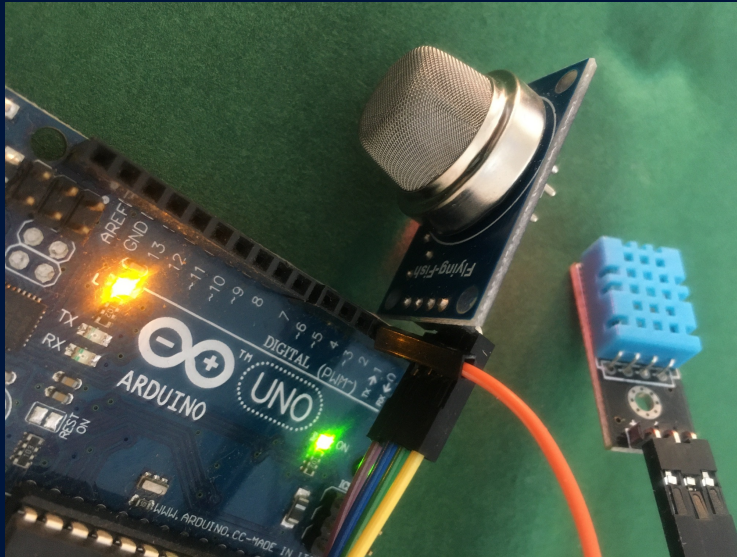


# COPENHAGEN BUSINESS ACADEMY



## DATA ENGINEERING



## FLOW 3 – Data Engineering

- cityflow.dk – API
  - Json-dataformat
  - https request
  - R request
  - File-persistence (json, csv)
  - Database-persistence
- Tooling
  - AWS EC2-instance (Ubuntu eller Windows?)

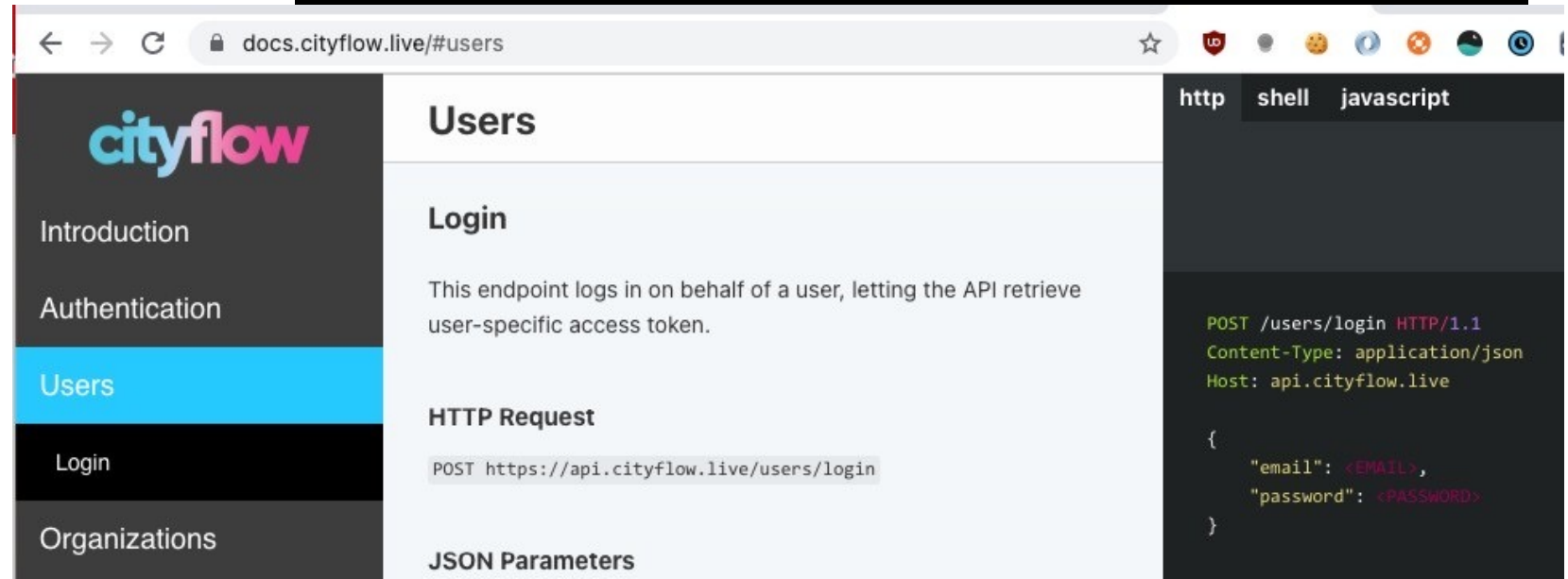
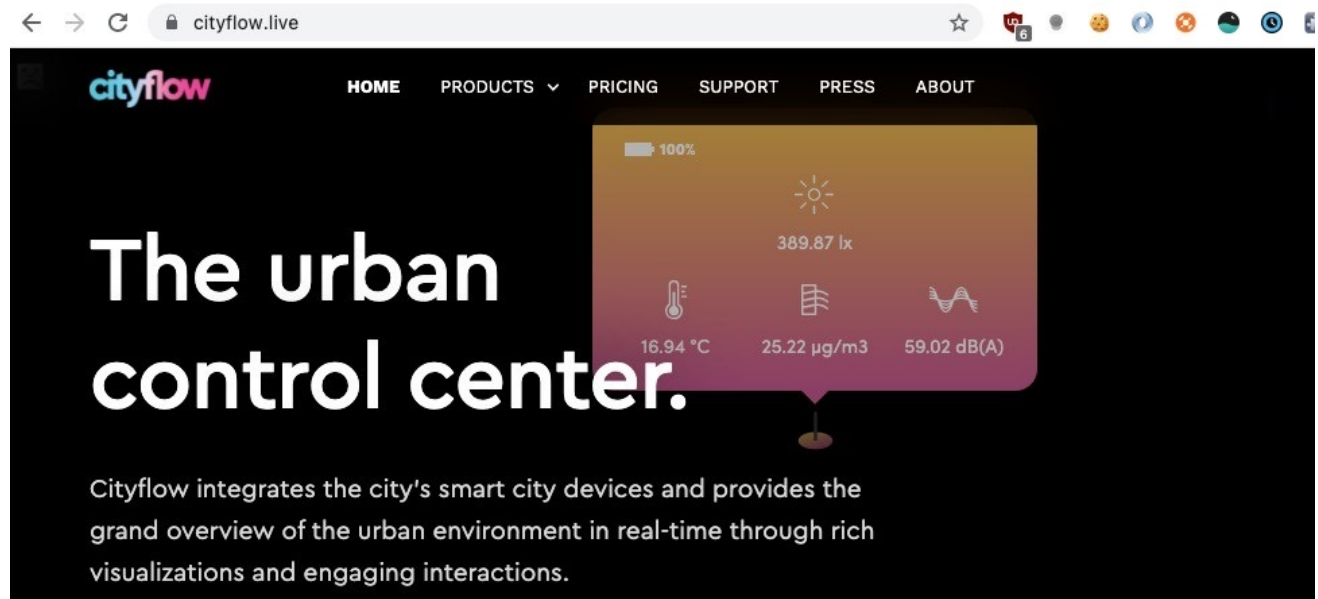
# FLOW 3 – Data Engineering

- Oplæg om de to cases i Flow 3
  - cityflow.dk – API
    - Json-dataformat
    - https request
    - python request
    - File-persistence (json, csv)
    - Database-persistence (MongoDB)
  - Bilbasen.dk – WebScraping
    - HTML/CSS/JavaScript primer
    - Scrape static (python bs4)
    - Scrape dynamic (python selenium)
- Tooling
  - GitHub / GitBash
  - Python (Anaconda) & jupyter notebook
  - AWS EC2-instance (Ubuntu med MongoDB og MySQL)
  - Arduino with serial-port driver

## FLOW 3 – User Stories

- Som underviser vil jeg gerne vide hvor meget den hvide bygning bliver brugt så jeg kan lægge aktiviteter i bygningen når den bliver mindst brugt
- Som studerende vil jeg gerne vide hvor det er bedst at lave en walk-and-talk i Århus i løbet af dagen på en hverdag
- Som ejer af en gammel Volvo vil jeg gerne kunne få en vurdering af hvor meget den er værd således at jeg kan planlægge hvornår jeg skal sælge den

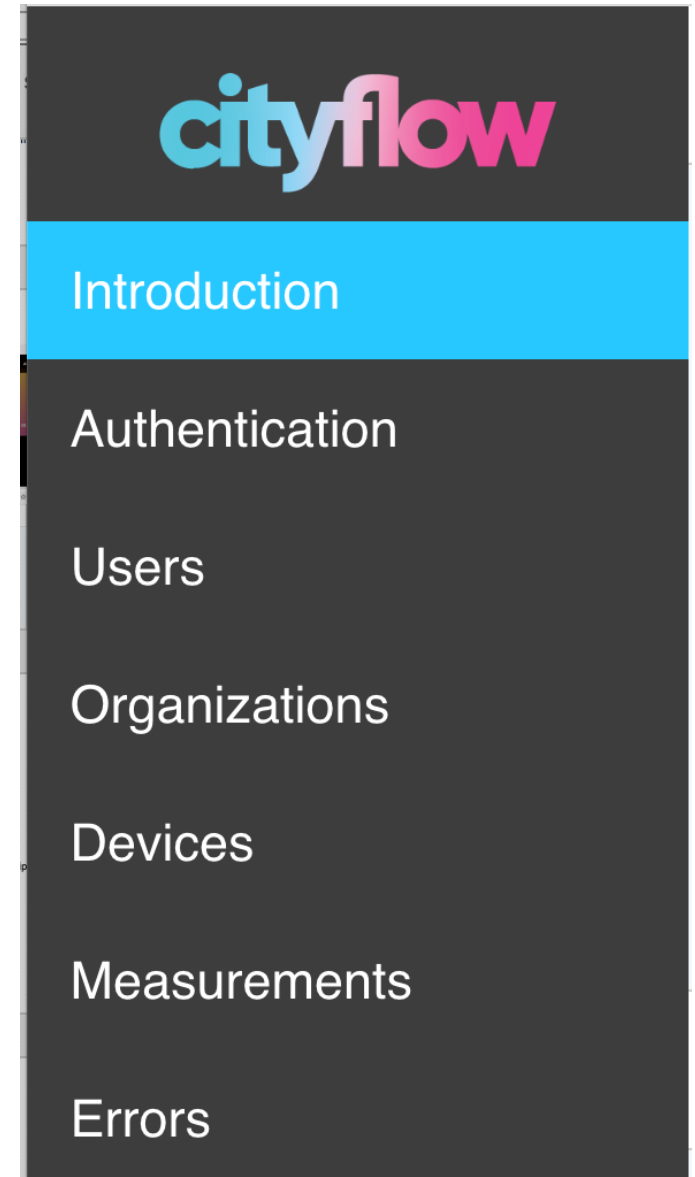
# Cityflow



# Cityflow – REST API



- What do I see at docs.cityflow.live?



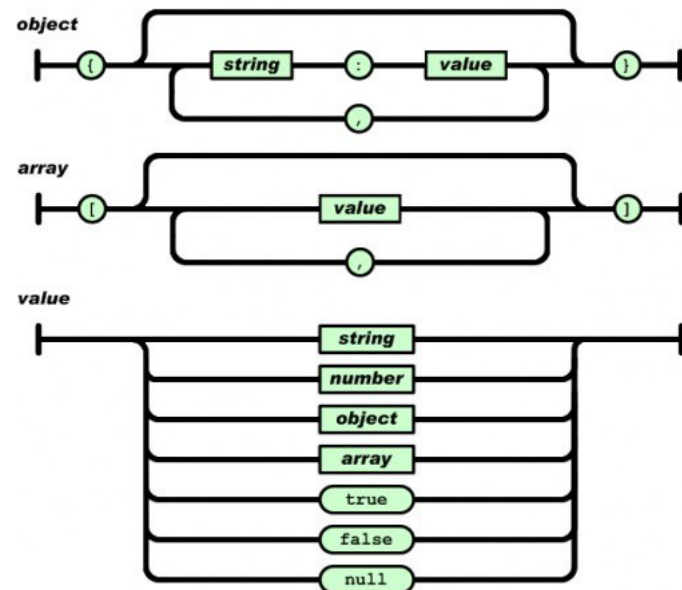
# Cityflow – REST API Data Format

- “The above command returns JSON structured like this”
- JSON – JavaScript Object Notation

JSONLint - The JSON Validator

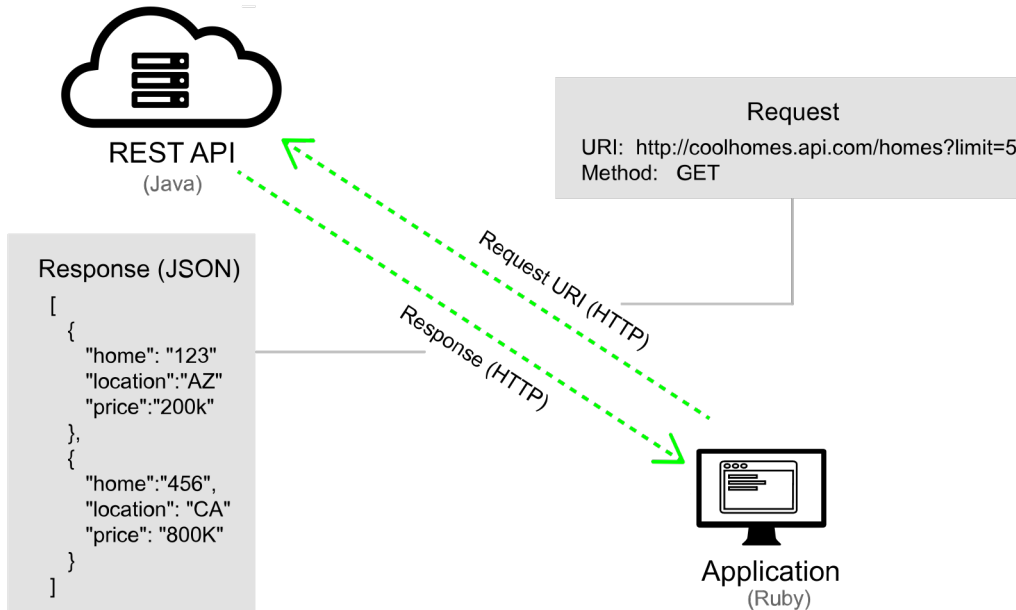
```

1 {
2   "id": "e00fce68c573b4acca2089ce",
3   "type": 150,
4   "location": 216,
5   "latitude": 56.1632767,
6   "longitude": 10.2105122,
7   "location_name": "Nørrebrogade",
8   "city": "Aarhus",
9   "country": "Denmark",
10  "roles": [
11    4
12  ],
13  "permissions": [],
14  "tags": [
15    "Randersvej"
16  ]
17 }
```



# Cityflow – REST API Request

## REST API model



- Http – hvad er det?
  - Chrome dev-tools
- Shell – hvad er det?
  - Intall Git Bash (windows)
  - Curl - hvad er det?

```
http  shell  javascript

curl --request GET \
  --url https://api.cityflow.live/devices \
  --header 'authorization: Bearer {BEARER_TOKEN}'
```

```
curl --request POST \
  --url https://api.cityflow.live/users/login \
  --header 'content-type: application/json' \
  --data '{
    "email": <EMAIL>,
    "password": <PASSWORD>
  }'
```

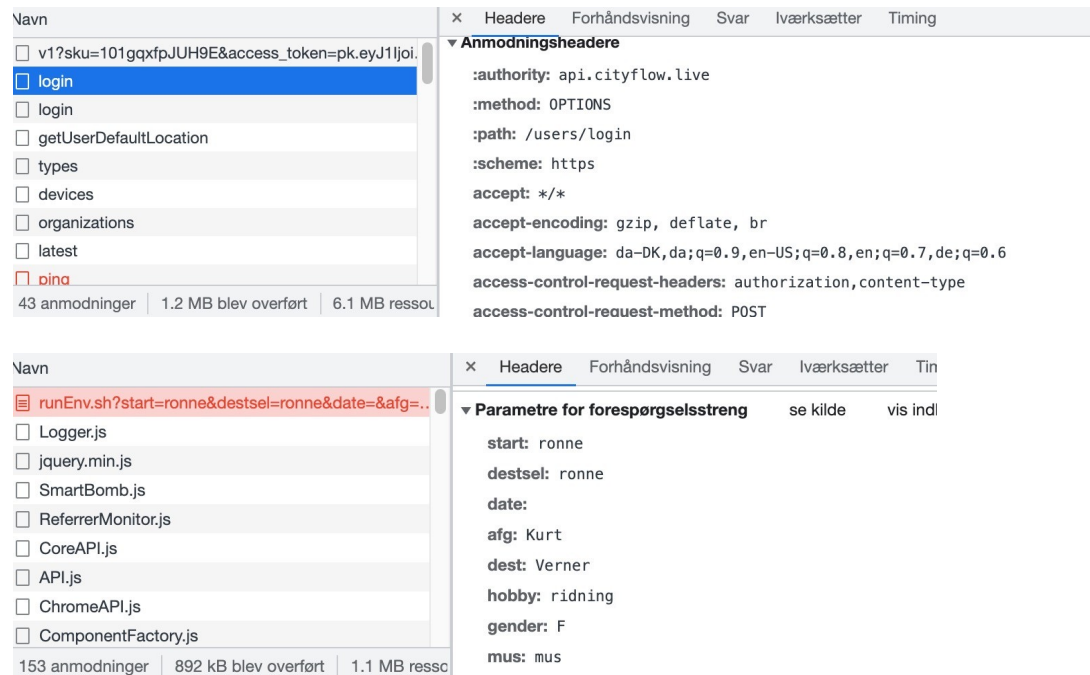


# Cityflow – REST API Request

1. Introduction
2. Protocols
3. Data Formats
4. Authentication, Part 1
5. Authentication, Part 2
6. API Design
7. Real-Time Communication
8. Implementation

To make a valid request, the client needs to include four things:

- 1 URL (Uniform Resource Locator) [1](#)
- 2 Method
- 3 List of Headers
- 4 Body



The screenshot displays a REST client interface with two requests listed on the left and their details on the right.

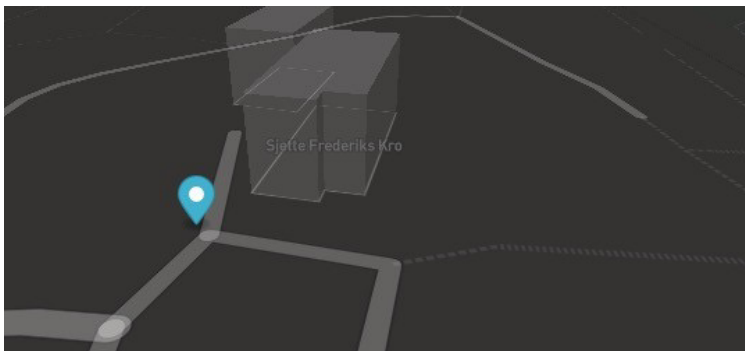
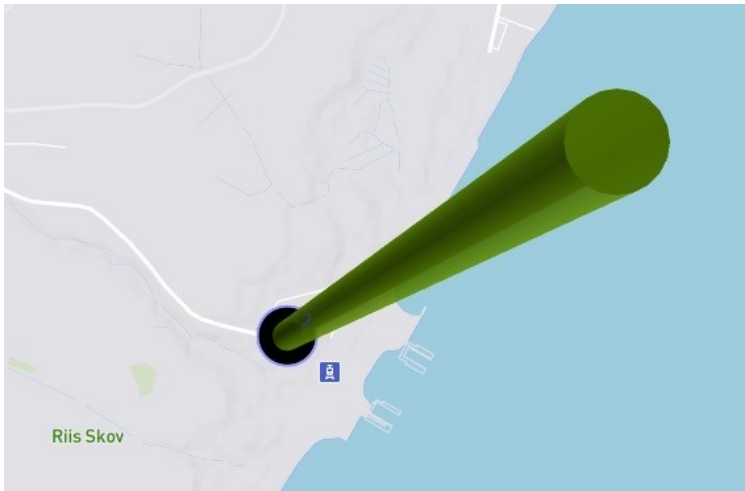
**Request 1:**

- Navn:** login
- URL:** v1?sku=101gqxfpJUH9E&access\_token=pk.eyJ1Ijoj...
- Headers (Anmodningsheadere):**
  - authority: api.cityflow.live
  - method: OPTIONS
  - path: /users/login
  - scheme: https
  - accept: \*/\*
  - accept-encoding: gzip, deflate, br
  - accept-language: da-DK, da;q=0.9, en-US;q=0.8, en;q=0.7, de;q=0.6
  - access-control-request-headers: authorization, content-type
  - access-control-request-method: POST
- Summary:** 43 anmodninger | 1.2 MB blev overført | 6.1 MB ressource

**Request 2:**

- Navn:** runEnv.sh?start=ronne&destsel=ronne&date=&afg=...
- URL:** runEnv.sh?start=ronne&destsel=ronne&date=&afg=...
- Parameters (Parametre for forespørgselsstreng):**
  - start: ronne
  - destsel: ronne
  - date:
  - afg: Kurt
  - dest: Verner
  - hobby: ridning
  - gender: F
  - mus: mus
- Summary:** 153 anmodninger | 892 kB blev overført | 1.1 MB ressource

# Cityflow – REST API Get Data I



http shell javascript

```
GET /devices HTTP/1.1
Authorization: Bearer {BEARER_TOKEN}
Host: api.cityflow.live
```

```
{
  id: e00fce689f02a96799f34fc2,
  type: 150,
  location: 190,
  latitude: 56.1770897,
  longitude: 10.2296247,
  location_name: Salonvejen,
  city: Risskov,
  country: Denmark,
  roles: [
    4
  ],
  permissions: [],
  tags: [
    Risskov
  ]
},
```

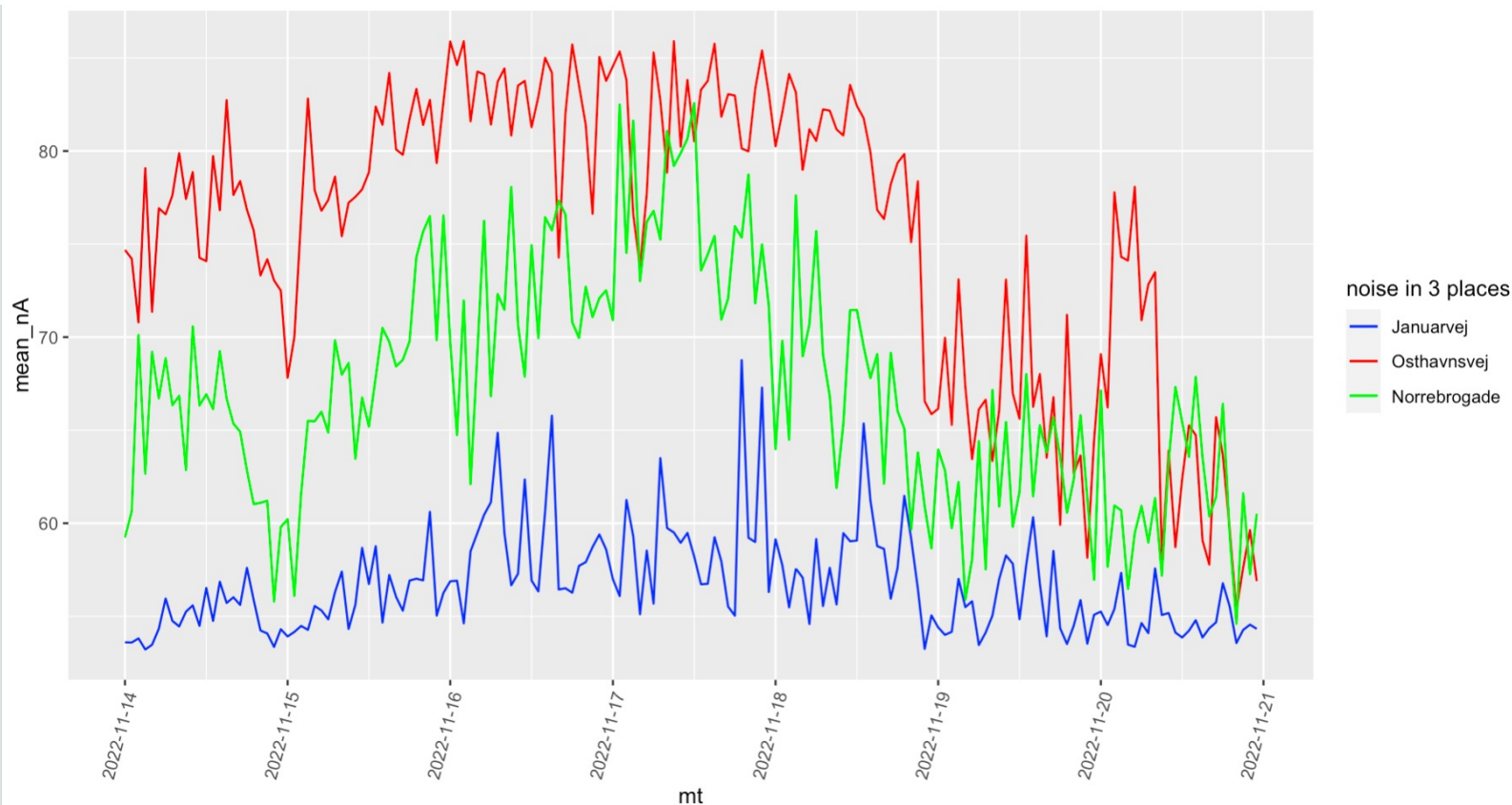
```
"time": "2019-11-28T07:23:19.346Z",  
"CityProbe 2",  
"Particle Pollution (PM2.5)",  
"Particle Pollution (PM4)",  
"Particle Pollution (PM10)",  
"Particle Pollution (PM1)",  
"Battery Level",  
"Temperature",  
"Humidity",  
"Luminosity",  
"Atmospheric Pressure",  
"Noise Average",  
"Noise Minimum",  
"Noise Maximum",  
"Noise Standard Deviation",  
"Rain Average",  
"Rain Minimum",  
"Rain Maximum",  
"Rain Standard Deviation",  
"Particulate Concentration (PC1)",  
"Particulate Concentration (PC2.5)",  
"Particulate Concentration (PC4)",  
"Particulate Concentration (PC10)",  
"Average Particle Size",  
"b": 19.05,  
"c": 163,  
"firmware_version": "49",  
"h": 0,  
"l": 3353,  
"location": "8",  
"n": 170,  
"p": 0,  
"p1": 0,  
"p2": 0,  
"r.avg": -1673.1,  
"r.max": -757,  
"r.min": -1772,  
"r.sd": 169.71,  
"r.var": null,  
"s.avg": 81.16,  
"s.max": 81.65,  
"s.min": 77.6,  
"s.sd": 0.89,  
"s.var": null,  
"seq": 1203,  
"t": 0,  
"uv": null,  
"device_id":  
"49004d000d50483553343720"
```

## Cityflow – hist measures

```
"time": "2019-11-27T08:20:00.000Z",  
  "mean_b": 32.98,  
  "mean_c": 144,  
  "mean_h": 0,  
  "mean_l": 502,  
  "mean_n": 172,  
  "mean_p": 0,  
  "mean_p1": 0,  
  "mean_p2": 0,  
  "mean_r.avg": -8.19,  
  "mean_r.max": 12,  
  "mean_r.min": -1311,  
  "mean_r.sd": 92.22,  
  "mean_r.var": null,  
  "mean_s.avg": 41.08,  
  "mean_s.max": 79.05,  
  "mean_s.min": 36.6,  
  "mean_s.sd": 3.19,  
  "mean_s.var": null,  
  "mean_seq": 1142,  
  "mean_t": 0,  
  "mean_uv": null,  
  "device_id":  
  "49004d000d50483553343720",  
  "location": "8"
```

# CityFlow - R

1. Get access-token
2. Get device-list
3. Get historic measurements from aarhus
  1. one week, 60 mins interval



# AWS Windows Server



**Microsoft Windows Server 2012 R2 Base** - ami-08a57643245353515

**Windows**

Free tier eligible

Microsoft Windows 2012 R2 Standard edition with 64-bit architecture. [English]

Root device type: ebs    Virtualization type: hvm    ENA Enabled: Yes

- Create Windows 2012 R2
- Connect via RDP
- Tilføj nøgle i Registry
- Åben powershell

### Select an existing key pair or create a new key pair

A key pair consists of a **public key** that AWS stores, and a **private key file** that you store. Together, they allow you to connect to your instance securely. For Windows AMIs, the private key file is required to obtain the password used to log into your instance. For Linux AMIs, the private key file allows you to securely SSH into your instance. Amazon EC2 supports ED25519 and RSA key pair types.

Note: The selected key pair will be added to the set of keys authorized for this instance. Learn more about [removing existing key pairs from a public AMI](#).

Choose an existing key pair

Select a key pair

dal | RSA

☐ I acknowledge that I have access to the corresponding private key file, and that without this file, I won't be able to log into my instance.

Cancel Launch Instances

policy

**Edit DWORD (32-bit) Value**

Value name:

SchUseStrongCrypto

Value data:

1

Base

☒ Hexadecimal

☐ Decimal

OK Cancel

BidInterface

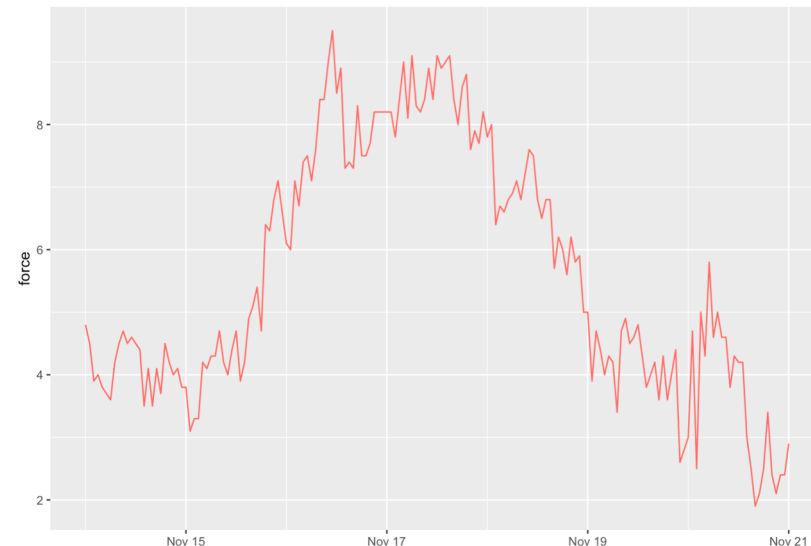
Computer\HKEY\_LOCAL\_MACHINE\SOFTWARE\Microsoft\.NETFramework\v4.0.30319

# CityFlow & DMI – vindretning og støj

Tesen er at støjen fra en vej er afhængig vindretning. Det skal undersøges.

1. Find locations i Aarhus som ligger umiddelbart vest for en større vej
2. Hent vind-data (retning og styrke) fra DMI's
  1. hvor vinden er fra vest og stærk men skifter på et tidspunkt
3. Hent fra samme tidsinterval støjdata fra cityflow

	mt	dir	force
1	2022-11-14 00:00:00	118	4.8
2	2022-11-14 01:00:00	119	4.5
3	2022-11-14 02:00:00	123	3.9
4	2022-11-14 03:00:00	115	4.0
5	2022-11-14 04:00:00	116	3.8
6	2022-11-14 05:00:00	107	3.7
7	2022-11-14 06:00:00	119	3.6
8	2022-11-14 07:00:00	107	4.2
9	2022-11-14 08:00:00	113	4.5
10	2022-11-14 09:00:00	109	4.7
11	2022-11-14 10:00:00	114	4.5
12	2022-11-14 11:00:00	112	4.6



# CityFlow & DMI – vindretning og støj

## Wind\_dir

The parameter wind\_dir returns the direction from which the wind blows in degrees, where north is given as 360.

Code	Value
0	calm

