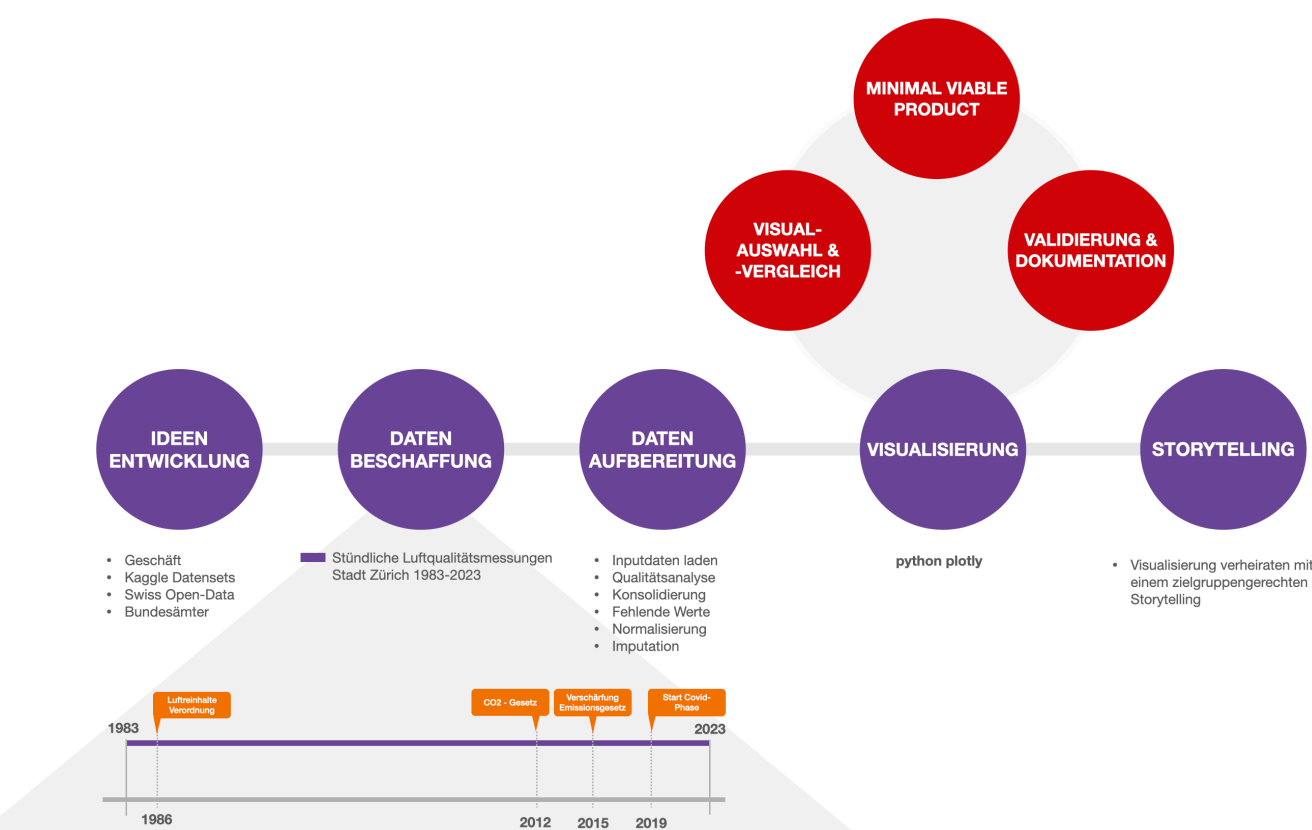


Luftqualitätsmessungen der Stadt Zürich 1983-2023

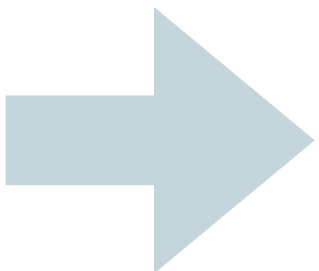
Ausgangslage für die Semesterarbeit ist ein Dataset von opendata.swiss zu stündlichen Luftqualitätsmessungen an ausgewählten Standorten in der Stadt Zürich von 1983-2023:
<https://opendata.swiss/de/dataset/stundlich-aktualisierte-luftqualitaetsmessungen-seit-19831/resource/9430220a-2254-42fa-89c3-563980073482>

direkte Verwendung nicht möglich -> Massnahmen

Datenaufbereitung



Alerts		
Intervall	has constant value ""	Constant
Datum	has a high cardinality: 351754 distinct values	High cardinality
Parameter	is highly overall correlated with Einheit	High correlation
Einheit	is highly overall correlated with Parameter	High correlation
Wert	has 126364 (2.6%) missing values	Missing
Dataset statistics		
Number of variables	7	
Number of observations	4812834	
Missing cells	126364	
Missing cells (%)	0.4%	
Duplicate rows	0	
Duplicate rows (%)	0.0%	
Total size in memory	257.0 MiB	
Average record size in memory	56.0 B	
Variable types		
Categorical	6	
Numeric	1	



Filterung & Pivotierung

- Filterung auf nur Werte mit Status „bereinigt“
- Datum, Standort und Parameter als Spalten, Einheit in Spaltenbeschreibung, restliche Spalten löschen

Imputation fehlender Werte

- Singuläre Imputation
- Substitution durch Lagemass (hier Mittelwert pro Standort)

Projektplanung



Daten-Profiling

(python ydata_profiling)

Einfache Struktur

```
df_air_quality.head(5)
```

	Datum	Standort	Parameter	Intervall	Einheit	Wert	Status
0	1987-01-01T03:00+0100	Zch_Stampfenbachstrasse	CO	h1	mg/m3	0.10	bereinigt
1	1987-01-01T03:00+0100	Zch_Stampfenbachstrasse	NO2	h1	µg/m3	33.42	bereinigt
2	1987-01-01T03:00+0100	Zch_Stampfenbachstrasse	NO	h1	µg/m3	6.18	bereinigt
3	1987-01-01T03:00+0100	Zch_Stampfenbachstrasse	NOx	h1	ppb	22.43	bereinigt
4	1987-01-01T03:00+0100	Zch_Stampfenbachstrasse	O3	h1	µg/m3	39.90	bereinigt

Datenbeschaffung und -exploration

(python pandas, requests)

Datenanalyse & Visualisierung

(python matplotlib, seaborn, plotly, Bokeh)

Storytelling

(Keynote/PDF vs. HTML/javascript/CSS)

Vorhandene Variablen:

- Stündliche Werte für die Luftqualität („Parameter“)
 - Ozon O3
 - Stickoxide NOx
 - Stickstoffmonoxid NO
 - Stickstoffdioxid NO2
 - Feinstaub
 - Kohlenmonoxid CO
 - Schwefeldioxid SO2
- für die Stationen („Standort“)
 - Stampfenbachstrasse
 - Schimmelstrasse
 - Rosengartenstrasse
 - Heubeeribüel
- inkl. Status
 - bereinigt / provisorisch