Technische Universität Berlin

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

Explanation of Air Pollution Using External Data Sources

Mahdi Esmailoghli, Sergey Redyuk, Ricardo Martinez, Ariane Ziehn,
Ziawasch Abedjan, Tilmann Rabl, Volker Markl (DIMA and BIGDAMA groups)

BTW Data Science Challenge

LuftDaten (pollution sensor data)

Challenges:

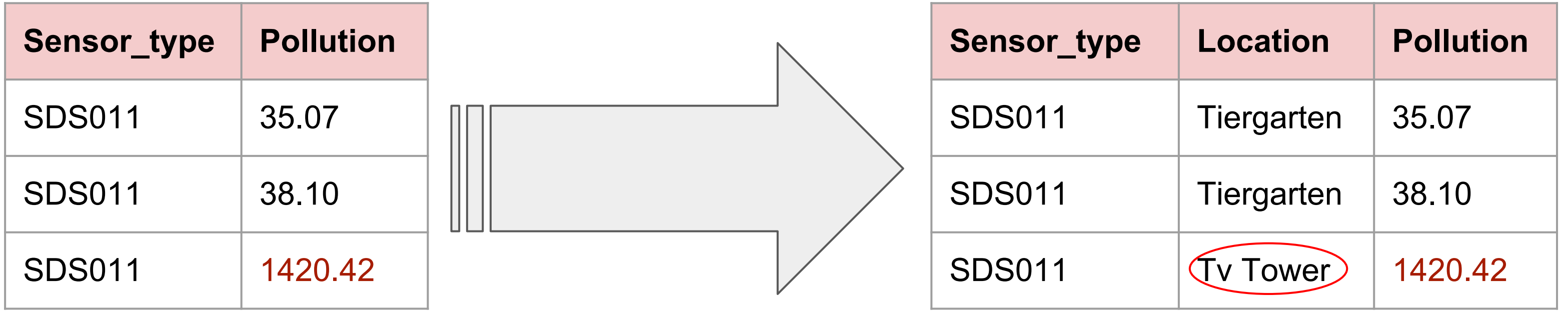
- Limited feature set
- Different schemas/sensors
- Malfunctioning sensors
- Stream nature of data

BTW Data Science Challenge - Our Goal

- Goal:
  - Explaining air pollution
  - Detecting the reasons of low air quality
- Problem:
  - Lack of information in provided data
  - Current ML algorithms cannot explain pollution based on provided data

BTW Data Science Challenge - Our Proposal

- Decision tree and Macrobase [Bailis’2017]*

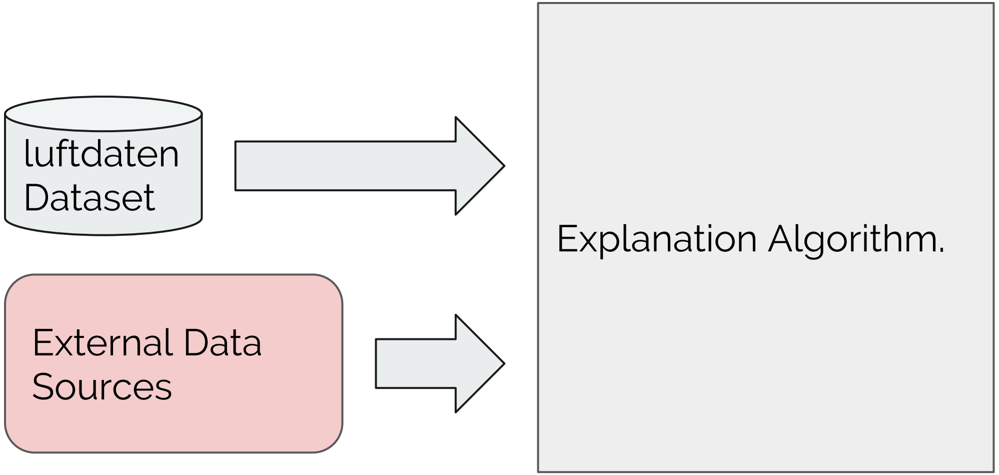


Sensor_type	Pollution
SDS011	35.07
SDS011	38.10
SDS011	1420.42

Sensor_type	Location	Pollution
SDS011	Tiergarten	35.07
SDS011	Tiergarten	38.10
SDS011	Tv Tower	1420.42

BTW Data Science Challenge - Our Proposal

- Enriching the main dataset (Luftdaten) with extra information
- Adding features that correlate with air pollution



luftdaten Dataset

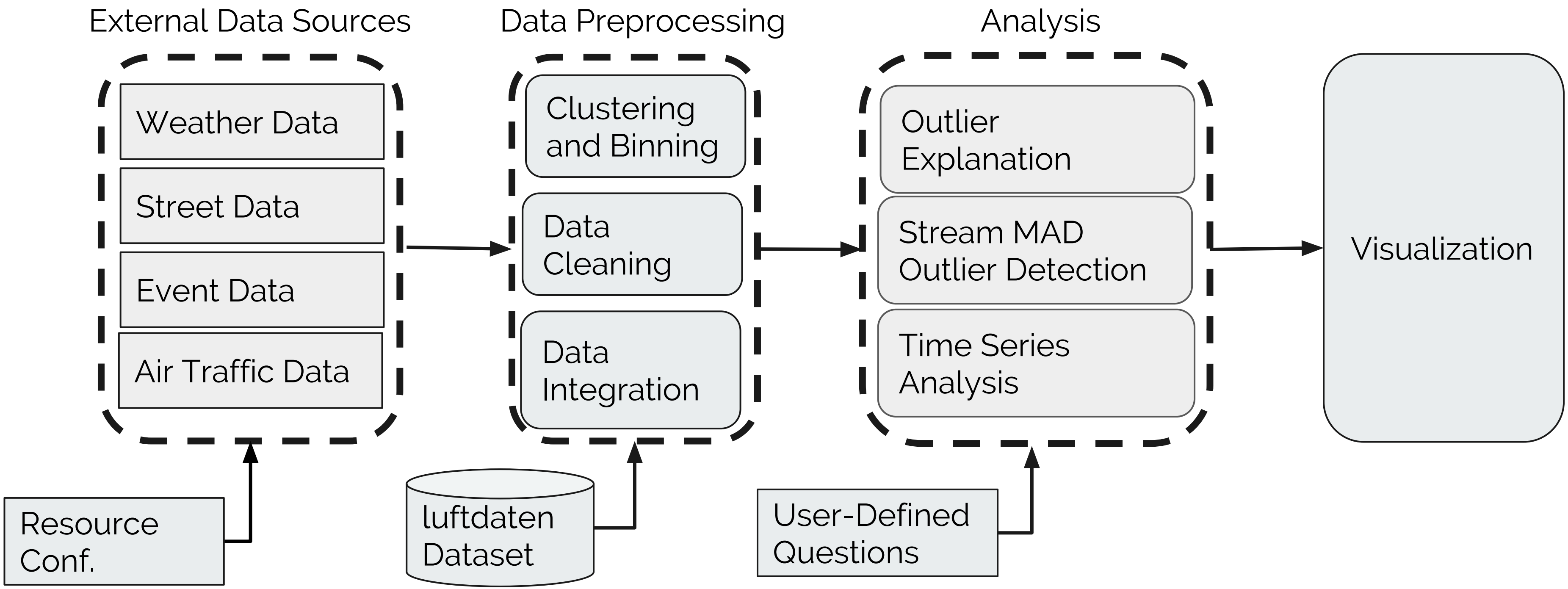
External Data Sources

Explanation Algorithm.

External Data Sources

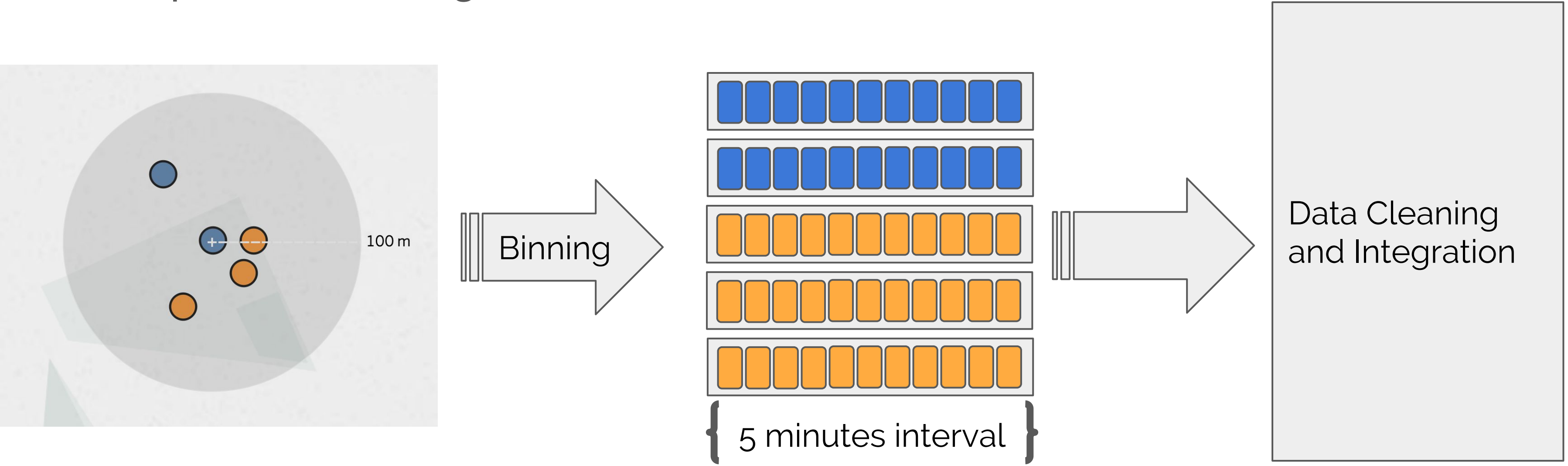
- Air traffic data
    - Airplanes’ route
- Event data
- Weather data
    - Wind (speed and direction)/Temperature/Precipitation
- Openstreetmap data
    - Number of crossroads and streets/Train stations

System Architecture



Clustering and Binning

- Spatial: clustering, 100-meter radius
- Temporal: binning, 5 minute-interval



Binning

5 minutes interval

Data Cleaning and Integration

Data Cleaning

- Wrong readings - malfunctioning sensors / network
- Deviating readings - outliers within the cluster / time slot

TimeStamp	P1
11:17:31	3.5
11:17:59	1.9
11:18:26	100012.7
11:20:44	3.2
11:21:58	2.4

Observation error

Data Integration

Time	P1
11:15:31	3.5
11:16:59	2.5
11:17:26	3.0
11:18:12	3.1
11:19:00	2.9

Time	Temp.
11:16:06	18.1
11:18:44	18.2

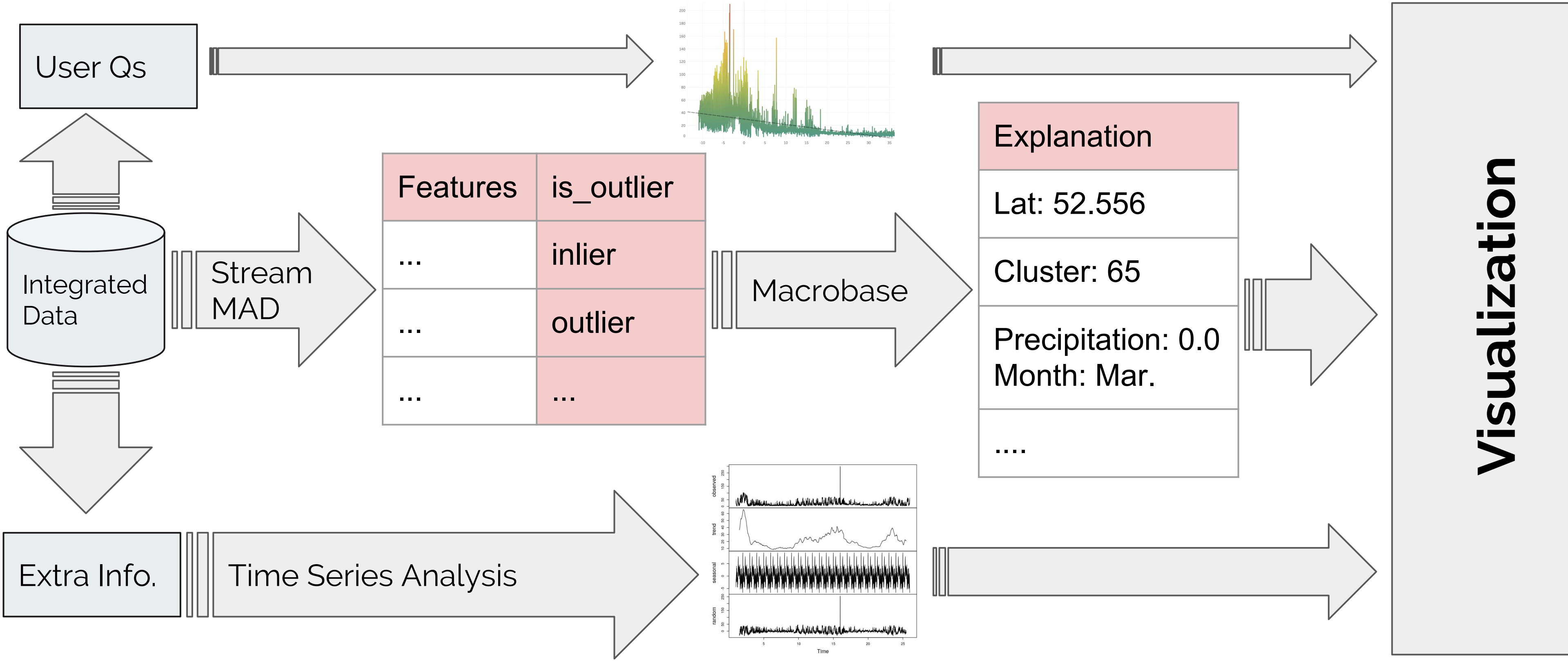
Time	Prec.
11:15:18	0.2
11:17:55	0.1
11:19:26	0.1

Time	Humid.
11:19:01	60%

Time	Wind	Degree
11:15:09	1.2	240
11:15:19	1.2	240
11:19:22	1.3	250

Time	P1	Temp.	Prec.	Humid.	Wind	Degree
11: [15 - 20]	3.0	18.15	0.1	60%	1.2	240

Analysis and Visualization



User Qs

Integrated Data

Stream MAD

Features	is_outlier
...	inlier
...	outlier
...	...

Macrobase

Explanation
Lat: 52.556
Cluster: 65
Precipitation: 0.0 Month: Mar.
....

Extra Info.

Time Series Analysis

Visualization

Results Based on External Data Sources

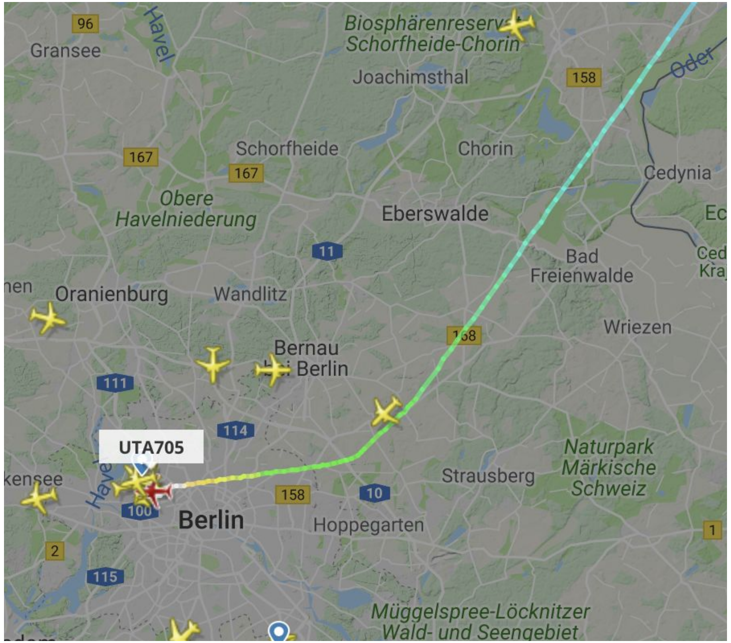
- Air traffic data
  - How does air traffic affect particulate matter pollution?
- Event data
  - Are there events that lead to short-term particulate matter pollution?
- Weather data
  - What is the correlation between weather data and air quality?
- Openstreetmap data
  - Do crossroads/roads/stations/diesel bans affect air pollution?

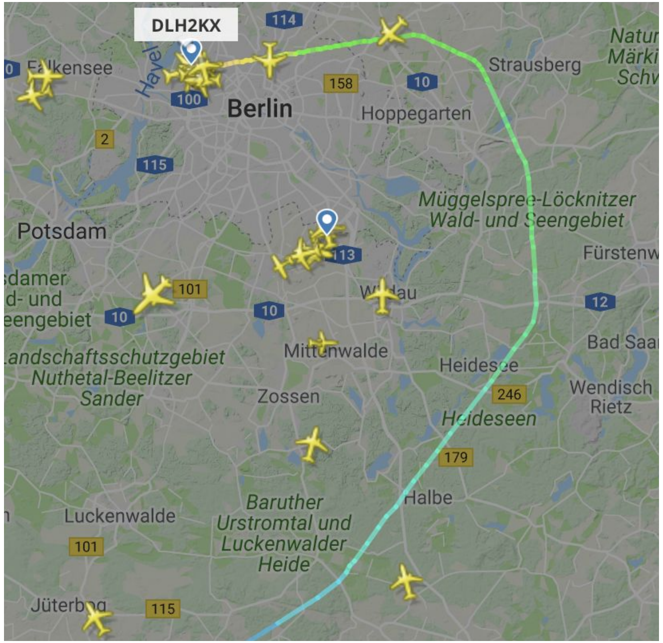
Berlin

Results (Air Traffic)

How Air Traffic Affects air quality?

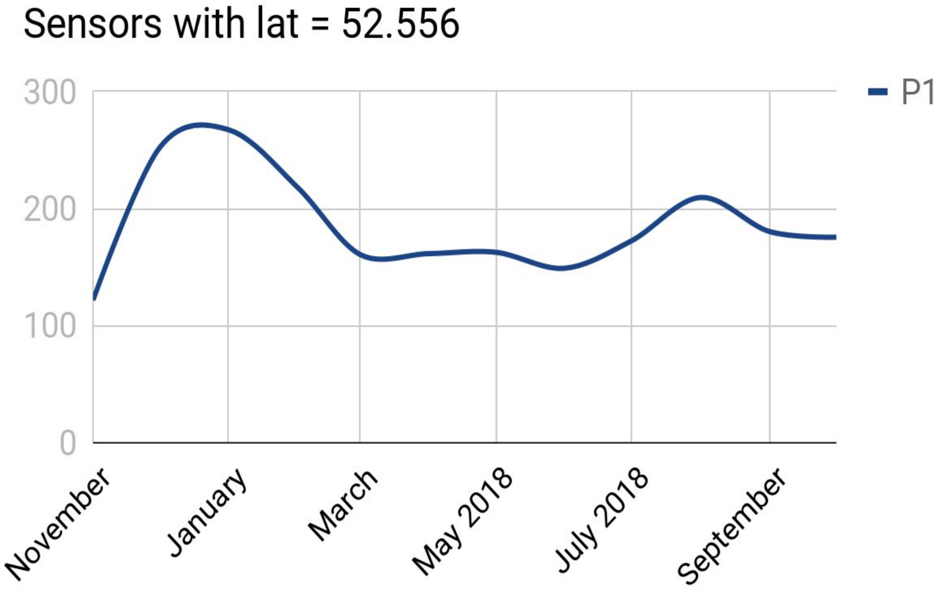
Explanation: Latitude: 52.556 (TXL Airport)

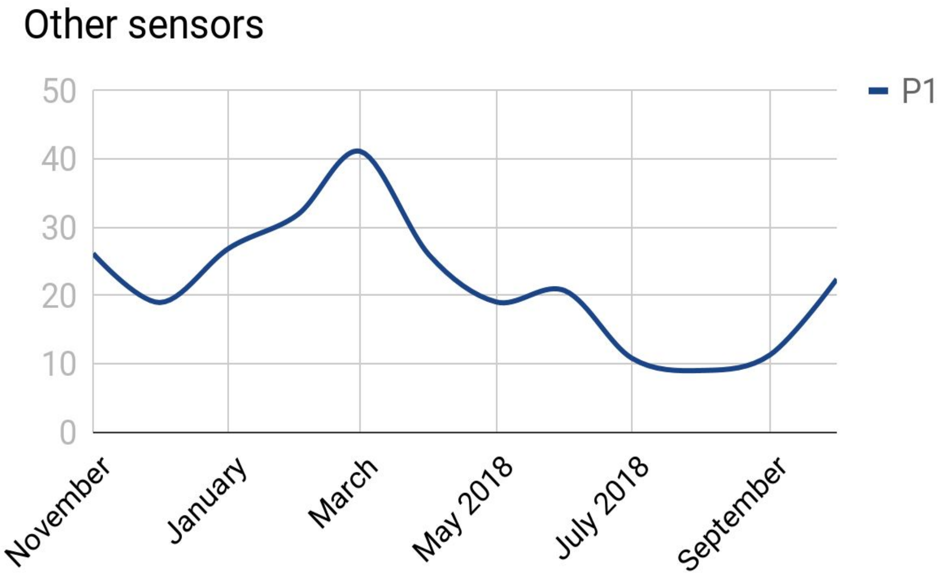




How Air Traffic Affects air quality?

Explanation: Latitude: 52.556 (TXL Airport)



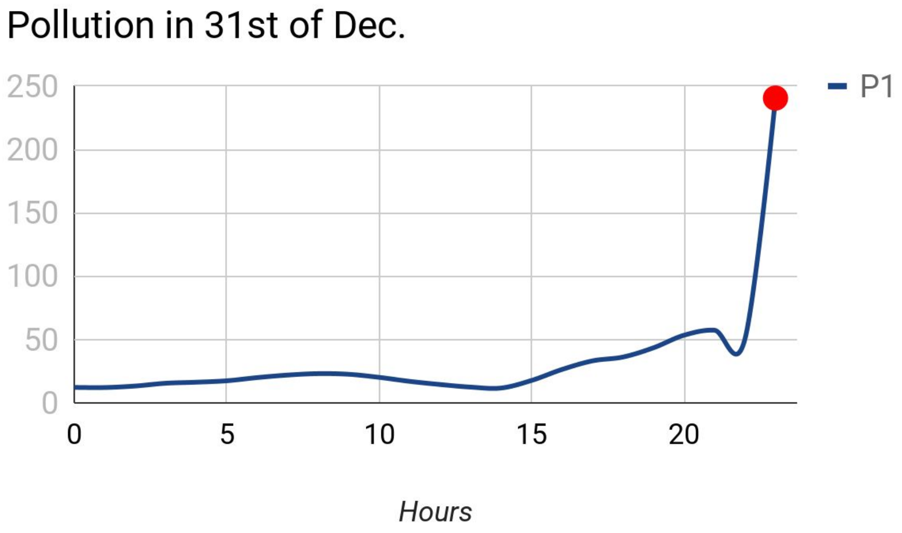


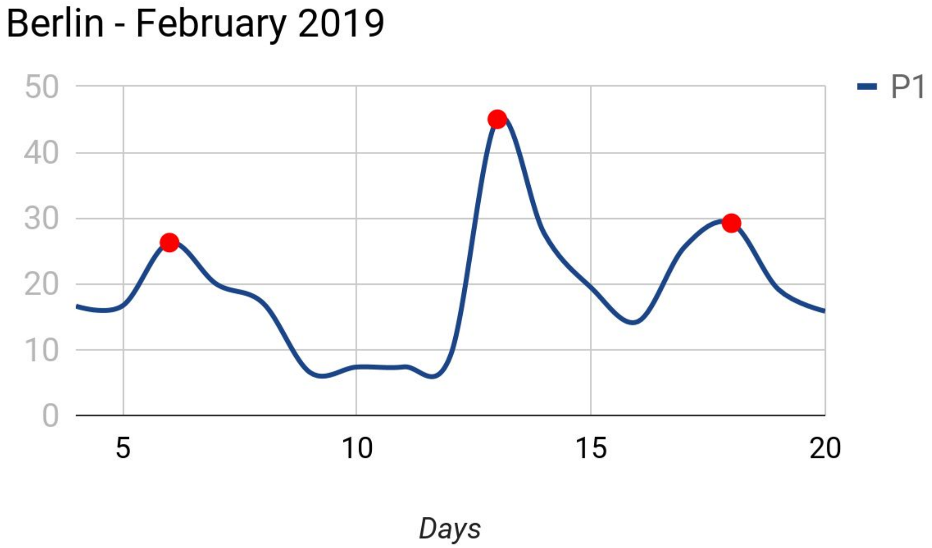
Results (Events)

How Events Play a Role in Pollution?

New Year's Eve

Berlin International Film Festival

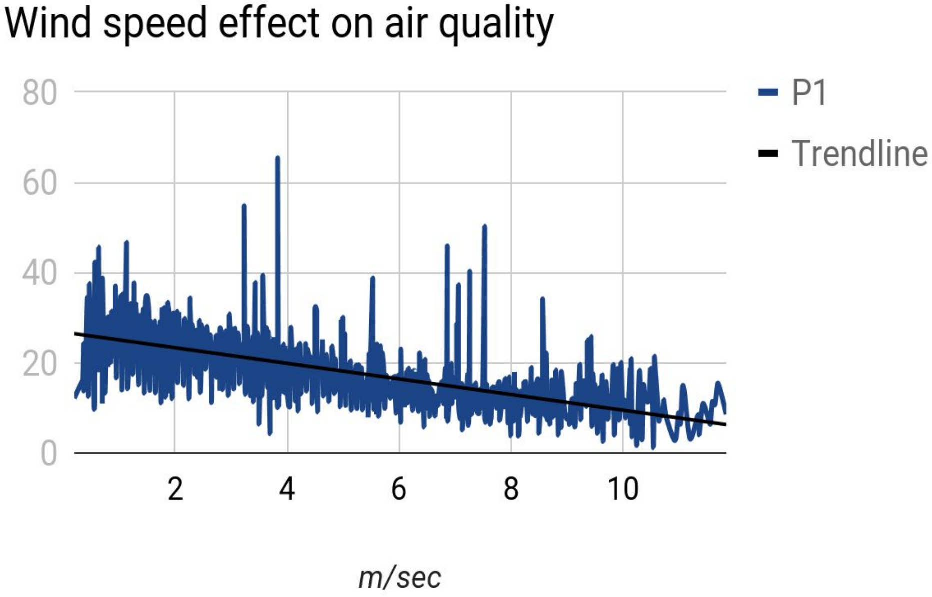


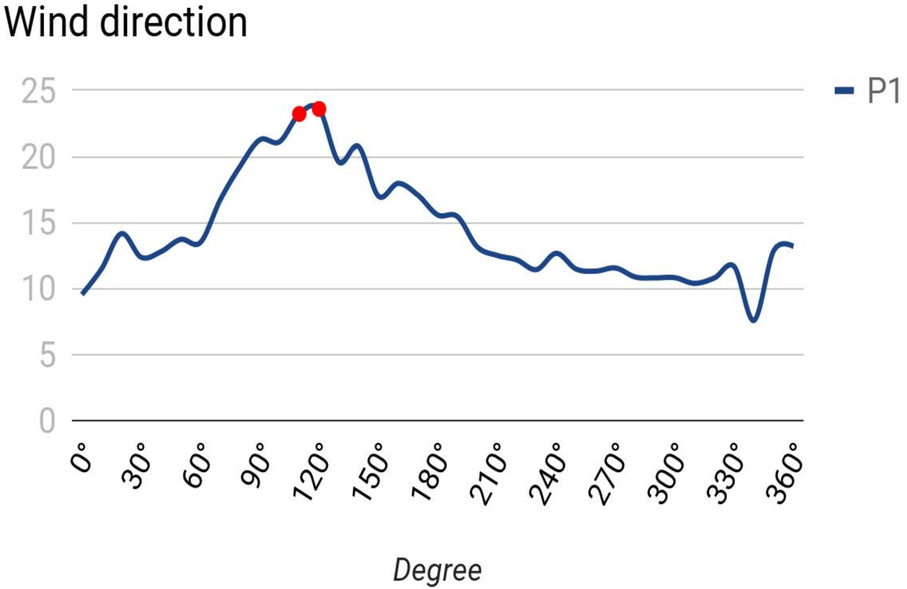


Results (Weather)

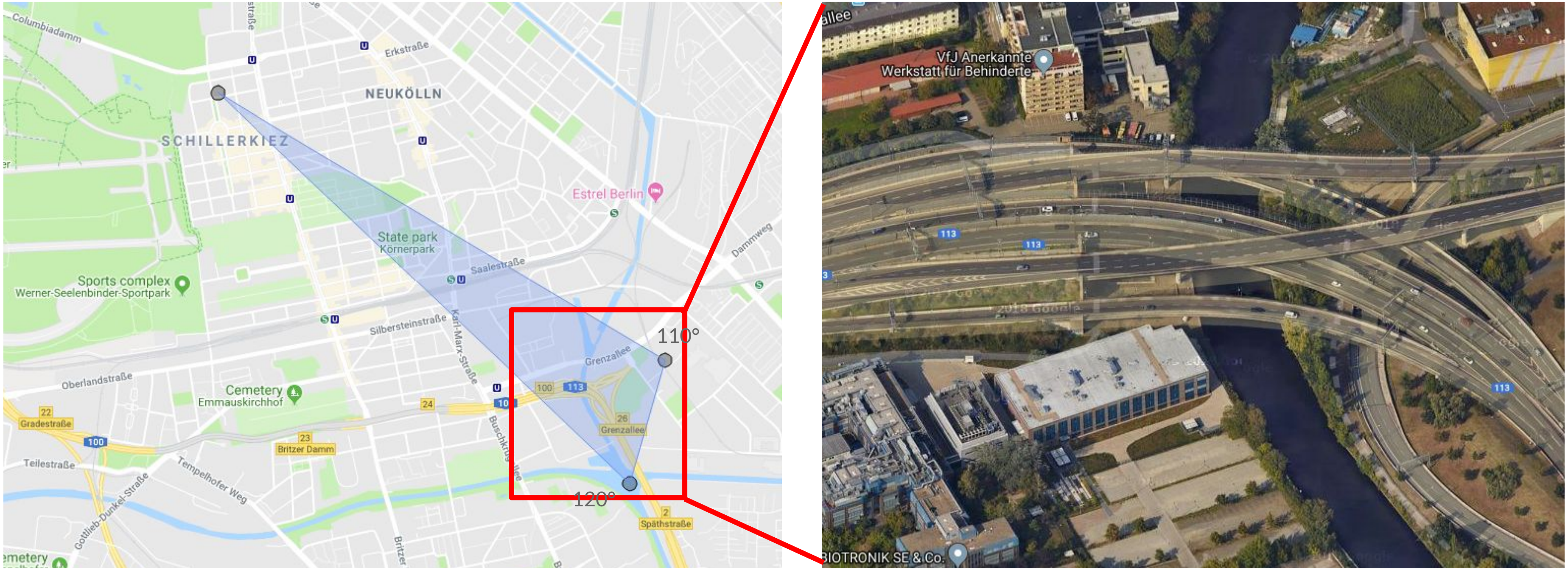
How Does The Weather Affect Air Pollution?

Explanation: Wind degree (cluster 104): 110 - 120





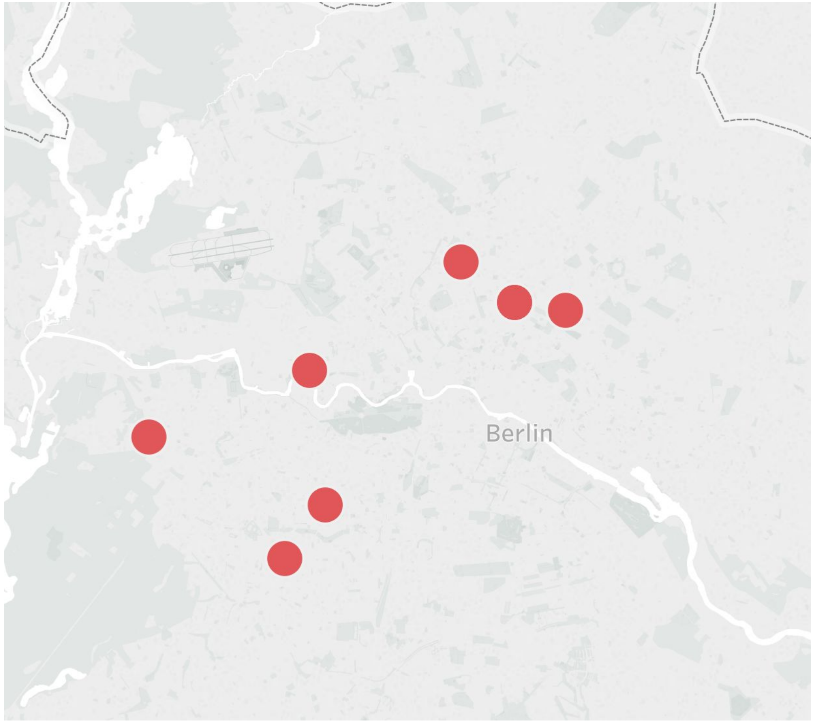
How Weather Data Affect Air Pollution?

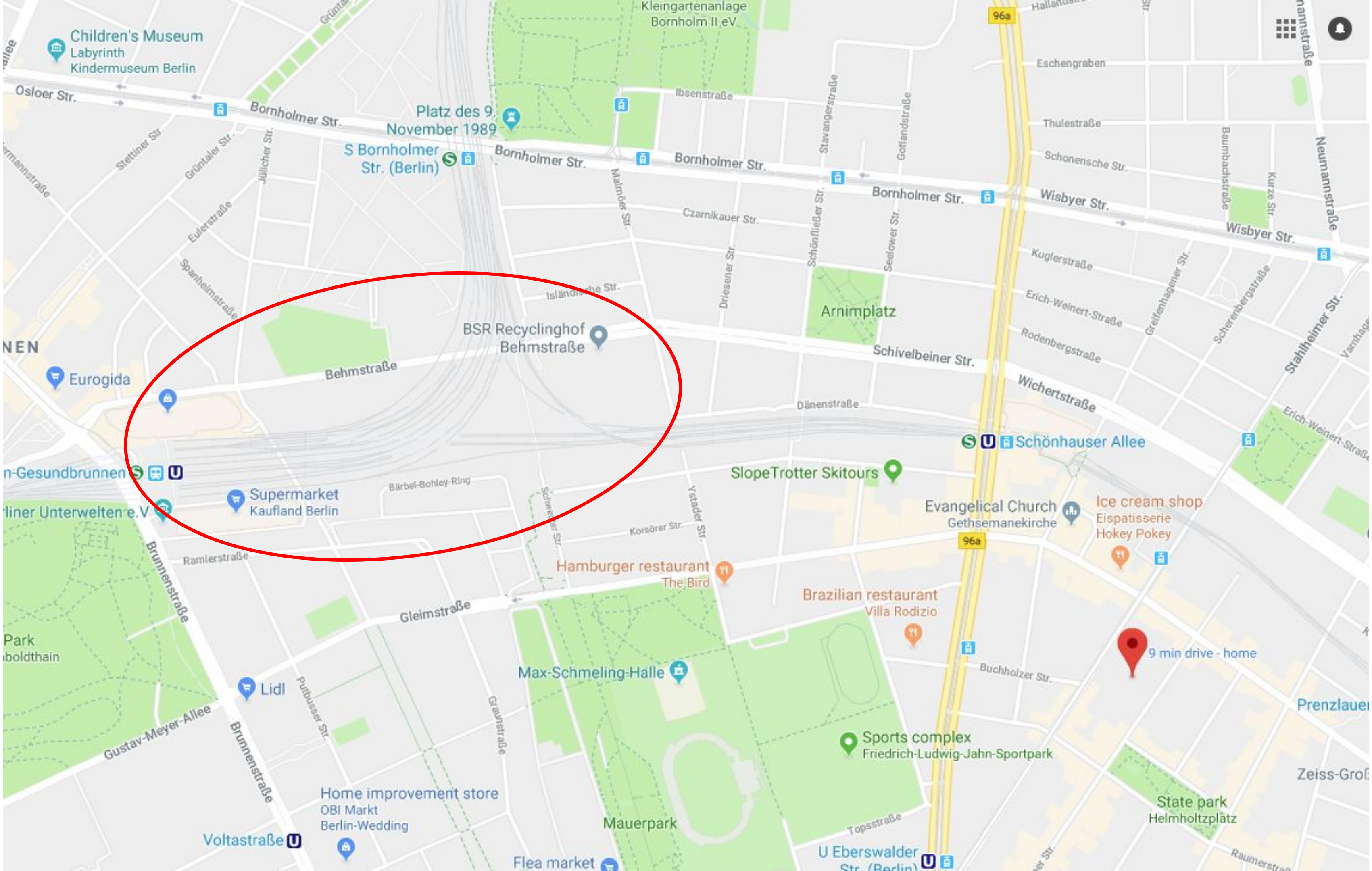


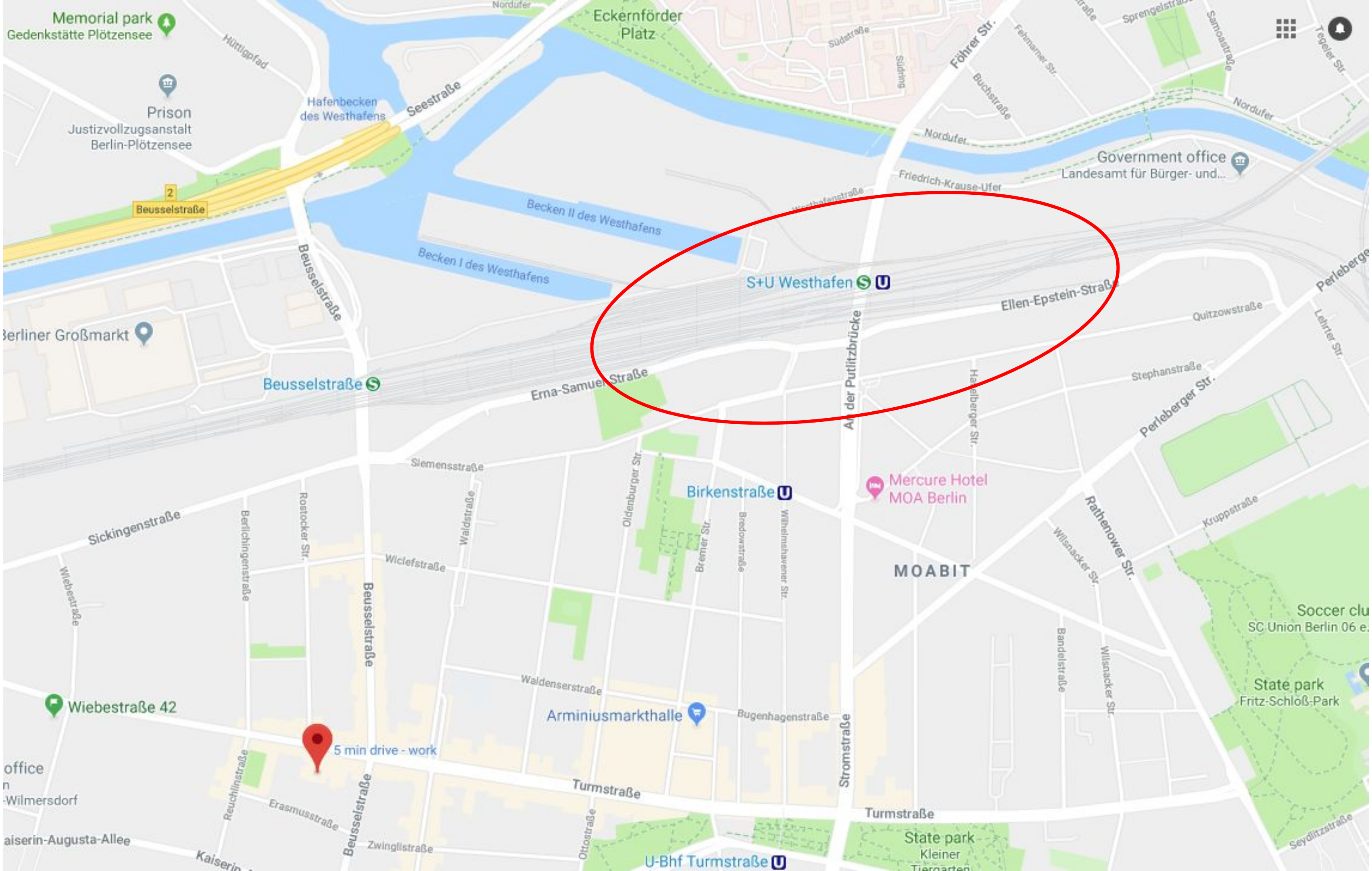
Results (OpenStreetMap)

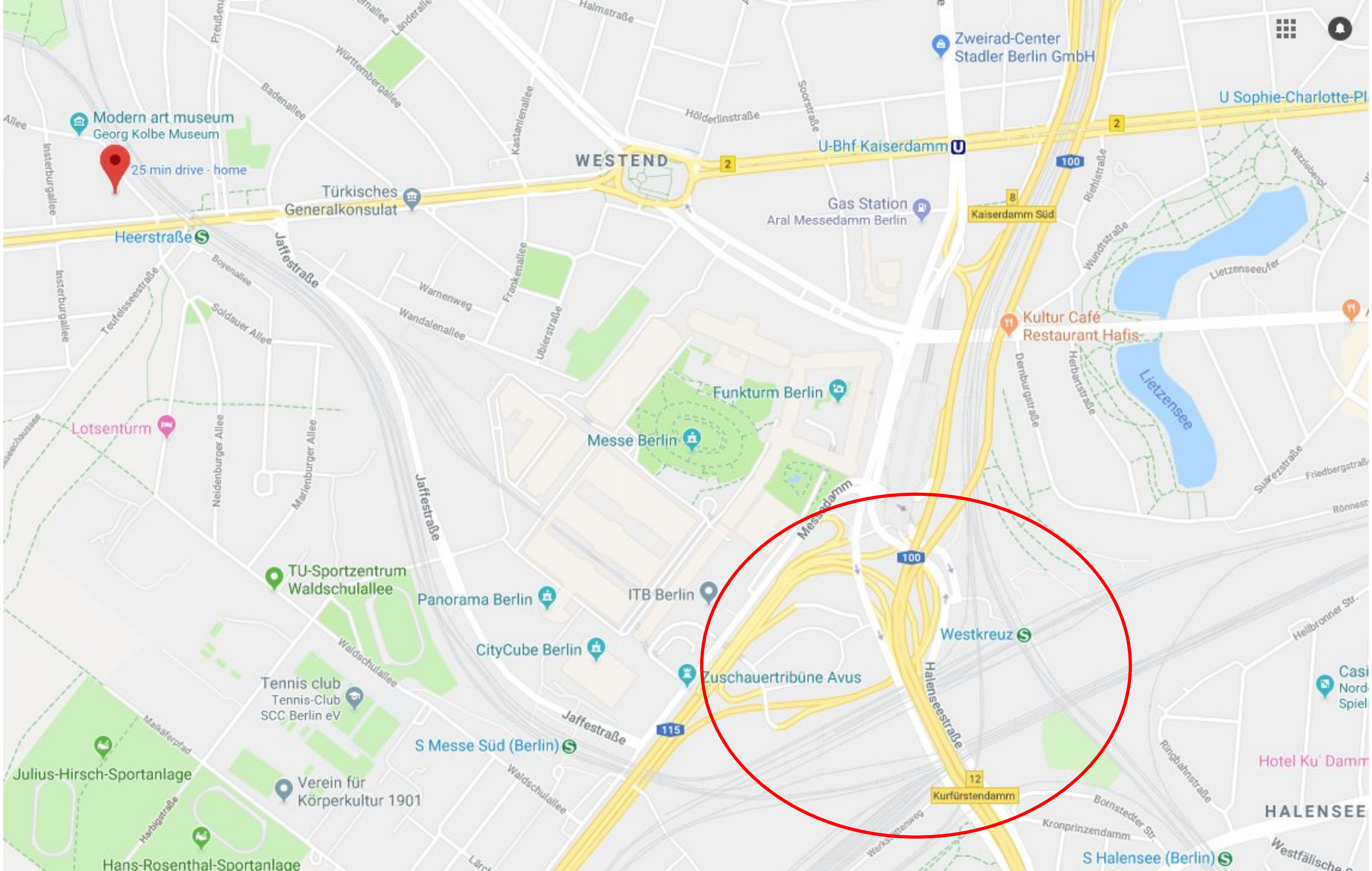
How Roads and Stations Affect Air Quality?

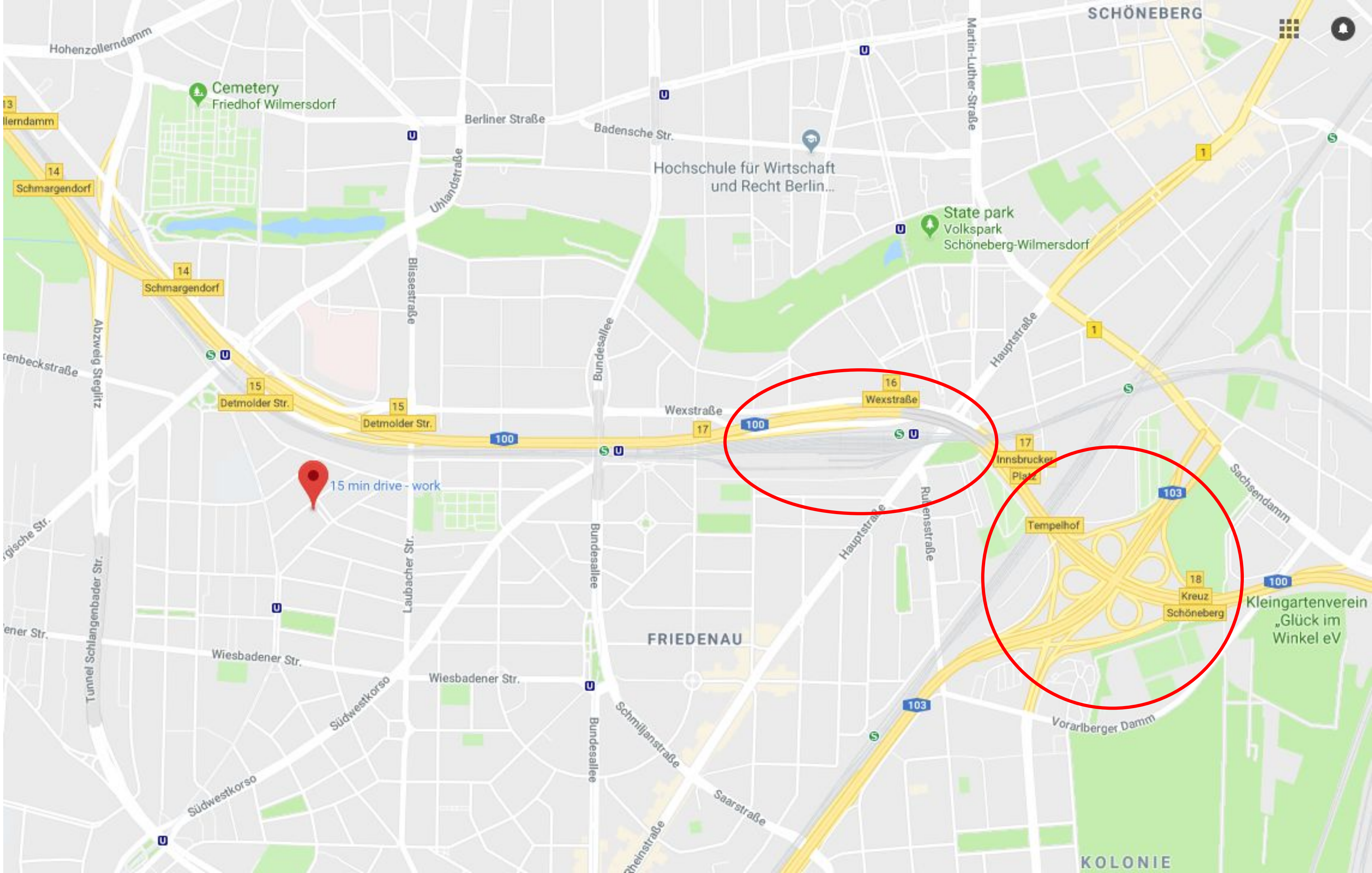
- The most polluted points are close to Ring or main S-Bahn stations in Berlin




Children's Museum
Labyrinth Kindermuseum Berlin
Osloer Str.
Bornholmer Str.
Platz des 9. November 1989
S Bornholmer Str. (Berlin)
Kleingartenanlage Bornholm II eV
Ibsenstraße
Stavangerstraße
Gotlandstraße
Eschengraben
Thulestraße
Schonensche Str.
Wisbyer Str.
Baumbachstraße
Kurze Str.
Neumannstraße
Stettiner Str.
Grüntaler Str.
Jülicher Str.
Eulerstraße
Malmöer Str.
Czarnikauer Str.
Schönfließer Str.
Seelower Str.
Kuglerstraße
Erich-Weinert-Straße
Greifenhagener Str.
Scherenbergstraße
Stahlheimer Str.
Spanheimstraße
Isländische Str.
Driesener Str.
Arnimplatz
Rodenbergstraße
BSR Recyclinghof Behmstraße
Behmstraße
Schivelbeiner Str.
Wichertstraße
Eurogida
Dänenstraße
Schönhauser Allee
n-Gesundbrunnen
SlopeTrotter Skitours
Supermarket
Kaufland Berlin
Bärbel-Bohley-Ring
Evangelical Church
Gethsemanekirche
Ice cream shop
Eispatisserie Hokey Pokey
liner Unterwelten e.V
Schwedter Str.
Ystader Str.
Korsörer Str.
96a
Ramlerstraße
Brunnenstraße
Hamburger restaurant
The Bird
Brazilian restaurant
Villa Rodizio
Gleimstraße
Park
boldthain
9 min drive - home
Buchholzer Str.
Prenzlauer
Max-Schmeling-Halle
Lidl
Putbusser Str.
Graunstraße
Gustav-Meyer-Allee
Sports complex
Friedrich-Ludwig-Jahn-Sportpark
Zeiss-Groß
Home improvement store
OBI Markt Berlin-Wedding
State park
Helmholtzplatz
Mauerpark
Topsstraße
Voltastraße
Flea market
U Eberswalder Str. (Berlin)
Raumerstraße

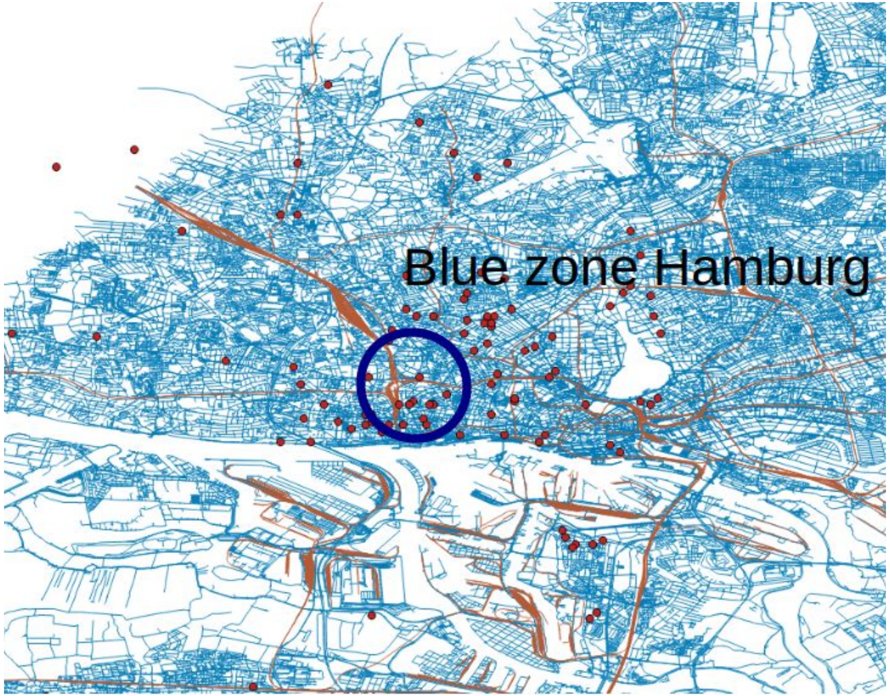

Memorial park
Gedenkstätte Plötzensee
Hüttigpfad
Prison
Justizvollzugsanstalt
Berlin-Plötzensee
Hafenbecken
des Westhafens
Seestraße
Norduler
Eckernförder
Platz
Südstraße
Südring
Föhrer Str.
Fehmarner Str.
Buchstraße
Sprengelstraße
Samoastraße
Tegeler Str.
Nordufer
Nordufer
Government office
Landesamt für Bürger- und...
Friedrich-Krause-Ufer
Westhafenstraße
2
Beusselstraße
Becken II des Westhafens
Becken I des Westhafens
Beusselstraße
S+U Westhafen
Ellen-Epstein-Straße
Quitzowstraße
Perleberger
Lehrter Str.
Berliner Großmarkt
Beusselstraße
Erna-Samuel-Straße
An der Putlitzbrücke
Havelberger Str.
Stephanstraße
Perleberger Str.
Siemensstraße
Oldenburger Str.
Birkenstraße
Mercure Hotel
MOA Berlin
Sickingenstraße
Berlichingenstraße
Rostocker Str.
Waldstraße
Wiclefstraße
Bremer Str.
Bredowstraße
Wilhelmshavener Str.
MOABIT
Wilsnacker Str.
Rathenower Str.
Kruppstraße
Wiebestraße
Beusselstraße
Soccer clu
SC Union Berlin 06 e.
Waldenserstraße
Bandelstraße
Wilsnacker Str.
State park
Fritz-Schloß-Park
Wiebestraße 42
Arminiusmarkthalle
Bugenhagenstraße
5 min drive - work
Turmstraße
Stromstraße
office
Wilmersdorf
Reuchlinstraße
Erasmusstraße
Beusselstraße
Turmstraße
aiserin-Augusta-Allee
Kaiserin
Zwinglistraße
Ottostraße
U-Bhf Turmstraße
State park
Kleiner
Tiergarten
Seydlitzstraße


Modern art museum
Georg Kolbe Museum
25 min drive - home
WESTEND
U-Bhf Kaiserdamm
U Sophie-Charlotte-Pl
Zweirad-Center Stadler Berlin GmbH
Türkisches Generalkonsulat
Heerstraße
Gas Station
Aral Messedamm Berlin
Kaiserdamm Süd
Kultur Café Restaurant Hafis
Lietzensee
Lietzenseeufer
Funkturm Berlin
Messe Berlin
Lotsenturm
TU-Sportzentrum Waldschulallee
Panorama Berlin
ITB Berlin
CityCube Berlin
Zuschauertribüne Avus
Westkreuz
Messedamm
Halenseestraße
Kurfürstendamm
Tennis club
Tennis-Club SCC Berlin eV
S Messe Süd (Berlin)
Julius-Hirsch-Sportanlage
Verein für Körperkultur 1901
Hans-Rosenthal-Sportanlage
Hotel Ku' Damm
HALENSEE
S Halensee (Berlin)
Casi
Nord
Spiel
Jaffestraße
Waldschulallee
Wandalenallee
Warnenweg
Soldauer Allee
Teufelsseestraße
Insterburgallee
Neidenburger Allee
Marienburger Allee
Frankenallee
Ubierstraße
Kastanienallee
Württembergallee
Badenallee
Hölderlinstraße
Soorstraße
Halmstraße
Preußenallee
Länderallee
Riehlstraße
Wundtstraße
Dernburgstraße
Herbartstraße
Witzlebenpl
Suarezstraße
Friedbergstraße
Rönnestr
Heilbronner Str.
Ringbahnstraße
Bornstedter Str
Kronprinzendamm
Westfälische
Werkstättenweg
Malkäferpfad
Harbigstraße
Boyenallee
100
115
2
8
12


SCHÖNEBERG
Hohenzollerndamm
Cemetery
Friedhof Wilmersdorf
Berliner Straße
Badensche Str.
Martin-Luther-Straße
Hochschule für Wirtschaft und Recht Berlin...
State park
Volkspark Schöneberg-Wilmersdorf
Uhlandstraße
Blissestraße
Bundesallee
Hauptstraße
14
Schmargendorf
15
Detmolder Str.
16
Wexstraße
17
Innsbrucker Platz
Tempelhof
18
Kreuz Schöneberg
100
103
1
Abzweig Steglitz
Wexstraße
15 min drive - work
Laubacher Str.
Rubensstraße
Sachsendamm
Kleingartenverein „Glück im Winkel eV
FRIEDENAU
Wiesbadener Str.
Südwestkorso
Tunnel Schlangenbader Str.
Schmiljanstraße
Saarstraße
Rheinstraße
Vorarlberger Damm
KOLONIE

How Do Diesel Bans Affect Pollution?

- 10% local decrease in pollution
- No global impact
- Berlin diesel ban (1st of April 2019)
- Affected streets: e.g. Friedrichstraße
- Due to the locality, diesel bans should address the most polluted roads



Conclusion

- Luftdaten is limited by its own
- Current solutions are not effective due to the dearth of information
- Idea of enriching main dataset with external data sources
- Detected causes of pollution: e.g. public events, weather, air traffic, and etc.
- We built a general pollution explanation system that can be applied on every city

Potential Future Directions

- Exploration of pollution causes
    a. Explore more dimensions, e.g., more cities, more influencing factors,
    b. Use other ML or statistical methods
- Research direction: automated selection additional sources
    a. What are effective heuristics to choose datasets that improve explanation experience?
    b. What types of indexing mechanisms are necessary to make this process efficient?