Python 程式教學-chapter 27

- NBA資料抓取分析
- 使用get+params
- 使用datetime套件

陳岳洋(Jerry老師)

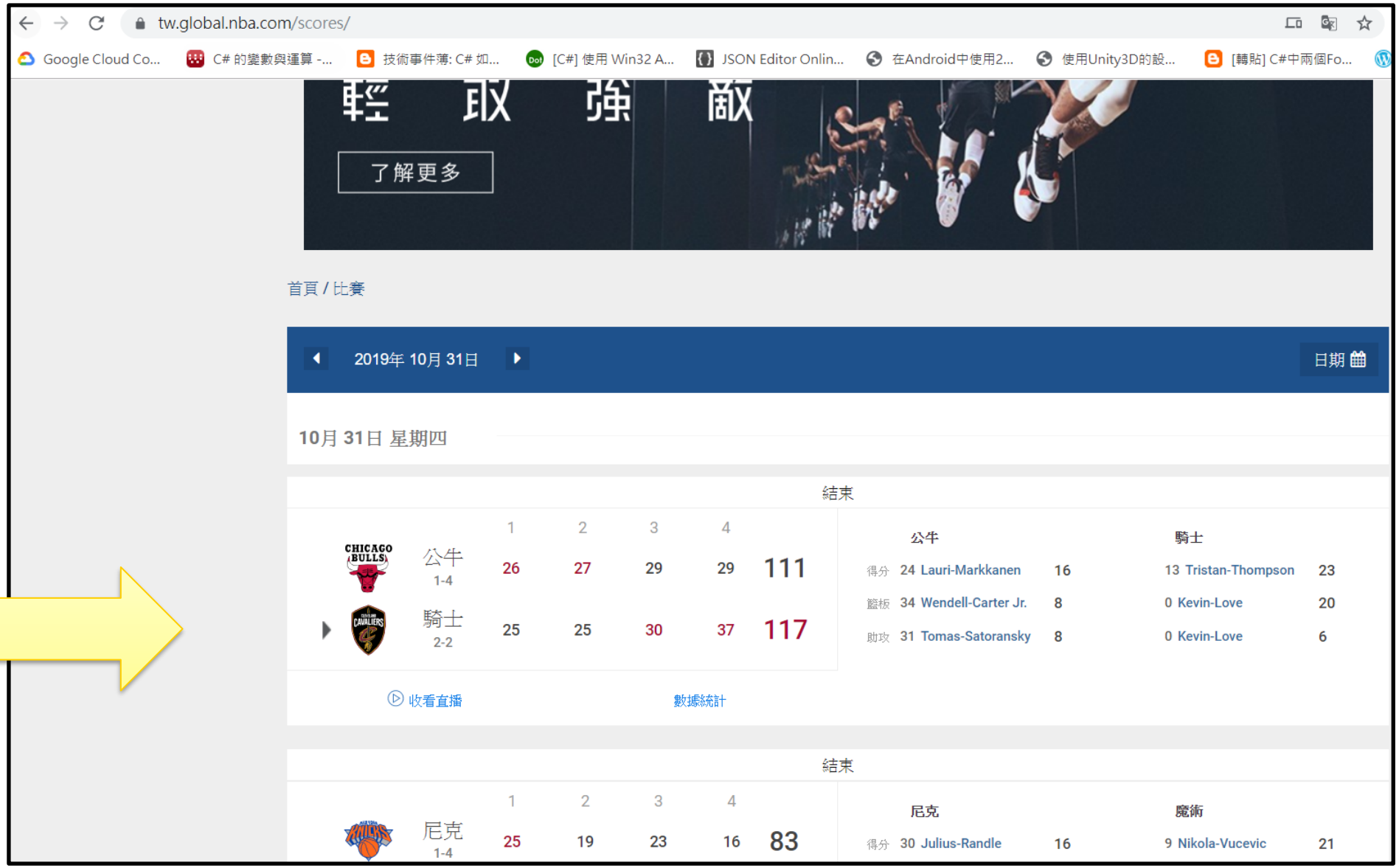
分析每局總和的勝率

籃球	足球	網球
不讓分	不讓分	不讓分
讓分	讓分	讓分
大小	大小	大小
單雙	單雙	單雙
單節不讓分	正確比數	單盤 不讓分
勝分差	總進球數	正確比數
單節大小	雙勝	
主客隊大小	第一球/下一球	
半場不讓分	主客隊大小	
半場讓分	半場不讓分	
半場大小	半場雙勝	
單節讓分	半全場	
得分較高半場	得分較高半場	
冠軍及特別項目	半場大小	
	兩隊是否都進球	
	半場正確比數	
	正確進球數	

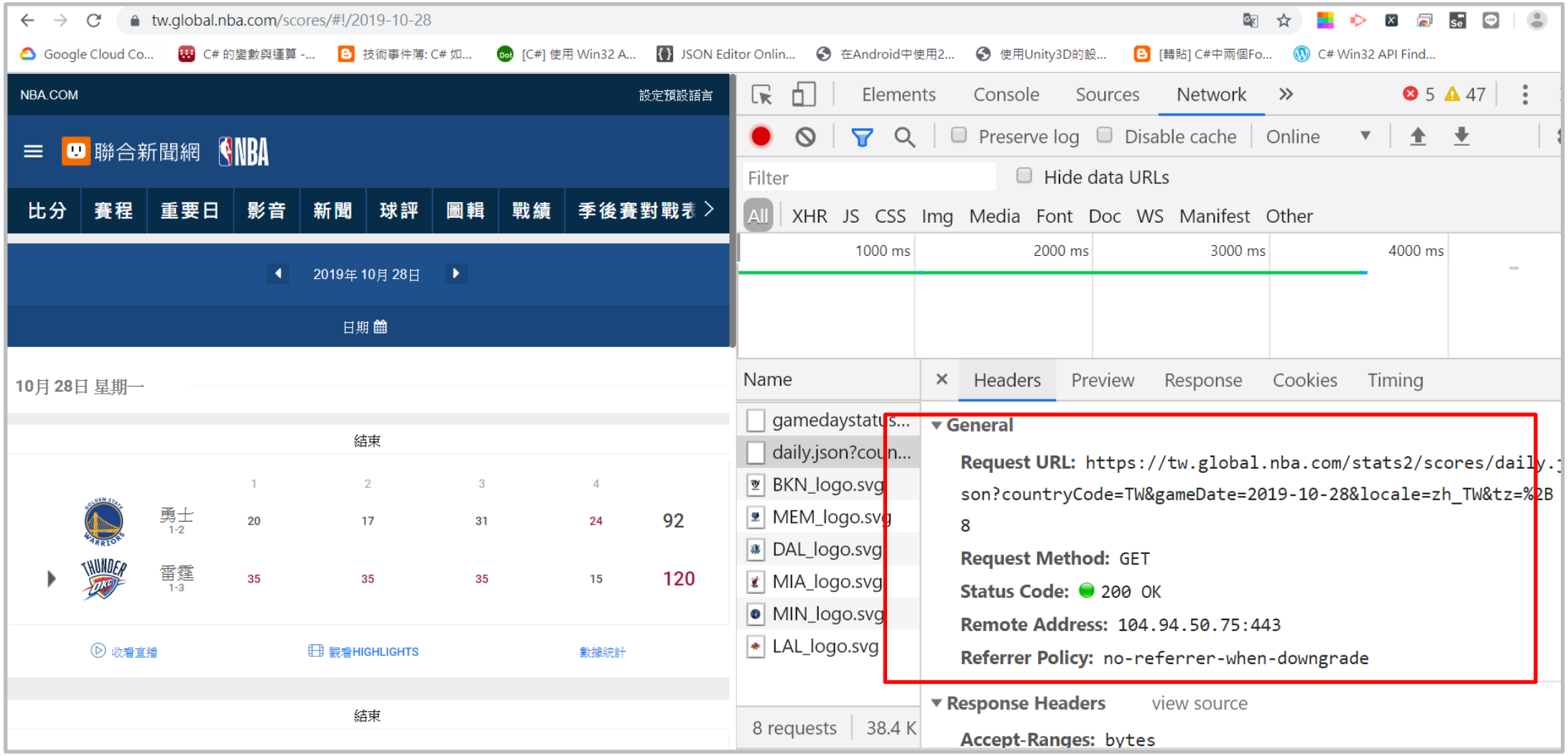
統計單雙機率

- https://tw.global.nba.com/scores/

每日每場資料



重點觀察資訊來源以及不同日期的取用方式



api+get(params)

- api➔

https://tw.global.nba.com/stats2/scores/daily.json?countryCode=TW&gameDate=2019-10-28&locale=zh_TW&tz=%2B8

- Params➔

Query String Parameters

countryCode: TW

gameDate: 2019-10-28

locale: zh_TW

tz: +8

使用 json套件

```
import json
resp=requests.get(api_url,params=from_data)

jsondata=json.loads(resp.text)
```

將 json轉成python
資料型態

```python
import requests
import json


api_url='https://tw.global.nba.com/stats2/scores/daily.json'

from_data={
    'countryCode': 'TW',
    'gameDate': '2019-10-28',
    'locale': 'zh_TW',
    'tz': '+8',
}

resp=requests.get(api_url,params=from_data)

if resp.status_code==200:
    jsondata=json.loads(resp.text)

    print(jsondata)
```


```
{'context': {'user': {'countryCode': '--', 'countryName': 'None', 'locale':
'zh_TW', 'timeZone': '+08:00', 'timeZoneCity': 'Australia/Perth'}, 'device':
{'clazz': None}}, 'error': {'detail': None, 'isError': 'false', 'message':
None}, 'payload': {'league': {'id': '00', 'name': 'NBA'}, 'season':
{'isCurrent': 'true', 'rosterSeasonType': 2, 'rosterSeasonYear': '2019',
'rosterSeasonYearDisplay': '2019-2020', 'scheduleSeasonType': 2,
'scheduleSeasonYear': '2019', 'scheduleYearDisplay': '2019-2020',
'statsSeasonType': 2, 'statsSeasonYear': '2019', 'statsSeasonYearDisplay':
'2019-2020', 'year': '2019', 'yearDisplay': '2019-2020'}, 'date': {'games':
[{'profile': {'arenaLocation': 'Cleveland, OH', 'arenaName': 'Rocket Mortgage
FieldHouse', 'awayTeamId': '1610612741', 'dateTimeEt': '2019-10-30T19:00',
'gameId': '0021900055', 'homeTeamId': '1610612739', 'number': '55',
'scheduleCode': None, 'seasonType': '2', 'sequence': '1', 'utcMillis':
'1572476400000'}, 'boxscore': {'attendance': '17,595', 'awayScore': 111,
'gameLength': '140', 'homeScore': 117, 'officialsDisplayName1': 'Rodney Mott',
'officialsDisplayName2': 'Derrick Collins', 'officialsDisplayName3': 'Phenizee
Ransom', 'period': '4', 'periodClock': '', 'status': '3', 'statusDesc': '結束',
'ties': '10'}, 'urls': [{'displayText': '收看直播', 'type': 'Live', 'value':
'http://watch.nba.com/packages?ls=leaguepass:localsite:tw:sib'}],
'broadcasters': [], 'homeTeam': {'profile': {'abbr': 'CLE', 'city': '克里夫蘭',
```


ch27_1.py

將respone資料進行解讀

將resp. text
進行解析

JSON Editor Online

New | Open ▾ | Save ▾ | Settings

powered by ace


```
公斤"},"statTotal":{"assists":3,"blocks":3,"defRebs":12,"fga":19
,"fgm":10,"fgpct":52.6,"fouls":3,"fta":6,"ftm":6,"ftpct":100.0
,"mins":32,"offRebs":2,"points":29,"rebs":14,"secs":29,"steals":1
,"tpa":5,"tpm":3,"tppct":60.0,"turnovers":5}}},"awayTeam"
:{"profile":{"abbr":"CHA","city":"夏洛特","cityEn":"Charlotte"
,"code":"hornets","conference":"Eastern","displayAbbr":"黃蜂"
,"displayConference":"東區","division":"東南組","id":"1610612766"
,"isAllStarTeam":false,"isLeagueTeam":true,"leagueId":"00","name"
:"黃蜂","nameEn":"Hornets"},"matchup":{"confRank":"12","divRank"
:"5","losses":"3","seriesText":null,"wins":"1"},"score":{"assists"
:19,"biggestLead":9,"blocks":4,"blocksAgainst":8,"defRebs":32
,"disqualifications":0,"ejections":0,"fastBreakPoints":26,"fga":91
,"fgm":35,"fgpct":38.5,"flagrantFouls":0,"fouls":20,"fta":23,"ftm"
:21,"ftpct":91.3,"fullTimeoutsRemaining":2,"mins":240,"offRebs":13
,"ot10Score":0,"ot1Score":0,"ot2Score":0,"ot3Score":0,"ot4Score":0
,"ot5Score":0,"ot6Score":0,"ot7Score":0,"ot8Score":0,"ot9Score":0
,"pointsInPaint":44,"pointsOffTurnovers":9,"q1Score":28,"q2Score"
:34,"q3Score":18,"q4Score":21,"rebs":45,"score":101,"seconds":0
,"shortTimeoutsRemaining":0,"steals":7,"technicalFouls":0,"tpa":35
,"tpm":10,"tppct":28.6,"turnovers":16},"pointGameLeader":{"profile"
:{"code":"miles_bridges","country":"美國","countryEn":"United
States","displayAffiliation":"Michigan State/United States"
,"displayName":"Miles Bridges","displayNameEn":"Miles Bridges"
,"dob":"890456400000","draftYear":"2018","experience":"1"
,"firstInitial":"Miles","firstName":"Miles","firstNameEn":"Miles"
,"height":"1.98","jerseyNo":"0","lastName":"Bridges","lastNameEn"
:"Bridges","leagueId":"00","playerId":"1628970","position":"前鋒"
,"schoolType":"College","weight":"102.1 公斤"},"statTotal"
:{"assists":3,"blocks":0,"defRebs":4,"fga":17,"fgm":8,"fgpct":47.1
,"fouls":3,"fta":5,"ftm":5,"ftpct":100.0,"mins":37,"offRebs":2
,"points":23,"rebs":6,"secs":15,"steals":0,"tpa":4,"tpm":2,"tppct"
:50.0,"turnovers":3}},"assistGameLeader":{"profile":{"code"
:"devonte_graham","country":"美國","countryEn":"United States"
```


Ln: 1 Col: 33485

Select a node...

- object {4}
  - context {2}
  - error {3}
  - payload {5}
  - timestamp : 1572487591050

另一種讀取方式

```
import requests
import json


api_url='''https://tw.global.nba.com/stats2/scores/daily.json?
countryCode=TW&gameDate={}&locale=zh_TW&tz=%2B8'''


#
#from_data={
#     'countryCode': 'TW',
#     'gameDate': '2019-10-28',
#     'locale': 'zh_TW',
#     'tz': '+8',
#}

resp=requests.get(api_url.format('2019-10-28'))

if resp.status_code==200:
    jsondata=json.loads(resp.text)

    print(jsondata)
```

ch27_2.py

找出比分資料

- payload {5}
  - league {2}
  - season {12}
  - date {3}
    - games [5]
      - 0 {6}
        - profile {11}
        - boxscore {12}
        - urls [2]
        - broadcasters [0]
        - homeTeam {6}
        - awayTeam {6}
          - profile {14}
          - matchup {5}
          - score {45}
          - pointGameLeader {2}
          - assistGameLeader {2}
          - reboundGameLeader {2}

- homeTeam {6}
  - profile {14}
  - matchup {5}
  - score {45}
  - pointGameLeader {2}
  - assistGameLeader {2}
  - reboundGameLeader {2}
- awayTeam {6}
  - profile {14}
  - matchup {5}
  - score {45}
  - pointGameLeader {2}
  - assistGameLeader {2}
  - reboundGameLeader {2}

```
▼ homeTeam {6}
  ▶ profile {14}
  ▶ matchup {5}
  ▶ score {45}
  ▶ pointGameLeader {2}
  ▶ assistGameLeader {2}
  ▶ reboundGameLeader {2}
▼ awayTeam {6}
  ▶ profile {14}
  ▶ matchup {5}
  ▶ score {45}
  ▶ pointGameLeader {2}
  ▶ assistGameLeader {2}
  ▶ reboundGameLeader {2}
```

```
q1Score : 35
q2Score : 35
q3Score : 35
q4Score : 15
rebs : 50
score : 120
seconds : 0
shortTimeoutsRemaining : 0
steals : 8
technicalFouls : 0
tpa  : 32
tpm  : 15
tppct : 46.9
turnovers : 24
```

故資料區段在playload➔date➔games➔homeTeam➔score

- 抓取主資料區段

```
game_data=jsondata['payload']['date']['games']

for game in game_data:
    print(game)
```

- ch27_3. py

```
game_data=jsondata['payload']['date']['games']


for game in game_data:

    #print(game)
    home_team=game['homeTeam']
    print(home_team['profile']['name'])
    print(home_team['score']['q1Score'],home_team['score']['q2Score'],
          home_team['score']['q3Score'],home_team['score']['q4Score'],
         home_team['score']['score'])


    away_team=game['awayTeam']

    print(away_team['profile']['name'])
    print(away_team['score']['q1Score'],away_team['score']['q2Score'],
          away_team['score']['q3Score'],away_team['score']['q4Score'],
         away_team['score']['score'])
```

```
雷霆
35 35 35 15 120
勇士
20 17 31 24 92
```

```python
import requests
import json


api_url='https://tw.global.nba.com/stats2/scores/daily.json'

form_data={
    'countryCode': 'TW',
    'gameDate': '2019-10-28',
    'locale': 'zh_TW',
    'tz': '+8',
}

resp=requests.get(api_url,params=form_data)

if resp.status_code==200:
    jsondata=json.loads(resp.text)

    print(jsondata)


game_data=jsondata['payload']['date']['games']


for game in game_data:

    #print(game)
    home_team=game['homeTeam']
    print(home_team['profile']['name'])
    print(home_team['score']['q1Score'],home_team['score']['q2Score'],
          home_team['score']['q3Score'],home_team['score']['q4Score'],
         home_team['score']['score'])


    away_team=game['awayTeam']

    print(away_team['profile']['name'])
    print(away_team['score']['q1Score'],away_team['score']['q2Score'],
          away_team['score']['q3Score'],away_team['score']['q4Score'],
         away_team['score']['score'])
```

```
2019-10-28
**************************************************
雷霆
35 35 35 15 120
勇士
20 17 31 24 92
**************************************************
灰熊
35 24 33 28 134
籃網
29 28 40 23 133
**************************************************
獨行俠
40 31 24 24 119
拓荒者
25 34 36 26 121
**************************************************
灰狼
36 19 22 39 116
熱火
23 34 27 25 109
**************************************************
```

程式練習

- 任意填入日期進行測試
- 輸入2011-01-01
  - 沒有比賽得處理方式
- 輸入2013-02-23
  - 增加延長賽紀錄["ot1Score"]

ch27_5.py

- 將抓取資料區段整理成函式
  - 傳入要抓取的日期
  - 每個日期合併成一個二維串列包含日期/隊伍名稱/每局比分 進行資料回傳

```
[['2013-02-23', '暴龍', 32, 16, 31, 21], ['2013-02-23', '尼克', 32, 19, 19, 28],
['2013-02-23', '巫師', 33, 31, 30, 25], ['2013-02-23', '金塊', 28, 36, 14, 35],
['2013-02-23', '山貓', 15, 22, 20, 18], ['2013-02-23', '公牛', 24, 27, 27, 27],
['2013-02-23', '溜馬', 34, 26, 30, 24], ['2013-02-23', '活塞', 17, 19, 16, 30],
['2013-02-23', '籃網', 31, 22, 22, 21], ['2013-02-23', '火箭', 34, 27, 23, 22],
['2013-02-23', '老鷹', 25, 36, 32, 29], ['2013-02-23', '國王', 33, 30, 25, 20],
['2013-02-23', '灰熊', 22, 21, 24, 21], ['2013-02-23', '魔術', 21, 17, 15, 29],
['2013-02-23', '黃蜂', 29, 29, 21, 21], ['2013-02-23', '獨行俠', 35, 18, 30, 21],
['2013-02-23', '雷霆', 33, 33, 34, 27], ['2013-02-23', '灰狼', 23, 35, 25, 28],
['2013-02-23', '太陽', 20, 21, 26, 21], ['2013-02-23', '塞爾蒂克', 28, 25, 28, 32],
['2013-02-23', '湖人', 30, 23, 27, 31], ['2013-02-23', '拓荒者', 28, 27, 27, 25],
['2013-02-23', '勇士', 25, 13, 29, 26], ['2013-02-23', '馬刺', 25, 14, 29, 25]]
```

將每天資料存成二維串列

```
list_data=[]
for game in game_data:

    home_team=game['homeTeam']
    away_team=game['awayTeam']

    list_data.append([date,home_team['profile']['name'],home_team['score']['q1Score'],home_team['score']['q2Score']
    ,home_team['score']['q3Score'],home_team['score']['q4Score']])

    list_data.append([date,away_team['profile']['name'],away_team['score']['q1Score'],away_team['score']['q2Score']
    ,away_team['score']['q3Score'],away_team['score']['q4Score']])
```

ch27_6.py

- 使用datetime套件

讓程式自己進行每日日期抓取

2011-12-26➔2011-12-27➔2011-12-28

```
import datetime


today=datetime.date.today()
print(today)
print(today.strftime("%Y-%m-%d"))

#取得昨天日期
yday = today+datetime.timedelta(days=-1)
#取得明天日期
tday = today+datetime.timedelta(days=1)

print(yday.strftime("%Y-%m-%d"),today.strftime("%Y-%m-%d"),
      tday.strftime("%Y-%m-%d"))
```

```
2019-10-31
2019-10-31
2019-10-30 2019-10-31 2019-11-01
```

ch27_7.py

設定起始日期跟終止日期

```
start_day=datetime.datetime(2019,10,1)

end_day=datetime.datetime(2019,10,31)

print(start_day,end_day)

while start_day<end_day:

    start_day+=datetime.timedelta(days=1)
    print(start_day.strftime("%Y-%m-%d"))
```

```
2019-10-02
2019-10-03
2019-10-04
2019-10-05
2019-10-06
2019-10-07
2019-10-08
2019-10-09
2019-10-10
2019-10-11
2019-10-12
2019-10-13
2019-10-14
2019-10-15
2019-10-16
2019-10-17
2019-10-18
2019-10-19
2019-10-20
2019-10-21
2019-10-22
2019-10-23
```

整合成連續抓取

```
import time

api_url='https://tw.global.nba.com/stats2/scores/daily.json'

start_day=datetime.datetime(2019,10,1)
end_day=datetime.datetime(2019,10,5)

list_data=[]

while start_day<end_day:
    print('擷取日期:%s'%(start_day.strftime("%Y-%m-%d")))
    list_data+=getNBAData(api_url,start_day.strftime("%Y-%m-%d"))
    start_day+=datetime.timedelta(days=1)
    time.sleep(1)
```

```
print(list_data)
```

```
擷取日期:2019-10-01
擷取日期:2019-10-02
擷取日期:2019-10-03
擷取日期:2019-10-04
[['2019-10-01', '火箭', 36, 41, 34, 29], ['2019-10-01', 'Sharks', 14, 20, 17, 20],
['2019-10-04', '快艇', 22, 18, 25, 31], ['2019-10-04', '火箭', 31, 26, 32, 20],
['2019-10-04', '國王', 39, 33, 25, 21], ['2019-10-04', '溜馬', 29, 30, 33, 26]]
```

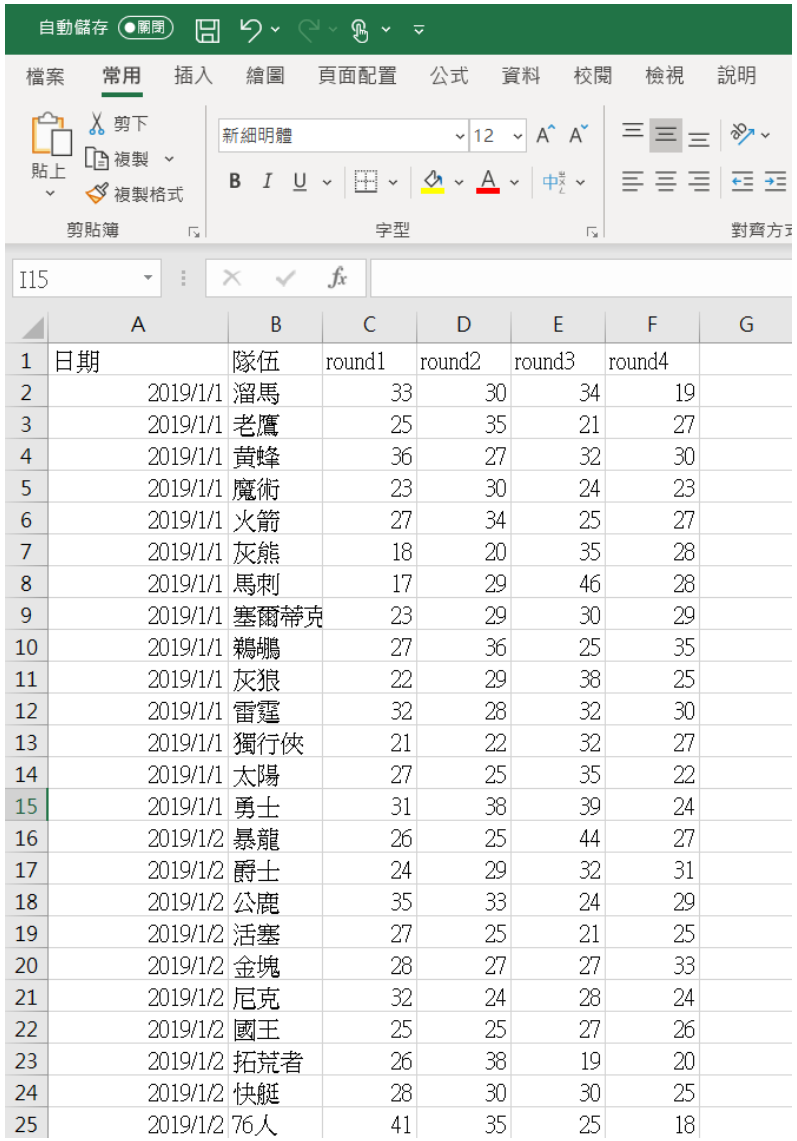
ch27_8.py

儲存成csv

```
# ==========================================================================
# 儲存csv
# ==========================================================================
import csv

today=datetime.datetime.today()
with open('nba_{}.csv'.format(today.strftime("%Y-%m-%d")),'w',newline='',
          encoding='utf-8-sig') as f:
    writer=csv.writer(f)

    writer.writerow(['日期','隊伍','round1','round2','round3','round4'])
    writer.writerows(list_data)
```



	A	B	C	D	E	F
1	日期	隊伍	round1	round2	round3	round4
2	2019/1/1	溜馬	33	30	34	19
3	2019/1/1	老鷹	25	35	21	27
4	2019/1/1	黃蜂	36	27	32	30
5	2019/1/1	魔術	23	30	24	23
6	2019/1/1	火箭	27	34	25	27
7	2019/1/1	灰熊	18	20	35	28
8	2019/1/1	馬刺	17	29	46	28
9	2019/1/1	塞爾蒂克	23	29	30	29
10	2019/1/1	鵜鶘	27	36	25	35
11	2019/1/1	灰狼	22	29	38	25
12	2019/1/1	雷霆	32	28	32	30
13	2019/1/1	獨行俠	21	22	32	27
14	2019/1/1	太陽	27	25	35	22
15	2019/1/1	勇士	31	38	39	24
16	2019/1/2	暴龍	26	25	44	27
17	2019/1/2	爵士	24	29	32	31
18	2019/1/2	公鹿	35	33	24	29
19	2019/1/2	活塞	27	25	21	25
20	2019/1/2	金塊	28	27	27	33
21	2019/1/2	尼克	32	24	28	24
22	2019/1/2	國王	25	25	27	26
23	2019/1/2	拓荒者	26	38	19	20
24	2019/1/2	快艇	28	30	30	25
25	2019/1/2	76人	41	35	25	18

ch27_9.py

統計資訊

```
round1
總場數:906
該局總和分數奇數:
該局總和分數偶數:
round2
總場數:906
該局總和分數奇數:
該局總和分數偶數:
round3
總場數:906
該局總和分數奇數:
該局總和分數偶數:
round4
總場數:906
該局總和分數奇數:
該局總和分數偶數:
```

出單還是出雙機率大?

ch27_10.py