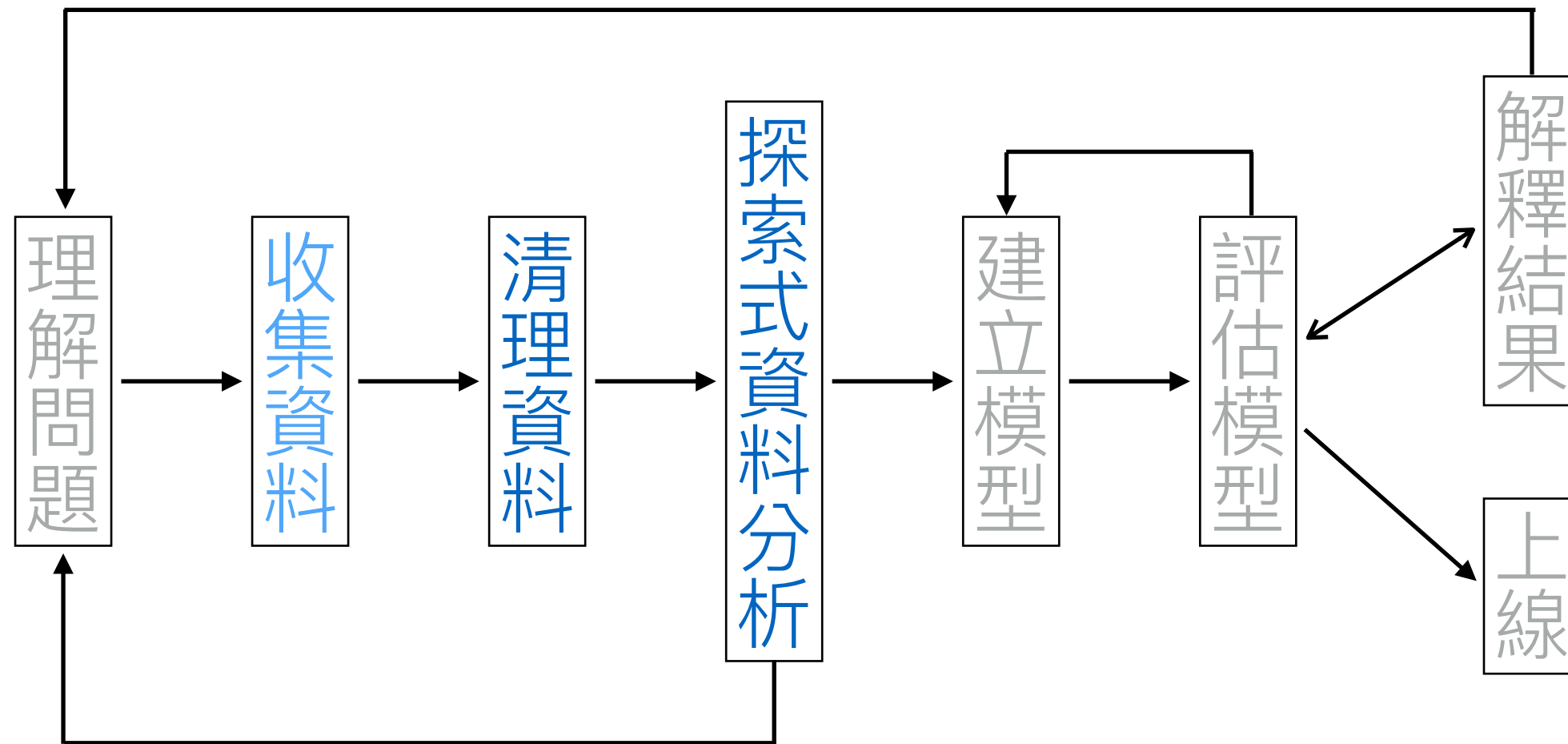


Pandas

Part I

Introduction to Pandas



Introduction to Pandas

- Input/Output (IO) data
- Prepare data
- Summarize data

Example: airbox data

1. Download data (zip format)
2. Read a csv file by pandas.read_csv
3. Convert a long table to a data matrix (wide table) by pandas.pivot_table (or pandas.DataFrame.pivot)

Read a csv file

```
In [1]: import pandas as pd
```

```
In [2]: filenm = 'iis_airbox_20200912.csv'  
df = pd.read_csv(filenm)  
df.head()
```

```
Out [2]:
```

	device_id	SiteName	PM25	timestamp
0	08BEAC0A0756	雲林縣縣立元長國小	15	2020-09-12 00:00:00
1	08BEAC09FFF4	臺中市立六寶國小	23	2020-09-12 00:00:00
2	08BEAC0A0404	臺中市立樂業國小	24	2020-09-12 00:00:00
3	74DA38F20DC4	高雄市市立後勁國小	9	2020-09-12 00:00:00
4	74DA38F20E78	苗栗縣縣立景山國小	24	2020-09-12 00:00:00



Prepare the data matrix

Long table轉wide table

```
In [3]: dm =pd.pivot_table(df, values='PM25', index='SiteName', columns='timestamp')
```

```
In [ ]: dm.to_csv('datamatrix.csv', na_rep='Null')
```

Long table vs wide table

Inspecting 'Python 3'				
	NAME	TYPE	SIZE	SHAPE
	df	DataFrame	13003296	406349 rows x 4 cols
	dm	DataFrame	2325906128	3458 rows x 84076 cols