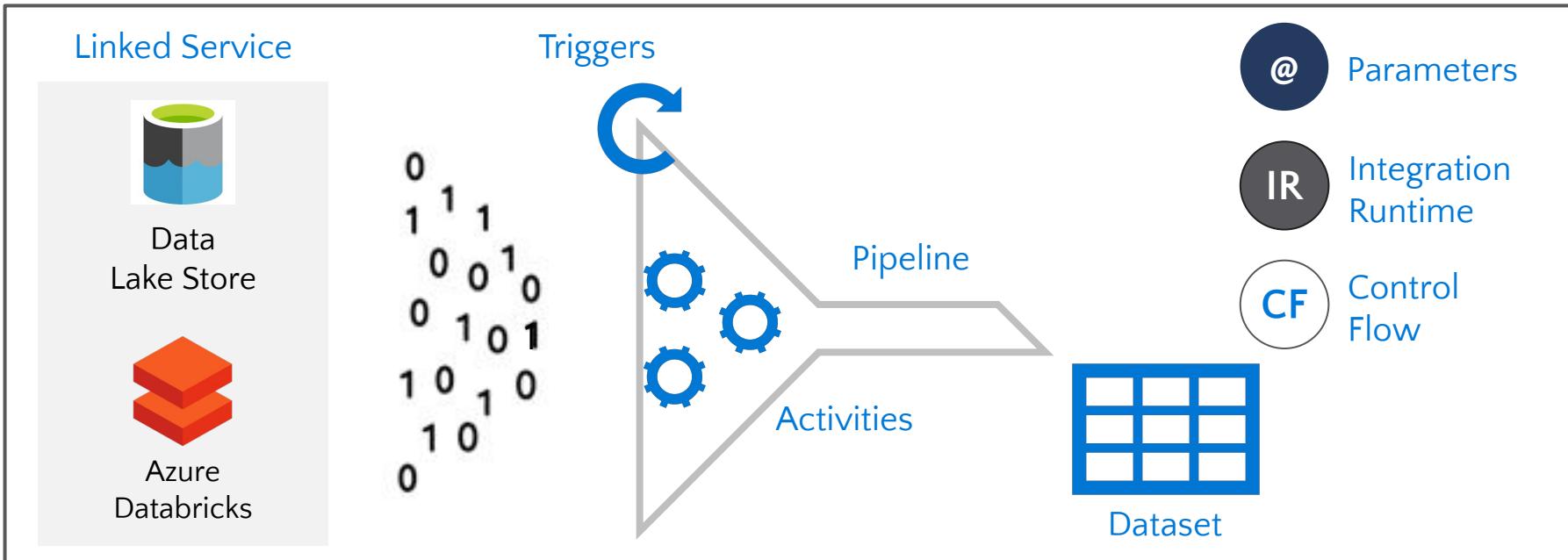


รองรับข้อมูลจากแหล่งข้อมูล
(data source) ที่หลากหลาย,
และนำไปเก็บได้หลากหลายที่

รองรับข้อมูลจากไฟล์ เข้าสู่ SQL
Server, SQL DW, Synapse
Analytics

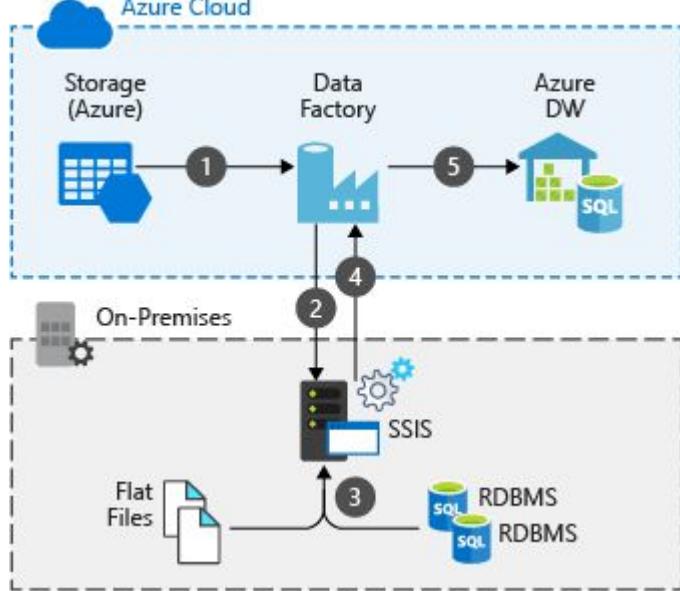
รองรับข้อมูลจากแหล่งข้อมูล
(data source) ที่หลากหลาย,
และนำไปเก็บได้หลากหลายที่

ອຮັບຍາຍ components ຂອງ Azure Data Factory



1222

ตัวอย่าง: ลองใช้งาน Azure Data Factory & Data Warehouse

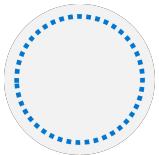


Azure Data Factory สามารถนำมาใช้ได้หลากหลาย
ตัวอย่างการใช้งาน

Hybrid ETL with Azure Data Factory

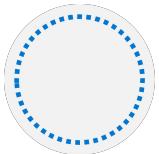
<https://docs.microsoft.com/en-us/azure/architecture/example-scenario/data/hybrid-etl-with-adf>

Lesson 2: Knowledge check



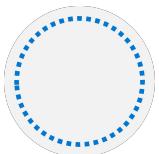
Which component of an Azure Data Factory can be triggered to run data ingestion tasks?

- CSV File
 - Pipeline
 - Linked service
-



When might you use PolyBase?

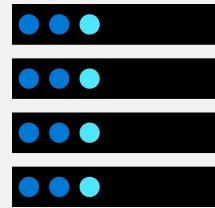
- To query data from external data sources from Azure SQL Database
 - To ingest streaming data using Azure Databricks
 - To orchestrate activities in Azure Data Factory
-



Which of these services can be used to ingest data into Azure Synapse Analytics?

- Azure Data Factory
- Power BI
- Azure Active Directory

Lesson 3: เรียนรู้การจัดเก็บข้อมูล และการใช้งานใน Azure



Lesson 3 objectiv es



อธิบายขั้นตอนการทำงานของ Data processing และตัวเลือกการ
ทำ analytics บน Azure



เรียนรู้ Azure Synapse Analytics

เรียนรู้ Azure Synapse Analytics



Synapse Pipelines



Synapse Link



Synapse Studio



Synapse SQL

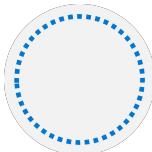


Azure Synapse
Analytics



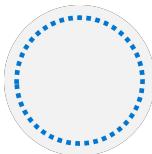
Synapse Spark

Lesson 3: Knowledge check



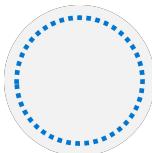
You have a large amount of data held in files in Azure Data Lake storage. You want to retrieve the data in these files and use it to populate tables held in Azure Synapse Analytics. Which processing option is most appropriate?

- Use Azure Synapse Link to connect to Azure Data Lake storage and download the data
 - Synapse SQL pool
 - Synapse Spark pool
-



Which of the components of Azure Synapse Analytics allows you to train AI models using AzureML?

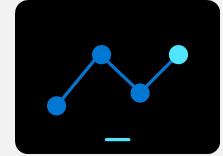
- Synapse Studio
 - Synapse Pipelines
 - Synapse Spark
-



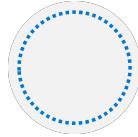
In Azure Databricks how do you change the language a cell uses?

- The first line in the cell is %language. For example, %scala
- Change the notebook language before writing the commands
- Wrap the command in the cell with ##language##

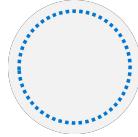
Lesson 4: เริ่มต้นการสร้าง data visualization ด้วย Power BI



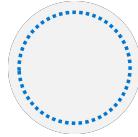
Lesson 4 objectiv es



เรียนรู้วิธีการทำงานร่วมกันของ Power BI services และแอพพลิเคชันอื่น ๆ

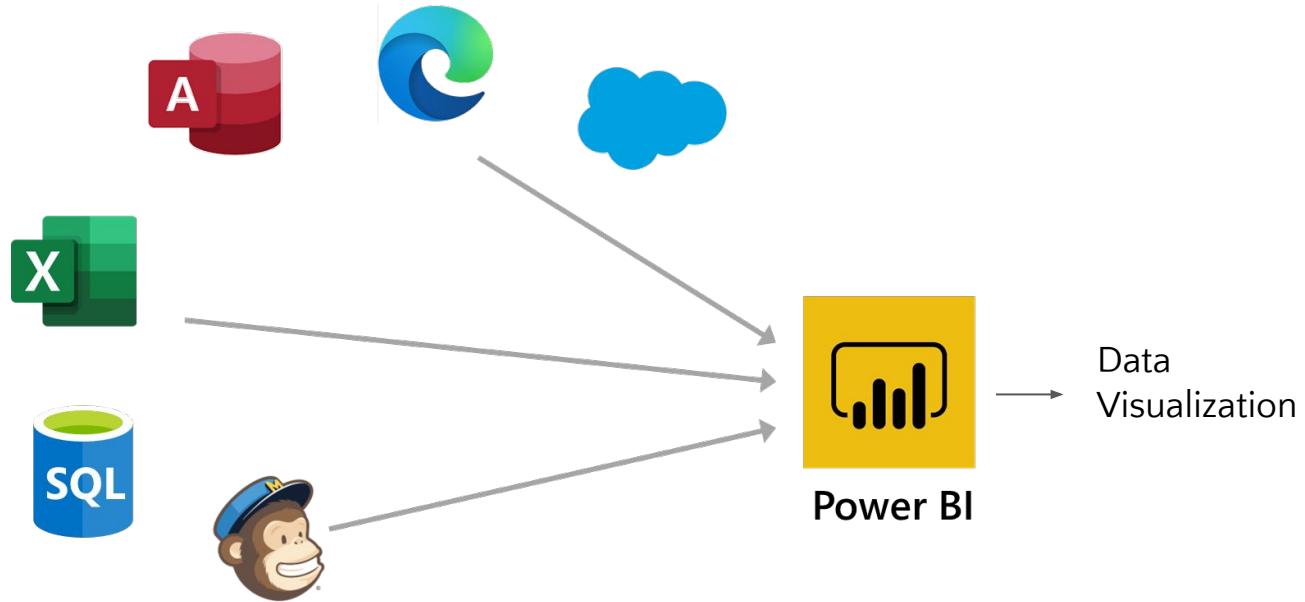


เรียนรู้การใช้งาน Power BI กับการสร้างผลประกอบการทางธุรกิจ



เรียนรู้วิธีการทำ Report ให้น่าสนใจและดึงดูด

Power BI คืออะไร?



ตัวอย่าง Data Visualization ที่สร้างจาก PowerBI



ตัวอย่าง Dashboard ที่สร้างจาก PowerBI:

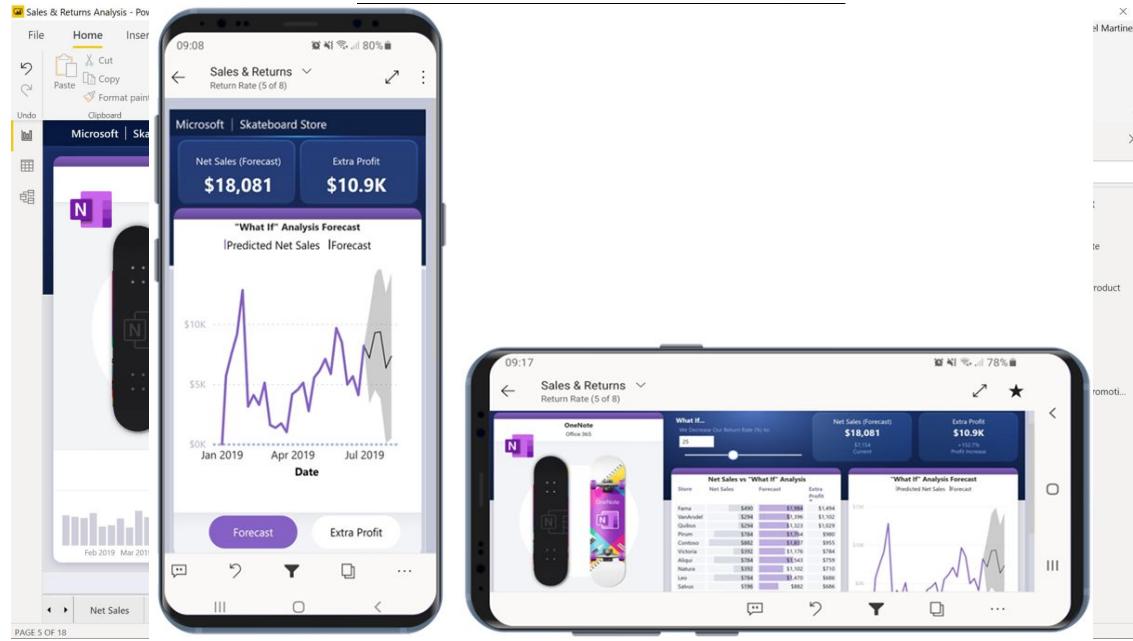
https://community.powerbi.com/t5/Galleries/ct-p/PBI_Comm_Galleries

เราจะใช้งาน Power BI ได้ผ่านช่องทางไหนบ้าง

1. Power BI Desktop
(Windows)

2. Power BI Service
(Website)

3. Power BI Mobile
(iOS, Android)



หน้าตาและส่วนประกอบของแอพ Power BI

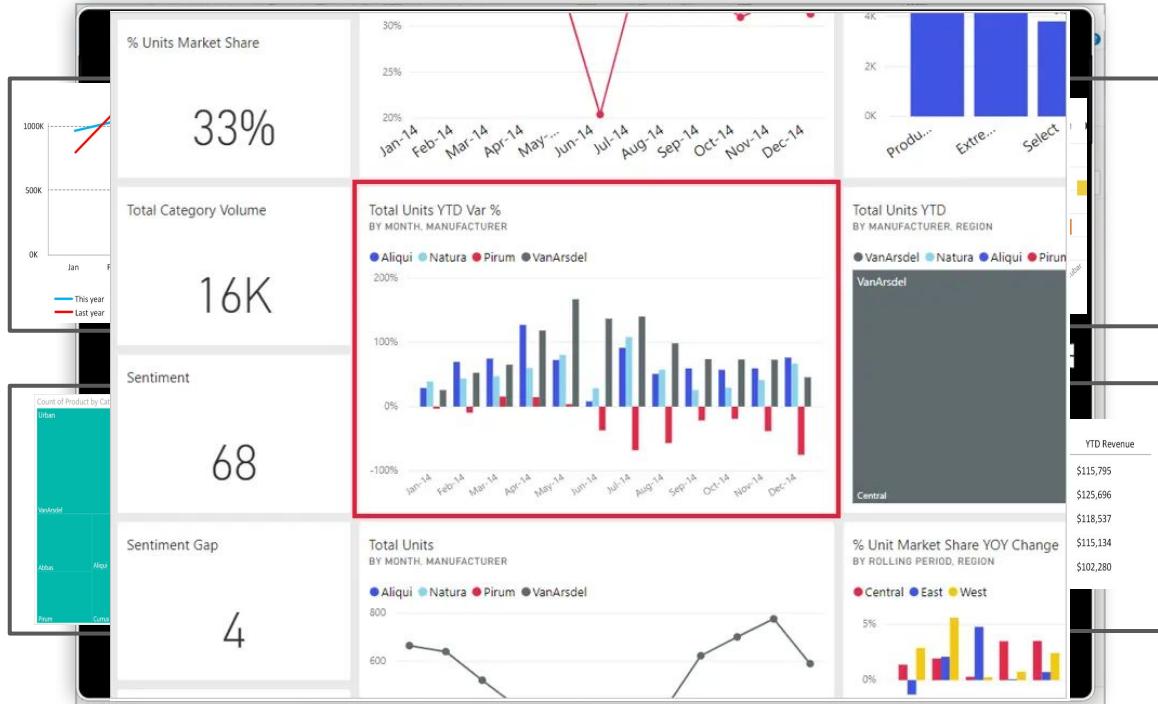
Visualizations

Datasets

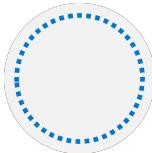
Reports

Dashboards

Tiles

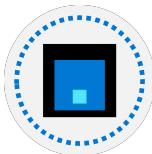


Lesson 4: Knowledge check



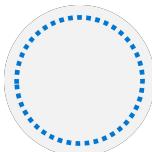
What is the common flow of activity in Power BI?

- Create a report in Power BI mobile, share it to the Power BI Desktop, view and interact in the Power BI service
 - Create a report in the Power BI service, share it to Power BI mobile, interact with it in Power BI Desktop
 - Bring data into Power BI Desktop and create a report, share it to the Power BI service, view and interact with reports and dashboards in the service and Power BI mobile
 - Bring data into Power BI mobile, create a report, then share it to Power BI Desktop
-



Which of the following are building blocks of Power BI?

- Tiles, dashboards, databases, mobile devices
 - Visualizations, datasets, reports, dashboards, tiles
 - Visual Studio, C#, and JSON files
-



A collection of ready-made visuals, pre-arranged in dashboards and reports is called what in Power BI?

- The canvas
- Scheduled refresh
- An app

ตารางเรียน Azure Data Fundamentals Training

CLASS 1

วันจันทร์ 17 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

- Core Data Concepts
Explore relational data in Azure



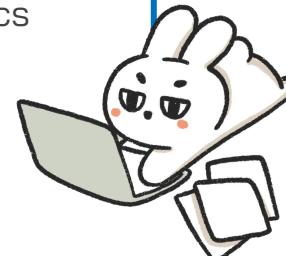
CLASS 2

วันพุธที่สุดวัน 20 Jan

เวลา 7:00 - 9:30 PM

เราจะเรียนอะไร?

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



↓
Next Class

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

