Introduction to Excel and Power BI

BOKSS IMHS Internal Training

30-May-2024 @ K83 & Zoom

Self-introduction

- Power BI Data Analyst Associate
- Google Analytics Certification
- Data Scientist and Data Analyst with R in DataCamp
- Re:Fresh Dashboard (password: refresh)
- Research Profile
- Front-end Website Example

Chapter 1

Basic Concept of Data Science

Data is everywhere

Type of Data

- Structured Data vs unstructured data
- Quantitative vs. qualitative data

Structured vs unstructured data

Structured data

- Organized
- Tabular format
- Predefined structure
- Text and numbers

Unstructured data

- Unorganized
- No specific format
- No predefined structure
- Text, images, audio, video

Structured vs unstructured data

Structured data

Name	Age	Gender
James	16	Male
Elizabeth	14	Female
Thomas	17	Male

Unstructured data

There are three students named James, Elizabeth, and Thomas. Their respective genders and ages are male 16, female 14, and male 17.

Quantitative vs. qualitative data

Quantitative data

- Numberical data
- Count, measure, percentage

Examples

- Age
- Temperature

Qualitative data

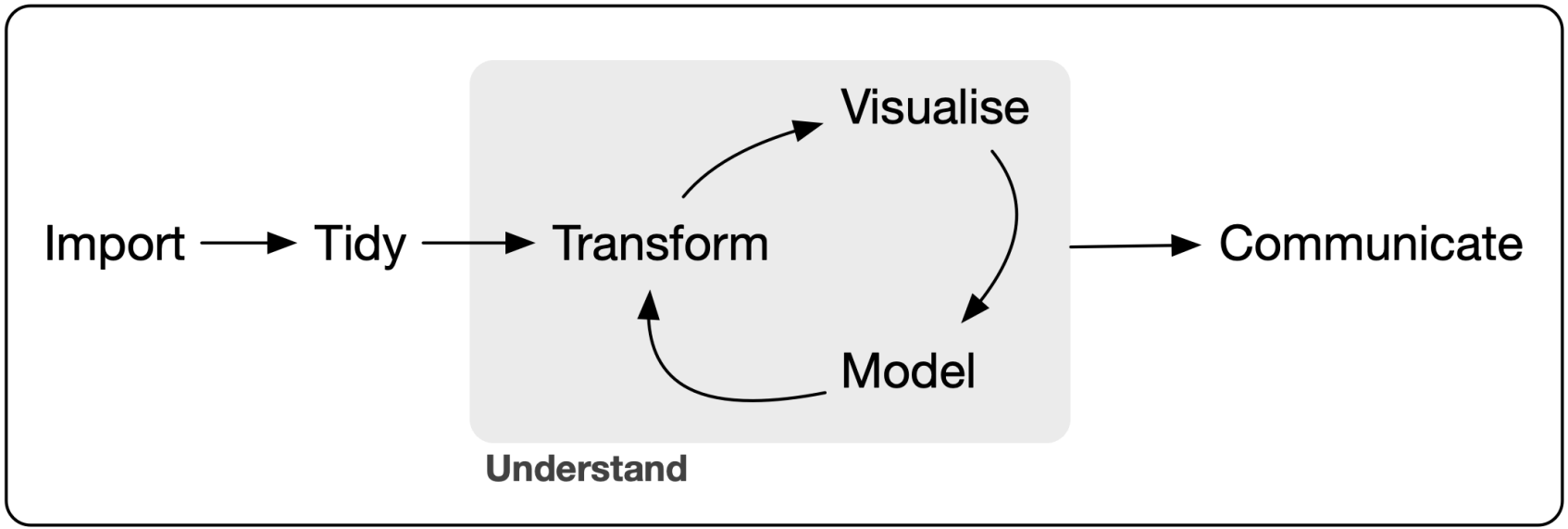
- Categorical data
- Group into categories

Examples

- Gender
- Holiday

Workflow of Data Science

"R for Data Science" (Wickham and Grolemund 2017)



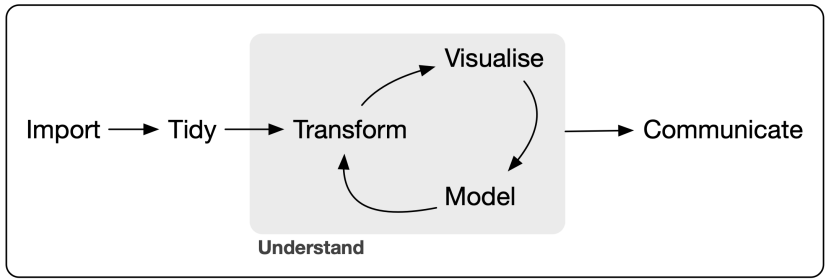
Workflow of Data Science

Import

- Open the file using one software
  - .csv / .tsv / .txt
  - .xlsx / .xls
  - .json

Tidy

- Format in tabular
  - Every column is a variable
  - Every row is an observation
  - Every cell has a single value



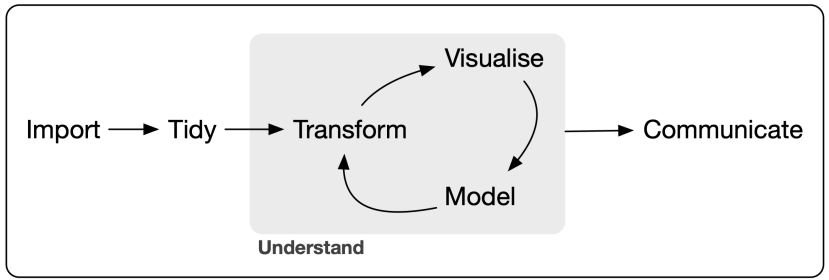
Workflow of Data Science

Transform

- Create new column
- Re-group categories
- Filter your sample
- Combine tables
- Handle missing data

Visualize

- Get the insight of your data
- Check the data quality
  - Missing (Completeness)
  - Validity / Accuracy



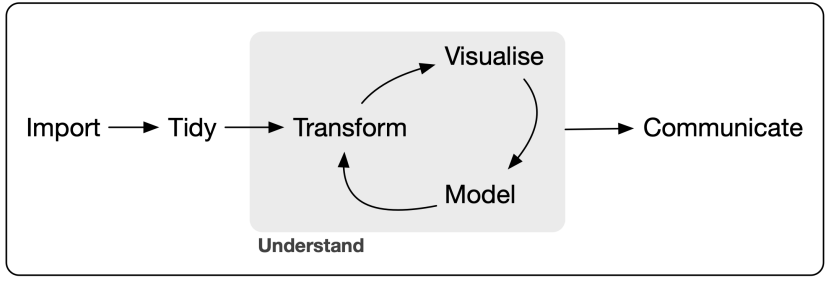
Workflow of Data Science

Modeling

- Descriptive (Business Intelligence)
- Out-of-scope
  - Predictive (Machine Learning)
  - Interference (Statistic)

Communicate

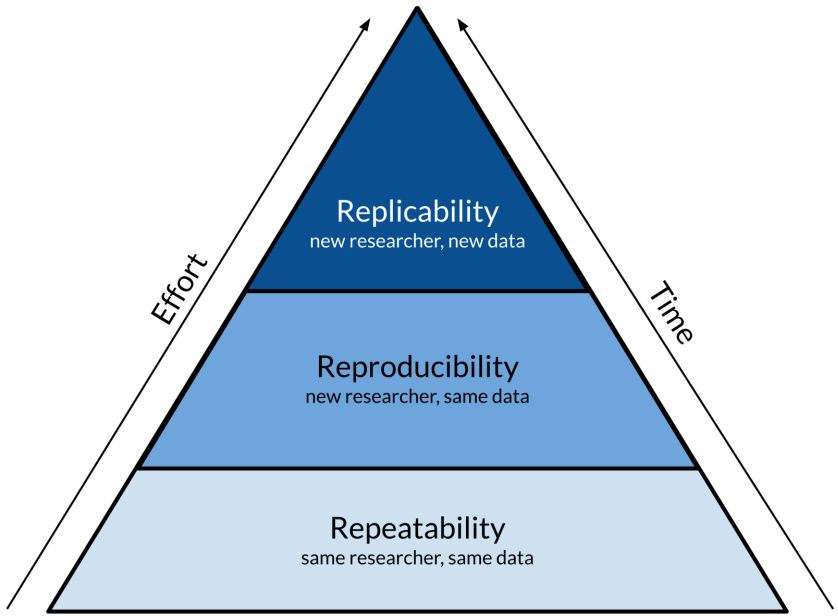
- Making (interactive) dashboard
  - Excel, Google sheet
  - Power BI, Looker, Tableau
  - R, Python



Common data cleaning challenges

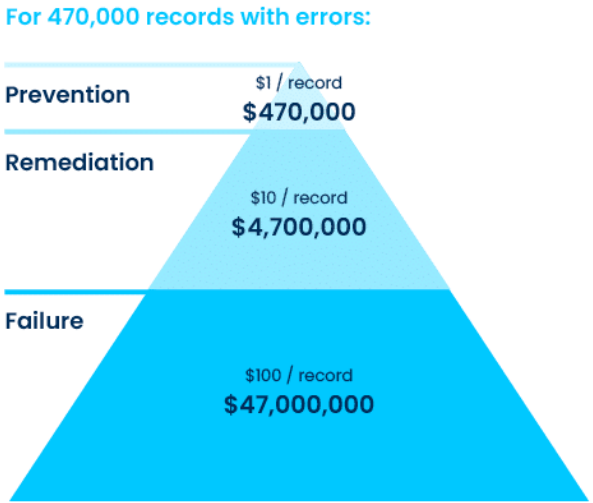
- Missing data
- Inconsistent
  - units
  - format
  - spelling variations
  - typo
- Outliers and errors
- Duplicate
- Unstructured
- Wrong data types

3-R's in data science



Based off of a figure from Essawy et al, 2020 https://doi.org/10.1016/j.envsoft.2020.104753

1-10-100 rule



reference

Chapter 2

Some Basic Tricks of Excel

Edit, Copy and Paste

Method 1

1. Select the cell
2. Make your edit, copy, or paste
3. When you are editing, **arrow** keyboard will select other cell

Method 2

1. Double click (F2) the cell
2. Make your edit, copy, or paste
3. When you are editing, **arrow** keyboard will be as usual

Remark

- You can only use the same method to copy and paste

Referencing

Absolute referencing

- **constant** when copied to other cell

```
=$A$2 / =A$2 / =$A2
```

[redacted]

Relative referencing

- **change** when copied to other cell

```
=A2 / =Sheet1!A2
```

[redacted]

Other topic

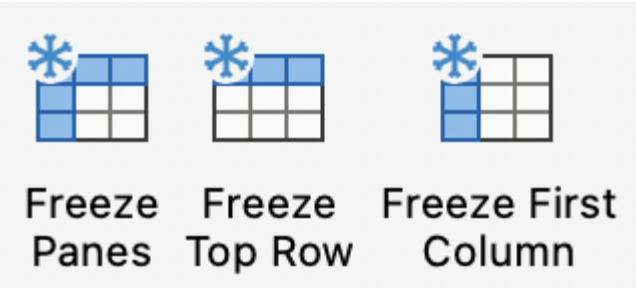
- Multiple lines in one cell
- Freeze cell
- Custom format
- F4 Hotkey
- Resize width and height
- Paste special

Multiple lines in one cell

- Press Alt + Enter

Freeze cell

- View ->
  - Freeze Panes (Custom the Freezing behavior)
  - Freeze Top Row
  - Freeze First Column



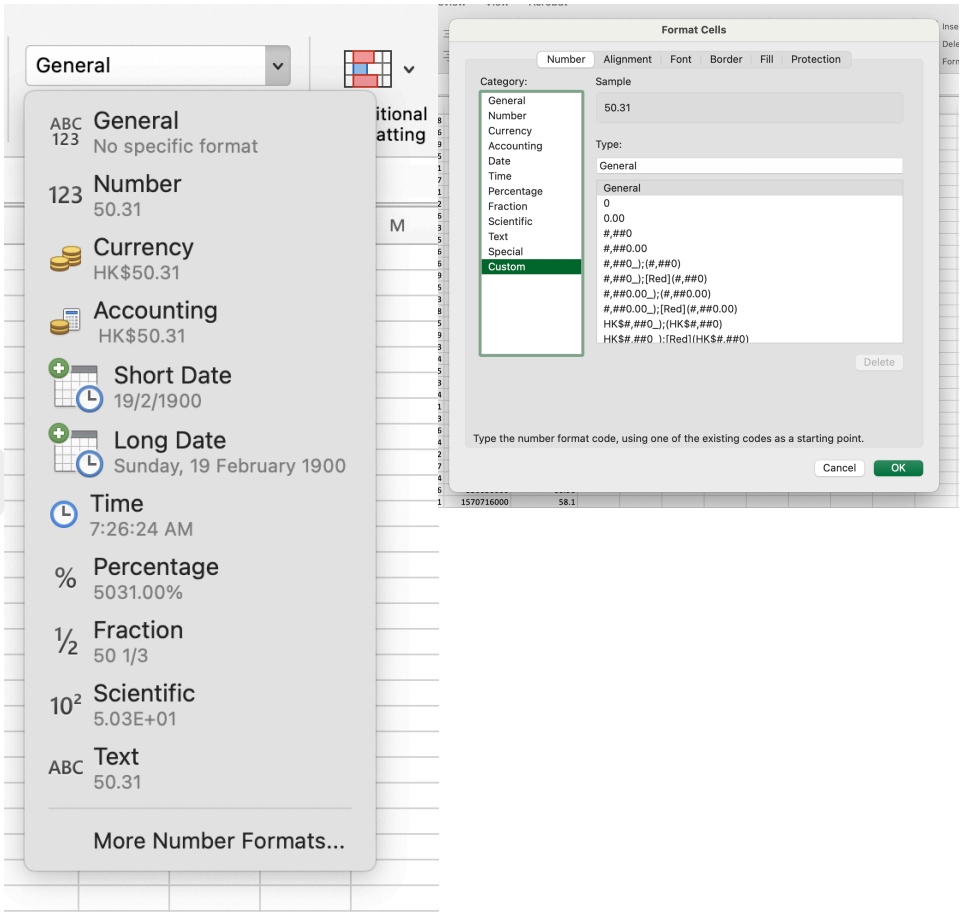
Formatting

- Build-in format
- Custom format
- Using formula

```
=TEXT(A2, "[$-404]aaaa")
```

reference-date-format

reference-other-format



F4 Hotkey

- Make previous action

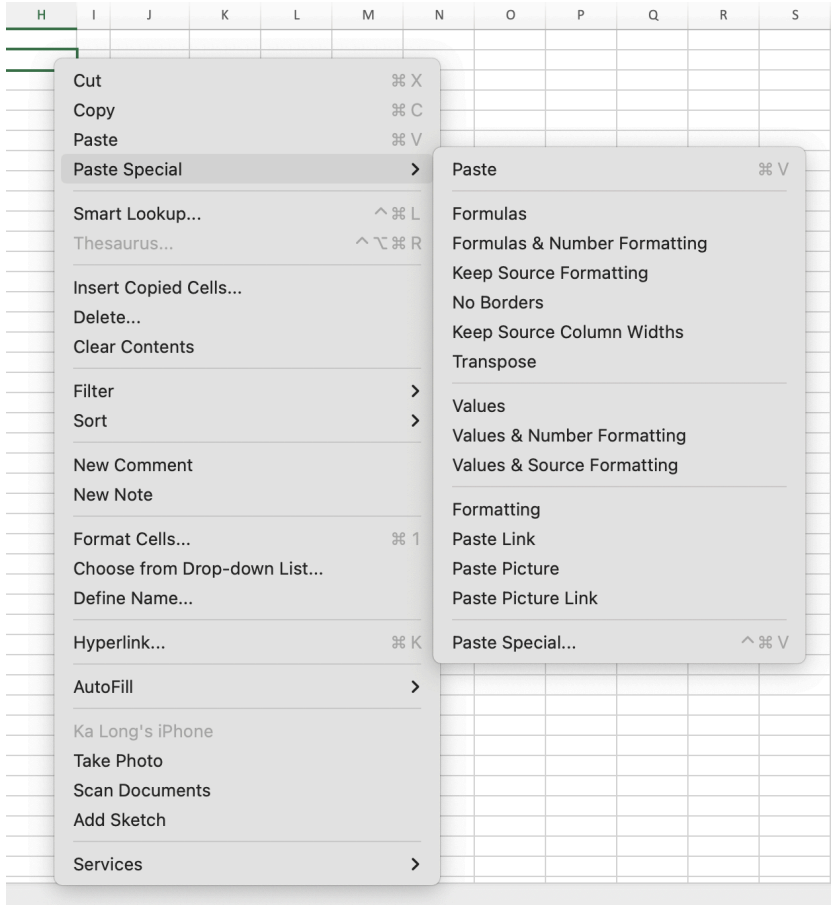
- Change referencing style

Resize width and height

1. Select all
2. Double click the height line or width line

Paste special

- Transpose
- Values



Exercise

Chapter 3

Using Table in Excel

Feature of Table in Excel

- Column header row
  - Unique header name
- Body
- Total row
  - Disable by default
  - Build-in function (SUM / AVERAGE)
  - Custom function

	A	B	C	D	E
1	OrderNo	SalesOrderLineKey	OrderQuantity	ItemCost	ItemPrice
2	SO43697	43697001	1	$2,171.29	$3,578.27
3	SO43698	43698001	1	$1,912.15	$3,399.99
4	SO43699	43699001	1	$1,912.15	$3,399.99
5	SO43700	43700001	1	$413.15	$699.10
6	SO43701	43701001	1	$1,912.15	$3,399.99
7	SO43702	43702001	1	$2,171.29	$3,578.27
8	SO43703	43703001	1	$2,171.29	$3,578.27
9	SO43704	43704001	1	$1,898.09	$3,374.99
10	SO43705	43705001	1	$1,912.15	$3,399.99
11	SO43706	43706001	1	$2,171.29	$3,578.27
12	SO43707	43707001	1	$2,171.29	$3,578.27
13	**Total**		**11**	**$20,816.29**	**$35,565.40**

Benefit of using Table

Structured referencing

- Automatically updates as data is added

```
=[@colName]
```

Data quality

- Automatically add new column
- Build-in filtering and sorting
- Data validation automatically updates
- Re-size table to add or remove data

Exercise

Chapter 4

Basic Functions in Excel

Type of functions in Excel

- Command button in Ribbon
  - Data validation
  - Conditional formatting
  - Remove duplicates
  - Flash fill
  - Split text to columns
- Operations
- Function
  - Aggregate
  - Text
  - Date & Time
  - Maths
  - Logical
  - Lookup

Operations

Symbol	Operation
+	Addition
-	Subtraction
*	Multiplication
/	Division
^	Exponentiation

Symbol	Operation
>	Greater than
<	Less than
>=	Greater than or equal to
<=	Less than or equal to
<>	Not equal
=	Equal to

Aggregate function

```
//Syntax
=SUM(cell1,[cell2],...)

//Example
=SUM(A2:A10)
=AVERAGE(A2:A10)
=MAX(A2:A10)
=MIN(A2:A10)
=COUNT(A2:A10)
=COUNTBLANK(A2:A10)
```

Text function

```excel
//Syntax
=LEFT(cell1, [number_character])

//Example
=LEFT("chankalong", 4) //Extract part of the string start from left
=RIGHT("chankalong", 4) //Extract part of the string start from right

//Syntax
=TRIM(cell1)

//Example
=TRIM() //Remove leading or padding space
=CLEAN() //Remove non-printed character
=UPPER() //CONVERT THE TEXT TO UPPERCASE
=LOWER() //convert the text to lowercase
=PROPER() //Convert The Text To Proper Case
=LEN() //Get the length of the text
```

Text function

```
//Syntax
=CONCAT(cell1, [cell2], ....)
=CONCATENATE("chankalong", "@", "bokss.org.hk") //=CONCAT() combine multiple cells, only on excel 2021

//Syntax and Example
=TEXTJOIN(delimiter, ignore_emphy_cell, cell1, [cell2], ....)
=TEXTJOIN(".", TRUE, "bokss", "org", "hk") //combine multiple cells, only on excel 2021

//Syntax and Example
=TEXTSPLIT(cell1, delimiter)
=TEXTSPLIT("chankalong@bokss.org.hk", "@") //split one cell into multiple cells

//Syntax and Example
=SUBSTITUTE(cell, old_text, new_text)
=SUBSTITUTE("chankalong", "long", "") //change specific character with another character

//Syntax and Example
=REPLACE(cell, start_num, num_character, new_text)
=REPLACE("chankalong", 7, 4, "") //change characters in a specified position of a text string
```

Date & Time function

Date

```excel
//Syntax and Example
=DATE(year, month, day)
=DATE(2024, 5, 30) //create date

//Syntax
=YEAR(date)

//Example
=YEAR("2024-05-30")
=MONTH("2024-05-30")
=DAY("2024-05-30")
=WEEKDAY("2024-05-30")
=WEEKNUM("2024-05-30")

//Syntax and Example
=DAYS(end_date, start_date)
=DAYS("2024-05-30", "2024-05-20")
```

Time

```excel
//Syntax and Example
=TIME(hour, minutes, second)
=TIME(10, 25, 0) //create time

//Syntax
=HOUR(time)

//Example
=HOUR("10:25 am")
=MINUTE("10:25 am")
=SECOND("10:25 am")
```

Current

```excel
=NOW() //get the current datetime
=TODAY() //get the current date
```

Maths function

```
//Syntax
=ROUND(cell, number_digits)

//Example
=ROUND(10.5, 0)
=ROUNDUP(10.5, 0)
=ROUNDDOWN(10.5, 0)

//Syntax
=ABS(cell)

//Example
=ABS(-10) //get the absolute value
=INT(-10.5) //get the integer (round down)
```

Logical function

```excel
//Syntax
=ISNUMBER(cell)

//Example
=ISNUMBER(100) //return TRUE if it is number
=ISERROR(100) //return TRUE if it is error
=ISERR(100) //return TRUE if it is error
=ISBLANK(100) //return TRUE if it is blank

//Syntax
=AND(logical1, [logical2], ...)

//Example
=AND("chankalong" = "chankalong", 1 = 1) //return TRUE if all logical value is TRUE
=OR("chankalong" <> "bokss", 1 > 2) //return TRUE if one of the logical value is TRUE

//Syntax and Example
=NOT(logical1)
=NOT(ISNUMBER(100)) //reverse logical value
```

Logical function

```
//Syntax and Example
=IF(logical, value_if_true, value_if_false)
=IF("chankalong" = "chankalong", "same people", "different people")

//Syntax and Example
=IFS(logical1, value_if_true1, [logical2], [value_if_true2], ...)
=IFS(A2 > 150, "A", A2 > 130, "B", A2 > 110, "C", TRUE, "D")
```

Logical + Aggregate function

[Aggregate][Logical]()

```
//Syntax
=SUMIF(logical_range, logical_criteria, [calculation_range])

//Example
=SUMIF(A1:A5, "=a", B1:B5)
=AVERAGEIF(A1:A5, "=a", B1:B5)
//if calculation_range is empty,
//calculation would be based on logical_range

//Syntax
=SUMIFS(calculation_range1, logical_range1, logical_criteria1,
[logical_range2, logical_criteria2], ...)

//Example
=SUMIFS(A1:A5, B1:B5, "=a", C1:C5, "=c")
=AVERAGEIFS(A1:A5, B1:B5, "= a", C1:C5, "= c")
```

D[Aggregate]()

```
//Syntax
=DMAX(cell_range, calculation_column,
criteria_table)

//Example
=DMAX(A1:E5, "value", E1:E2)
=DMIN()
```

Lookup function

```
//Syntax
=VLOOKUP(loop_up_value, look_up_table, column_index)

//Example
=VLOOKUP(A1, E1:G10, 2)
//loop_up_value should be located in the first column of look_up_table
=HLOOKUP()
```

Other useful function

```excel
//Syntax
=SEARCH(find_text, search_text)

//Example
=SEARCH("a", "chankalong") //case-insensitive and can use wildcard, return the position
=SEARCH("a?o", "chankalong") //case-insensitive and can use wildcard, return the position
=FIND("a", "chankalong") //case-sensitive and can use wildcard, return the position

//Syntax and Example
=INDIRECT(reference_text)
=INDIRECT("A1") //get the value in A1

//Syntax and Example
=INDEX(range, row_number)
=INDEX(A1:A5, 2) //get the value of second row in the A1:A5
=MATCH(lookup_value, look_up_table)
=MATCH("e", A1:A5) //get the location of "e" in the A1:A5
```

Other useful function

```excel
//Syntax and Example
=ROW() //get the row number
=COLUMN() //get the column number
=ADDRESS(row_number, column_number)
=ADDRESS(2, 3) //return "$C$2$"
=OFFSET(reference_cell, row_offset_number, column_offset_number)
=OFFSET(A1, 2, 3) //get the value offset from A1

//Syntax and Example
=LARGE(range, nth)
=LARGE(A1:A10, 3) //get the nth largest value in a range
=SMALL(range, nth)
=SMALL(A1:A10, 2) //get the nth smallest value in a range

//Syntax
=ROWS(range)

//Example
=ROWS(A1:A20) //get number of rows in the range
=COLUMNS(A1:P1) //get number of columns in the range
```

Other useful function only in Excel 2021 or Excel Web

```
=SORT()
=SORTBY()
=UNIQUE()
=FILTER()
=XLOOKUP()
=XMATCH()
=SWITCH()
```

Wildcards in Excel

Use to match pattern in function

- Find and Replace
- Conditional Formatting
- Filter
- SEARCH()
- XMATCH()
- [Aggregate][Logical]()

Symbol	Meaning
?	1 character
*	0 or more character
~	Escape whilcards

Pivot Tables

Benefit

- Aggregate and organize data in dynamic tables
- Transform rows to columns, or vice versa
- Group, filter, aggregate without need to make changes

Use case

- Collapse groups
- Combine two table (Vlook() + Column())
- Data validation with function
- Conditional formatting with function

- some example

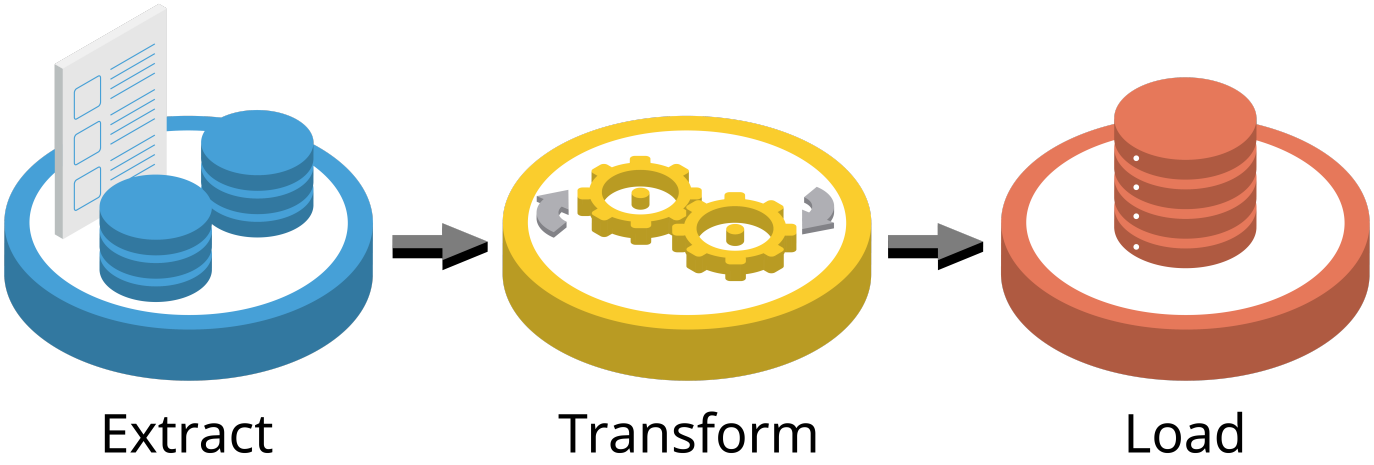
```
=ISNUMBER(SEARCH("gmail.com", A1)) //check whether the cell have "gmail.com"
=AND(LEN(A2) = 8, ISNUMBER(A2)) //check is it a phone number (e.g.: 37515493)
=LEFT(A3, SEARCH("-", A3)-1) //extract the text before "-"
=INDIRECT("Table1[gender]") //use in data validation to get the list in the gender column
=VLOOKUP(A2, $A$2:$F$100, COLUMN()-1, FALSE)
```

Exercise

Chapter 5

Power Query in Excel

Big data era

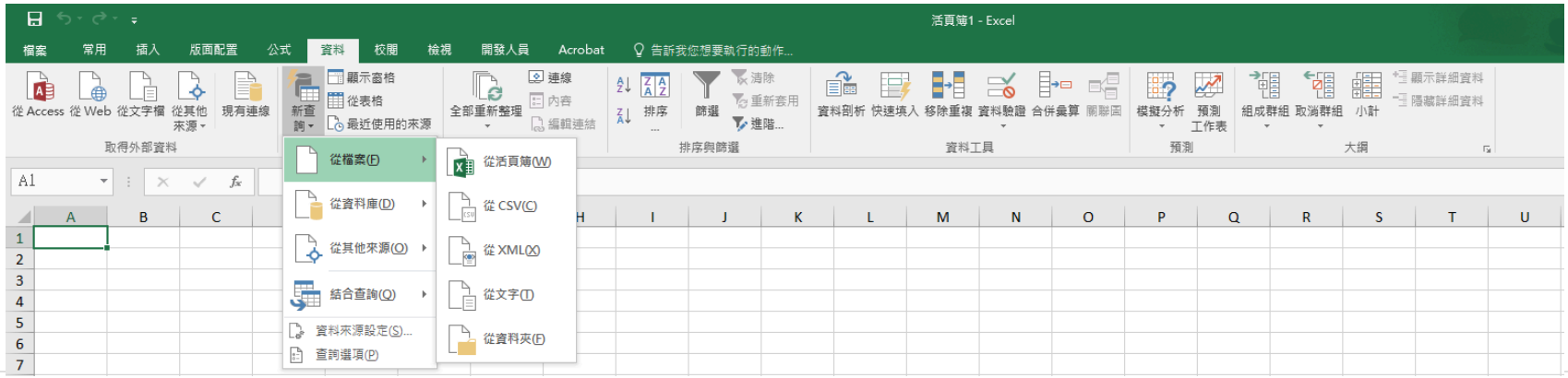


Benefit of power query

- Combine data from different source
- Keep track of every step in the workflow
- Easy to do every step in workflow
- Update the analytics when the data source update
- Ensure the reproducibility

Import data

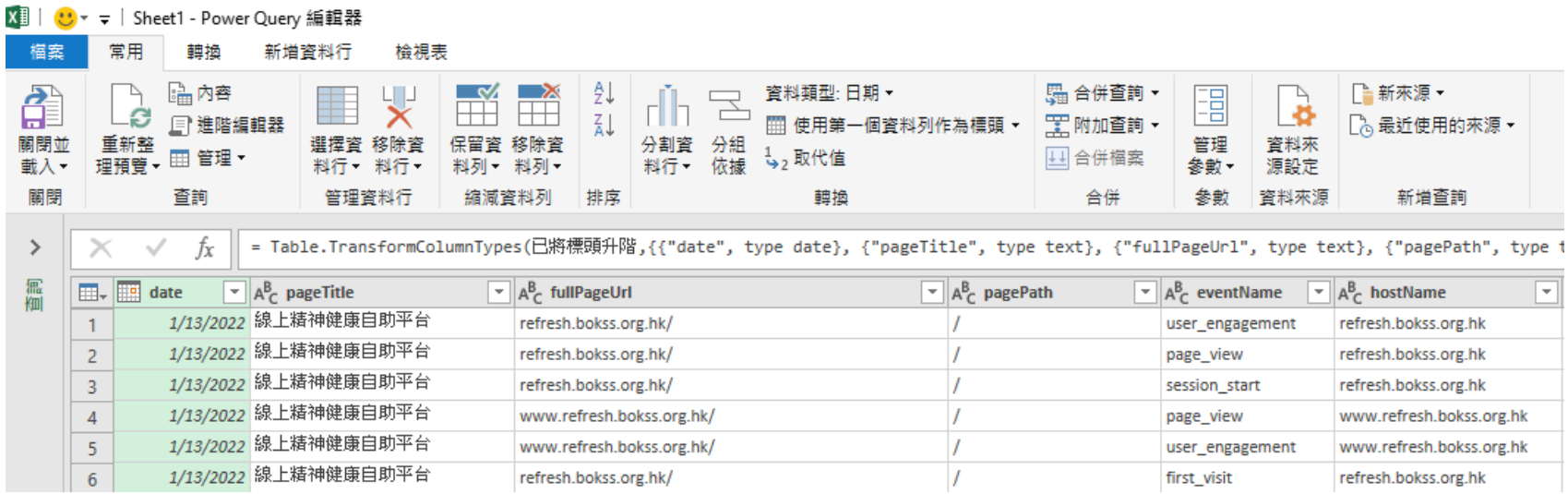
- Single file
- Multiple files with same structure
- Multiple sheets in Excel with same structure (= Excel.CurrentWorkbook())
- File from Onedrive Business (BOKSS)



Multiple sheets in one Excel

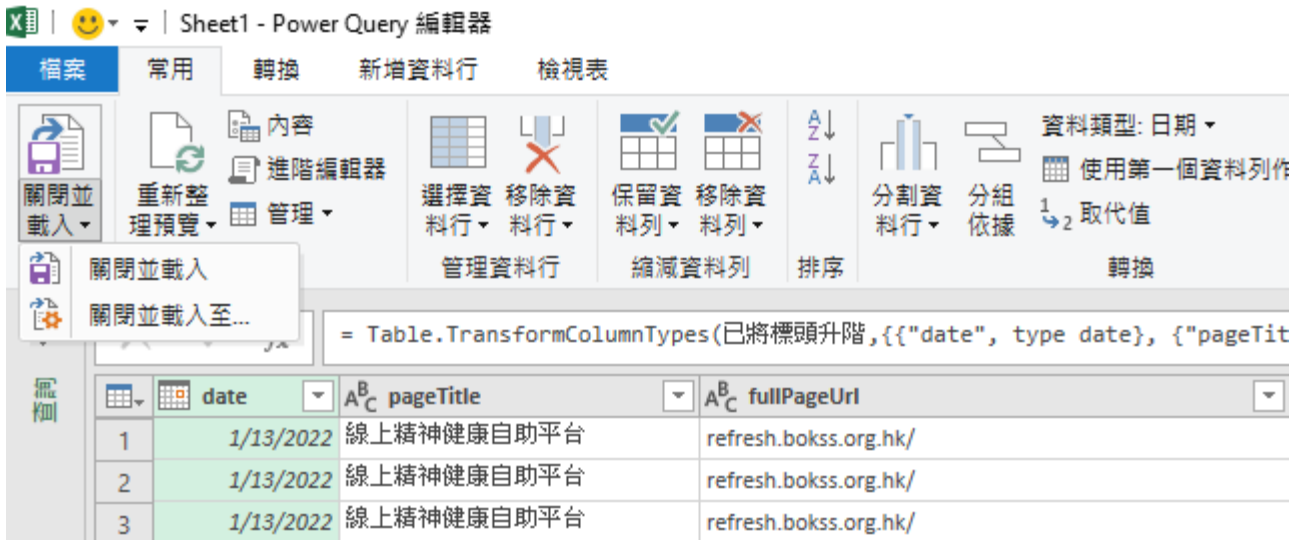
Transform

- Modify current column
- Add new column



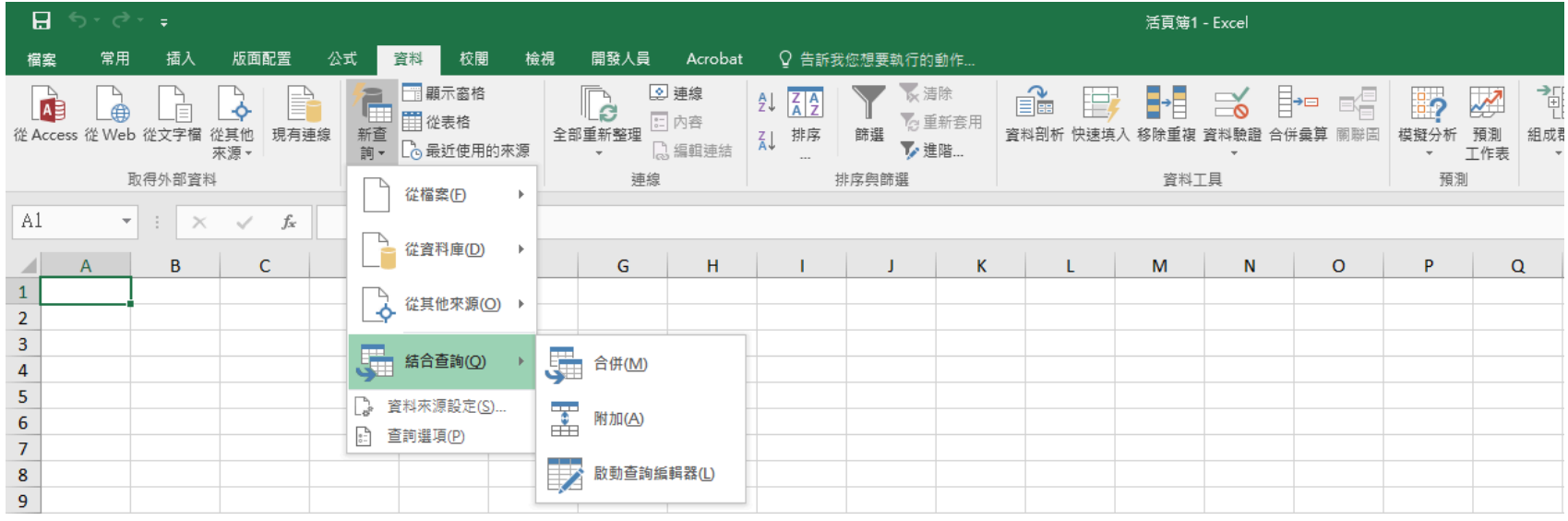
Load

- Load to excel sheet
- Load to excel connection



Combine query

- Merge (Join or map)
- Append



Different type of join

- Full join
- Left join or right join
- Inner join
- Anti join

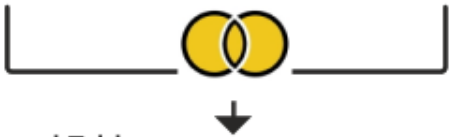
Full join

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	2	35

Right Table

ID	Country
1	USA
2	Canada
3	Panama
4	Spain



Merged Table

Date	CountryID	Units	Country
1/1/2020	1	40	USA
1/2/2020	1	25	USA
1/4/2020	2	35	Canada
1/3/2020	3	30	Panama
null	4	*null*	Spain

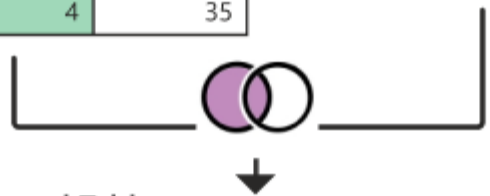
Left join (Right join)

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	4	35

Right Table

ID	Country
1	USA
2	Canada
3	Panama



Merged Table

Date	CountryID	Units	Country
1/1/2020	1	40	USA
1/2/2020	1	25	USA
1/3/2020	3	30	Panama
1/4/2020	4	35	*null*

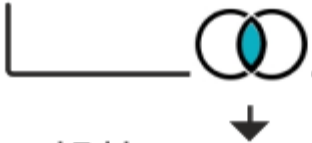
Inner join

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	2	35

Right Table

ID	Country
3	Panama
4	Spain



Merged Table

Date	CountryID	Units	Country
1/3/2020	3	30	Panama

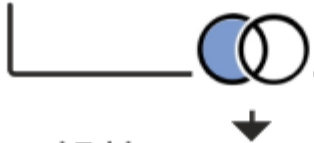
Anti join

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	2	35

Right Table

ID	Country
3	Panama
4	Spain



Merged Table

Date	CountryID	Units	Country
1/1/2020	1	40	*null*
1/2/2020	1	25	*null*
1/4/2020	2	35	*null*

Exercise

1. Append the Meeting calendar for 2016-2019 and 2020-2023
2. Clean the date and time column
3. remove the missing row for district column
4. Merge the Ambulance with calendar (filter Ambulance only)
5. how many emphy row in the Count column?

Chapter 6

Power BI Navigation

Report View

1. Canvas
2. Filters Pane
3. Visualization Pane
   - Build visual
   - Format page
4. Data Pane
5. Page Overview

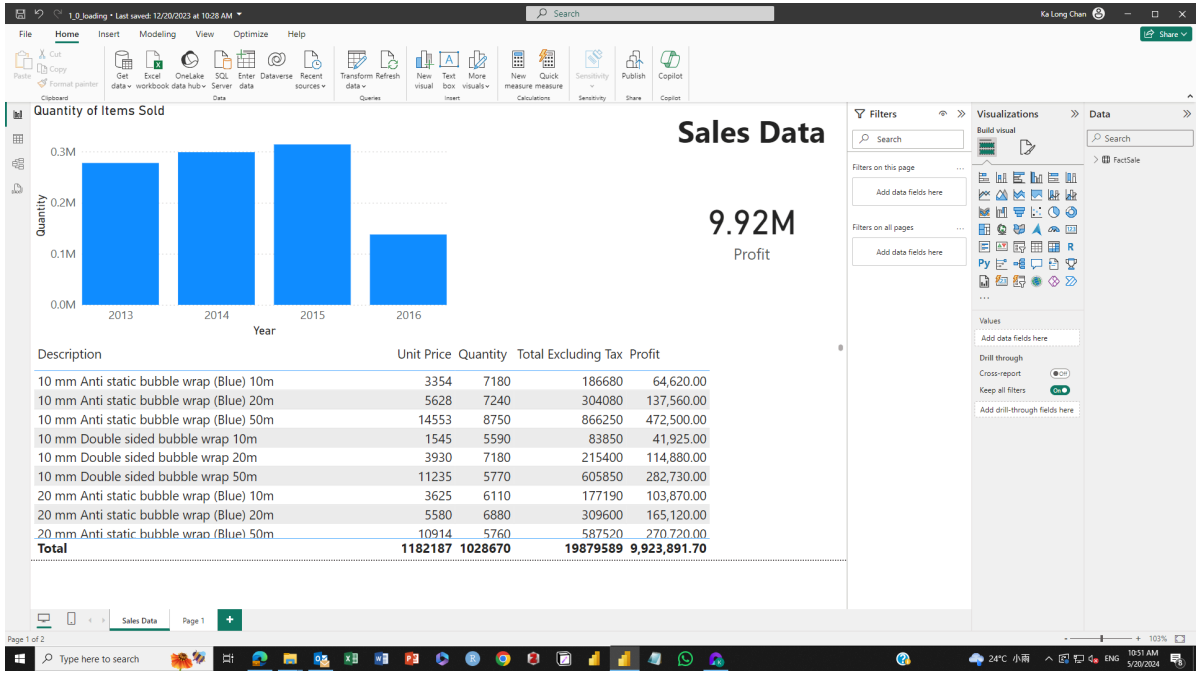
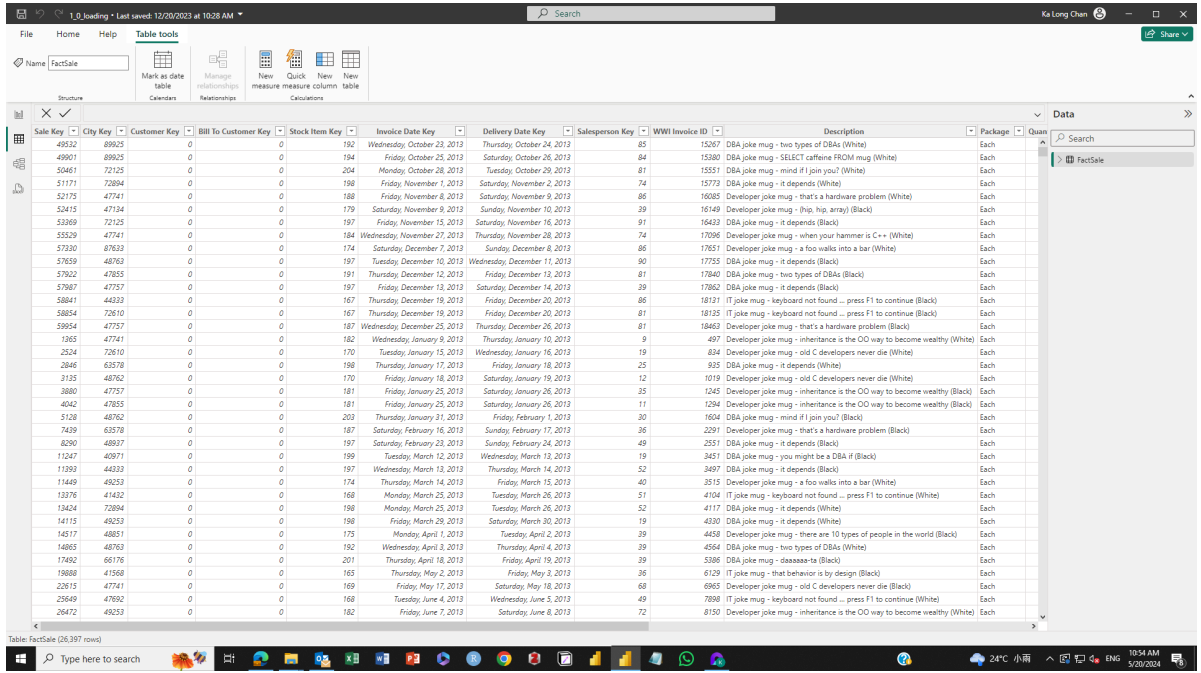


Table View

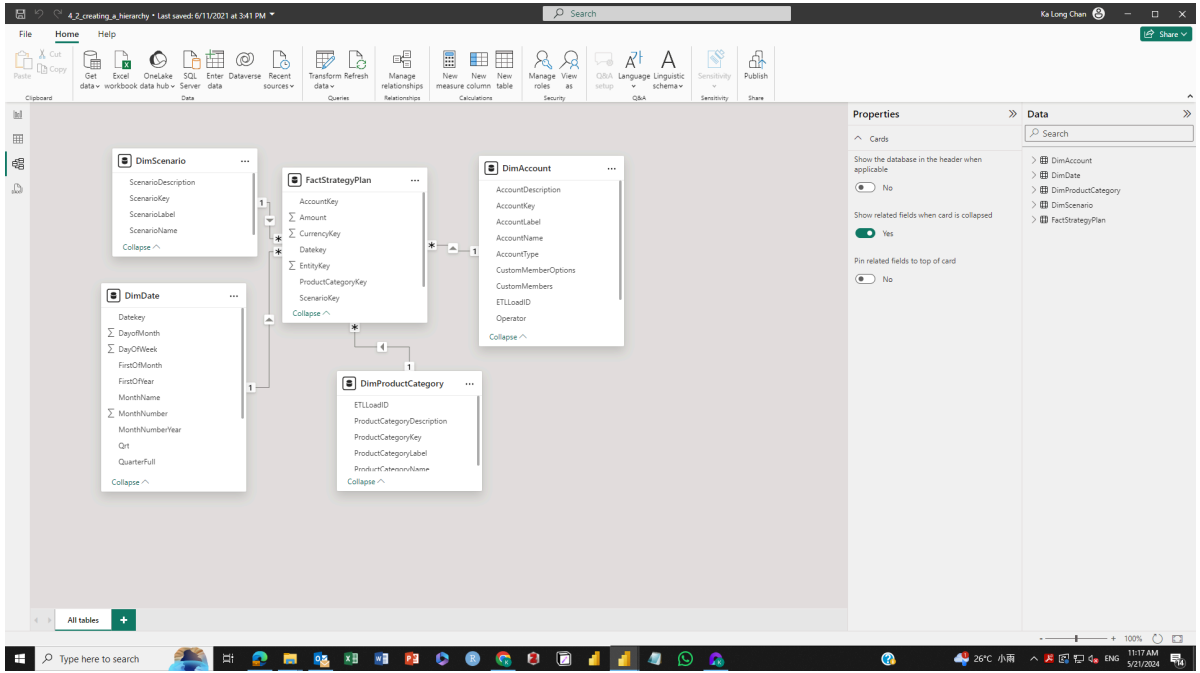
1. Data Grid
2. Data Pane
3. DAX Formula bar
4. Table Tools



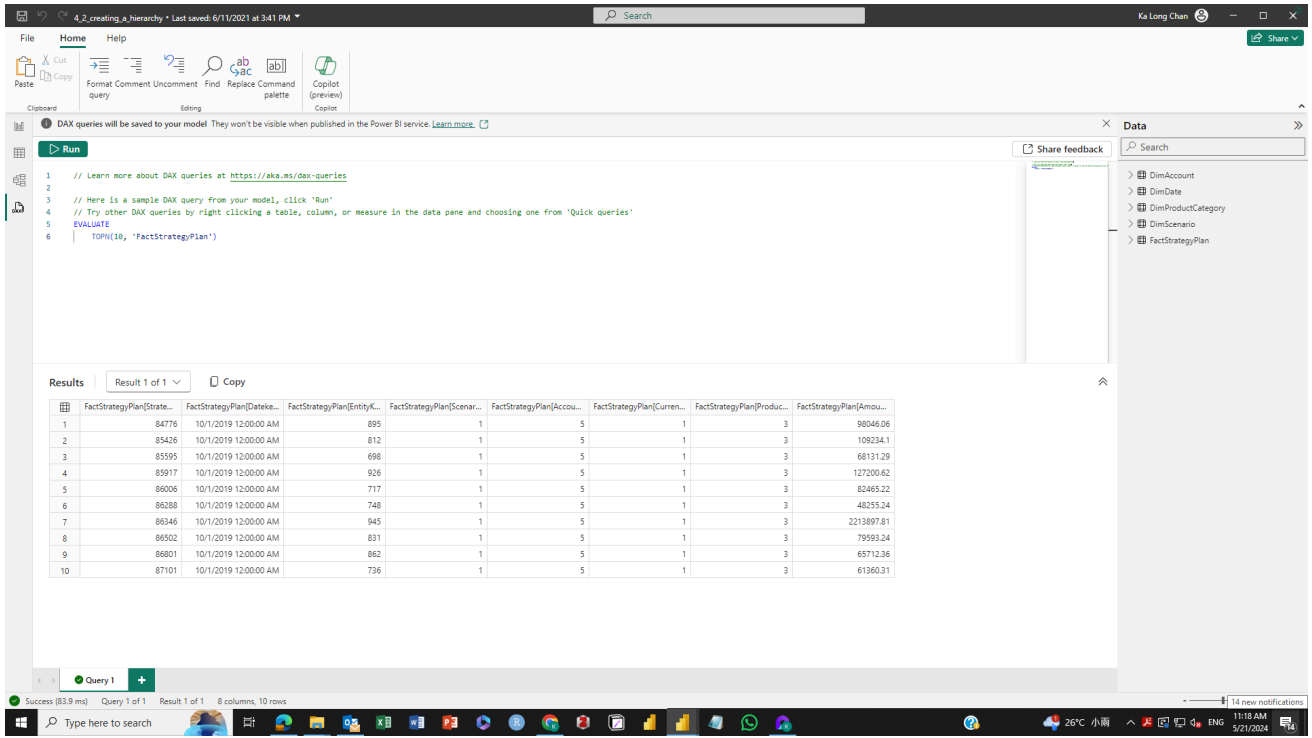
Sale Key	City Key	Customer Key	Bill To Customer Key	Stock Item Key	Invoice Date Key	Delivery Date Key	Salesperson Key	WWI Invoice ID	Description	Package	Quan
49532	89925	0	0	192	Wednesday, October 23, 2013	Thursday, October 24, 2013	85	15267	DBA joke mug - two types of DBAs (White)	Each	
49901	89925	0	0	194	Friday, October 25, 2013	Saturday, October 26, 2013	84	15380	DBA joke mug - SELECT caffeine FROM mug (White)	Each	
50461	72125	0	0	204	Monday, October 28, 2013	Tuesday, October 29, 2013	81	15551	DBA joke mug - mind if I join you? (White)	Each	
51171	72894	0	0	198	Friday, November 1, 2013	Saturday, November 2, 2013	74	15773	DBA joke mug - it depends (White)	Each	
52175	47741	0	0	188	Friday, November 8, 2013	Saturday, November 9, 2013	86	16085	Developer joke mug - that's a hardware problem (White)	Each	
52415	47134	0	0	179	Saturday, November 9, 2013	Sunday, November 10, 2013	39	16149	Developer joke mug - (hip, hip, array) (Black)	Each	
53369	72125	0	0	197	Friday, November 15, 2013	Saturday, November 16, 2013	91	16433	DBA joke mug - it depends (Black)	Each	
55529	47741	0	0	184	Wednesday, November 27, 2013	Thursday, November 28, 2013	74	17096	Developer joke mug - when your hammer is C++ (White)	Each	
57330	87633	0	0	174	Saturday, December 7, 2013	Sunday, December 8, 2013	86	17651	Developer joke mug - a foo walks into a bar (White)	Each	
57659	48763	0	0	197	Tuesday, December 10, 2013	Wednesday, December 11, 2013	90	17755	DBA joke mug - it depends (Black)	Each	
57922	47855	0	0	191	Thursday, December 12, 2013	Friday, December 13, 2013	81	17840	DBA joke mug - two types of DBAs (Black)	Each	
57987	47757	0	0	197	Friday, December 13, 2013	Saturday, December 14, 2013	39	17862	DBA joke mug - it depends (Black)	Each	
58841	44333	0	0	167	Thursday, December 19, 2013	Friday, December 20, 2013	86	18121	IT joke mug - keyboard not found ... press F1 to continue (Black)	Each	
58854	72610	0	0	167	Thursday, December 19, 2013	Friday, December 20, 2013	81	18135	IT joke mug - keyboard not found ... press F1 to continue (Black)	Each	
59954	47757	0	0	187	Wednesday, December 25, 2013	Thursday, December 26, 2013	81	18463	Developer joke mug - that's a hardware problem (Black)	Each	
1365	47741	0	0	182	Wednesday, January 9, 2013	Thursday, January 10, 2013	9	497	Developer joke mug - inheritance is the OO way to become wealthy (White)	Each	
2524	72610	0	0	170	Tuesday, January 15, 2013	Wednesday, January 16, 2013	19	834	Developer joke mug - old C developers never die (White)	Each	
2846	63578	0	0	198	Thursday, January 17, 2013	Friday, January 18, 2013	25	935	DBA joke mug - it depends (White)	Each	
3135	48762	0	0	170	Friday, January 18, 2013	Saturday, January 19, 2013	12	1019	Developer joke mug - old C developers never die (White)	Each	
3880	47757	0	0	181	Friday, January 25, 2013	Saturday, January 26, 2013	35	1245	Developer joke mug - inheritance is the OO way to become wealthy (Black)	Each	
4042	47855	0	0	181	Friday, January 25, 2013	Saturday, January 26, 2013	11	1294	Developer joke mug - inheritance is the OO way to become wealthy (Black)	Each	
5128	48762	0	0	203	Thursday, January 31, 2013	Friday, February 1, 2013	30	1604	DBA joke mug - mind if I join you? (Black)	Each	
7439	63578	0	0	187	Saturday, February 16, 2013	Sunday, February 17, 2013	36	2291	Developer joke mug - that's a hardware problem (Black)	Each	
8290	48937	0	0	197	Saturday, February 23, 2013	Sunday, February 24, 2013	49	2551	DBA joke mug - it depends (Black)	Each	
11247	40971	0	0	199	Tuesday, March 12, 2013	Wednesday, March 13, 2013	19	3451	DBA joke mug - you might be a DBA if (Black)	Each	
11393	44333	0	0	197	Wednesday, March 13, 2013	Thursday, March 14, 2013	52	3497	DBA joke mug - it depends (Black)	Each	
11449	49253	0	0	174	Thursday, March 14, 2013	Friday, March 15, 2013	40	3515	Developer joke mug - a foo walks into a bar (White)	Each	
13376	41432	0	0	168	Monday, March 25, 2013	Tuesday, March 26, 2013	51	4104	IT joke mug - keyboard not found ... press F1 to continue (White)	Each	
13424	72894	0	0	198	Monday, March 25, 2013	Tuesday, March 26, 2013	52	4117	DBA joke mug - it depends (White)	Each	
14115	49253	0	0	198	Friday, March 29, 2013	Saturday, March 30, 2013	19	4330	DBA joke mug - it depends (White)	Each	
14517	48851	0	0	175	Monday, April 1, 2013	Tuesday, April 2, 2013	39	4458	Developer joke mug - there are 10 types of people in the world (Black)	Each	
14865	48763	0	0	192	Wednesday, April 3, 2013	Thursday, April 4, 2013	39	4564	DBA joke mug - two types of DBAs (White)	Each	
17492	66176	0	0	201	Thursday, April 18, 2013	Friday, April 19, 2013	39	5386	DBA joke mug - daaaaaa-ta (Black)	Each	
19888	41568	0	0	165	Thursday, May 2, 2013	Friday, May 3, 2013	36	6129	IT joke mug - that behavior is by design (Black)	Each	
22615	47741	0	0	169	Friday, May 17, 2013	Saturday, May 18, 2013	68	6965	Developer joke mug - old C developers never die (Black)	Each	
25649	47692	0	0	168	Tuesday, June 4, 2013	Wednesday, June 5, 2013	49	7898	IT joke mug - keyboard not found ... press F1 to continue (White)	Each	
26472	49253	0	0	182	Friday, June 7, 2013	Saturday, June 8, 2013	72	8150	Developer joke mug - inheritance is the OO way to become wealthy (White)	Each	

Model View

1. Data Model
2. Data Pane
3. Properties Pane

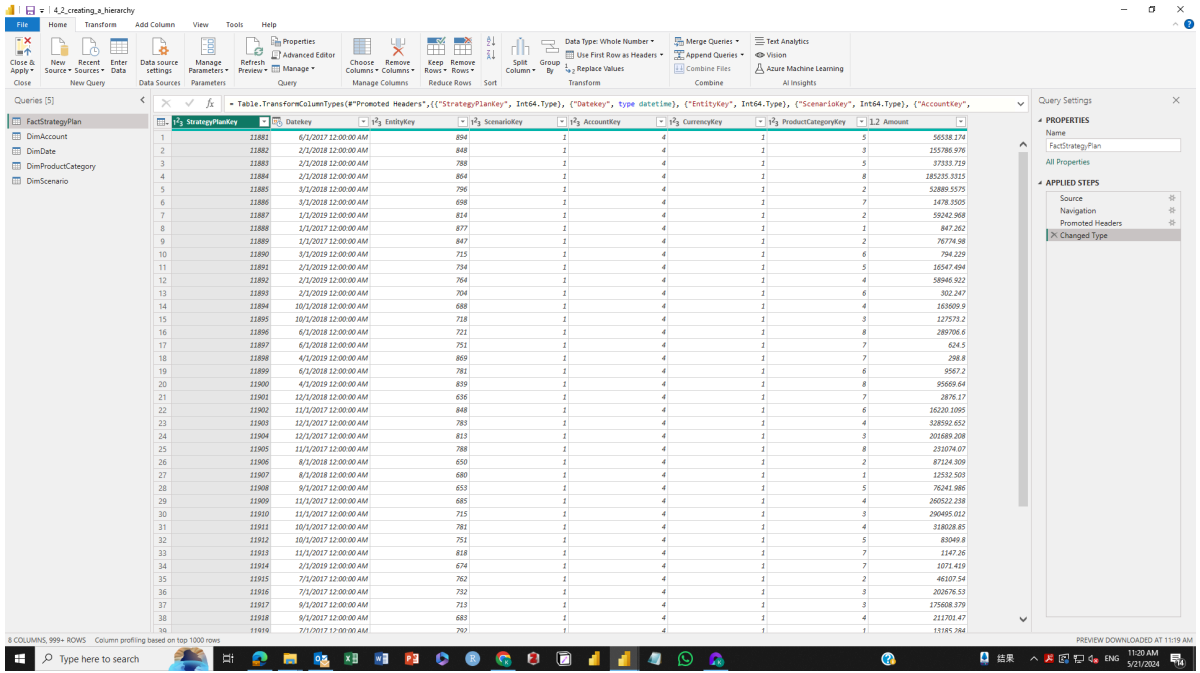


DAX Query View



Power Query Editor

1. Query Editor Ribbon
2. Queries Pane
3. Query Setting
4. Status Bar
5. Table Pane

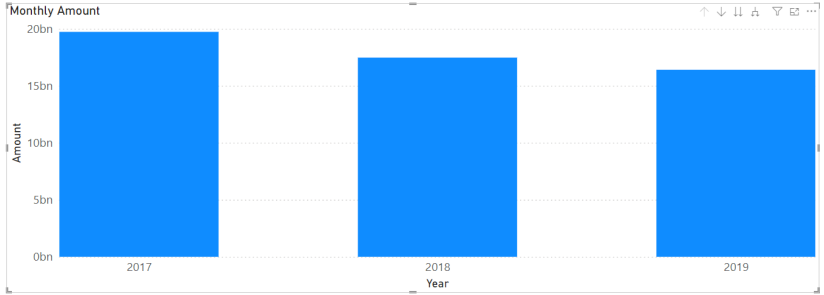


Drilling Down

1. Date
   - Auto create Date Table
   - Mark as Date Table
   - Using DAX
   - Using Power Query (M)

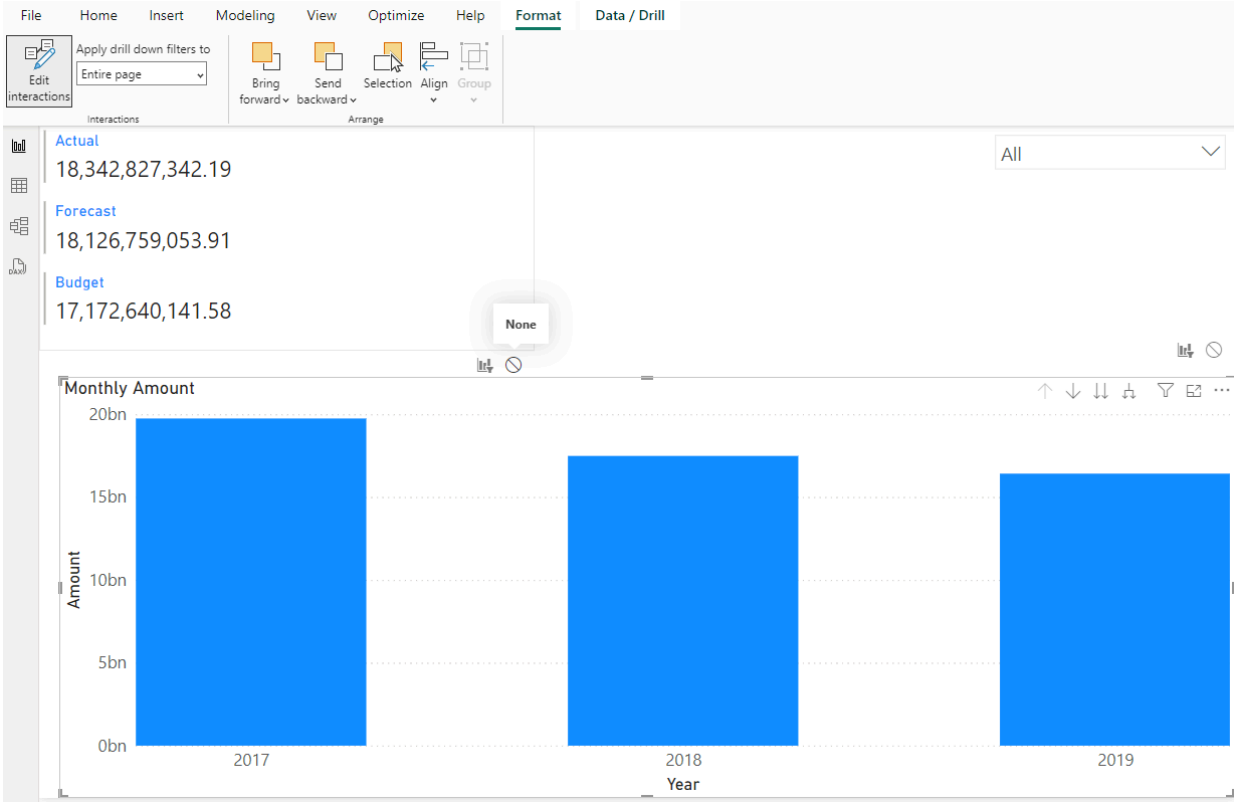
```
=List.Dates(#date(2020, 1, 1), 365*7, #duration(1, 0, 0, 0))
```

2. Custom Hierarchies
   - Department -> Unit -> Team



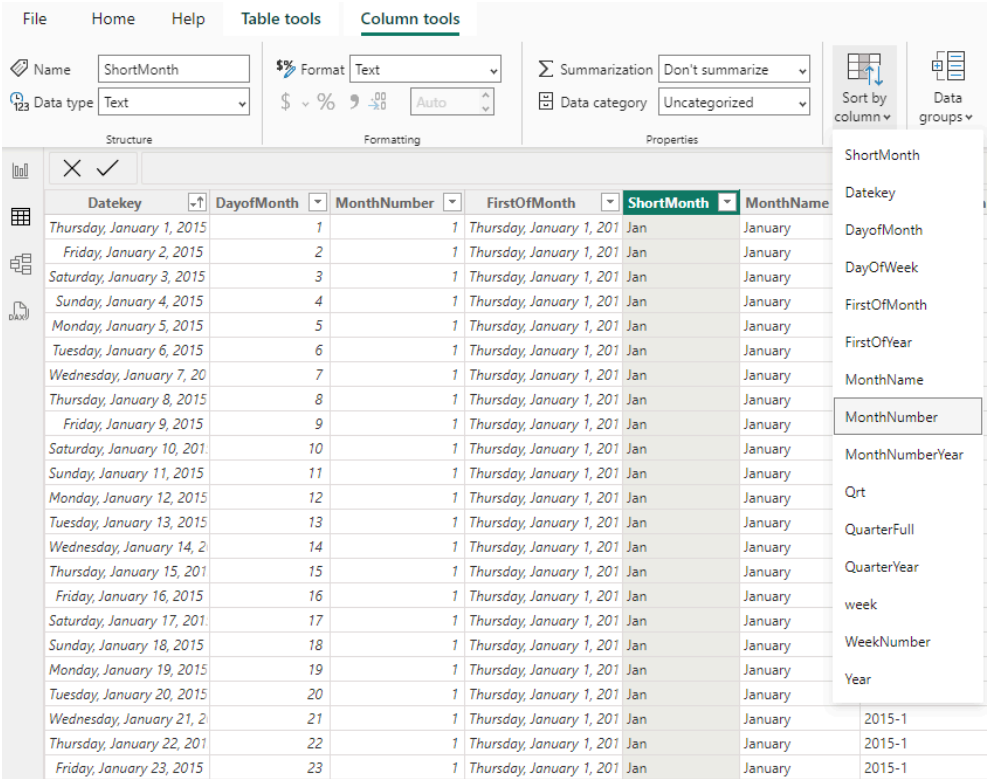
Filter

1. Visual Level
2. Page Level
3. Report Level
4. Slicer
5. Interactions



Sorting

1. Sort in Table View
2. sort in Visual
3. Sort by column



Visual and Format

Exercise

1. Create a date table
2. Create date hierarchy
3. Add a visual and drill-down of date hierarchy
4. what is the average profit on 2015 Quarter 1 for Packet Package? (Answer with 2 decimal digits)

Chapter 7

DAX in Power BI

What is DAX

- **D**ata **A**nalysis e**X**pressions
- Based on Excel formulas and functions
- Create calculated columns, measures, tables
- DAX reference
  - https://docs.microsoft.com/en-us/dax/dax-function-reference

Calculated columns

- Calculate at row level
- Add new column to existing table
- Based on Excel formulas and functions
- Calculate when data is refreshed

```
new_column = Sales[Price] + Sales[Tax]
```

Calculated measures

- Aggregates multiple rows
- Add a new field for visualization
- Calculated when you interact with the visuals
- Two ways
  - from scratch
  - use Quick Measure

Context of DAX

Aspect	Custom column	Calculate column	Calculated table	Measure	Visual calculation
Language	M	DAX	DAX	DAX	DAX
Computed at	Data refresh	Data refresh	Data refresh	On demand	On demand
Persistence	Results saved	Results saved	Results saved	Calculated as required	Calculated as required
Context	Row	Row	Row	Filter	Visual
Stored in	Table	Model	Model	Model	Visual
Changes with user interaction in report	No	No	No	Yes	Yes
Usage	Slicers, filters, rows, columns	Slicers, filters, rows, columns	In a measure, calculated column, or visual calculation definition	Value in a visual and visual level filter	Value in a visual and visual level filter

reference

Row Context

- Use current row (all row)
- Custom column (M language)
- Calculated columns

```
Sales[Price] * Sales[Tax]
```

Filter Context

- Filter before calculation is carried out
- Calculated measures
  - Aggregates on the calculated columns

```
SUM(Sales[Profit])

SUMX(<table>, <expression>)
SUMX(Sales, Sales[Price] * Sales[Tax])
SUMX(FILTER(Sales, Sales[Region]="EMEA"), Sales[Price] * Sales[Tax])

CALCULATE(<expression>, <filter1>, ...other filter conditions)
CALCULATE(SUM(Sales[Price]), Sales[Region]="EMEA")
```

VAR and RETURN

- Useful for complax calculation

```
Sales YoY Growth % =
DIVIDE(
    ([Sales] - CALCULATE([Sales], PARALLELPERIOD('Date'[Date], -12, MONTH))),
    CALCULATE([Sales], PARALLELPERIOD('Date'[Date], -12, MONTH))
)
```

```
Sales YoY Growth % =
VAR SalesPriorYear =
    CALCULATE([Sales], PARALLELPERIOD('Date'[Date], -12, MONTH))
RETURN
    DIVIDE(([Sales] - SalesPriorYear), SalesPriorYear)
```

Exercise

1. Create three filter calculated measure of total amount among "Actual", "Forecast", "Budget"
2. create the calculated measure of actual / budget
3. what is the percentage of actual / budget for Expense account type in cell phone?

Chapter 8

M language in Power BI

What is M language

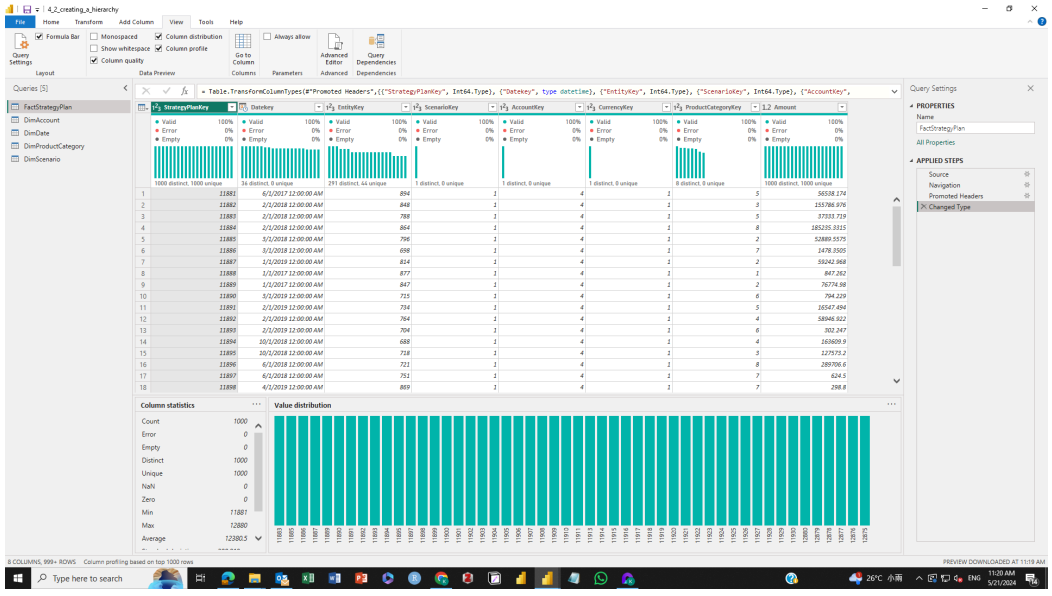
- programme language use in Power Query
- all action in Power Query will be convert to M language
- You can write M language manually in
  - Custom Column
  - Formula bar
  - Advanced Editor
- M language reference
  - https://learn.microsoft.com/en-us/powerquery-m/

Chapter 9

Power Query in Power BI

Data preview

- Quick analyze the data in power query
- Helps diagnose errors and inconsistencies
- Helps you decide what transformation(s) to use



Data Transform

- Pivot column : reshape data for report
- unpivot column: reshape data for data analysis
- Transpose: swap column and row name
- Group by

Unpivot and Pivot Column

Exercise

1. import and clean dataset
2. What is the average value of Southern?

Chapter 10

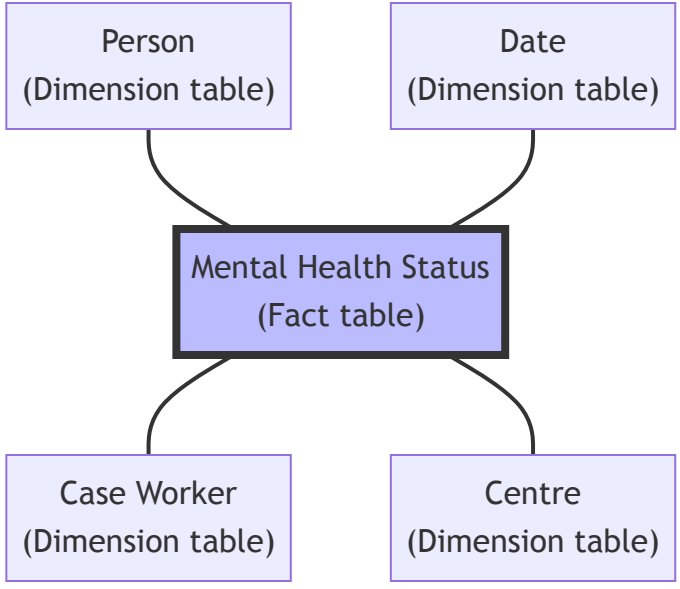
Data Model in Power BI

What is data model

- Tables
- Columns
- Data types
- Relationship between tables
- Keys

Star Schema

- Facts table: metrics
  - Facts (measures)
  - Keys: use to build relationship between tables
- Dimensions table: context
- Benefit
  - Reduce file size
  - Reduce redundant data
  - Easy to manage



Fact table

- Make up of
  - Facts (measures)
  - Keys: use to build relationship between tables
- Fact tables are long and narrow
  - Lots of rows
  - Fewer columns

id	centre_id	depression_scure
1	1	10
2	1	15
3	2	12
4	2	19
5	3	20

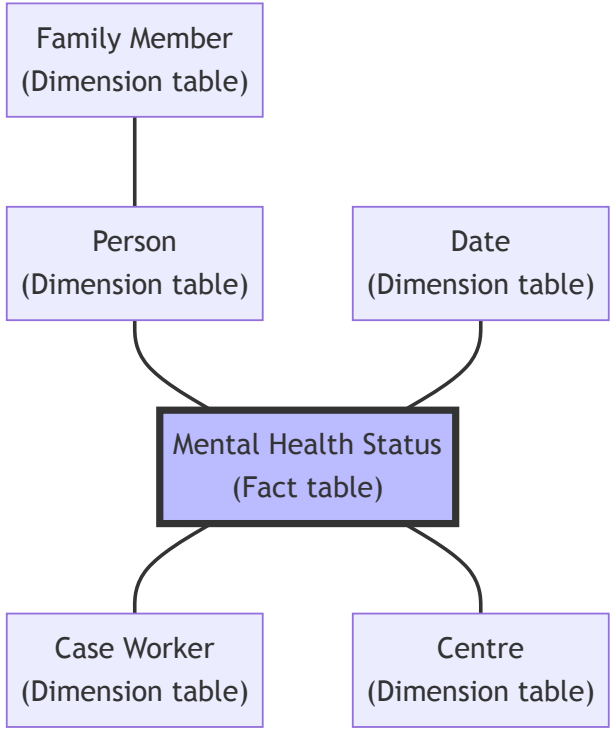
Dimension table

- Provide context
- Shared concepts
- Contain static data
- Dimension tables are short and wide
  - Few of rows
  - Lots of columns

centre_id	name_centre	district_centre
1	ICCMW (Kwai Tsing)	KT
2	ICCMW (Wan Chai)	WC
3	ICCMW (Eastern)	HKE

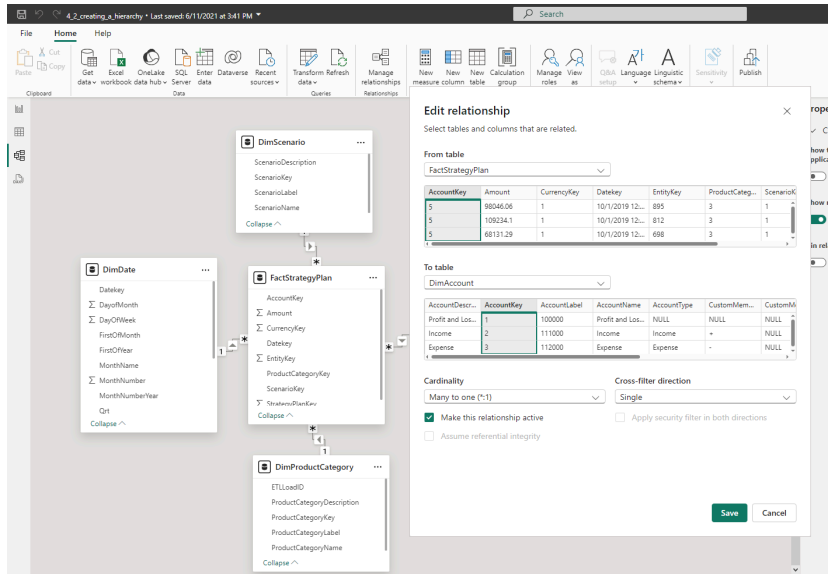
Snowflake Schema

- Allows relationships between dimensions
- Fact table(s) remain the same
- Benefit
  - Reduce file size
  - Reduce redundant data
  - Easy to manage



Build data model in Power BI

- Power BI will define relationship automatically
- You can also define relationship between tables in Model View
  - One-to-many relationship
  - Many-to-many relationship
  - One-to-one relationship
- Each pair of table can have more than one relationship



Create star schema or snowflake schema

- Modify the source file (not recommend)
- Use M language in Power Query

1. import file
2. use advanced editor to use "[]" to include all code
3. reference to the original query

Build data model in Power BI

- If more than one relationships
  - Set the active relationship in Model View
  - using DAX to activate the relationship when calculating a measure

```
= CALCULATE(SUM(InternetSales[SalesAmount]), USERELATIONSHIP(InternetSales[ShippingDate], DateTime[Date]))
```

Final Exercise

1. create dimension table from the code manual
2. create a Power BI report
   - one visual of single variables and one visual including two variables
   - one calculated measure
   - a slicer for filter
   - create conditional formatting for one visual
   - do not use the default theme
   - make a mobile view
3. publish the Power BI and share your Power BI link

Chapter 11

Advanced use case in Power BI

Advanced use case in Power BI

- Bookmark
- Selection
- Button and action
- Publish
- Mobile View
- Drill-through
- Customize tooltip

Bookmark

- save the filter you had applied

Selection

- manage the visibility of visual
- manage the order of visual
- can combine with Bookmark

Button and action

- Create a App-like experience
- Add action to buttons
- Actions
  - Nagative between pages
  - Filter
  - Bookmark

Publish

- Publish to Power BI Service
- Data model will also be published
- If data source connect to Onedrive, report would be update automatically

Mobile View

- Make a mobile layout for mobile user

Drill-through

- Go to another page and apply filter automatically

1. create a new page
2. add a new visual
3. add a filter into drill-through fields
4. turn on the cross-report under drill through
5. go to the original page and right click a plot in visual
6. select drill-through

Tooltip

- Allow to customize the mouseover information

1. create a new page
2. change the page to tooltip
3. change the tooltip of a visual to the newly create page

Chapter 12

Integration with Other Microsoft Product

Integration

- Microsoft Power Automate
- Microsoft Forms
- Microsoft Fabric
- Microsoft Power App (Less use in our case)

Microsoft Forms + Power Automate

1. Collect data using Forms
2. Data of Microsoft Forms -> Excel (Power Automate)
3. Add workflow to the Forms
    - Send email
    - Make approval
    - Add new column
    - Conditional control

Microsoft Fabric

1. Excel data -> Lakehouse (Dataflow)
2. Power BI or Notebook
3. Create Scorecard / Alert
4. Trigger another workflow using Power Automate

Chapter 13

Advance Usage for Form

Form platform

- Google Form (Google Form add-on formfacade)
- SurveyJS (Case library)
- MS Form (Less powerful)

Pre-fill

- Google Form
- MS Form
- Formfacade: can hide the pre-fill question

Regular expression

- Google Form
- SurveyJS (Case library)

Regular expression

symbol	rule	example	result
.	single character	b.y	boy / buy
?	zero or one of character	bo?y	boy / by
*	zero or more of character	bo*y	booy / boy / by
+	one or more of character	bo+y	booy / boy
^	start of the string	bo+y	booy / boy
$	end of the string	bo+y	booy / boy

Regular expression

symbol	rule	example	result
[a-zA-Z]	character in the range: a-z or A-Z	b[a-zA-Z]y	boy / buy
[0-9]	Any digits	b[0-9]y	boy / by
|	or	b[o|u]y	buy / boy
\	match symbol literally	\.com	.com
{}	repeat how many times	[0-9]{8}	37515493

online testing

Exercise

- Use google form to create

1. email
2. mobile (8 digits with "+852")
3. HKID (e.g.: A123)
4. 12-hour time (e.g.: 10:30 AM)

Evaluation Form

MS Form

Future Study

- Microsoft Certified: Power BI Data Analyst Associate
- Google Analytics and Google Tag Manager (For online service)
- Using Fabric Notebook (R / Python)
- Power Platform (Power App / Power Automate / Power Page)