Power Query Workshop

By วิศวกรรีพอร์ต

9 มกราคม 2022

วิศวกรรีพอร์ต

ชื่อ	**ณัฐวุฒิ ฉัตรวิริยะเจริญ** (บิว)
ประสบการณ์ทำงาน	2017 – ปัจจุบัน: Managing Director: Wisawakorn Report
	2012 – 2017: Business Analysis Manager: National Starch & Chemical
	2010 – 2011: Business Intelligence Manager: Schneider Electric
	2006 – 2009: Marketing Analyst: Siam City Cement
	2002 – 2006: Regional Sales Representative: Siam City Cement
การศึกษา	ปริญญาโท: MBA Young Executive จุฬาลงกรณ์มหาวิทยาลัย
	ปริญญาตรี: วิศวกรรมศาสตร์ ภาควิชาวิศวกรรมสิ่งแวดล้อม จุฬาลงกรณ์มหาวิทยาลัย
ผลงาน	www.reportingengineer.com เผยแพร่บทความและแนวคิดสร้างสรรค์รีพอร์ต www.facebook.com/reportingengineer มีผู้ติดตามมากกว่า 100,000 คน

Concept ของ Power Query คือ ETL

Extract **T**ranform **L**oad

- ข้อมูลจากหลาย Platform
- แบบไฟล์เดียว
- แบบทั้งโฟลเดอร์
- ใส่เงื่อนไขในการเลือกไฟล์
- ดึงข้อมูลจากอินเทอร์เน็ต

แปลงให้ได้ฟอร์แมตที่ต้องการ

Pivot Table ต้องการฐานข้อมูลที่เป็นแนวตั้ง

	A	B	C	D	E	F	G
1	Product	Thailand	Indonesia	Malaysia	Philippines	Singapore	Vietnam
2	Product01	66	37	86	76	39	80
3	Product02	81	37	38	77	50	29
4	Product03	75	76	27	61	81	51
5	Product04	39	25	51	26	89	26
6	Product05	62	56	76	21	58	70
7	Product06	52	44			52	44
8	Product07	91	26			28	45
9	Product08	25	53	64	98	75	95
10	Product09	26	85	95	84	61	27
11	Product10	47	40	93	52	99	62

1

- ข้อมูลแนวขวาง (แบบที่ 1) ไม่สามารถใช้เป็นฐานข้อมูลของ Pivot Table ได้
- ต้องเปลี่ยนข้อมูลแนวขวาง ให้เป็นแนวตั้ง (แบบที่ 2, Tabular Format)
  - ใช้ VBA คำสั่งวนลูป
  - ใช้ **Power Query** (คำสั่ง Unpivot Columns)
    - Power Query เป็นฟีเจอร์มาตรฐานตั้งแต่ Excel 2016
    - Excel 2010/2013 ต้องลง Add-Ins เพิ่ม

	A	B	C
1	Product	Country	Volume
2	Product01	Indonesia	37
3	Product01	Malaysia	86
4	Product01	Philippines	76
5	Product01	Singapore	39
6	Product01	Thailand	66
7	Product01	Vietnam	80
8	Product02	Indonesia	37
9	Product02	Malaysia	38
10	Product02	Philippines	77
11	Product02	Singapore	50
12	Product02	Thailand	81
13	Product02	Vietnam	29
14	Product03	Indonesia	76
15	Prod	Malaysia	27
16	Prod	Philippines	61
17	Product03	Sing	81
18	Product03	Thaila	75
19	Product03	Vietnam	51
20	Product04	Indonesia	25

2

ควรเซ็ต Locale ใน Power Query เป็น English US

Excel 2010/ 2013:
Power Query/ Options

Microsoft 365, Excel 2016/ 2019:
Data/ Get Data/ Query Options

ค่าดีฟอลต์คือ English (United States)

ควรเซ็ต Locale เป็น English (United States) เพื่อป้องกันปัญหาปลีกย่อย

รูปแบบวันที่ของ English (United States) คือ **MM/DD/YYYY** ซึ่งอาจไม่ตรงกับรูปแบบที่คนไทยถนัด (แนะนำให้เซ็ตแบบนี้ไปก่อน)

เมื่อ Load ข้อมูลใส่ Excel แล้ว สามารถปรับรูปแบบให้เป็นแบบที่ต้องการ (dd/MMM/yy) ใน Excel ได้

แหล่งข้อมูลและบทความส่วนใหญ่มักเซ็ต Locale เป็น English (United States)

Data Type ใน Power Query Editor

1.2 Decimal Number	ตัวเลขที่มีทศนิยม (ทศนิยมสูงสุด 15 ตำแหน่ง)
$ Fixed decimal number	ตัวเลขที่มีทศนิยม (4 ตำแหน่ง) หรือตัวเลขทั้งหมดไม่เกิน 19 หลัก (รวมตัวเลขทศนิยม)
123 Whole Number	จำนวนเต็ม (สูงสุด 19 หลัก)
% Percentage	ข้อมูลที่เป็นเปอร์เซ็นต์
Date/Time	วันที่และเวลา (ถ้ามีแค่วันที่จะแสดงเวลาเที่ยงคืน)
Date	วันที่ (รูปแบบเหมือน Excel)
Time	เวลา (รูปแบบเหมือน Excel)
Date/Time/Timezone	วันที่/ เวลา/ Timezone
Duration	ช่วงเวลา (24 ชั่วโมง = 1.0, day.hour:minute:second เช่น 2.22:30:58)
ABC Text	ตัวหนังสือ
True/False	TRUE/ FALSE
Binary	ข้อมูลที่เป็นไฟล์ เช่น ไฟล์ csv, txt, xlsx
Using Locale...	เช่น ใช้ในกรณีเปลี่ยน locale จาก US เป็น UK

Date.From : แปลงให้เป็น Date

Date.From(value)

- แปลงวันที่ในรูปแบบ text (date_text) ให้เป็นวันที่
- value คือ ค่าใด ๆ ที่สื่อถึงวันที่ โดยอาจเป็นวันที่ในรูปแบบ text
  - 25-Jan-2019
  - Jan-25, 2019
  - Jan-25/ 2019
  - 2019-01-25
  - 20190125

<u>ตัวอย่าง</u>

- Date.From([Month]&"-1") แปลงให้เป็นวันที่ 1 ของเดือนนั้น ๆ ในปีปัจจุบัน
- **Date.From([Month]&"-1, 2022")** แปลงให้เป็นวันที่ 1 ของเดือนนั้น ๆ ในปี 2022

<u>หมายเหตุ</u>:
รูปแบบของวันที่ (วัน-เดือน-ปี, เดือน-วัน-ปี) ขึ้นกับรูปแบบวันที่ที่เซ็ตใน Windows
อาจใช้ฟังก์ชัน Date.FromText ได้เช่นกัน

Text.From : แปลงให้เป็น Text

Text.From(value)

- แปลง value ให้เป็น Text
- value สามารถเป็น Data Type แบบใดก็ได้
  - Number
  - Date
  - DateTime
  - Logical
  - Duration
  - Binary (หมายถึงไฟล์ เช่น txt, csv, xlsx)

<u>ตัวอย่าง</u>

- Text.From([Year]) แปลง Data Type ของคอลัมน์ Year ให้เป็นตัวเลข
- Date.From("1-"&[Month]&"-"&Text.From([Year])) แปลงให้เป็นวันที่ 1 ของเดือนและปีนั้น ๆ

<u>หมายเหตุ</u>: ถ้าต้องการแปลงเป็นตัวเลข ใช้ฟังก์ชัน Number.From

รูปแบบการแสดงผลขึ้นกับการเซ็ตใน Windows

การเซ็ตค่าใน Windows 10
(พิมพ์คำค้นหาใน search box)
Date & time settings
Region/ Change data formats

ถ้าเป็นไปได้ควรเซ็ต Region format เป็นภาษาอังกฤษเพื่อป้องกันปัญหาปลีกย่อย
(ทั้งนี้ ขึ้นกับแนวทางของบริษัทนั้น ๆ)

ถ้าเลือกเป็น **English (United States)** จะได้การแสดงผลเดือนเป็น Jan, Feb, Mar

(ถ้าเลือกเป็น Thai จะได้การแสดงผลเดือนเป็น ม.ค., ก.พ., มี.ค.)

สามารถเลือกฟอร์แมตของ Short date ให้เป็นแบบที่ถนัด เช่น dd-MMM-yy (วัน-เดือน-ปี)

<u>Note</u>: การเปลี่ยนรูปแบบภาษา และ วันที่ มีผลต่อโปรแกรมอื่นด้วย

ถ้าเปลี่ยนไฟล์ ต้องเปลี่ยน Data Source ทุกครั้ง

ถ้าเป็น Excel 2010/ 2013 คลิก
Power Query/ Data source settings

Data Source จะระบุ Drive ด้วย ดังนั้นถ้าใช้ไฟล์นี้หลายคนต้อง Map Drive ให้ตรงกัน

เลือกไฟล์

ถ้าเปลี่ยนชื่อไฟล์ก็ต้องเปลี่ยน Data Source เช่นกัน

สร้างคอลัมน์คำนวณ: Calculated Fields

Calculated Fields ใช้ได้เฉพาะการคำนวณง่ายๆ เช่น บวก ลบ หาร

กรณีการหาร (เช่น %Margin, Avg Price) มักใช้ Calculated Fields

เมื่อสร้าง Calculated Fields แล้วจะเกิดฟิลด์ใหม่ใน Pivot Cache

ถ้ามีเงื่อนไขใน Calculated Fields เงื่อนไขนั้นต้องใช้ฟิลด์ที่เป็นตัวเลขเท่านั้น ใช้กับฟิลด์ที่เป็นตัวหนังสือไม่ได้
(เช่น IF(Channel = "Export" ใช้ไม่ได้)
(ต้องใช้ Power Pivot)

ควรตั้งชื่อฟิลด์คำนวณให้สื่อความหมาย เข้าใจง่าย

ใช้กับฟิลด์ที่เป็นตัวเลข

ปรับรูปแบบตัวเลขตามต้องการ

1 **คลิกขวา**ที่ฟิลด์ตัวเลข

เลือกรูปแบบที่ต้องการ

ปรับความผิดพลาด (Error) ให้เป็นค่าว่าง

คลิกขวาเซลล์ใดๆใน Pivot Table

ตารางเดียวที่มีทุกคอลัมน์ ไม่เวิร์ก !!

	OrderDetailID	OrderID	OrderDate	Quantity	Channel	Country	CountryGroup	ProductName	Color	ProductSubcategory
1	1	43659	01-Jul-17	1	Store	United States	North America	Mountain-100 Black, 42	Black	Mountain Bikes
2	2	43659	01-Jul-17	3	Store	United States	North America	Mountain-100 Black, 44	Black	Mountain Bikes
3	3	43659	01-Jul-17	1	Store	United States	North America	Mountain-100 Black, 48	Black	Mountain Bikes
4	4	43659	01-Jul-17	1	Store	United States	North America	Mountain-100 Silver, 38	Silver	Mountain Bikes
5	5	43659	01-Jul-17	1	Store	United States	North America	Mountain-100 Silver, 42	Silver	Mountain Bikes
6	10	43659	01-Jul-17	6	Store	United States	North America	Mountain Bike Socks, M	White	Socks
7	6	43659	01-Jul-17	2	Store	United States	North America	Mountain-100 Silver, 44	Silver	Mountain Bikes
8	7	43659	01-Jul-17	1	Store	United States	North America	Mountain-100 Silver, 48	Silver	Mountain Bikes
9	12	43659	01-Jul-17	4	Store	United States	North America	Sport-100 Helmet, Blue	Blue	Helmets
10	8	43659	01-Jul-17	3	Store	United States	North America	Long-Sleeve Logo Jersey, M	Multi	Jerseys
11	11	43659	01-Jul-17	2	Store	United States	North America	AWC Logo Cap	Multi	Caps
12	9	43659	01-Jul-17	1	Store	United States	North America	Long-Sleeve Logo Jersey, XL	Multi	Jerseys
13	13	43660	01-Jul-17	1	Store	United States	North America	Road-650 Red, 44	Red	Road Bikes
14	14	43660	01-Jul-17	1	Store	United States	North America	Road-450 Red, 52	Red	Road Bikes
15	15	43661	01-Jul-17	1	Store	Canada	North America	HL Mountain Frame - Black, 48	Black	Mountain Frames
16	16	43661	01-Jul-17	1	Store	Canada	North America	HL Mountain Frame - Black, 42	Black	Mountain Frames
17	17	43661	01-Jul-17	2	Store	Canada	North America	HL Mountain Frame - Black, 38	Black	Mountain Frames
18	18	43661	01-Jul-17	4	Store	Canada	North America	AWC Logo Cap	Multi	Caps
19	19	43661	01-Jul-17	4	Store	Canada	North America	Long-Sleeve Logo Jersey, L	Multi	Jerseys

- ไฟล์ใหญ่โดยไม่จำเป็น
- คอลัมน์นึงมีข้อมูลซ้ำกันเยอะ
- คำนวณช้ามาก
- บางข้อมูลอาจยังไม่มีในตารางนี้ แต่จำเป็นต้องแสดงขึ้นมาและแสดงตัวเลขเป็น 0
- ไม่ได้ใช้ประโยชน์จาก Data Model
- ใช้ DAX ตระกูล Time Intelligence ไม่ได้

Data Model = ยก Access มาไว้ใน Excel

ไม่ต้องใช้ VLOOKUP ดึงข้อมูลมารวมกันที่ตารางเดียวเหมือน Pivot Table

สร้าง Pivot Table ได้จากทุกตารางและทุกคอลัมน์ใน Data Model

คำนวณเร็วกว่า และขนาดไฟล์เล็กกว่า Pivot Table แบบธรรมดา

Engine ที่ใช้เรียกว่า Columnar Database (Vertipaq, X Velocity)

Power Pivot คืออะไร?

Power Pivot

- ลบข้อจำกัดของ Pivot Table แบบธรรมดา
- ฐานข้อมูลคือ **Data Model** (ไม่ใช่ตารางเพียงตารางเดียว)
- สร้างสูตรคำนวณ และเพิ่มเงื่อนไขในการคำนวณได้ (Calculated Fields ใน Pivot Table ไม่สามารถใส่เงื่อนไขได้)
- ภาษาของสูตรคำนวณที่ใช้เรียกว่า **DAX** (Data Analysis eXpression)
- DAX ใช้ได้ใน Power Pivot, Power BI และ SSAS

=

Data Model

- ตารางหลายตารางที่มีความสัมพันธ์กัน
- รองรับข้อมูลมากกว่า 1 ล้านบรรทัด (ขีดจำกัดของ Excel)

+

Pivot Table

- ใช้งานง่าย
- Interface ที่ทุกคนคุ้นเคย

อาจเรียก Pivot Table ที่สร้างจาก Power Pivot ว่า **Data Model Pivot Table** ก็ได้

เทคนิคการดึงข้อมูลจาก Table ไปที่ Data Model

การดึงข้อมูลจากตารางใน Excel (ไฟล์ที่มี Data Model)
ต้องปรับฟอร์แมตให้เป็น Table ก่อน (Ctrl+T)
(ถ้าไม่ปรับเป็น Table จะดึงข้อมูลเข้า Data Model ไม่ได้)

จากนั้นคลิกเซลล์ใด ๆ ใน Table

Power Query: Load To Data Model

ไม่จำเป็นต้อง export ข้อมูลออกมาเป็น Table (option แรก)

การไม่ export เป็น Table มีข้อดีคือ จำนวนข้อมูลสามารถ **เกิน 1 ล้านบรรทัดได้**

เลือกให้ export ข้อมูลไปที่ Data Model เพื่อใช้ Power Pivot สรุปข้อมูล

ถ้าตารางนั้นไม่ได้อยู่ในไฟล์ที่ใช้สร้าง Data Model การใช้ Power Query เชื่อมต่อข้อมูลแล้ว Load To Data Model คือวิธีที่ง่ายที่สุด **แต่จะเปลี่ยนชื่อตารางภายหลังไม่ได้**

ถ้าตารางนั้นอยู่ในไฟล์ที่ใช้สร้าง Data Model ใช้วิธี Add To Data Model

ดูความสัมพันธ์ของตารางใน Data Model

คลิกริบเบิน Power Pivot/ Manage เมื่อเข้ามาใน Power Pivot คลิก Home/ **Diagram View**

หมายเหตุ: ถ้าไม่มีริบเบิน Power Pivot ให้ใช้ริบเบิน Data/ Manage Data Model (Office365)

สร้าง Pivot Table จาก Data Model

สร้างแบบวิธีปกติ

ทำได้ 2 วิธีคือ
1 สร้างแบบวิธีปกติ
2 สร้างจากหน้าต่าง Power Pivot

เลือกออปชัน Use this workbook's Data Model
ถ้าไฟล์นั้นมี Data Model แล้ว ออปชันนี้จะเป็น Default

สร้างจากหน้าต่าง Power Pivot

Power Pivot ดึงข้อมูลได้จากทุกตาราง

ถ้าสร้าง Pivot Table จาก Data Model สามารถสร้างจากตารางที่มีความสัมพันธ์กันมากกว่า 1 ตารางได้

Pivot Table ที่สร้างจาก Data Model ก็คือ **Power Pivot** นั่นเอง !

Power Pivot สามารถสร้าง **Measure** ซึ่งใช้สูตร DAX และยืดหยุ่นมากกว่า Calculated Fields

อย่า! ลากคอลัมน์ตัวเลขมาที่ Value

การดึงคอลัมน์ตัวเลขมาที่ Value ใน Pivot Table คือการสร้าง **Implicit Measure**

Implicit Measure จะแสดงใน Measure Grid ด้วย
(เปิด option: Advanced/Show Implicit Measure)

หลีกเลี่ยงการใช้ Implicit Measure

- ปรับฟอร์แมตตัวเลขไม่ได้
- เปลี่ยนชื่อไม่ได้
  - ได้แค่ในตารางนั้น
  - ใช้ชื่อที่เปลี่ยนกับตารางใหม่ไม่ได้
- คำนวณได้เฉพาะแบบง่าย ๆ (SUM, COUNT)
- ใส่เงื่อนไขไม่ได้

ให้ใช้ Explicit Measure แทน

สร้าง measure เอง

DAX คืออะไร

DAX = Data Analysis eXpressions

DAX คือภาษาที่ใช้ใน 3 โปรแกรม
1. Excel (ใช้ผ่าน Power Pivot)
2. Power BI Desktop
3. SSAS (SQL Server Analysis Service)

- DAX เกิดมาพร้อมกับ Power Pivot ในปี 2009
- DAX เป็นภาษาที่มีลักษณะการเขียนแบบฟังก์ชัน (**Functional Language**)
  - ลักษณะคล้ายฟังก์ชันใน Excel (เขียนซ้อนกันไปมา)
    - บางฟังก์ชันมีใน Excel
    - บางฟังก์ชันไม่มีใน Excel
  - ไม่ได้เขียนเป็นบรรทัดเหมือนภาษาคอมพิวเตอร์ทั่วไป
  - ไม่อ้างอิงข้อมูลในลักษณะของ cell (A50, B200) แต่อ้างอิงในลักษณะตาราง (Table) และคอลัมน์ (อ้างอิงบรรทัดไม่ได้)
- DAX ออกแบบมาเพื่อสรุปข้อมูลใน Data Model
- **DAX คือหัวใจสำคัญของ Power Pivot, Power BI**

DAX: DISTINCTCOUNT

DISTINCTCOUNT (Column)

- **นับค่าที่เป็น unique** หรือค่าที่ไม่ซ้ำ ถ้ามีค่าซ้ำกันนับเป็น 1
- ไม่ต้องสร้างคอลัมน์ 1/COUNTIFS() เหมือน Pivot Table แบบปกติ
- ใช้ใน Power Pivot (Power BI)
- #ProductSold =DISTINCTCOUNT(Sales[ProductCode])
  - นับว่ามีสินค้าถูกขาย (ในตาราง Sales) กี่ตัว
- #StoreSold = DISTINCTCOUNT(Sales[StoreCode])
  - นับว่ามีร้านค้ากี่ร้าน (ในตาราง Sales) ที่มีการซื้อขายเกิดขึ้น
- #Day = DISTINCTCOUNT(Sales[Date])
  - นับว่ามีวันที่เกิดการซื้อขาย (ในตาราง Sales) ทั้งหมดกี่วัน

<u>หมายเหตุ</u>: [ProductCode], [StoreCode], [Date] คือชื่อคอลัมน์ในตาราง Sales