

Wandervogel

June 10, 2017

```
In [1]: # Import wichtiger Python-Module für Datenanalyse
```

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from pandas import DataFrame, Series

import re
```

```
In [2]: # Einlesen der Wandervogel-Metadaten
```

```
wandervogel_df = pd.read_csv('metadatenwandervogelf5000.csv')
```

```
# Einlesen der Liste der Koordinaten von Schweizer Ortschaften
```

```
koord_df = pd.read_csv('PLZO_CSV_LV03.csv', sep=';')
```

```
In [3]: # Sample Liste Wandervogel
```

```
wandervogel_df
```

```
Out[3]:
```

	Signatur	Titel
0	Sozarch_F_5000-Dx-431	Kreistanz um ein loderndes Feuer
1	Sozarch_F_5000-Dx-432	Prellen: Spiel im Freien - Kinder und Erwachsene
2	Sozarch_F_5000-Dx-433	Vier Wandervögel sitzen auf einer Holzbank vor
3	Sozarch_F_5000-Dx-434	Kreistanz um ein loderndes Feuer
4	Sozarch_F_5000-Dx-435	Mann Gitarre spielen
5	Sozarch_F_5000-Dx-436	Gruppe von Wandervögeln im Freien: Mädchen Gitarre
6	Sozarch_F_5000-Dx-437	Zwei Personen sitzen am Tisch und essen (Mahlzeit)
7	Sozarch_F_5000-Dx-438	Winterstimmung: Junge Männer auf der Strasse und
8	Sozarch_F_5000-Dx-439	Junge Männer beim Essen
9	Sozarch_F_5000-Dx-440	Junge Männer bei einer Rast auf einem Brunnen
10	Sozarch_F_5000-Fx-001	Gruppenporträt vor der Wandervogel-Landsgemeinde
11	Sozarch_F_5000-Fx-002	Versammelte Wandervögel an der Wandervogel-Landsgemeinde
12	Sozarch_F_5000-Fx-003	Tanzende Wandervögel an der Wandervogel-Landsgemeinde
13	Sozarch_F_5000-Fx-004	Tanzende Wandervögel an der Wandervogel-Landsgemeinde
14	Sozarch_F_5000-Fx-006	Wandervogel-Landsgemeinde auf dem Bantiger bei
15	Sozarch_F_5000-Fx-007	Wandervogel-Landsgemeinde auf dem Bantiger bei
16	Sozarch_F_5000-Fx-008	Wandervogel-Landsgemeinde auf dem Bantiger bei
17	Sozarch_F_5000-Fx-009	Wandervogel-Landsgemeinde auf dem Bantiger bei
18	Sozarch_F_5000-Fx-010	Seilziehen an der Wandervogel-Landsgemeinde (1.

19	Sozarch_F_5000-Fx-011	Wandervogel-Landsgemeinde auf dem Bantiger bei.
20	Sozarch_F_5000-Fx-012	Wandervogel-Landsgemeinde auf dem Bantiger bei.
21	Sozarch_F_5000-Fx-013	Wandervogel-Landsgemeinde bei Steinmaur (Heitl.
22	Sozarch_F_5000-Fx-014	Wandervogel-Landsgemeinde bei Steinmaur (Heitl.
23	Sozarch_F_5000-Fx-015	Wandervogel-Landsgemeinde bei Steinmaur (Heitl.
24	Sozarch_F_5000-Fx-016	Wandervogel-Landsgemeinde bei Steinmaur (Heitl.
25	Sozarch_F_5000-Fx-018	Wandervogel-Landsgemeinde bei Steinmaur (Heitl.
26	Sozarch_F_5000-Fx-019	Wandervogel-Landsgemeinde beim Schloss Buonas .
27	Sozarch_F_5000-Fx-020	Wandervogel-Landsgemeinde beim Schloss Buonas .
28	Sozarch_F_5000-Fx-021	Wandervogel-Landsgemeinde beim Schloss Buonas .
29	Sozarch_F_5000-Fx-022	Wandervogel-Landsgemeinde beim Schloss Buonas .
..	...	
568	Sozarch_F_5000-Nx-446	Ostschweizer Wandervögel unterhalten sich (ver.
569	Sozarch_F_5000-Nx-447	Eine Fahrt zusammen mit den Flawilern auf den .
570	Sozarch_F_5000-Nx-448	Eine Fahrt zusammen mit den Flawilern auf den .
571	Sozarch_F_5000-Nx-449	Eine Fahrt zusammen mit den Flawilern auf den .
572	Sozarch_F_5000-Nx-450	Eine Fahrt zusammen mit den Flawilern auf den .
573	Sozarch_F_5000-Nx-451	Skihütte Ober Altberg Jbergereg (sic) - Altwä.
574	Sozarch_F_5000-Nx-452	Teilansicht der Skihütte Ober Altberg Jbergere.
575	Sozarch_F_5000-Nx-454	Treffen von Wandervögeln am Walensee - Mahlzei.
576	Sozarch_F_5000-Nx-455	Hütten-Usrumete Prestenburg - Wandervögel vor .
577	Sozarch_F_5000-Nx-456	Hütten-Usrumete Prestenburg - Wandervögel vor .
578	Sozarch_F_5000-Nx-457	Treffen von Wandervögeln am Walensee - Wanderv.
579	Sozarch_F_5000-Px-005	Wandervogel-Landsgemeinde: Auflistung der Vera.
580	Sozarch_F_5000-Px-274	Erster Briefkopf des Schweizer Wandervogels (1.
581	Sozarch_F_5000-Px-480	Abgebildete "Wandervögel" beim Haar waschen am.
582	Sozarch_F_5000-Zx-073	Schweizer Wandervogel Landsgemeinde Beinwiel [.
583	Sozarch_F_5000-Zx-481	Zeichnung aus der Sammlung von Gottfried Gries.
584	Sozarch_F_5000-Zx-562	Abbildung eines Wandervogels Blockflöte spiele.
585	Sozarch_F_5000-Zx-590	Drucke/Linolschnitte Wandervogel: Abbildung de.
586	Sozarch_F_5000-Zx-591	Drucke/Linolschnitte Wandervogel: Abbildung ei.
587	Sozarch_F_5000-Zx-592	Drucke/Linolschnitte Wandervogel: Abbildung de.
588	Sozarch_F_5000-Zx-593	Drucke/Linolschnitte Wandervogel: Abbildung ei.
589	Sozarch_F_5000-Zx-594	Drucke/Linolschnitte Wandervogel: Abbildung ei.
590	Sozarch_F_5000-Zx-595	Drucke/Linolschnitte Wandervogel: Abbildung ei.
591	Sozarch_F_5000-Zx-596	Drucke/Linolschnitte Wandervogel: Abbildung ei.
592	Sozarch_F_5000-Zx-597	Drucke/Linolschnitte Wandervogel: Abbildung fl.
593	Sozarch_F_5000-Zx-598	Drucke/Linolschnitte Wandervogel: Abbildung ei.
594	Sozarch_F_5000-Zx-599	Drucke/Linolschnitte Wandervogel: Abbildung ei.
595	Sozarch_F_5000-Zx-600	Drucke/Linolschnitte Wandervogel: Abbildung de.
596	Sozarch_F_5000-Zx-601	Drucke/Linolschnitte Wandervogel: Abbildung ei.
597	Sozarch_F_5000-Zx-602	Drucke/Linolschnitte Wandervogel: Abbildung fü.

		Zeit \
0	Neuzeit;20. Jh.;1901-1950;1921-1930	/Neuzeit;2...
1	Neuzeit;20. Jh.;1901-1950;1921-1930	/Neuzeit;2...
2	Neuzeit;20. Jh.;1901-1950;1921-1930	/Neuzeit;2...
3	Neuzeit;20. Jh.;1901-1950;1921-1930	/Neuzeit;2...

4 Neuzeit;20. Jh.;1901-1950;1921-1930 /Neuzeit;2...
 5 Neuzeit;20. Jh.;1901-1950;1921-1930 /Neuzeit;2...
 6 Neuzeit;20. Jh.;1901-1950;1921-1930 /Neuzeit;2...
 7 Neuzeit;20. Jh.;1901-1950;1921-1930 /Neuzeit;2...
 8 Neuzeit;20. Jh.;1901-1950;1921-1930 /Neuzeit;2...
 9 Neuzeit;20. Jh.;1901-1950;1921-1930 /Neuzeit;2...
 10 Neuzeit;20. Jh.;1901-1950;1901-1910;1910
 11 Neuzeit;20. Jh.;1901-1950;1901-1910;1910
 12 Neuzeit;20. Jh.;1901-1950;1911-1920;1913
 13 Neuzeit;20. Jh.;1901-1950;1911-1920;1913
 14 Neuzeit;20. Jh.;1901-1950;1911-1920;1914
 15 Neuzeit;20. Jh.;1901-1950;1911-1920;1914
 16 Neuzeit;20. Jh.;1901-1950;1911-1920;1914
 17 Neuzeit;20. Jh.;1901-1950;1911-1920;1914
 18 Neuzeit;20. Jh.;1901-1950;1911-1920;1914 /Neuz...
 19 Neuzeit;20. Jh.;1901-1950;1911-1920;1914
 20 Neuzeit;20. Jh.;1901-1950;1911-1920;1914
 21 Neuzeit;20. Jh.;1901-1950;1911-1920;1915
 22 Neuzeit;20. Jh.;1901-1950;1911-1920;1915
 23 Neuzeit;20. Jh.;1901-1950;1911-1920;1915
 24 Neuzeit;20. Jh.;1901-1950;1911-1920;1915
 25 Neuzeit;20. Jh.;1901-1950;1911-1920;1915
 26 Neuzeit;20. Jh.;1901-1950;1911-1920;1916
 27 Neuzeit;20. Jh.;1901-1950;1911-1920;1916
 28 Neuzeit;20. Jh.;1901-1950;1911-1920;1916
 29 Neuzeit;20. Jh.;1901-1950;1911-1920;1916

 568 Neuzeit;20. Jh.;1901-1950;1931-1940
 569 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 570 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 571 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 572 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 573 Neuzeit;20. Jh.;1901-1950;1941-1950;1944
 574 Neuzeit;20. Jh.;1901-1950;1941-1950;1944
 575 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 576 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 577 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 578 Neuzeit;20. Jh.;1901-1950;1931-1940;1938
 579 Neuzeit;20. Jh.;1901-1950;1911-1920;1911
 580 Neuzeit;20. Jh.;1901-1950;1911-1920;1920
 581 Neuzeit;20. Jh.;1901-1950;1911-1920;1918
 582 Neuzeit;20. Jh.;1901-1950;1931-1940;1939
 583 Neuzeit;20. Jh.;1901-1950;1911-1920;1918
 584 Neuzeit;20. Jh.;1901-1950;1911-1920;1920
 585 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 586 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 587 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 588 Neuzeit;20. Jh.;1951-2000;1951-1960;1954

589 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 590 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 591 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 592 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 593 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 594 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 595 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 596 Neuzeit;20. Jh.;1951-2000;1951-1960;1954
 597 Neuzeit;20. Jh.;1951-2000;1951-1960;1954

	Ort
0	Europa;Schweiz
1	Europa;Schweiz
2	Europa;Schweiz
3	Europa;Schweiz
4	Europa;Schweiz
5	Europa;Schweiz
6	Europa;Schweiz
7	Europa;Schweiz
8	Europa;Schweiz
9	Europa;Schweiz
10	Europa;Schweiz;Zürich, Kanton;Kyburg
11	Europa;Schweiz;Zürich, Kanton;Kyburg
12	Europa;Schweiz;Thurgau;Märstetten
13	Europa;Schweiz;Thurgau;Märstetten
14	Europa;Schweiz;Bern, Kanton;Bolligen
15	Europa;Schweiz;Bern, Kanton;Bolligen
16	Europa;Schweiz;Bern, Kanton;Bolligen
17	Europa;Schweiz;Bern, Kanton;Bolligen
18	Europa;Schweiz
19	Europa;Schweiz;Bern, Kanton;Bolligen
20	Europa;Schweiz;Bern, Kanton;Bolligen
21	Europa;Schweiz;Zürich, Kanton;Steinmau
22	Europa;Schweiz;Zürich, Kanton;Steinmau
23	Europa;Schweiz;Zürich, Kanton;Steinmau
24	Europa;Schweiz;Zürich, Kanton;Steinmau
25	Europa;Schweiz;Zürich, Kanton;Steinmau
26	Europa;Schweiz;Zug, Kanton;Risch
27	Europa;Schweiz;Zug, Kanton;Risch
28	Europa;Schweiz;Zug, Kanton;Risch
29	Europa;Schweiz;Zug, Kanton;Risch
..	...
568	Europa;Schweiz
569	Europa;Schweiz;St. Gallen, Kanton;Flaw
570	Europa;Schweiz;St. Gallen, Kanton;Flaw
571	Europa;Schweiz;St. Gallen, Kanton;Flaw
572	Europa;Schweiz;St. Gallen, Kanton;Flaw
573	Europa;Schweiz;Schwyz, Kanton

```

574      Europa;Schweiz;Schwyz, Kanton
575      Europa;Schweiz;St. Gallen, Kanton
576      Europa;Schweiz
577      Europa;Schweiz
578      Europa;Schweiz;St. Gallen, Kanton
579      Europa;Schweiz;St. Gallen, Kanton;Rapp
580      Europa;Schweiz
581      Europa;Schweiz
582      Europa;Schweiz;Aargau;Beinwil am See
583      Europa;Schweiz
584      Europa;Schweiz
585      Europa;Schweiz
586      Europa;Schweiz
587      Europa;Schweiz
588      Europa;Schweiz
589      Europa;Schweiz
590      Europa;Schweiz
591      Europa;Schweiz
592      Europa;Schweiz
593      Europa;Schweiz
594      Europa;Schweiz
595      Europa;Schweiz
596      Europa;Schweiz
597      Europa;Schweiz

```

```
[598 rows x 4 columns]
```

```
In [4]: # Sample Liste Ortschaften
        koord_df
```

```

Out[4]:
      Ortschaftsname  PLZ  Zusatzziffer  Gemeindename  BFS-M
0      Aeugst am Albis  8914           0      Aeugst am Albis
1      Aeugstertal    8914           2      Aeugst am Albis
2      Zwillikon     8909           0  Affoltern am Albis
3  Affoltern am Albis  8910           0  Affoltern am Albis
4      Bonstetten    8906           0      Bonstetten
5      Hausen am Albis  8915           0      Hausen am Albis
6      Ebertswil     8925           0      Hausen am Albis
7      Hedingen     8908           0      Hedingen
8      Kappel am Albis  8926           0      Kappel am Albis
9      Hauptikon     8926           2      Kappel am Albis
10     Uerzlikon     8926           3      Kappel am Albis
11      Knonau       8934           0      Knonau
12     Maschwanden   8933           0     Maschwanden
13     Mettmenstetten  8932           0     Mettmenstetten
14     Obfelden     8912           0     Obfelden
15     Ottenbach    8913           0     Ottenbach
16     Rifferswil    8911           0     Rifferswil

```

17	Stallikon	8143	0	Stallikon	1
18	Uetliberg	8143	2	Stallikon	1
19	Wettswil	8907	0	Wettswil am Albis	1
20	Adlikon b. Andelfingen	8452	0	Adlikon	2
21	Benken ZH	8463	0	Benken (ZH)	2
22	Berg am Irchel	8415	0	Berg am Irchel	2
23	Gräslikon	8415	1	Berg am Irchel	2
24	Buch am Irchel	8414	0	Buch am Irchel	2
25	Dachsen	8447	0	Dachsen	2
26	Dorf	8458	0	Dorf	2
27	Feuerthalen	8245	0	Feuerthalen	2
28	Langwiesen	8246	0	Feuerthalen	2
29	Flaach	8416	0	Flaach	2
...
4125	Montenol	2884	0	Clos du Doubs	680
4126	Epauvillers	2885	0	Clos du Doubs	680
4127	Epiquerez	2886	0	Clos du Doubs	680
4128	Seleute	2888	0	Clos du Doubs	680
4129	Ocourt	2889	0	Clos du Doubs	680
4130	Chevenez	2906	0	Haute-Ajoie	680
4131	Réclère	2912	2	Haute-Ajoie	680
4132	Roche-d'Or	2912	3	Haute-Ajoie	680
4133	Damvant	2914	0	Haute-Ajoie	680
4134	Miécourt	2946	0	La Baroche	681
4135	Charmoille	2947	0	La Baroche	681
4136	Fregiécourt	2953	2	La Baroche	681
4137	Pleujouse	2953	3	La Baroche	681
4138	Asuel	2954	0	La Baroche	681
4139	Vaduz	9490	0	Vaduz	700
4140	Triesen	9495	0	Triesen	700
4141	Balzers	9496	0	Balzers	700
4142	Triesenberg	9497	0	Triesenberg	700
4143	Schaan	9494	0	Schaan	700
4144	Planken	9498	0	Planken	700
4145	Nendeln	9485	0	Eschen	700
4146	Eschen	9492	0	Eschen	700
4147	Schaanwald	9486	0	Mauren	700
4148	Mauren FL	9493	0	Mauren	700
4149	Gamprin-Bendern	9487	0	Gamprin	700
4150	Ruggell	9491	0	Ruggell	701
4151	Schellenberg	9488	0	Schellenberg	701
4152	Thunersee	9999	1	Thunersee	907
4153	Brienzersee	9999	2	Brienzersee	908
4154	Bielersee	9999	0	Bielersee (BE)	914

	Kantonskürzel	E	N
0	ZH	679435	235653
1	ZH	679815	237410

2	ZH	675104	237891
3	ZH	676439	236967
4	ZH	677798	241025
5	ZH	682996	232962
6	ZH	684168	231177
7	ZH	676518	239054
8	ZH	682261	231285
9	ZH	680023	231590
10	ZH	680340	230665
11	ZH	677402	230822
12	ZH	674785	232061
13	ZH	677613	232938
14	ZH	674501	235214
15	ZH	673059	237325
16	ZH	680092	233046
17	ZH	679550	242200
18	ZH	679541	244825
19	ZH	678320	243535
20	ZH	694409	270774
21	ZH	691309	278717
22	ZH	687248	269506
23	ZH	687807	268053
24	ZH	689122	267148
25	ZH	688269	280254
26	ZH	690995	269699
27	ZH	690270	282950
28	ZH	691949	282202
29	ZH	687597	269831
...
4125	JU	578339	244708
4126	JU	575584	242753
4127	JU	571221	242047
4128	JU	575284	246188
4129	JU	572746	244626
4130	JU	566980	249085
4131	JU	561107	247330
4132	JU	563394	246376
4133	JU	559098	247002
4134	JU	580216	252840
4135	JU	582573	252385
4136	JU	581565	251434
4137	JU	582806	251379
4138	JU	582684	250006
4139	LI	758143	224107
4140	LI	758755	219339
4141	LI	756750	214756
4142	LI	759752	220689
4143	LI	756734	226749

4144	LI	759689	228165
4145	LI	759518	229631
4146	LI	757709	230755
4147	LI	760824	231314
4148	LI	759540	231915
4149	LI	756594	231421
4150	LI	758240	234339
4151	LI	759966	233616
4152	BE	621182	170795
4153	BE	640931	175396
4154	BE	580262	215169

[4155 rows x 8 columns]

```
In [5]: # Das genauste Datum scheint immer das Datum nach dem letzten Strichpunkt zu sein
# Doch als erstes müssen alle NaNs im DataFrame durch leere Strings ersetzt werden
wandervogel_df.fillna('', inplace=True)

set([l[-1] for l in wandervogel_df.Zeit.str.split(';')])
```

```
Out[5]: {'',
'1901-1910',
'1906',
'1907',
'1908',
'1909',
'1910',
'1911',
'1911-1920',
'1912',
'1913',
'1914',
'1915',
'1916',
'1917',
'1918',
'1919',
'1920',
'1921',
'1921-1930',
'1922',
'1923',
'1924',
'1925',
'1926',
'1927',
'1928',
'1929',
```



```
'1930',
'1931',
'1931-1940',
'1932',
'1933',
'1935',
'1936',
'1937',
'1938',
'1939',
'1941',
'1942',
'1944',
'1945',
'1949',
'1951-1960',
'1954',
'1955'}
```

```
In [6]: # Die These scheint zu stimmen. Uns interessiert nur die genaueste Zeitangabe
# also ersetzen den jeweiligen String durch diese
wandervogel_df.Zeit.replace('.*;', '', regex=True, inplace=True)
```

```
In [7]: # Als nächstes soll die Ortsspalte in der Wandervogel-Tabelle in Kontinent,
# Als Trenner zwischen diesen Entitäten dient offenbar der Strichpunkt.
# Nicht alle Einträge haben offenbar alle vier Kategorien Kontinent, Land,
# zuerst soll aber noch überprüft werden, ob es Einträge gibt, welche mehr
vector = [len(l) > 4 for l in wandervogel_df.Ort.str.split(';')]
wandervogel_df[vector]
```

```
Out [7]:
```

	Signatur	Titel
121	Sozarch_F_5000-Fx-115	Wandervögel der Ortsgruppe Aarau-Stadt - Sing-
125	Sozarch_F_5000-Fx-119	Wandervögel der Ortsgruppe Aarau-Stadt - "Drec-
126	Sozarch_F_5000-Fx-120	Wandervögel der Ortsgruppe Aarau-Stadt - "Drec-
134	Sozarch_F_5000-Fx-128	Burgruine Oberreinach (Römerswil): Wandervogel.
140	Sozarch_F_5000-Fx-134	Fahrt der Mädchen der Wandervogelortsgruppe St.
184	Sozarch_F_5000-Fx-178	Giacomopass - Grosse Fahrt von Emmi Bloch (188.
185	Sozarch_F_5000-Fx-179	Abends bei den Morasthütten - Wandervögel sitz.
190	Sozarch_F_5000-Fx-184	Lagerleben: Wandervögel im Tessin oder im Wall.
349	Sozarch_F_5000-Fx-347	Gruppenporträt der Wandervögel auf dem Pizzo L.
388	Sozarch_F_5000-Fx-386	Zusammenkünfte der OG Aarau und Olten - Gruppe.
389	Sozarch_F_5000-Fx-387	Zusammenkünfte der OG Aarau und Olten - Gruppe.

	Zeit	Ort
121	1913	Europa;Schweiz;Aargau;Aarau /Europa;Sc
125	1913	Europa;Schweiz;Aargau;Aarau /Europa;Sc
126	1913	Europa;Schweiz;Aargau;Aarau /Europa;Sc
134	1914	Europa;Schweiz;Aargau;Seon /Europa;Sch

```

140  1919  Europa;Schweiz;Thurgau;Arbon /Europa;S
184  1910  Europa;Schweiz;Tessin /Europa;Schweiz;
185  1910  Europa;Schweiz;Tessin /Europa;Schweiz;
190  1910  Europa;Schweiz;Tessin /Europa;Schweiz;
349  1912  Europa;Schweiz;Tessin /Europa;Schweiz;
388  1911  Europa;Schweiz;Aargau;Aarau /Europa;Sc
389  1911  Europa;Schweiz;Aargau;Aarau /Europa;Sc

```

```

In [8]: # Diese Einträge müssen bereinigt werden, indem der letzte Teil gelöscht wird
wandervogel_df.replace(' ?/Europa;S.*$', '', regex=True, inplace=True)

```

```

In [9]: # Wir erstellen ein neues Dataframe mit den Orten nach Kontinent, Land, Kanton
orte = pd.DataFrame(data=wandervogel_df.Ort.str.split(';').tolist(), columns=['Kontinent', 'Land', 'Kanton', 'Ort'])
orte.fillna('', inplace=True)

```

```

In [10]: # Als nächstes überprüfen wir, ob es noch weitere Unregelmässigkeiten in der Liste der vorhandenen Ortsangaben
print(orte.Kontinent.unique())
print(orte.Land.unique())
print(orte.Kanton.unique())
print(orte.Ort.unique())

```

```

['Europa' '']
['Schweiz' ' ' 'Deutschland' 'Italien' 'Frankreich']
[' ' 'Zürich, Kanton' 'Thurgau' 'Bern, Kanton' 'Zug, Kanton' 'Aargau'
 'Glarus, Kanton' 'Appenzell Ausserrhoden' 'Baselland' 'Jura, Kanton'
 'Baselstadt' 'Neuchâtel, Kanton' 'St. Gallen, Kanton' 'Luzern, Kanton'
 'Solothurn, Kanton' 'Schaffhausen, Kanton' 'Obwalden' 'Baden-Württemberg'
 'Graubünden' 'Lombardia' 'Graubünden /Europa' 'Tessin' 'Wallis'
 'Nidwalden' 'Uri' 'Schwyz, Kanton' 'Baden-Württemberg / ' 'Sachsen'
 'Glarus, Kanton /Europa' 'Bayern']
[' ' 'Kyburg' 'Märstetten' 'Bolligen' 'Steinmau' 'Risch' 'Bäretswil'
 'Biberstein' 'Hombrech' 'Illnau-E' 'Kyburg / ' 'Glarus S' 'Seengen'
 'Hüttwilen' 'Rümlang' 'Brütten' 'Nürensdorf' 'Bellikon' 'Ormalingen'
 'Rekingen' 'Aarau' 'Erlinsbach (Aargau)' 'Basel, Stadt' 'Bern, Stadt' 'Biel'
 'Frauenfeld' 'Horgen / ' 'Kreuzlingen' 'Küsnacht' 'Le Locle' 'Lenzburg' 'Lichtenstein'
 'Luzern,' 'Olten' 'Zürich,' 'Schaffhausen' 'Seon' 'St. Gallen' 'Arbon' 'Engelberg'
 'Wädenswil' 'Winterthur' 'Steckborn /Euro' 'Morbegno' 'Livigno /Euro' 'Ital'
 'Poschiavo' 'Binn' 'Blenio' 'Aeschi bei' 'Krauchthal' 'Grub' 'Ebnet'
 'Orvin' 'Richters' 'Lugano' 'Hallwil' 'Reichenbach' 'Riemenst' 'Birrwil'
 'Tomils' 'Oberglatt' 'Glarus N' 'Chur' 'Sumvitg' 'Milano /Europ' 'Savognin'
 'Avers' 'Flum' 'Frauenfeld /Eur' 'Gold' 'Tenna' 'Staufen' 'Evilard'
 'Fischbach-Göslik' 'Dietikon' 'Wohlen (AG)' 'Niederwil (AG)' 'Seuzach'
 'Otelfingen' 'Rapp' 'Val-d'Aoste' 'Leipzig' 'Alpthal' 'Stäfa' 'Islisberg'
 'Embrach' 'Knonau' 'Anwil' 'Rüti (ZH)' 'Rüschlikon' 'Greifensee' 'Brunegg' 'Bassersdorf'
 'Cavigliano /Euro' 'Centovalli' 'Ascona' 'Aranno' 'Kronach' 'Furna'
 'Zernez' 'Flawil' 'Beinwil am See']

```

```

In [11]: # Die Länder werden durch ihre Kürzel ersetzt
laender_map = {'Schweiz': 'CH', 'Deutschland': 'GER', 'Italien': 'IT', 'Fr
for k, v in laender_map.items():
    orte.Land.replace(k, v, inplace=True)

In [12]: # Bei den Kantonen scheint es zwei Probleme zu geben: Einerseits das Postl
# gefolgt allenfalls von "Europa".
orte.Kanton.replace(', Kanton$', '', inplace=True, regex=True)
orte.Kanton.replace(' /.*$', '', inplace=True, regex=True)

In [13]: # Die Kantone werden ebenfalls durch ihre Kürzel ersetzt
kantone_map = {'Zürich': 'ZH', 'Thurgau': 'TG', 'Bern': 'BE', 'Zug': 'ZG',
               'Glarus': 'GL', 'Appenzell Ausserrhoden': 'AR', 'Baselland',
               'Baselstadt': 'BS', 'Neuchâtel': 'NE', 'St. Gallen': 'SG', 'I
               'Schaffhausen': 'SH', 'Obwalden': 'OW', 'Graubünden': 'GR',
               'Wallis': 'VS', 'Nidwalden': 'NW', 'Uri': 'UR', 'Schwyz': '
for k, v in kantone_map.items():
    orte.Kanton.replace(k, v, inplace=True)

In [14]: # Bei den Ortschaften scheint es folgende Probleme zu geben: Schrägstriche
# sehr kurze Namen sowie abgeschnittene Namen
orte.Ort.replace(' /.*$', '', inplace=True, regex=True)
orte.Ort.replace(', Stad.*$', '', inplace=True, regex=True)
orte.Ort.replace(' \(.*$', '', inplace=True, regex=True)

In [15]: # Bevor das Problem der sehr kurzen und abgeschnittenen Namen angegangen w
# vereinigt werden
wandervogel_df.drop('Ort', axis=1, inplace=True)
wandervogel_neu = pd.concat([wandervogel_df, orte], axis=1)
wandervogel_neu

```

```

Out[15]:
          Signatur
0    Sozarch_F_5000-Dx-431    Kreistanz um ein loderndes Fe
1    Sozarch_F_5000-Dx-432    Prellen: Spiel im Freien - Kinder und Erwachse
2    Sozarch_F_5000-Dx-433    Vier Wandervögel sitzen auf einer Holzbank vor
3    Sozarch_F_5000-Dx-434    Kreistanz um ein loderndes Fe
4    Sozarch_F_5000-Dx-435    Mann Gitarre spiel
5    Sozarch_F_5000-Dx-436    Gruppe von Wandervögeln im Freien: Mädchen Git
6    Sozarch_F_5000-Dx-437    Zwei Personen sitzen am Tisch und essen (Mahlz
7    Sozarch_F_5000-Dx-438    Winterstimmung: Junge Männer auf der Strasse u
8    Sozarch_F_5000-Dx-439    Junge Männer beim Es
9    Sozarch_F_5000-Dx-440    Junge Männer bei einer Rast auf einem Brun
10   Sozarch_F_5000-Fx-001    Gruppenporträt vor der Wandervogel-Landsgemein
11   Sozarch_F_5000-Fx-002    Versammelte Wandervögel an der Wandervogel-Lar
12   Sozarch_F_5000-Fx-003    Tanzende Wandervögel an der Wandervogel-Landsq
13   Sozarch_F_5000-Fx-004    Tanzende Wandervögel an der Wandervogel-Landsq
14   Sozarch_F_5000-Fx-006    Wandervogel-Landsgemeinde auf dem Bantiger bei
15   Sozarch_F_5000-Fx-007    Wandervogel-Landsgemeinde auf dem Bantiger bei
16   Sozarch_F_5000-Fx-008    Wandervogel-Landsgemeinde auf dem Bantiger bei

```

17	Sozarch_F_5000-Fx-009	Wandervogel-Landsgemeinde auf dem Bantiger bei
18	Sozarch_F_5000-Fx-010	Seilziehen an der Wandervogel-Landsgemeinde (1
19	Sozarch_F_5000-Fx-011	Wandervogel-Landsgemeinde auf dem Bantiger bei
20	Sozarch_F_5000-Fx-012	Wandervogel-Landsgemeinde auf dem Bantiger bei
21	Sozarch_F_5000-Fx-013	Wandervogel-Landsgemeinde bei Steinmaur (Heitl
22	Sozarch_F_5000-Fx-014	Wandervogel-Landsgemeinde bei Steinmaur (Heitl
23	Sozarch_F_5000-Fx-015	Wandervogel-Landsgemeinde bei Steinmaur (Heitl
24	Sozarch_F_5000-Fx-016	Wandervogel-Landsgemeinde bei Steinmaur (Heitl
25	Sozarch_F_5000-Fx-018	Wandervogel-Landsgemeinde bei Steinmaur (Heitl
26	Sozarch_F_5000-Fx-019	Wandervogel-Landsgemeinde beim Schloss Buonas
27	Sozarch_F_5000-Fx-020	Wandervogel-Landsgemeinde beim Schloss Buonas
28	Sozarch_F_5000-Fx-021	Wandervogel-Landsgemeinde beim Schloss Buonas
29	Sozarch_F_5000-Fx-022	Wandervogel-Landsgemeinde beim Schloss Buonas
..	...	
568	Sozarch_F_5000-Nx-446	Ostschweizer Wandervögel unterhalten sich (ver
569	Sozarch_F_5000-Nx-447	Eine Fahrt zusammen mit den Flawilern auf den
570	Sozarch_F_5000-Nx-448	Eine Fahrt zusammen mit den Flawilern auf den
571	Sozarch_F_5000-Nx-449	Eine Fahrt zusammen mit den Flawilern auf den
572	Sozarch_F_5000-Nx-450	Eine Fahrt zusammen mit den Flawilern auf den
573	Sozarch_F_5000-Nx-451	Skihütte Ober Altberg Jbergeregg (sic) - Altw
574	Sozarch_F_5000-Nx-452	Teilansicht der Skihütte Ober Altberg Jbergere
575	Sozarch_F_5000-Nx-454	Treffen von Wandervögeln am Walensee - Mahlzei
576	Sozarch_F_5000-Nx-455	Hütten-Usrumete Prestenburg - Wandervögel vor
577	Sozarch_F_5000-Nx-456	Hütten-Usrumete Prestenburg - Wandervögel vor
578	Sozarch_F_5000-Nx-457	Treffen von Wandervögeln am Walensee - Wanderv
579	Sozarch_F_5000-Px-005	Wandervogel-Landsgemeinde: Auflistung der Vera
580	Sozarch_F_5000-Px-274	Erster Briefkopf des Schweizer Wandervogels (1
581	Sozarch_F_5000-Px-480	Abgebildete "Wandervögel" beim Haar waschen am
582	Sozarch_F_5000-Zx-073	Schweizer Wandervogel Landsgemeinde Beinwiel
583	Sozarch_F_5000-Zx-481	Zeichnung aus der Sammlung von Gottfried Gries
584	Sozarch_F_5000-Zx-562	Abbildung eines Wandervogels Blockflöte spiele
585	Sozarch_F_5000-Zx-590	Drucke/Linolschnitte Wandervogel: Abbildung de
586	Sozarch_F_5000-Zx-591	Drucke/Linolschnitte Wandervogel: Abbildung ei
587	Sozarch_F_5000-Zx-592	Drucke/Linolschnitte Wandervogel: Abbildung de
588	Sozarch_F_5000-Zx-593	Drucke/Linolschnitte Wandervogel: Abbildung ei
589	Sozarch_F_5000-Zx-594	Drucke/Linolschnitte Wandervogel: Abbildung ei
590	Sozarch_F_5000-Zx-595	Drucke/Linolschnitte Wandervogel: Abbildung ei
591	Sozarch_F_5000-Zx-596	Drucke/Linolschnitte Wandervogel: Abbildung ei
592	Sozarch_F_5000-Zx-597	Drucke/Linolschnitte Wandervogel: Abbildung fl
593	Sozarch_F_5000-Zx-598	Drucke/Linolschnitte Wandervogel: Abbildung ei
594	Sozarch_F_5000-Zx-599	Drucke/Linolschnitte Wandervogel: Abbildung ei
595	Sozarch_F_5000-Zx-600	Drucke/Linolschnitte Wandervogel: Abbildung de
596	Sozarch_F_5000-Zx-601	Drucke/Linolschnitte Wandervogel: Abbildung ei
597	Sozarch_F_5000-Zx-602	Drucke/Linolschnitte Wandervogel: Abbildung fü

	Zeit	Kontinent	Land	Kanton	Ort
0	1931-1940	Europa	CH		
1	1931-1940	Europa	CH		

2	1931-1940	Europa	CH		
3	1931-1940	Europa	CH		
4	1931-1940	Europa	CH		
5	1931-1940	Europa	CH		
6	1931-1940	Europa	CH		
7	1931-1940	Europa	CH		
8	1931-1940	Europa	CH		
9	1931-1940	Europa	CH		
10	1910	Europa	CH	ZH	Kyburg
11	1910	Europa	CH	ZH	Kyburg
12	1913	Europa	CH	TG	Märstetten
13	1913	Europa	CH	TG	Märstetten
14	1914	Europa	CH	BE	Bolligen
15	1914	Europa	CH	BE	Bolligen
16	1914	Europa	CH	BE	Bolligen
17	1914	Europa	CH	BE	Bolligen
18	1915	Europa	CH		
19	1914	Europa	CH	BE	Bolligen
20	1914	Europa	CH	BE	Bolligen
21	1915	Europa	CH	ZH	Steinmau
22	1915	Europa	CH	ZH	Steinmau
23	1915	Europa	CH	ZH	Steinmau
24	1915	Europa	CH	ZH	Steinmau
25	1915	Europa	CH	ZH	Steinmau
26	1916	Europa	CH	ZG	Risch
27	1916	Europa	CH	ZG	Risch
28	1916	Europa	CH	ZG	Risch
29	1916	Europa	CH	ZG	Risch
..
568	1931-1940	Europa	CH		
569	1938	Europa	CH	SG	Flaw
570	1938	Europa	CH	SG	Flaw
571	1938	Europa	CH	SG	Flaw
572	1938	Europa	CH	SG	Flaw
573	1944	Europa	CH	SZ	
574	1944	Europa	CH	SZ	
575	1938	Europa	CH	SG	
576	1938	Europa	CH		
577	1938	Europa	CH		
578	1938	Europa	CH	SG	
579	1911	Europa	CH	SG	Rapp
580	1920	Europa	CH		
581	1918	Europa	CH		
582	1939	Europa	CH	AG	Beinwil am See
583	1918	Europa	CH		
584	1920	Europa	CH		
585	1954	Europa	CH		
586	1954	Europa	CH		

587	1954	Europa	CH
588	1954	Europa	CH
589	1954	Europa	CH
590	1954	Europa	CH
591	1954	Europa	CH
592	1954	Europa	CH
593	1954	Europa	CH
594	1954	Europa	CH
595	1954	Europa	CH
596	1954	Europa	CH
597	1954	Europa	CH

[598 rows x 7 columns]

```
In [16]: # Das Problem der sehr kurzen Namen
vector = [(len(l) < 3) and (len(l) > 0) for l in wandervogel_neu['Ort']]
wandervogel_neu[vector]
```

```
Out[16]:
```

	Signatur	Titel
130	Sozarch_F_5000-Fx-124	Wandervogelfahrt der Ortsgruppe Schaffhausen:
131	Sozarch_F_5000-Fx-125	Wandervogelfahrt der Ortsgruppe Schaffhausen:
158	Sozarch_F_5000-Fx-152	Hohen Krähen - Gruppenporträt der Wandervögel
498	Sozarch_F_5000-Fx-525	Vogelperspektiveaufnahme: Versammelte Wandervögel

	Zeit	Kontinent	Land	Kanton	Ort
130	1914	Europa	CH	SH	Sc
131	1914	Europa	CH	SH	Sc
158	1912	Europa	GER	Baden-Württemberg	M
498	1929	Europa	CH	SH	Rü

```
In [17]: # Orte werden manuell verbessert
wandervogel_neu.iat[130,6] = 'Schaffhausen'
wandervogel_neu.iat[131,6] = 'Schaffhausen'
wandervogel_neu.iat[158,6] = 'Mühlhausen'
wandervogel_neu.iat[498,6] = 'Rüdlingen'
```

```
In [18]: # Zeige Ortschaften an, die einen partiellen Match im Titel haben (Hinweis)
def_column_width= pd.options.display.max_colwidth
def_rows_displayed = pd.options.display.max_rows
pd.options.display.max_colwidth = 140
pd.options.display.max_rows = 100
vector = [True if re.search(o+'\w+', t) and o is not '' else False for t, o in zip(wandervogel_neu['Titel'], wandervogel_neu['Ort'])]
pd.options.display.max_colwidth = def_column_width
pd.options.display.max_rows = def_rows_displayed
```

```
In [19]: orte_ignoriert = [462, 463, 464]
```

```

    for i, e in enumerate(vector):
        if e and i not in orte_ignoriert:
            wandervogel_neu['Ort'].iat[i] = re.search(wandervogel_neu['Ort'].iat[i]).group(1)

In [20]: orte_ungueltig = [150, 151, 152, 153, 154, 156, 161, 198, 229, 275, 276, 334]

    for i in orte_ungueltig:
        wandervogel_neu['Kanton'].iat[i] = ''
        wandervogel_neu['Ort'].iat[i] = ''

In [21]: orte_ersetzt = {145: 'Bisital', 146: 'Fluhberg', 147: 'Schwarzenberg', 148: 'Thurgau',
                        155: 'Thurmündung', 162: 'Pfannenstiel', 163: 'Seelis', 220: 'Thurgau',
                        228: 'Engadin', 230: 'Bözberg', 231: 'Morcote', 232: 'Hintere Rhodan',
                        334: 'Sonnenberg', 569: 'Speer', 570: 'Speer', 571: 'Speer'}

    for k, v in orte_ersetzt.items():
        wandervogel_neu['Kanton'].iat[k] = ''
        wandervogel_neu['Ort'].iat[k] = v

In [22]: koords = DataFrame(columns=['E', 'N'])
empty_df = DataFrame(columns=['E', 'N'])
empty_df.loc[0] = ['', '']

    for i, (kanton, ort) in enumerate(zip(wandervogel_neu.Kanton, wandervogel_neu.Ort)):
        vector = [(ort == o) and (kanton == k) for o, k in zip(koord_df.Ortschiff, koord_df.Kanton)]
        koord_temp = koord_df.loc[vector, ['E', 'N']]
        koords = pd.concat([koords, koord_temp[0:1]] if koord_temp.shape[0] > 0 else [])
    wandervogel_final = pd.concat([wandervogel_neu, koords.reset_index(drop=True)])

In [23]: wandervogel_final.to_csv('output.csv', encoding='utf-8')

```