

AUTOMATIC DATA SCRAPING BY BEAUTIFUL SOUP



Pic by : Aren Vandenburg

Who am I ?

- Sophomore at Computer Science, CU
- Want to talk with me?
 - ML
 - Quantum
 - Python
 - Node.js (React & Express)
 - Cat
 - Investing



“การลงทุนมีความเสี่ยง
ผู้ลงทุนโปรดศึกษาข้อมูลก่อนการลงทุน”

“หุ้นที่มีใน slide เป็นการสุ่ม และ
ไม่ได้ชี้นำการลงทุนแต่อย่างใด”





เกี่ยวกับ ตลาด. | กฎเกณฑ์/การกำกับ | สินค้า/บริการ | ข้อมูลการซื้อขาย | ข้อมูลบริษัท/หลักทรัพย์ | หน่วยงาน | ความรู้การลงทุน | ข่าว/กิจกรรม

Get Quote



หาซื้อ요

หน้าแรก > ข้อมูลบริษัท/หลักทรัพย์ > ตลาดสารทุน > ข้อมูลรายบริษัท/หลักทรัพย์

E-mail or Username

.....

ลงชื่อเข้าใช้

 Log In

สมัครสมาชิก | ลืมรหัสผ่าน
☒ ให้ฉันอยู่ในระบบ

ข้อมูลบริษัท/หลักทรัพย์

ข่าวบริษัท/หลักทรัพย์ ▾

ตลาดสารทุน ^

ข้อมูลรายบริษัท/หลักทรัพย์

ปฏิทินหลักทรัพย์

หุ้นไอพีโอ

บริษัท/หลักทรัพย์จดทะเบียน

เข้าใหม่

รายชื่อบริษัท/หลักทรัพย์

เลวรหัสหลักทรัพย์สากล

รายชื่อบริษัทที่รายงานว่าได้รับ

สัมปทานจากรัฐ

ตลาดสารหนี้ ▾

Opportunity Day &

Company Highlights ▾

C-Update

ข้อมูลรายบริษัท/หลักทรัพย์

ค้นหาบริษัท / หลักทรัพย์จดทะเบียน

สรุปข้อสังเกตบริษัทจดทะเบียน

บริษัท/หลักทรัพย์

งบการเงิน/ผลประกอบการ

ผู้ถือหุ้นรายใหญ่

ข้อมูลสิทธิประโยชน์

ข่าว

ราคาวันนี้

ราคาย้อนหลัง

ผลประกอบการสำคัญ (ตาราง)

ผลประกอบการสำคัญ (กราฟ)

งบการเงินล่าสุด

(หน่วย: ล้านบาท)

งวดงบการเงิน ณ วันที่	งบปี 58 31/12/2558	งบปี 59 31/12/2559	งบปี 60 31/12/2560	งบปี 61 31/12/2561	ไตรมาส1/62 31/03/2562
บัญชีทางการเงินที่สำคัญ					
สินทรัพย์รวม	21,297.52	21,329.86	23,483.96	24,748.81	26,029.62
หนี้สินรวม	8,197.03	6,541.92	6,779.00	6,004.38	6,217.23
ส่วนของผู้ถือหุ้น	12,735.69	14,473.85	16,402.38	18,454.61	19,522.67
มูลค่าหุ้นที่เรียกชำระแล้ว	730.05	730.05	730.05	730.05	730.05
รายได้รวม	17,941.96	18,128.01	18,530.13	18,541.03	4,733.02
กำไรสุทธิ	3,435.83	3,626.17	3,943.89	4,151.89	1,081.40
กำไรต่อหุ้น (บาท)	4.72	4.98	5.41	5.70	1.48
อัตราส่วนทางการเงินที่สำคัญ					
ROA(%)	21.94	21.55	22.01	21.57	20.25

Related Info.

- ราคาดัชนีตลาดหลักทรัพย์
- ราคาหลักทรัพย์
- สรุปฐานะการเงิน ผลการดำเนินงานรายอุตสาหกรรม
- รายชื่อบริษัทสมาชิกตลาดหลักทรัพย์แห่งประเทศไทย
- บริการข้อมูลหลักทรัพย์

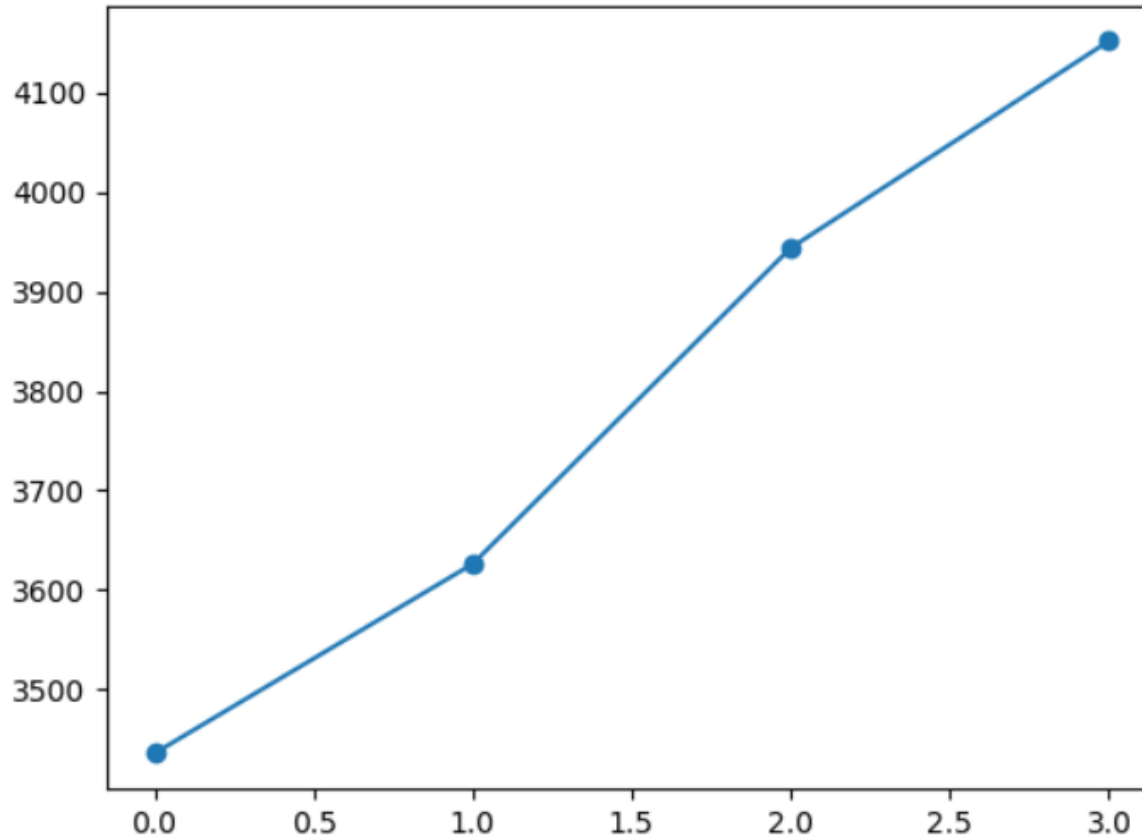
- ปฏิทินการจัดประชุมผู้ถือหุ้น
- คำศัพท์น่ารู้
- เว็บไซต์ ก.ล.ด.


www.settrade.com

ข้อมูลบริษัท/หลักทรัพย์

[จำนวนบริษัท/หลักทรัพย์](#)
[ตราสารทุน](#)
[ข้อมูลรายบริษัท/หลักทรัพย์](#)
[ปฏิทินหลักทรัพย์](#)
[หุ้นไอพีโอ](#)
[บริษัท/หลักทรัพย์จดทะเบียน](#)
[เข้าใหม่](#)
[รายชื่อบริษัท/หลักทรัพย์](#)
[เลวรหัสหลักทรัพย์สากล](#)
[รายชื่อบริษัทที่รายงานว่าได้รับ](#)
[สัมปทานจากรัฐ](#)
[ตราสารหนี้](#)
[Opportunity Day &](#)
[Company Highlights](#)
[C-Update](#)

NPM



Related Info.

- ราคาดัชนีตลาดหลักทรัพย์
- ราคาหลักทรัพย์
- สรุปฐานะการเงิน ผลการดำเนินงานรายไตรมาส
- รายชื่อบริษัทสมาชิกตลาดหลักทรัพย์แห่งประเทศไทย
- บริการข้อมูลหลักทรัพย์

- ปฏิทินการจัดประชุมผู้ถือหุ้น
- คำศัพท์น่ารู้
- เว็บไซต์ ก.ล.ด.

www.settrade.com

กำไรสุทธิ	3,435.83	3,626.17	3,943.89	4,151.89	1,081.40
-----------	----------	----------	----------	----------	----------

กำไรต่อหุ้น (บาท)	4.72	4.98	5.41	5.70	1.48
-------------------	------	------	------	------	------

อัตราส่วนทางการเงินที่สำคัญ

ROA(%)	21.94	21.55	22.01	21.57	20.25
--------	-------	-------	-------	-------	-------

MORE THAN 700 STOCKS



รายได้รวม	17,941.96	18,128.01	18,530.13	18,541.03	4,733.02
กำไรสุทธิ	3,435.83	3,626.17	3,943.89	4,151.89	1,081.40
กำไรต่อหุ้น (บาท)	4.72	4.98	5.41	5.70	1.48

```

<tr >
  <td style="text-align:left;" height="15">รายได้รวม</td>
    <td >17,941.96&nbsp;&nbsp;&nbsp;</td>
    <td >18,128.01&nbsp;&nbsp;&nbsp;</td>
    <td >18,530.13&nbsp;&nbsp;&nbsp;</td>
    <td >18,541.03&nbsp;&nbsp;&nbsp;</td>
    <td style="background-color: #EAF0FE;" >4,733.02&nbsp;&nbsp;&nbsp;</td>
</tr>
<tr >
  <td style="text-align:left;" height="15">กำไรสุทธิ</td>
    <td >3,435.83&nbsp;&nbsp;&nbsp;</td>
    <td >3,626.17&nbsp;&nbsp;&nbsp;</td>
    <td >3,943.89&nbsp;&nbsp;&nbsp;</td>
    <td >4,151.89&nbsp;&nbsp;&nbsp;</td>
    <td style="background-color: #EAF0FE;" >1,081.40&nbsp;&nbsp;&nbsp;</td>
</tr>
<tr >
  <td style="text-align:left;" height="15">กำไรต่อหุ้น (บาท)</td>
    <td >4.72&nbsp;&nbsp;&nbsp;</td>
    <td >4.98&nbsp;&nbsp;&nbsp;</td>
    <td >5.41&nbsp;&nbsp;&nbsp;</td>
    <td >5.70&nbsp;&nbsp;&nbsp;</td>
    <td style="background-color: #EAF0FE;" >1.48&nbsp;&nbsp;&nbsp;</td>

```

BEAUTIFULSOUP

WEB
PAGE



SOURCE
CODE



OBJ.



REQUESTS

PARSE

```
pip install beautifulsoup4
pip install requests
```



Nakarin S.

`soup.findAll('tag')`

Eg: `soup.findAll('ul')`

```
<div class="container">
    <ul>
        <li>text 1</li>
        <li>text 2</li>
        <li>text 3</li>
        <li>text 4</li>
    </ul>
</div>
```

```
<p class="foo" id="bar">  
    zoo  
</p>
```

```
soup.findAll('p')
```

```
soup.findAll(class_='foo')
```

```
soup.findAll(id='bar')
```

```
soup.findAll(string='zoo')
```

.parent

Eg: `soup.find('ul').parent`

```
<div class="container">
  <ul>
    <li>text 1</li>
    <li>text 2</li>
    <li>text 3</li>
    <li>text 4</li>
  </ul>
</div>
```

Going up

.contents

Eg: `soup.find('ul').contents`

```
<div class="container">  
  <ul>  
    <li>text 1</li>  
    <li>text 2</li>  
    <li>text 3</li>  
    <li>text 4</li>  
  </ul>  
</div>
```

Going down

.string

Eg: `soup.find('li').string`

```
<div class="container">  
  <ul>  
    <li>text 1</li>  
    <li>text 2</li>  
    <li>text 3</li>  
    <li>text 4</li>  
  </ul>  
</div>
```

Going down

`.next_sibling` and `.previous_sibling`

Eg: `soup.find('li').next_sibling`

```
<div class="container">
  <ul>
    <li>text 1</li>
    <li>text 2</li>
    <li>text 3</li>
    <li>text 4</li>
  </ul>
</div>
```

Going sideways

- text 1
- text 2
- text 3
- text 4

```
div class="container">  
  <ul>  
    <li>text 1</li>  
    <li>text 2</li>  
    <li>text 3</li>  
    <li>text 4</li>  
  </ul>  
</div>
```

```
import requests
from bs4 import BeautifulSoup

url =
'http://128.0.0.1:5500/foo_1.html'
text = requests.get(url).text
soup = BeautifulSoup(text, 'lxml')
li_s = (soup.findAll('li'))
print(li_s)
for li in li_s:
    print(li.string)
```



```
import requests
from bs4 import BeautifulSoup

url =
'http://128.0.0.1:5500/foo_1.html'
text = requests.get(url).text
soup = BeautifulSoup(text, 'lxml')
li_s = (soup.findAll('li'))
print(li_s)
for li in li_s:
    print(li.string)
```

Output :

```
<!DOCTYPE html><html lang="en"> <head> <meta charset="UTF-8" /> <meta name="viewport"
content="width=device-width, initial-scale=1.0" /> <meta http-equiv="X-UA-Compatible"
content="ie=edge" /> <title>Document</title> <link rel="stylesheet"
href="https://stackpath.bootstrapcdn.com/bootstrap/4.3.1/css/bootstrap.min.css"
integrity="sha384-....."
```

```
import requests
from bs4 import BeautifulSoup

url =
'http://128.0.0.1:5500/foo_1.html'
text = requests.get(url).text
soup = BeautifulSoup(text, 'lxml')
li_s = (soup.findAll('li'))
print(li_s)
for li in li_s:
    print(li.string)
```

Output :

[text_1,text_2,text_3]

```
import requests
from bs4 import BeautifulSoup

url =
'http://128.0.0.1:5500/foo_1.html'
text = requests.get(url).text
soup = BeautifulSoup(text, 'lxml')
li_s = (soup.findAll('li'))
print(li_s)
for li in li_s:
    print(li.string)
```

Output :

Text_1

Text_2

Text_3

Nakarin S.

Application

- Data Mining
- Chatbot
- Report Generator
- Valuation Model

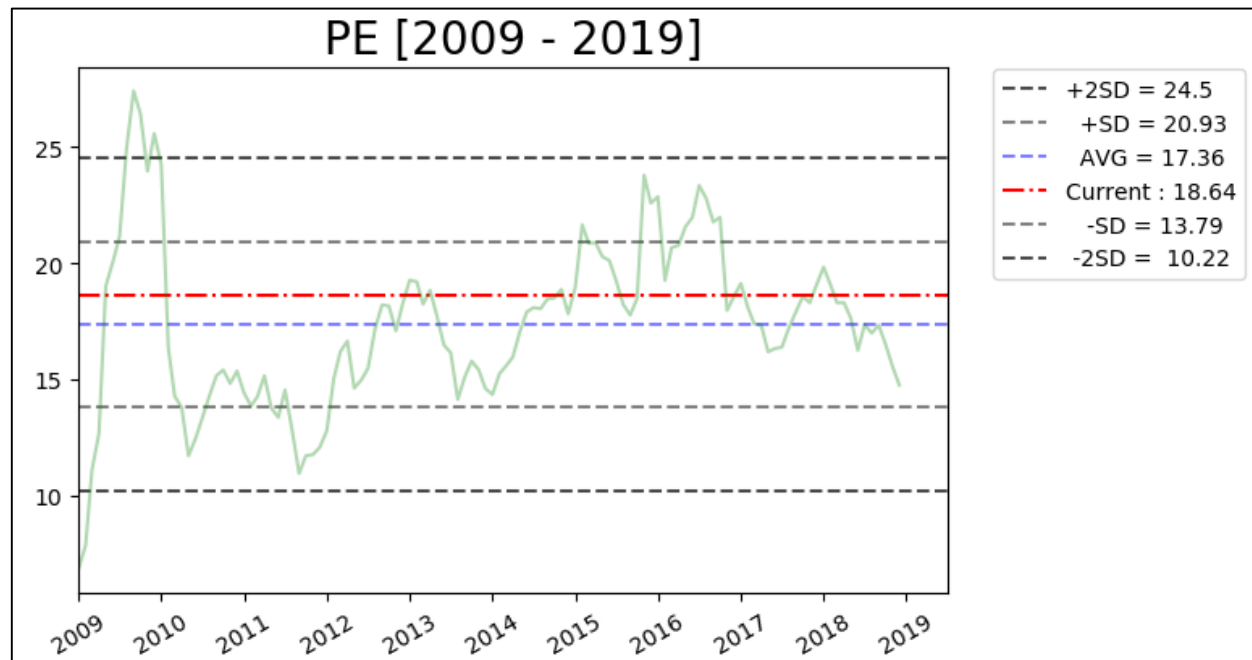
LINE BOT

Libraries

- BeautifulSoup
- Flask
- Requests
- Linebot-SDK



Market Valuation Report



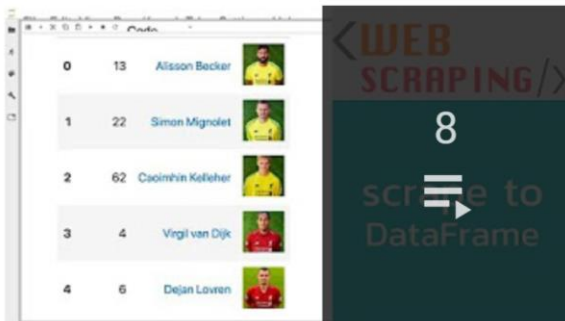
Libraries

- BeautifulSoup
- Requests
- Pandas
- Matplotlib

Reference

<https://www.crummy.com/software/BeautifulSoup/bs4/doc/>

Further Study



สอนทำ web scraping ด้วย Python

prasertcbs

[VIEW FULL PLAYLIST](#)

- PrasertCBS@Youtube
<https://bit.ly/2y2zH3L>
- <https://scrapy.org/>
 - My Recommended Playlist
<https://bit.ly/2Y3ggSZ>



Slide

Thank you <3

- GitHub: <https://github.com/nakarinstk>
- Email: [nakarin.chr98@outlook.com](mailto:nakarinstk@outlook.com)

