

ใบงานที่ 7

เรื่อง ฝึกการทำ ETL โดยใช้ Power Query

วันที่ปฏิบัติงาน 13 มกราคม 2566

กำหนดส่งงาน 13 มกราคม 2566

- วัตถุประสงค์**
1. เพื่อให้นักศึกษาฝึกกระบวนการ ETL
 2. เพื่อให้นักศึกษาฝึกการ Transformation
 3. เพื่อให้นักศึกษาได้ฝึกใช้งาน Power Query

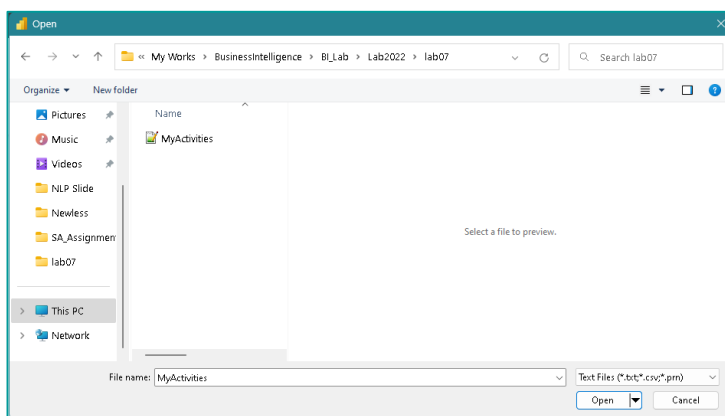
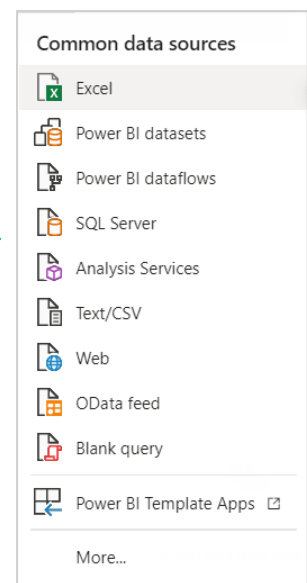
รายการ Software และไฟล์ข้อมูลที่ต้องใช้ในการฝึก

1. Microsoft Power BI Desktop
2. MyActivities.csv (download จาก Microsoft Teams วิชา Business Intelligence ใน Assignment ใบงานที่ 7)

เนื้อหา

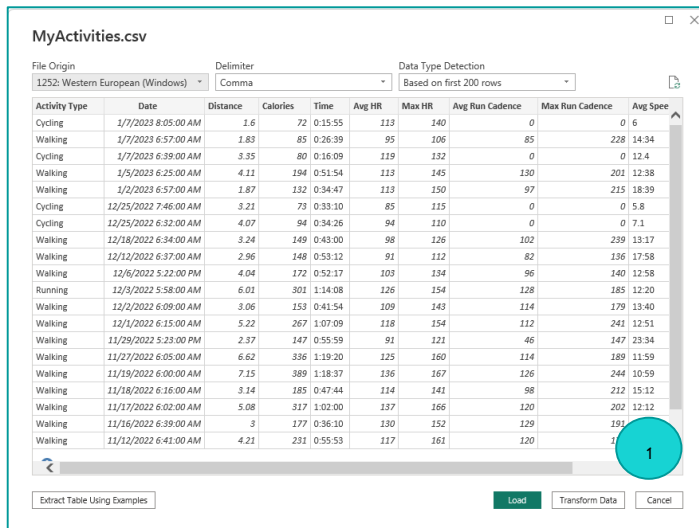
1. การนำข้อมูลเข้า Microsoft Power BI

- 1.1 ไปที่แถบเมนู data แล้วเลือก Get data
- 1.2 จะปรากฏ เมนูย่อย Common data sources ขึ้นมาเพื่อให้เลือกประเภทของข้อมูลที่จะนำเข้า
- 1.3 เลือกประเภทของข้อมูลที่ต้องการนำเข้า จะปรากฏหน้าจอ ดังภาพที่ 1



ภาพที่ 1 หน้าต่างให้เลือกไฟล์ที่จะนำเข้าสู่ Power BI

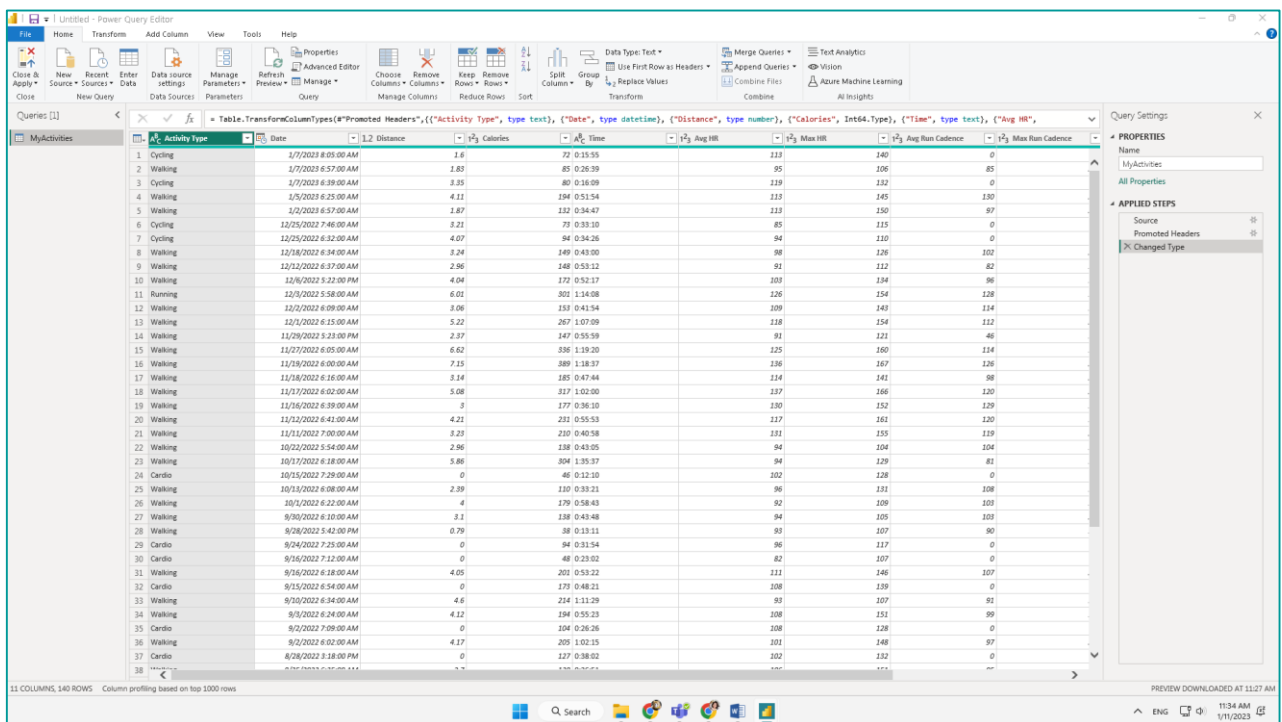
- 1.4 เลือกไฟล์ข้อมูลที่ต้องการนำเข้า power BI จากนั้น เลือก Open
- 1.5 จะปรากฏหน้าจอ ดังภาพที่ 2



ภาพที่ 2 หน้าต่างให้เลือกไฟล์ที่จะนำเข้าสู่ Power BI

1.6 กดปุ่ม Transform data

จากนั้นจะปรากฏหน้าจอ Power Query ดังภาพที่ 3



ภาพที่ 3 หน้าต่าง Power Query ที่ปรากฏขึ้น หลังเลือก Transform ข้อมูล

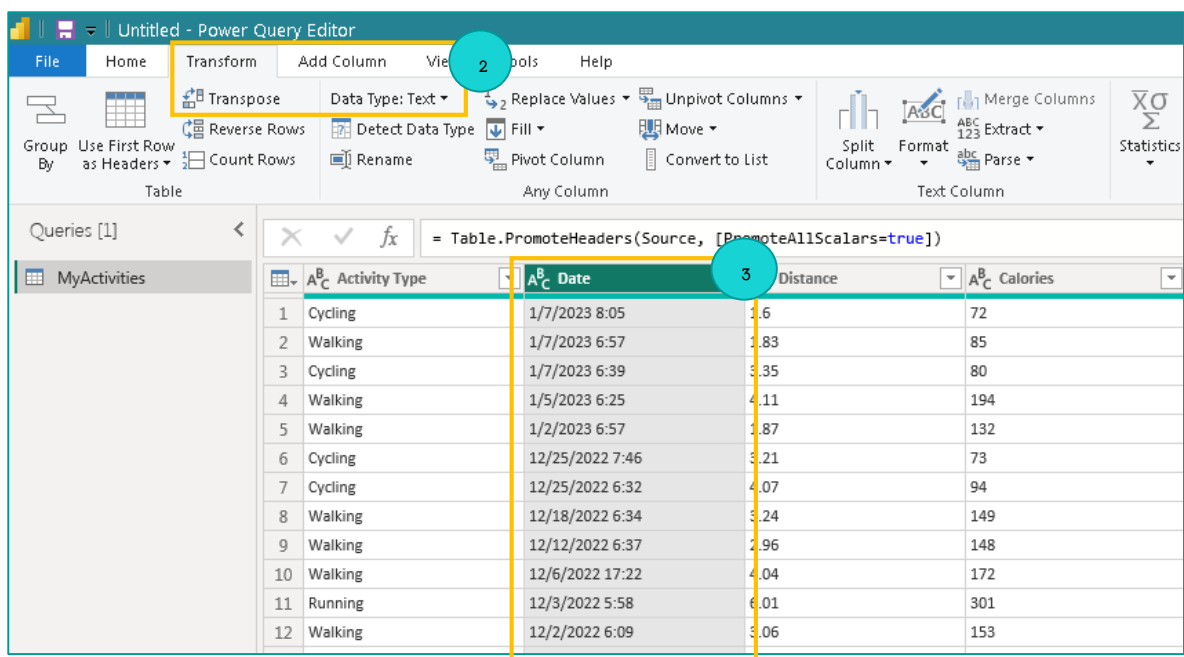
2. ชนิดข้อมูลใน Power BI

Data type	Icon	Description
Text	ABC	ข้อมูลที่เป็นตัวอักษร ข้อความ ตัวเลขและวันที่ที่นำเสนอในรูปแบบของ text format ความยาวสูงสุดของข้อความ คือ 268,435,456 Unicode characters หรือ 536,870,912 bytes.
True/False	✓✗	ค่า Boolean (True/False)
Decimal number	1.2	ข้อมูลตัวเลขจำนวนทศนิยม 64-bit (eight-byte) ซึ่งสามารถเก็บตัวเลขได้ 15 หลัก
Fixed decimal number	\$	ข้อมูลตัวเลขจำนวนทศนิยม ที่เป็น Currency type ที่จะ มี , ขึ้นระหว่างหลักร้อยกับพัน หลักล้านกับหลักแสน ค่าตัวเลขที่เก็บได้คือ 922,337,203,685,477.5807 (positive or negative)

Data type	Icon	Description
Whole number		ตัวเลขจำนวนเต็มขนาด 64-bit (eight-byte) เก็บตัวเลขได้สูงสุด 19 digits โดยมีค่าอยู่ในช่วง $-9,223,372,036,854,775,807$ ($-2^{63}+1$) ถึง $9,223,372,036,854,775,806$ ($2^{63}-2$)
Percentage		ตัวจำนวนทศนิยม ที่มี format ต้องมี % เนื่องจากเป็น percentage
Date/Time		ข้อมูลวันและเวลา โดยที่แสดงทั้งวันและเวลา โดยค่าของ Date/Time จะถูกเก็บในลักษณะ Decimal Number type ดังนั้นเราสามารถ convert ค่าได้ การเก็บเวลาจะถูกเก็บเป็นตัวเลขเศษส่วน 1/300 seconds (3.33 ms) และ Dates จะมีปี ค.ศ. ระหว่าง 1900 – 9999
Date		ข้อมูลวันและเวลา โดยที่แสดงแค่ทั้งวันที่ เดือน ปี
Time		ข้อมูลวันและเวลา โดยที่แสดงเพียงเวลา
Date/Time/Timezone		จะเก็บ UTC Date/Time พร้อมกับ time-zone
Duration		นำเสนอระยะเวลา โดยอยู่ในรูปแบบของจุดทศนิยม
Binary		ข้อมูล Binary data ถูกเก็บในรูปแบบ binary format.
Any		ข้อมูลประเภทอื่น ที่ไม่ทราบสามารถบอกชนิดได้อย่างชัดเจน

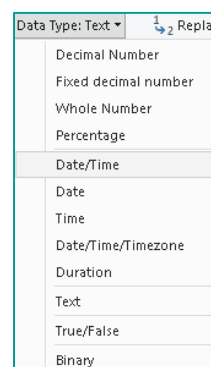
3. การเปลี่ยนชนิดข้อมูล

3.1 ไปที่เมนู Transform แล้วไปที่ Data Type

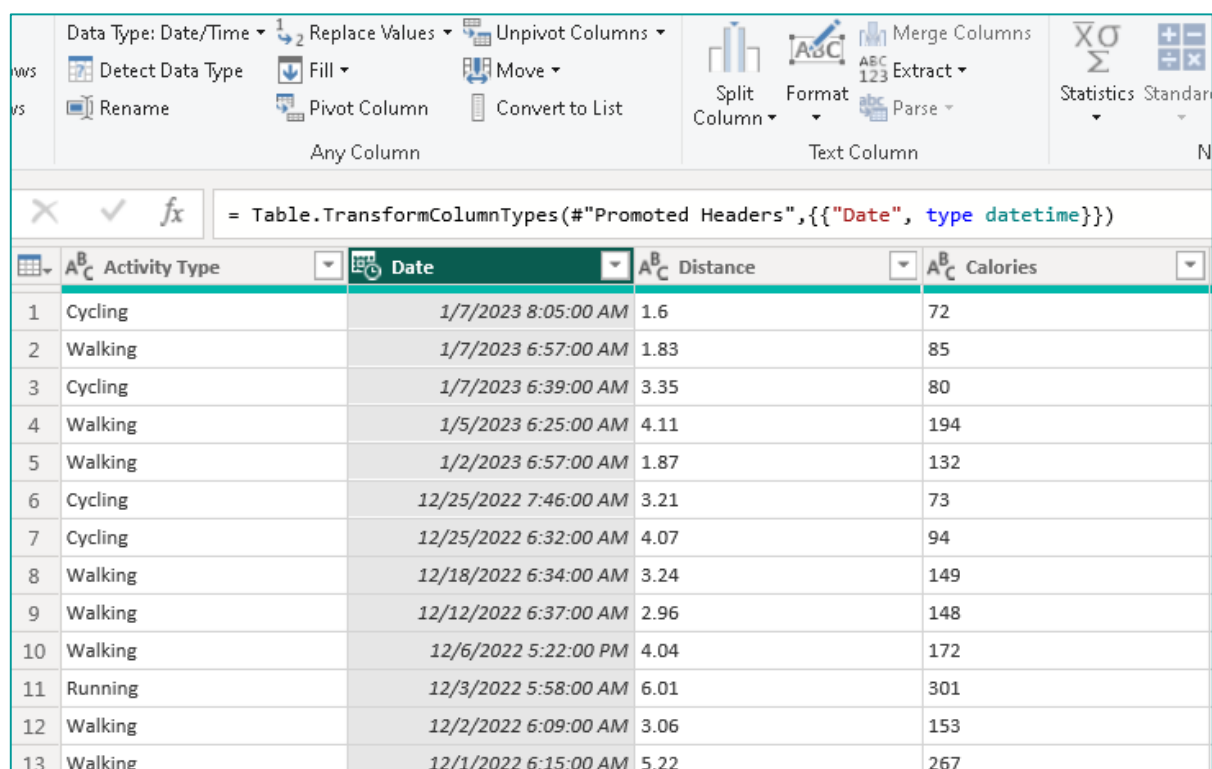


ภาพที่ 4 หน้าต่าง Power Query ที่ปรากฏขึ้น หลังเลือก Transform ข้อมูล

3.2 Click เลือก Data type จะปรากฏ menu list ขึ้นมา ดังภาพ



3.3 Click เลือก Date/Time ชนิดข้อมูลใน Column Date จะเปลี่ยนเป็นชนิด Date/Time

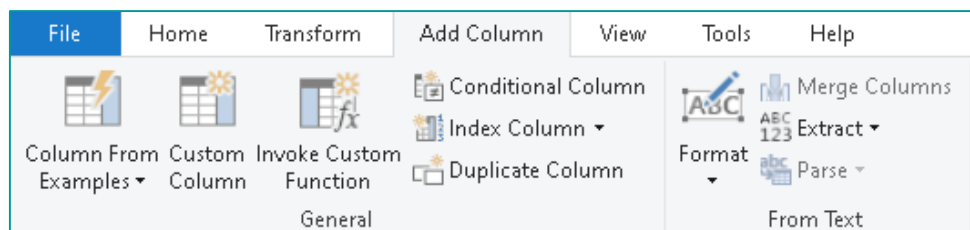


	Activity Type	Date	Distance	Calories
1	Cycling	1/7/2023 8:05:00 AM	1.6	72
2	Walking	1/7/2023 6:57:00 AM	1.83	85
3	Cycling	1/7/2023 6:39:00 AM	3.35	80
4	Walking	1/5/2023 6:25:00 AM	4.11	194
5	Walking	1/2/2023 6:57:00 AM	1.87	132
6	Cycling	12/25/2022 7:46:00 AM	3.21	73
7	Cycling	12/25/2022 6:32:00 AM	4.07	94
8	Walking	12/18/2022 6:34:00 AM	3.24	149
9	Walking	12/12/2022 6:37:00 AM	2.96	148
10	Walking	12/6/2022 5:22:00 PM	4.04	172
11	Running	12/3/2022 5:58:00 AM	6.01	301
12	Walking	12/2/2022 6:09:00 AM	3.06	153
13	Walking	12/1/2022 6:15:00 AM	5.22	267

ภาพที่ 5 แสดงชนิดข้อมูลใน Column Date เปลี่ยนจาก Text เป็น Date/Time

4 การเพิ่ม Column ใหม่ ในกรณีที่เราต้องการข้อมูลเพื่อการวิเคราะห์เพิ่มเติม

4.1 เลือกเมนู Add Column



ภาพที่ 6 แสดงเมนู Add Column และ Submenu

4.1.1 เลือกกรุป Column ที่ต้องการสร้าง

- Column From Examples: คือ การสร้าง column ใหม่ พร้อมข้อมูลโดยดูจากข้อมูลตัวอย่าง
- Custom Column: คือ การสร้าง column ใหม่ โดยที่เราสามารถกำหนดข้อมูลได้ตามต้องการจากสูตร
- Condition Column คือ การสร้าง column ใหม่ พร้อมข้อมูลตามเงื่อนไขที่เรากำหนด
- Index Column คือ การ Column ใหม่เพื่อเป็นดัชนี
- Duplicate Column คือ สำเนา Column ที่เราต้องการซึ่งเป็น Column ที่มีอยู่แล้วในตาราง มาเพิ่มอีก 1Column โดย Column ที่เป็นสำเนาจะมีชนิดข้อมูลและข้อมูลเหมือนกับต้นฉบับ

4.1.1.1 การสร้าง Column From Examples

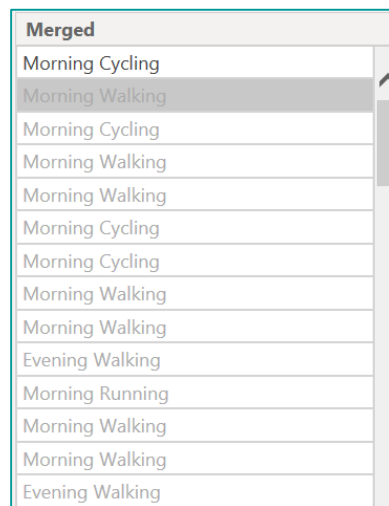
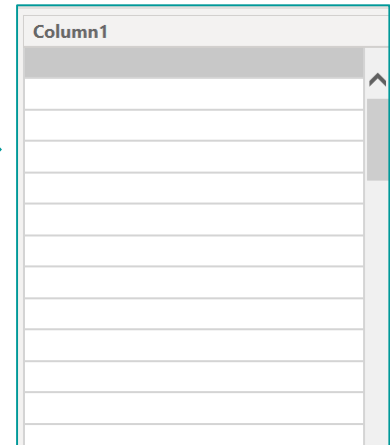
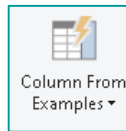
4.1.1.1.1 เลือกเมนู Add Column

4.1.1.1.2 เลือก Column From Examples

จะปรากฏ Column1 เปล่าๆ ขึ้นมา ดังภาพ

4.1.1.1.3 พิมพ์ตัวอย่างข้อมูลที่ยากให้สร้างขึ้นอัตโนมัติ

ลงที่แถวแรกของ Column1 ดังภาพ



หมายเหตุ:

ตัวอย่างข้อมูลที่ยากให้สร้าง

ขึ้นอัตโนมัติ ควรเป็นข้อมูลที่

เคยมีปรากฏอยู่แล้วใน Column อื่น

4.1.1.2 การสร้าง Custom Column

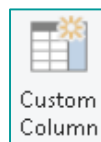
4.1.1.2.1 เลือกเมนู Add Column

4.1.1.2.2 เลือก Custom Column

4.1.1.2.3 ระบุชื่อ Column ใหม่ที่ต้องการสร้างที่ช่องหมายเลข 1

4.1.1.2.4 เขียน Function ลงใน Custom column formula ที่ช่อง

หมายเลข 2 เช่น ต้องการให้แสดงชื่อวัน



Custom Column

Add a column that is computed from the other columns.

New column name

Custom

Custom column formula

= Date.DayOfWeekName([Date])

Available columns

Activity Type

Date

Distance

Calories

Time

Avg HR

Max HR

<< Insert

Learn about Power Query formulas

✓ No syntax errors have been detected.

OK

Cancel

Function ที่ทำงานเกี่ยวกับวันที่และเวลา

- แสดงวันที่อย่างเดียวน Date.Day(ระบุข้อมูลวันเดือนปี(เช่น Date.Day("1/15/2023") จะได้วันที่ 15
- แสดงเดือนอย่างเดียวน Date.Month(ระบุข้อมูลวันเดือนปี(เช่น Date.Month("1/15/2023") จะได้วันที่ 1
- แสดงปี Date.Year(ระบุข้อมูลวันเดือนปี(เช่น Date.Year("1/15/2023") จะได้ปี ค.ศ. 2023
- แสดงชื่อวันในสัปดาห์ Date.DayOfWeekName(ระบุข้อมูลวันเดือนปี(เช่น Date.DayOfWeekName("1/15/2023") จะได้ Sunday
- แสดงชื่อเดือน Date.MonthName(ระบุข้อมูลวันเดือนปี(เช่น Date.MonthName("1/15/2023") จะได้ January

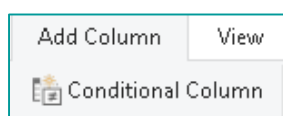
4.1.1.3 การสร้าง Conditional Column

4.1.1.3.1 เลือกเมนู Add Column

4.1.1.3.2 เลือก Conditional Column

จากนั้นจะปรากฏหน้าจอ
ดังภาพ

4.1.1.3.3 ตั้งชื่อ Column ที่หมายเลข 1

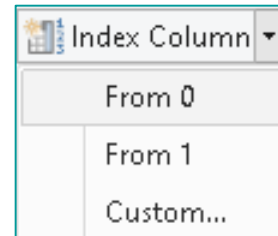
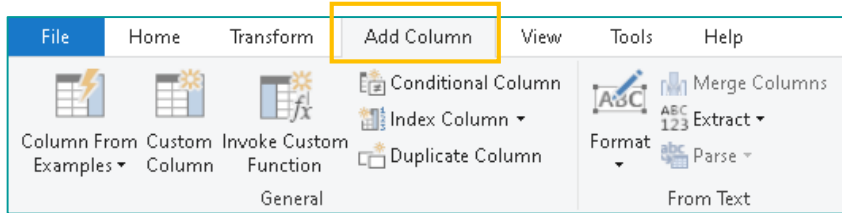


4.1.1.3.4 กำหนดเงื่อนไขที่ต้องการสร้างข้อมูล ที่หมายเลข 2 โดยระบุชื่อ Column ที่จะใช้ข้อมูลใน Column นั้นมาเป็นข้อมูลตั้งต้นที่นำได้ตรวจสอบว่าตรงกับเงื่อนไขไหม ระบุ Operation ระบุ Value (ค่าที่ใช้เป็นเงื่อนไขตรวจสอบ) และระบุ Output ที่จะสร้างถ้าเงื่อนไขเป็นจริง ดังภาพที่ 7

ภาพที่ 7 แสดงการระบุเงื่อนไขให้กับ Conditional Column ที่ต้องการสร้างขึ้น

4.1.1.4 การสร้าง Index Column

4.1.1.4.1 เลือกเมนู Add Column



4.1.1.4.2 กดเลือก Index Column จากนั้นจะปรากฏ dropdowns list ดังภาพ

4.1.1.4.3 เลือกตัวเลขเริ่มต้นของ index ที่จะสร้าง ว่าจะเริ่มจาก 0

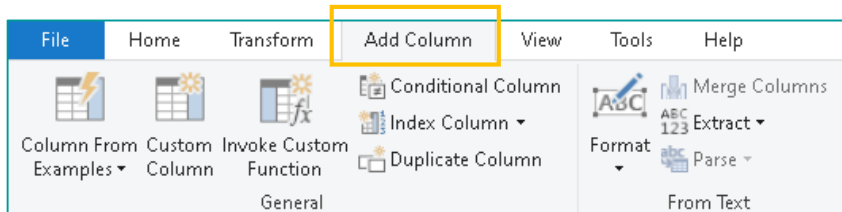
หรือ 1 หรือ เลขอื่นตามต้องการ

จากนั้นจะปรากฏ Column ใหม่ชื่อ Index ดังภาพ

1 ² 3 Index
1
2
3
4
5
6
7

4.1.1.5 การสร้าง Duplicate Column

4.1.1.5.1 เลือกเมนู Add Column



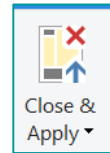
Date
1/7/2023 8:05:00 AM
1/7/2023 6:57:00 AM
1/7/2023 6:39:00 AM
1/5/2023 6:25:00 AM
1/2/2023 6:57:00 AM
12/25/2022 7:46:00 AM
12/25/2022 6:32:00 AM
12/18/2022 6:34:00 AM
12/12/2022 6:37:00 AM

4.1.1.5.2 กดเลือก Column ต้นฉบับที่ต้องการสำเนา

4.1.1.5.3 เลือก Duplicate Column

Date - Copy
1/7/2023 8:05:00 AM
1/7/2023 6:57:00 AM
1/7/2023 6:39:00 AM
1/5/2023 6:25:00 AM
1/2/2023 6:57:00 AM
12/25/2022 7:46:00 AM
12/25/2022 6:32:00 AM
12/18/2022 6:34:00 AM
12/12/2022 6:37:00 AM

5 บันทึกข้อมูลที่ Transform ให้เรียบร้อย โดยการกด Close & Apply



คำสั่ง

1. Download file MyActivities.csv จาก Microsoft Teams วิชา Business Intelligence ใน Assignment ใบบงานที่ 7
2. Get data จากไฟล์ MyActivities.csv เข้า Power BI เพื่อเป็นตาราง Fact ชื่อว่า MyActivities
3. Get data จากไฟล์ MyActivities.csv เข้า Power BI เพื่อเป็นตาราง Dimension ชื่อว่า ActivityDateTime โดยตารางนี้จะเน้นเรื่องเวลา โดยการ Transform ข้อมูลที่นำเข้ามาจาก MyActivities.csv ให้เป็นไปตามข้อกำหนดในตารางด้านล่าง ด้วย Power Query

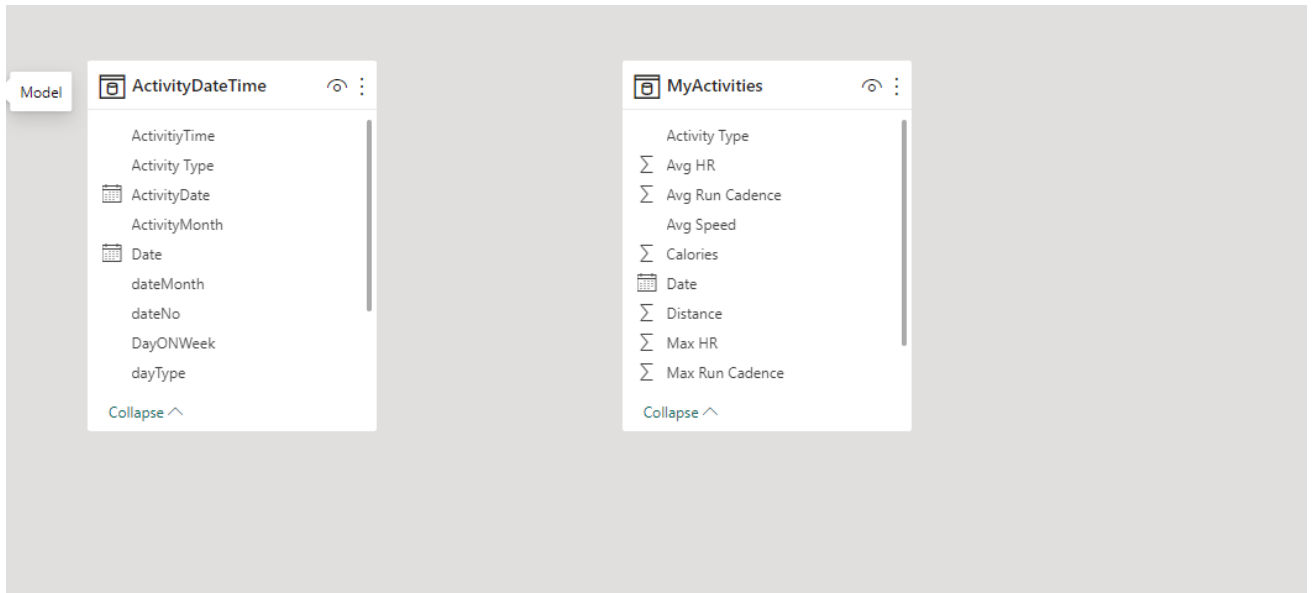
Column	Data type	ข้อมูลที่แสดง	
Date	Date/Time	วันที่ เดือน ปีและเวลา	
ActivityDate	Date	วัน เดือน ปี เช่น 1/14/2023	Duplicate column & Change data type
ActivityTime	Time	เวลา เช่น 5:30 AM	Duplicate column & Change data type
dateNo	Text	วันที่ เช่น 14	Custom column + เขียนสูตร
monthNo	Text	เดือนที่ เช่น 1	Custom column + เขียนสูตร
Year	Text	ปี ค.ศ. เช่น 2023	Custom column + เขียนสูตร
DayOfWeek	Text	ชื่อวันใน 1 สัปดาห์ เช่น Sunday	Custom column + เขียนสูตร
activityMonth	Text	ชื่อเดือน เช่น January	Custom column + เขียนสูตร
TimeOfDay	Text	ช่วงเวลา เช่น Morning	Conditional column โดยมีเงื่อนไขดังนี้ Morning: ก่อน 10:59 AM Afternoon: 11:00 AM – 4:59 PM Evening: เวลาหลัง 9:00 PM
Season	Text	ฤดู เช่น Summer	Conditional column โดยมีเงื่อนไขดังนี้ Morning: ก่อน 10:59 AM Afternoon: 11:00 AM – 4:59 PM Evening: เวลาหลัง 9:00 PM
dayType	Text	ประเภทของวัน เช่น Weekend Workday	Conditional column โดยมีเงื่อนไขดังนี้ Weekend: Saturday, Sunday Workday: Monday – Friday
Merged	Text	ชื่อการออกกำลังกาย และช่วงเวลาที่ออกกำลังกาย เช่น Morning Walking	Column From Examples

จากนั้นลบ Column อื่นๆ ที่นอกเหนือจากในตารางทั้ง

6 บันทึกข้อมูลที่ Transform ให้เรียบร้อย โดยการกด Close & Apply



- ไปที่แถบเปลี่ยนมุมมอง และ Click ที่ปุ่ม เพื่อสลับมาแสดงหน้า Relations จากนั้นเชื่อมความสัมพันธ์ระหว่างตาราง MyActivities กับตาราง ActivityDateTime
- Capture หน้าจอ Relation ที่ปรากฏลงในพื้นที่ว่างด้านล่าง



- ไปที่ Power BI แล้วมาที่แถบเปลี่ยนมุมมอง และ Click ที่ปุ่ม เพื่อสลับมาแสดงหน้า Data จากนั้น ไปที่ Field Tab ที่อยู่ทางด้านขวาของจอแล้วลองเลือกตารางที่จะตาราง จากนั้น Capture หน้าจอที่เปลี่ยนไปในแต่ละตารางลงในพื้นที่ว่างด้านล่าง

Activity Type	Date	ActivityDate	ActivityTime	dateNo	dateMonth	Year	DayONWeek	ActivityMonth	TimeOfDay	Season	dayType	Merged
Walking	12/18/2022 6:34:00 AM	Sunday, December 18, 2022	6:34:00 AM	18	12	2022	Sunday	December	Morning	Winter	Weekend	Morning Walking
Walking	12/12/2022 6:37:00 AM	Monday, December 12, 2022	6:37:00 AM	12	12	2022	Monday	December	Morning	Winter	Weekday	Morning Walking
Walking	12/2/2022 6:09:00 AM	Friday, December 2, 2022	6:09:00 AM	2	12	2022	Friday	December	Morning	Winter	Weekday	Morning Walking
Walking	12/2/2022 6:15:00 AM	Thursday, December 2, 2022	6:15:00 AM	1	12	2022	Thursday	December	Morning	Winter	Weekday	Morning Walking
Walking	11/27/2022 6:05:00 AM	Sunday, November 27, 2022	6:05:00 AM	27	11	2022	Sunday	November	Morning	Winter	Weekend	Morning Walking
Walking	11/19/2022 6:00:00 AM	Saturday, November 19, 2022	6:00:00 AM	19	11	2022	Saturday	November	Morning	Winter	Weekend	Morning Walking
Walking	11/18/2022 6:18:00 AM	Friday, November 18, 2022	6:18:00 AM	18	11	2022	Friday	November	Morning	Winter	Weekday	Morning Walking
Walking	11/17/2022 6:02:00 AM	Thursday, November 17, 2022	6:02:00 AM	17	11	2022	Thursday	November	Morning	Winter	Weekday	Morning Walking
Walking	11/16/2022 6:09:00 AM	Wednesday, November 16, 2022	6:09:00 AM	16	11	2022	Wednesday	November	Morning	Winter	Weekday	Morning Walking
Walking	11/12/2022 6:41:00 AM	Saturday, November 12, 2022	6:41:00 AM	12	11	2022	Saturday	November	Morning	Winter	Weekend	Morning Walking
Walking	11/12/2022 7:00:00 AM	Friday, November 11, 2022	7:00:00 AM	11	11	2022	Friday	November	Morning	Winter	Weekday	Morning Walking
Walking	10/22/2022 5:54:00 AM	Saturday, October 22, 2022	5:54:00 AM	22	10	2022	Saturday	October	Morning	Rainy	Weekend	Morning Walking
Walking	10/17/2022 6:18:00 AM	Monday, October 17, 2022	6:18:00 AM	17	10	2022	Monday	October	Morning	Rainy	Weekday	Morning Walking
Walking	10/12/2022 6:08:00 AM	Thursday, October 13, 2022	6:08:00 AM	13	10	2022	Thursday	October	Morning	Rainy	Weekday	Morning Walking
Walking	10/1/2022 6:22:00 AM	Saturday, October 1, 2022	6:22:00 AM	1	10	2022	Saturday	October	Morning	Rainy	Weekend	Morning Walking
Walking	9/30/2022 6:10:00 AM	Friday, September 30, 2022	6:10:00 AM	30	9	2022	Friday	September	Morning	Rainy	Weekday	Morning Walking
Walking	9/16/2022 6:18:00 AM	Friday, September 16, 2022	6:18:00 AM	16	9	2022	Friday	September	Morning	Rainy	Weekday	Morning Walking
Walking	9/10/2022 6:54:00 AM	Saturday, September 10, 2022	6:54:00 AM	10	9	2022	Saturday	September	Morning	Rainy	Weekend	Morning Walking
Walking	9/9/2022 6:24:00 AM	Saturday, September 9, 2022	6:24:00 AM	9	9	2022	Saturday	September	Morning	Rainy	Weekend	Morning Walking
Walking	9/2/2022 6:02:00 AM	Friday, September 2, 2022	6:02:00 AM	2	9	2022	Friday	September	Morning	Rainy	Weekday	Morning Walking
Walking	8/25/2022 6:35:00 AM	Thursday, August 25, 2022	6:35:00 AM	25	8	2022	Thursday	August	Morning	Rainy	Weekday	Morning Walking
Walking	8/20/2022 6:10:00 AM	Saturday, August 20, 2022	6:10:00 AM	20	8	2022	Saturday	August	Morning	Rainy	Weekend	Morning Walking
Walking	8/12/2022 7:01:00 AM	Friday, August 12, 2022	7:01:00 AM	12	8	2022	Friday	August	Morning	Rainy	Weekday	Morning Walking
Walking	8/11/2022 6:11:00 AM	Thursday, August 11, 2022	6:11:00 AM	11	8	2022	Thursday	August	Morning	Rainy	Weekday	Morning Walking
Walking	8/6/2022 6:20:00 AM	Saturday, August 6, 2022	6:20:00 AM	6	8	2022	Saturday	August	Morning	Rainy	Weekend	Morning Walking
Walking	8/5/2022 6:07:00 AM	Friday, August 5, 2022	6:07:00 AM	5	8	2022	Friday	August	Morning	Rainy	Weekday	Morning Walking
Walking	7/26/2022 6:20:00 AM	Thursday, July 26, 2022	6:20:00 AM	18	7	2022	Thursday	July	Morning	Rainy	Weekday	Morning Walking
Walking	7/23/2022 6:44:00 AM	Saturday, July 23, 2022	6:44:00 AM	23	7	2022	Saturday	July	Morning	Rainy	Weekend	Morning Walking
Walking	7/22/2022 6:35:00 AM	Friday, July 22, 2022	6:35:00 AM	22	7	2022	Friday	July	Morning	Rainy	Weekday	Morning Walking
Walking	7/17/2022 6:30:00 AM	Sunday, July 17, 2022	6:30:00 AM	17	7	2022	Sunday	July	Morning	Rainy	Weekend	Morning Walking
Walking	7/15/2022 6:04:00 AM	Friday, July 15, 2022	6:04:00 AM	15	7	2022	Friday	July	Morning	Rainy	Weekday	Morning Walking
Walking	7/14/2022 6:06:00 AM	Thursday, July 14, 2022	6:06:00 AM	14	7	2022	Thursday	July	Morning	Rainy	Weekday	Morning Walking
Walking	7/8/2022 5:58:00 AM	Friday, July 8, 2022	5:58:00 AM	8	7	2022	Friday	July	Morning	Rainy	Weekday	Morning Walking
Walking	7/2/2022 5:54:00 AM	Saturday, July 2, 2022	5:54:00 AM	2	7	2022	Saturday	July	Morning	Rainy	Weekend	Morning Walking
Walking	6/25/2022 5:59:00 AM	Saturday, June 25, 2022	5:59:00 AM	25	6	2022	Saturday	June	Morning	Summer	Weekend	Morning Walking
Walking	6/18/2022 6:28:00 AM	Thursday, June 18, 2022	6:28:00 AM	18	6	2022	Thursday	June	Morning	Summer	Weekday	Morning Walking
Walking	6/18/2022 6:20:00 AM	Saturday, June 18, 2022	6:20:00 AM	18	6	2022	Saturday	June	Morning	Summer	Weekend	Morning Walking
Walking	6/14/2022 6:38:00 AM	Tuesday, June 14, 2022	6:38:00 AM	14	6	2022	Tuesday	June	Morning	Summer	Weekday	Morning Walking

- Save และส่งไฟล์ Lab07_xxx.pbix) xxx คือ รหัส ตัวทำของนัก 3ศึกษา(เข้ามาใน Microsoft Teams พร้อมไฟล์ใบงานนี้