

Making Open Weather Data More Accessible: Extracting Seasonal Trends from Singapore Weather Station Data

Presented by:
Ong Chin Hwee (@ongchinhwee)
1 December 2019

About me

Ong Chin Hwee 王敬惠

- Data Engineer @ ST Engineering
- Background in aerospace engineering + computational modelling
- Loves (and contributes to) pandas
- Mentor team at BigDataX, one of the Top 10 Data Communities in Singapore



Singapore 新加坡:
1°17'22.81"N, 103°51'0.25"E
北纬1度, 经纬103度





A year is usually split into four
seasons...

常說一年划分为四季。。

春
Spring



春
Spring



夏
Summer

春
Spring



秋
Autumn



夏
Summer

@ongchinhwee

春

Spring



秋

Autumn



夏

Summer



冬

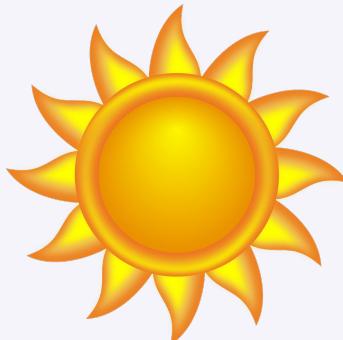
Winter



...how about my home country?
。。。新加坡有多少季節？

Are the “four seasons” just...

Are the “four seasons” just...



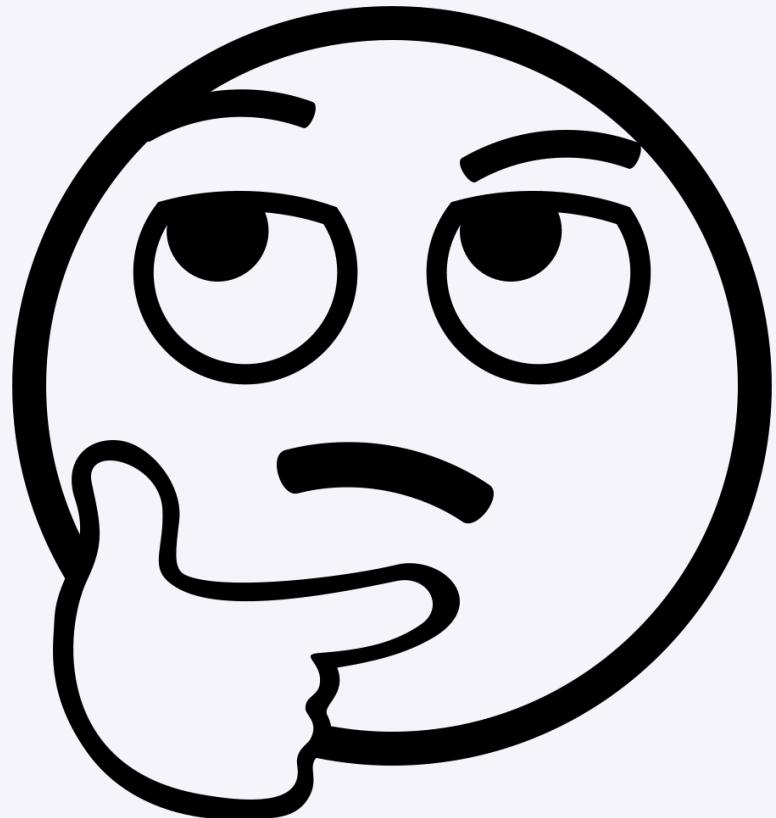
1. Hot
2. Very Hot
3. Extremely Hot
4. Rain?

熱

很熱

超熱

雨



Let's discover seasons through **data**.
讓我們以**數據**來發掘季節。

Extracting Weather Data



Realtime Weather Readings across Singapore

FILES IN THIS DATASET

Air Temperature across Singapore



< > Embed Chart

API View

Rainfall across Singapore



Relative Humidity across
Singapore



Wind Direction across Singapore



GET <https://api.data.gov.sg/v1/environment/air-temperature> Get air temperature readings across Singapore

- Has per-minute readings from NEA
- Use the `date_time` parameter to retrieve the latest available data at that moment in time
- Use the `date` parameter to retrieve all of the readings for that day.

Parameters

Try it out

Name	Description
People	

Data.gov.sg - Singapore's Open Data Portal

新加坡開源數據平台

Realtime Weather Readings across Singapore

Real-time API (Application Programming Interface) on
Data.gov.sg (Singapore's open data portal)

即時應用程式介面

Realtime Weather Readings across Singapore

Real-time API on Data.gov.sg (Singapore's open data portal)

Open government data available under the Singapore Open Data License

政府開放資料

Realtime Weather Readings across Singapore

Real-time API on Data.gov.sg (Singapore's open data portal)

Open government data available under the Singapore Open Data License

Minute-by-minute weather station readings

“Let’s try to scrap weather data for a
specific weather station!”

“Let’s try to scrap weather data for a
specific weather station!”

“How about we scrap month-long data from
the API?”

date_time YYYY-MM-DD[T]HH:mm:ss (SGT)
string
(query)

date YYYY-MM-DD
string
(query)

Execute

Clear

Responses

Curl

```
curl -X GET "https://api.data.gov.sg/v1/environment/air-temperature?date=2019-11-23" -H "accept: application/json"
```

Request URL

```
https://api.data.gov.sg/v1/environment/air-temperature?date=2019-11-23
```

Responses

Curl

```
curl -X GET "https://api.data.gov.sg/v1/environment/air-temperature?date=2019-11-23" -H "accept: application/json"
```

Request URL

```
https://api.data.gov.sg/v1/environment/air-temperature?date=2019-11-23
```

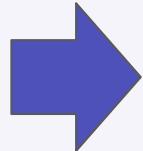
```
        "id": "S100",
        "device_id": "S100",
        "name": "Woodlands Road",
        "location": {
            "latitude": 1.4172,
            "longitude": 103.74855
        }
    },
    {
        "id": "S115",
        "device_id": "S115",
        "name": "Tuas South Avenue 3",
        "location": {
            "latitude": 1.29377,
            "longitude": 103.61843
        }
    }
],
"reading_type": "DBT 1M F",
"reading_unit": "deg C"
},
"items": [
{
    "timestamp": "2019-11-23T00:01:00+08:00",
    "readings": [
        {
            "station_id": "S109",
            "value": 25.2
        },
        {
            "station_id": "S117",
            "value": 26
        },
        {
            "station_id": "S115",
            "value": 26.5
        }
    ]
}
```

@ongchinhwee

```
        "id": "S100",
        "device_id": "S100",
        "name": "Woodlands Road",
        "location": {
            "latitude": 1.4172,
            "longitude": 103.74855
        }
    },
    {
        "id": "S115",
        "device_id": "S115",
        "name": "Tuas South Avenue 3",
        "location": {
            "latitude": 1.29377,
            "longitude": 103.61843
        }
    }
],
"reading_type": "DBT 1M F",
"reading_unit": "deg C"
},
"items": [
{
    "timestamp": "2019-11-23T00:01:00+08:00",
    "readings": [
        {
            "station_id": "S109",
            "value": 25.2
        },
        {
            "station_id": "S117",
            "value": 26
        },
        {
            "station_id": "S118",
            "value": 26.5
        }
    ]
}
```

Nested JSON format!

```
        "id": "S100",
        "device_id": "S100",
        "name": "Woodlands Road",
        "location": {
            "latitude": 1.4172,
            "longitude": 103.74855
        }
    },
    {
        "id": "S115",
        "device_id": "S115",
        "name": "Tuas South Avenue 3",
        "location": {
            "latitude": 1.29377,
            "longitude": 103.61843
        }
    }
],
"reading_type": "DBT 1M F",
"reading_unit": "deg C"
},
"items": [
{
    "timestamp": "2019-11-23T00:01:00+08:00",
    "readings": [
        {
            "station_id": "S109",
            "value": 25.2
        },
        {
            "station_id": "S117",
            "value": 26
        },
        {
            "station_id": "S118",
            "value": 27
        }
    ]
}
```



@ongchinhwee



Open Data = Available
開放資料 = 任何人都能取得

Open Data = Accessible? 開放資料 = 任何人都能使用？

*“The data must be available **as a whole** and at no more than a reasonable reproduction cost...The data must also be available **in a convenient and modifiable form.**”*

- Open Knowledge Foundation

Open Data - Available, Readable, Accessible

Available: Open data is **freely available** for everyone to access and use without restrictions or exorbitant costs.

Open Data - Available, Readable, Accessible

Available: Open data is **freely available** for everyone to access and use without restrictions or exorbitant costs.

Readable: Open data is **more usable** when it is shared in a machine-readable format such as JSON

Open Data - Available, Readable, Accessible

Available: Open data is **freely available** for everyone to access and use without restrictions or exorbitant costs.

Readable: Open data is **more usable** when it is shared in a machine-readable format such as JSON

Accessible: Open data is only **useful** if it's shared in a way that everyone can easily discover and understand.

If not everyone can understand
how to use the available data, is
the data really **open**?





I ended up creating my own
scraping tool.

 Code

 Issues 0

 Pull requests 0

 Projects 0

 Security

 Insights

Join GitHub today

Dismiss

GitHub is home to over 40 million developers working together to host and review code, manage projects, and build software together.

 Sign up

Scraping Meteorological Data from Data.gov.sg APIs

 24 commits

 2 branches

 0 packages

 0 releases

 2 contributors

Branch: master ▾

New pull request

Find file

Clone or download ▾

 hweecat Merge pull request #8 from hweecat/airtemp_rain ...

Latest commit 823bc3b 5 days ago

 .gitignore

update timezone code for pandas 0.25

22 days ago

 API_scraping_datagovsg_(airtemp_rainfall).py

add try-except logic for null-data-for-date case

7 days ago

Data.gov.sg Weather Data API Scraping

Scraping weather data from APIs via Python “Requests” library

“Requests”:

Python library for humans to send HTTP requests



Data.gov.sg Weather Data API Scraping

Scraping weather data from APIs via **Python “Requests” library**

Currently supported Data.gov.sg APIs:

1. **Air Temperature (in degree Celsius)** -- 17 Weather Stations
2. **Rainfall (in mm)** -- 53 Weather Stations

Data.gov.sg Weather Data API Scraping

Scraping weather data from APIs via **Python “Requests” library**

Currently supported Data.gov.sg APIs:

1. Air Temperature (in degree Celsius)
2. Rainfall (in mm)

Scrap data for **continuous time range** + **specific weather station**

API Scraping Demo



Singapore's Weather Station Positions (Source: Weather.gov.sg)

新加坡气象站位置

@ongchinhwee

Time Series Analysis of Singapore Weather Station Data



Selected **weather station**:

Changi Weather Station (**ID: S24**)

Analysis **timeframe**:

2 Dec 2016 to 30 Nov 2019 (~3 years)

Objective:

- Extract trend and seasonality from minute-by-minute weather time series

@ongchinhwee

Time Series Analysis Demo

Key Takeaways

1. Open data needs to be **accessible to anyone** to be **useful**.
 - Putting data out on the web with open data license is **usually not enough** - it's not just developers using open data!

Key Takeaways

1. Open data needs to be **accessible to anyone** to be **useful**.
 - Putting data out on the web with open data license is **usually not enough** - it's not just developers using open data!
2. If an open data source is **not intuitive enough** to use, build a tool to make the data **more accessible** for everyone!

Reach out to me!



: ongchinhwee



: @ongchinhwee



: hweecat

<https://ongchinhwee.me>

And check out my project on:



[hweecat/api-scraping-nea-datasets](https://github.com/hweecat/api-scraping-nea-datasets)