Summary

Content

- Data
  - Main data: *download*
  - Extra data:
    - AgERA5: *download*
    - Traffic data: understand
  - How to use the data?
- Algorithms
  - Labels?
  - Features?
  - train/test?

Main data – download

- Duration: **01.01.2010 → 31.12.2021** (4383 days)
- 8 parameters
- Output file: *main_data.csv*

```
#    Column                Non-Null Count
---  ------                --------------
0    Graz-DB (LUFE)        4374 non-null
1    Graz-DB (LUTE)        4357 non-null
2    Graz-DB (NO)          4368 non-null
3    Graz-DB (NO2)         4368 non-null
4    Graz-DB (NOX)         4368 non-null
5    Graz-DB (PM10)        4327 non-null
6    Graz-DB (SO2)         4355 non-null
7    Graz-DB (STBK 10K)    4310 non-null
8    Graz-N (LUFE)         4365 non-null
9    Graz-N (LUTE)         4340 non-null
10   Graz-N (NO)           4368 non-null
11   Graz-N (NO2)          4368 non-null
12   Graz-N (NOX)          4368 non-null
13   Graz-N (PM10)         2683 non-null
14   Graz-N (SO2)          4355 non-null
15   Graz-N (STBK 10K)     4305 non-null
16   Graz-O P (LUFE)       1676 non-null
17   Graz-O P (LUTE)       1680 non-null
18   Graz-O P (NO)         4345 non-null
19   Graz-O P (NO2)        4347 non-null
20   Graz-O P (NOX)        4347 non-null
21   Graz-O P (PM10)       2251 non-null
22   Graz-O P (STBK 10K)   4340 non-null
23   Graz-S (LUFE)         4356 non-null
24   Graz-S (LUTE)         4332 non-null
25   Graz-S (NO)           4352 non-null
26   Graz-S (NO2)          4352 non-null
27   Graz-S (NOX)          4352 non-null
28   Graz-S (PM10)         4350 non-null
29   Graz-S (SO2)          4371 non-null
30   Graz-S (STBK 10K)     4345 non-null
31   Graz-W (LUFE)         4348 non-null
32   Graz-W (LUTE)         4369 non-null
33   Graz-W (NO)           4321 non-null
34   Graz-W (NO2)          4321 non-null
35   Graz-W (NOX)          4321 non-null
36   Graz-W (PM10)         4311 non-null
37   Graz-W (STBK 10K)     4311 non-null
38   Graz-W (SO2)          62 non-null
```

Locations
- 'Graz-DB'
- 'Graz-N'
- 'Graz-O'
- 'Graz-S'
- 'Graz-W'

Parameters
- 'LUFE'
- 'LUTE'
- 'NO'
- 'NO2'
- 'NOX'
- 'PM10'
- 'SO2'
- 'STBK 10K'

```
komponente = {
    'SO2': 1,
    'NO': 3,
    'NO2': 4,
    'NOX': 5,
#    'Schwebstaub': 7,
    'Lufttemperatur (LUTE)': 8,
    'Relative Luftfeuchte (LUFE)': 9,
    'PM10 Kont. (STBK 10G)': 114,
#    'PM10 Grav. (STBK 10G)': 119,
#    'PM2,5 Grav. (STBK 25G)': 122,
    'Feinstabu (PM10)': 125
}
```

AgERA5: *download*

- Duration: **01.01.2010** → **31.12.2021** (4383 days)
- 10 parameters
- Output file: *main_data.csv*

```
parameter = [('cloud_cover','24_hour_mean'),
            ('2m_temperature','day_time_maximum'),
            ('2m_temperature','night_time_minimum'),
            ('10m_wind_speed','24_hour_mean'),

            ('liquid_precipitation_duration_fraction',''),
            ('solar_radiation_flux',''),
            ('precipitation_flux',''),
            ('solid_precipitation_duration_fraction', ''),
            ('2m_dewpoint_temperature','24_hour_mean'),
            ('vapour_pressure', '24_hour_mean'),

            ]
```

Traffic data: understand

Time Check

transformed

	#rows	#cols
BvSuttner-Brücke	336	10
Keplerbrücke	336	76
Karlauer Gürtel	336	72
Rösselmühlgasse	336	76
Don Bosco	336	76
Glacis	336	76
Kärntner Straße	336	63

	#rows	#cols	2019	2020	2021	total
BvSuttner-Brücke	2688	3	7	49	NaN	56
Keplerbrücke	24864	3	14	186	318.0	518
Karlauer Gürtel	23520	3	14	158	318.0	490
Rösselmühlgasse	24864	3	14	186	318.0	518
Don Bosco	24864	3	14	186	318.0	518
Glacis	24864	3	14	186	318.0	518
Kärntner Straße	20496	3	14	95	318.0	427

See slides *20220111_traffic_data_michael*

Output Transformed file: *Ver_transformed.csv*

year	2019				2020				2021			
	actual	duration	max	min	actual	duration	max	min	actual	duration	max	min
BvSuttner-Brücke	7	7	2019-04-07	2019-04-01	49	91	2020-04-19	2020-01-20	NaN	NaN	NaT	NaT
Keplerbrücke	14	49	2019-05-19	2019-04-01	186	347	2020-12-31	2020-01-20	318	318	2021-11-14	2021-01-01
Karlauer Gürtel	14	49	2019-05-19	2019-04-01	158	347	2020-12-31	2020-01-20	318	318	2021-11-14	2021-01-01
Rösselmühlgasse	14	49	2019-05-19	2019-04-01	186	347	2020-12-31	2020-01-20	318	318	2021-11-14	2021-01-01
Don Bosco	14	49	2019-05-19	2019-04-01	186	347	2020-12-31	2020-01-20	318	318	2021-11-14	2021-01-01
Glacis	14	49	2019-05-19	2019-04-01	186	347	2020-12-31	2020-01-20	318	318	2021-11-14	2021-01-01
Kärntner Straße	14	49	2019-05-19	2019-04-01	95	347	2020-12-31	2020-01-20	318	318	2021-11-14	2021-01-01

How to use the data?

- Combine AgERA5 to main_data

- If we want to use traffic data, only can combine year 2019 → 2020

Thank you!