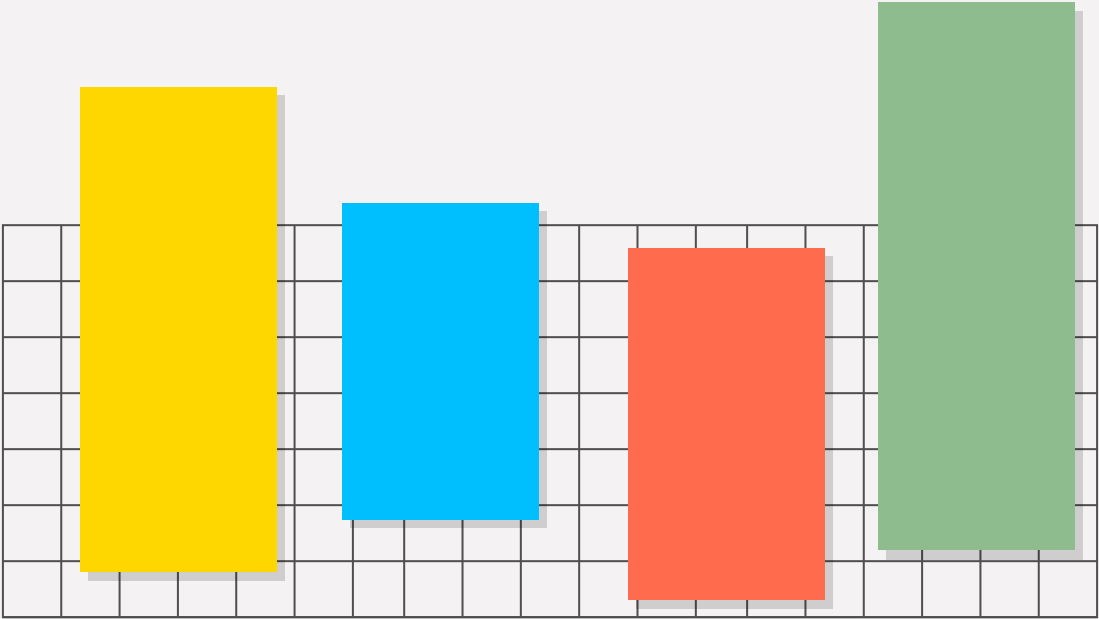




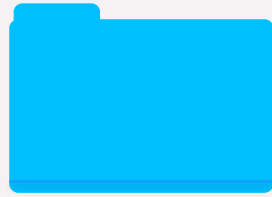
The Discovery of Handwashing

> Rekonstruktion der Forschung von Ignaz Semmelweis anhand einer datenbasierten Analyse mit Python





Annahme



Vorgehen

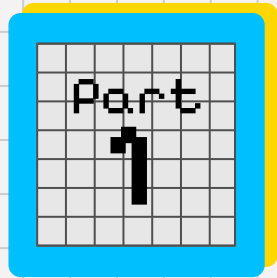


Auswertung



Fazit

Annahme



> Es wird angenommen, dass es eine erhöhte Sterblichkeitsrate in einer der beiden Kliniken, aufgrund der direkten Arbeit in der Autopsie, ohne Händewaschen gibt.

Vorgehen



(12, 4)

48

<class 'pandas.core.frame.DataFrame'>

RangeIndex: 12 entries, 0 to 11

Data columns (total 4 columns):

Column Non-Null Count Dtype

0 year 12 non-null int64
1 births 12 non-null int64
2 deaths 12 non-null int64
3 clinic 12 non-null object

dtypes: int64(3), object(1)

memory usage: 512.0+ bytes

None

	year	births	deaths	clinic
0	1841	3036	237	clinic 1
1	1842	3287	518	clinic 1
2	1843	3060	274	clinic 1
3	1844	3157	260	clinic 1
4	1845	3492	241	clinic 1

Dataset einlesen:

file_path = "yearly_deaths_by_clinic.csv"

df_yearly = pd.read_csv(file_path)

```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

notwendige
libraries



year	births	deaths	clinic
1841	3036	237	clinic 1
1842	3287	518	clinic 1
1843	3060	274	clinic 1
1844	3157	260	clinic 1
1845	3492	241	clinic 1
1846	4010	459	clinic 1
1841	2442	86	clinic 2
1842	2659	202	clinic 2
1843	2739	164	clinic 2
1844	2956	68	clinic 2
1845	3241	66	clinic 2
1846	3754	105	clinic 2

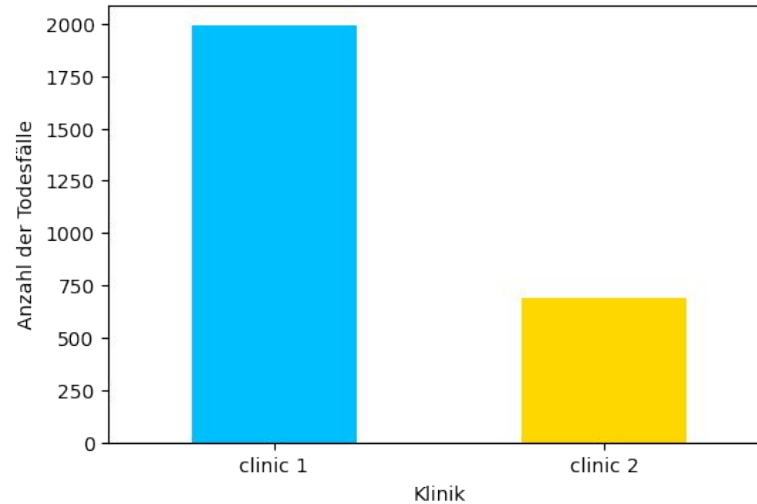
Todesfälle pro Klinik

Klinik 1: 1989

Klinik 2: 691



Gesamte Todesfälle pro Klinik

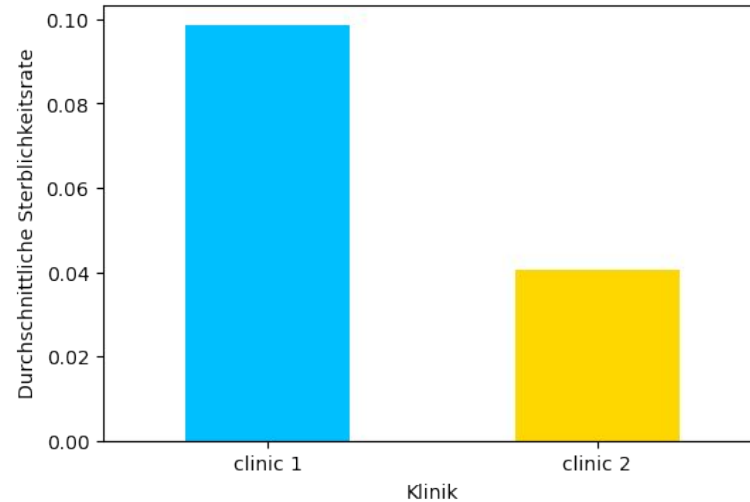


Sterblichkeitsrate

Klinik 1: 0.098505

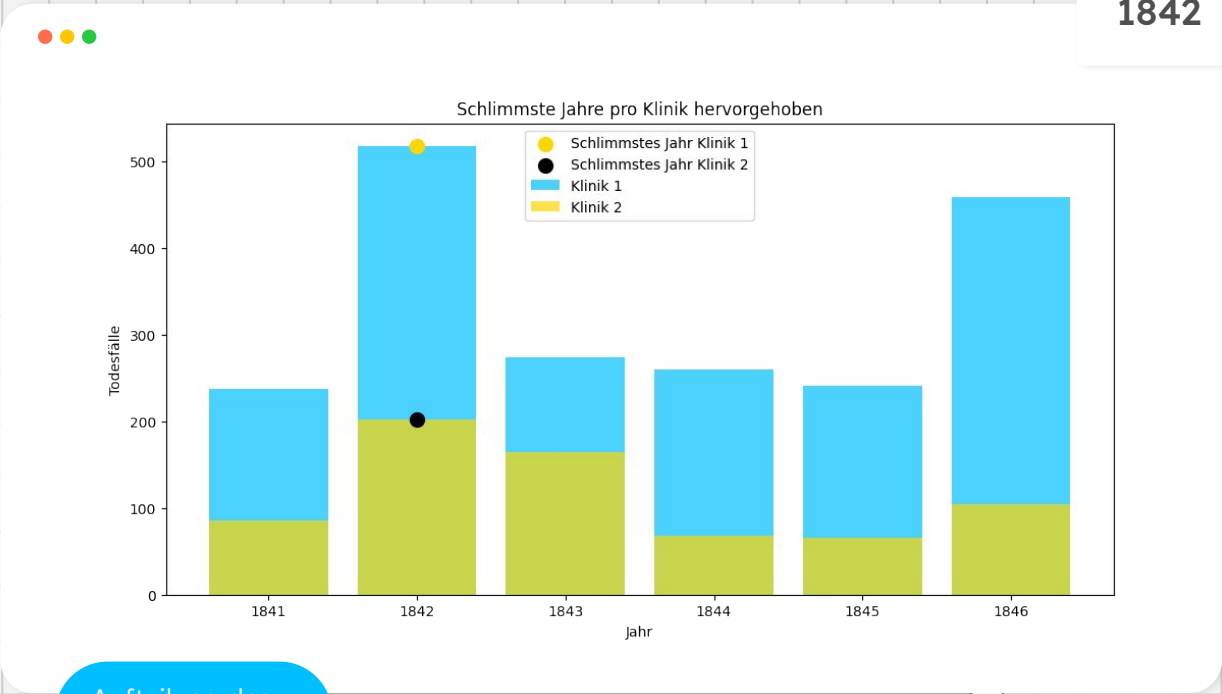
Klinik 2: 0.040400

Durchschnittliche Sterblichkeitsrate pro Klinik



Most hectic years

```
worst_year_clinic_1 = clinic_1.loc[clinic_1["deaths"].idxmax()]\nworst_year_clinic_2 = clinic_2.loc[clinic_2["deaths"].idxmax()]
```



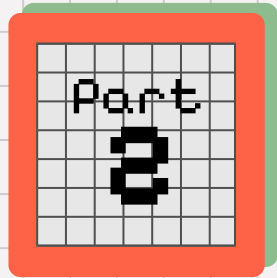
Worst Year

1842

Aufteilung des Datasets

```
clinic_1 = df_yearly[df_yearly["clinic"] == "clinic 1"]\nclinic_2 = df_yearly[df_yearly["clinic"] == "clinic 2"]
```


Annahme



> Es wird angenommen, dass sich die Todesrate, durch die Einführung des Händewaschens in beiden Kliniken, verringert.

Vorgehen

> Dataset einlesen:

```
file_path = "monthly_deaths.csv"
monthly_data = pd.read_csv(file_path)
```

Tode pro Monat berechnen:

```
monthly_data["death_proportion"] = monthly_data["deaths"] /
monthly_data["births"]
```

Änderung vom Datum von string zu datetime:

```
monthly_data["date"] = pd.to_datetime(monthly_data["date"])
```

Festlegung von HW-Startdatum:

```
handwashing_start = pd.to_datetime("1847-06-01")
```

1841-01-01 00:00:00	254	37	0.14566929133858267	False
1841-02-01 00:00:00	239	18	0.07531380753138076	False
1841-03-01 00:00:00	277	12	0.04332129963898917	False
1841-04-01 00:00:00	255	4	0.01568627450980392	False
1841-05-01 00:00:00	255	2	0.00784313725490196	False
1841-06-01 00:00:00	200	10	0.05	False
1841-07-01 00:00:00	190	16	0.08421052631578947	False
1841-08-01 00:00:00	222	3	0.013513513513513514	False
1841-09-01 00:00:00	213	4	0.018779342723004695	False
1841-10-01 00:00:00	236	26	0.11016949152542373	False
1841-11-01 00:00:00	235	53	0.225531914893617	False
1842-01-01 00:00:00	307	64	0.20846905537459284	False
1842-02-01 00:00:00	311	38	0.12218649517684887	False
1842-03-01 00:00:00	264	27	0.10227272727272728	False
1842-04-01 00:00:00	242	26	0.10743801652892562	False
1842-05-01 00:00:00	310	10	0.03225806451612903	False
1842-06-01 00:00:00	273	18	0.06593406593406594	False
1842-07-01 00:00:00	231	48	0.2077922077922078	False
1842-08-01 00:00:00	216	55	0.25462962962962965	False
1842-09-01 00:00:00	223	41	0.18385650224215247	False
1842-10-01 00:00:00	242	71	0.29338842975206614	False
1842-11-01 00:00:00	209	48	0.22966507177033493	False
1842-12-01 00:00:00	239	75	0.3138075313807531	False
1843-01-01 00:00:00	272	52	0.19117647058823528	False
1843-02-01 00:00:00	263	42	0.1596958174904943	False
1843-03-01 00:00:00	266	33	0.12406015037593984	False
1843-04-01 00:00:00	285	34	0.11929824561403508	False
1843-05-01 00:00:00	246	15	0.06097560975609756	False
1843-06-01 00:00:00	196	8	0.04081632653061224	False
1843-07-01 00:00:00	191	1	0.005235602094240838	False
1843-08-01 00:00:00	193	3	0.015544041450777202	False
1843-09-01 00:00:00	221	5	0.02262443438914027	False
1843-10-01 00:00:00	250	44	0.176	False
1843-11-01 00:00:00	252	18	0.07142857142857142	False
1843-12-01 00:00:00	236	19	0.08050847457627118	False
1844-01-01 00:00:00	244	37	0.15163934426229508	False
1844-02-01 00:00:00	257	29	0.11284046692607004	False
1844-03-01 00:00:00	276	47	0.17028985507246377	False
1844-04-01 00:00:00	208	36	0.17307692307692307	False
1844-05-01 00:00:00	240	14	0.058333333333333334	False
1844-06-01 00:00:00	224	6	0.026785714285714284	False
1844-07-01 00:00:00	206	9	0.043689320388349516	False
1844-08-01 00:00:00	269	17	0.06319702602230483	False
1844-09-01 00:00:00	245	3	0.012244897959183673	False
1844-10-01 00:00:00	248	8	0.03225806451612903	False
1844-11-01 00:00:00	245	27	0.11020408163265306	False
1844-12-01 00:00:00	256	27	0.10546875	False

Sterblichkeitsraten vor und nach dem Händewaschen

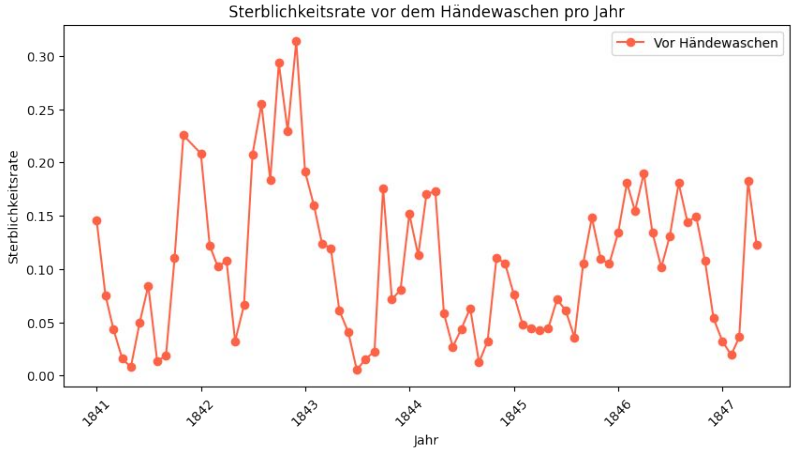
Maximum vor HW:
0.3138075313807531
am 1842-12-01

vor und nach dem Händewaschen

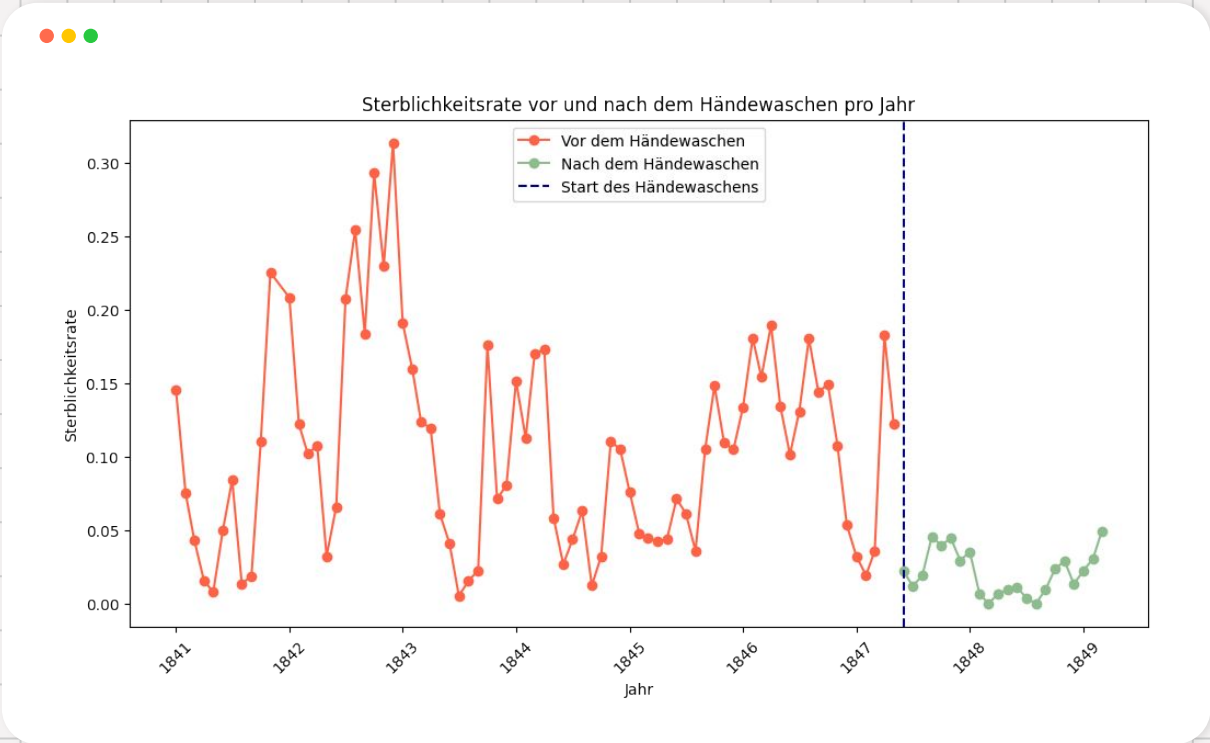
```
before_handwashing = monthly_data[monthly_data["handwashing"] == 0]
```



Maximum nach HW:
0.04926108374384237
am 1849-03-01

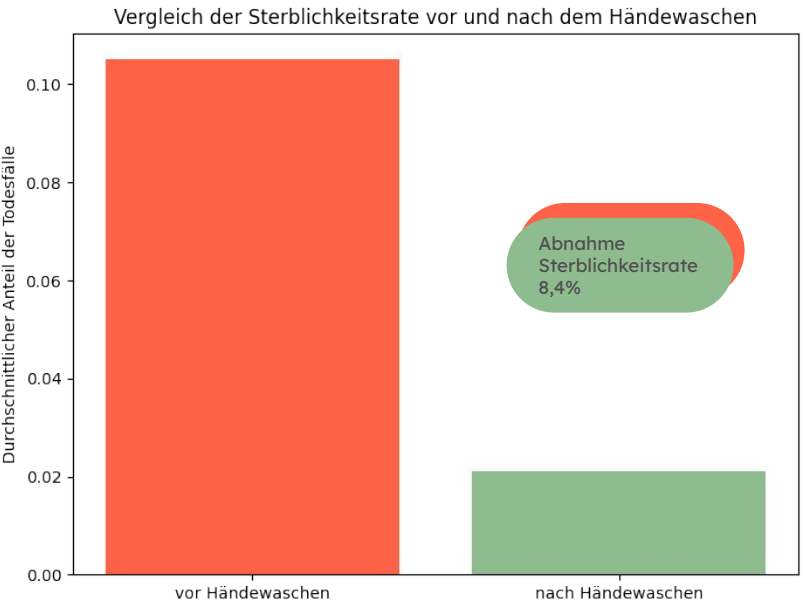


Beide Graphen in einem Diagramm zur Veranschaulichung



Peaks in Prozent:
Vor HW = 31,4 %
Nach HW = 4,9 %

Auswertung

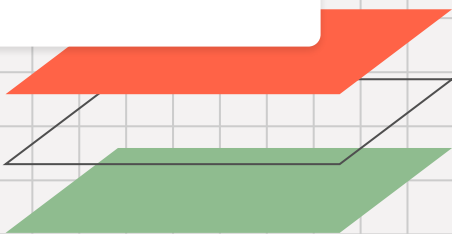


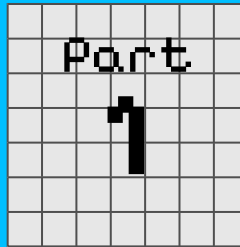
Durchschnitte & Abnahme berechnen:

```
durchschnittliche_Todesfälle_vor = before_handwashing["death_proportion"].mean()  
durchschnittliche_Todesfälle_nach = after_handwashing["death_proportion"].mean()  
Abnahme_durchschnittliche_Todesfälle = durchschnittliche_Todesfälle_vor - durchschnittliche_Todesfälle_nach
```



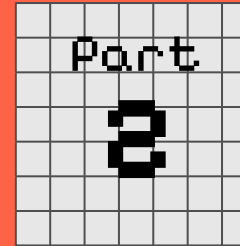
Durchschnitt. vor HW: 0.105 => 10,5 %
Durchschnitt. nach HW: 0.021 => 2,1 %
Abnahme durchschnitt. Sterblichkeitsrate:
0.084





Fazit

Die Analyse des Datasets bestätigt, die Annahme, dass durch die Arbeit in der Autopsie in Klinik 1, mehr gebärende Frauen mit dem Kindbettfieber infiziert wurden. Dies ist der Grund für mehr Todesfälle in Klinik 1 im Vergleich zu Klinik 2.



Fazit

Die Analyse des Datasets bestätigt, die Annahme, dass die Todesrate durch die Einführung des Händewaschens in beiden Kliniken deutlich zurückgegangen ist.

+

Cam Do

c.do@tu-braunschweig.de

+

Anna Grünewald

a.gruenewald@tu-braunschweig.de

+

Fabian Rothberger

f.rothberger@tu-braunschweig.de

+

Flemming Schur

flemming.schur@tu-braunschweig.de

