Visualisasi dengan Python

August 17, 2019

1 Data Visualization with Python

1.1 Table of Contents

1.1.1 Dataset: Immigration to Canada from 1980 to 2013

Dataset sumber: International migration flows to and from selected countries - The 2015 revision.

Dataset ini berisi data tahunan imigran internasional yang dicatat berdasarkan negara tujuan. Data ini berisi masuk dan keluarnya imigran berdasarkan tempat kelahiran, kewarganegaraan. Total ada 45 negara

1.1.2 Pandas Dasar

[2]:
```
# import librari numpy dan pandas

import numpy as np
import pandas as pd
```

[3]:
```
# Read data dari url

df_can = pd.read_excel('https://s3-api.us-geo.objectstorage.softlayer.net/
 ↪cf-courses-data/CognitiveClass/DV0101EN/labs/Data_Files/Canada.xlsx',
                       sheet_name='Canada by Citizenship',
                       skiprows=range(20),
                       skipfooter=2)

print ('Data read into a pandas dataframe!')
```

```
Data read into a pandas dataframe!
```

[4]:
```
# Menampilkan 5 data awal

df_can.head()
```

[4]:
```
          Type    Coverage          OdName  AREA AreaName   REG  \
0  Immigrants  Foreigners     Afghanistan   935     Asia  5501
1  Immigrants  Foreigners         Albania   908   Europe   925
2  Immigrants  Foreigners         Algeria   903   Africa   912
3  Immigrants  Foreigners  American Samoa   909  Oceania   957
4  Immigrants  Foreigners         Andorra   908   Europe   925

            RegName  DEV             DevName  1980  ...  2004  2005  2006  \
0    Southern Asia  902  Developing regions    16  ...  2978  3436  3009
1  Southern Europe  901   Developed regions     1  ...  1450  1223   856
2  Northern Africa  902  Developing regions    80  ...  3616  3626  4807
3        Polynesia  902  Developing regions     0  ...     0     0     1
4  Southern Europe  901   Developed regions     0  ...     0     0     1

   2007  2008  2009  2010  2011  2012  2013
0  2652  2111  1746  1758  2203  2635  2004
1   702   560   716   561   539   620   603
2  3623  4005  5393  4752  4325  3774  4331
3     0     0     0     0     0     0     0
4     1     0     0     0     0     1     1

[5 rows x 43 columns]
```

[5]:
```
# Menampilkan info type data

df_can.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 195 entries, 0 to 194
Data columns (total 43 columns):
Type        195 non-null object
Coverage    195 non-null object
OdName      195 non-null object
AREA        195 non-null int64
AreaName    195 non-null object
REG         195 non-null int64
RegName     195 non-null object
DEV         195 non-null int64
DevName     195 non-null object
1980        195 non-null int64
1981        195 non-null int64
1982        195 non-null int64
1983        195 non-null int64
1984        195 non-null int64
1985        195 non-null int64
1986        195 non-null int64
1987        195 non-null int64
1988        195 non-null int64
```

```
1989        195 non-null int64
1990        195 non-null int64
1991        195 non-null int64
1992        195 non-null int64
1993        195 non-null int64
1994        195 non-null int64
1995        195 non-null int64
1996        195 non-null int64
1997        195 non-null int64
1998        195 non-null int64
1999        195 non-null int64
2000        195 non-null int64
2001        195 non-null int64
2002        195 non-null int64
2003        195 non-null int64
2004        195 non-null int64
2005        195 non-null int64
2006        195 non-null int64
2007        195 non-null int64
2008        195 non-null int64
2009        195 non-null int64
2010        195 non-null int64
2011        195 non-null int64
2012        195 non-null int64
2013        195 non-null int64
dtypes: int64(37), object(6)
memory usage: 65.6+ KB
```

[6]:
```
# Menampilkan header dari atribut atau columns

df_can.columns.values
```

[6]:
```
array(['Type', 'Coverage', 'OdName', 'AREA', 'AreaName', 'REG', 'RegName',
       'DEV', 'DevName', 1980, 1981, 1982, 1983, 1984, 1985, 1986, 1987,
       1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998,
       1999, 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009,
       2010, 2011, 2012, 2013], dtype=object)
```

[7]:
```
# Menampilkan parameter index dar data

df_can.index.values
```

[7]:
```
array([  0,   1,   2,   3,   4,   5,   6,   7,   8,   9,  10,  11,  12,
        13,  14,  15,  16,  17,  18,  19,  20,  21,  22,  23,  24,  25,
        26,  27,  28,  29,  30,  31,  32,  33,  34,  35,  36,  37,  38,
        39,  40,  41,  42,  43,  44,  45,  46,  47,  48,  49,  50,  51,
        52,  53,  54,  55,  56,  57,  58,  59,  60,  61,  62,  63,  64,
        65,  66,  67,  68,  69,  70,  71,  72,  73,  74,  75,  76,  77,
        78,  79,  80,  81,  82,  83,  84,  85,  86,  87,  88,  89,  90,
```

```
 91,  92,  93,  94,  95,  96,  97,  98,  99, 100, 101, 102, 103,
104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116,
117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129,
130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142,
143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155,
156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168,
169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181,
182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194])
```

[8]:
```
# Tipe class bukan list

print(type(df_can.columns))
print(type(df_can.index))
```

```
<class 'pandas.core.indexes.base.Index'>
<class 'pandas.core.indexes.range.RangeIndex'>
```

[9]:
```
# Meengubah kolom dan index menjadi list

df_can.columns.tolist()
df_can.index.tolist()

print (type(df_can.columns.tolist()))
print (type(df_can.index.tolist()))
```

```
<class 'list'>
<class 'list'>
```

[10]:
```
# Melihat dimensi data

df_can.shape
```

[10]:
```
(195, 43)
```

[13]:
```
# Mengubah nama atribut header

df_can.rename(columns={'OdName':'Country', 'AreaName':'Continent', 'RegName':
 ↪'Region'}, inplace=True)
df_can.columns
```

[13]:
```
Index([  'Country', 'Continent',    'Region',   'DevName',        1980,
              1981,        1982,        1983,        1984,        1985,
              1986,        1987,        1988,        1989,        1990,
              1991,        1992,        1993,        1994,        1995,
              1996,        1997,        1998,        1999,        2000,
              2001,        2002,        2003,        2004,        2005,
              2006,        2007,        2008,        2009,        2010,
              2011,        2012,        2013],
      dtype='object')
```

[16]:
```
# Menambahkan kolom 'Total', Nilai dari total imigran berdasarkan negara dari␣
↪tahun 1980 - 2013

df_can['Total'] = df_can.sum(axis=1)
```

[15]:
```
# Melihat data yang bernilai null

df_can.isnull().sum()
```

[15]:
```
Country      0
Continent    0
Region       0
DevName      0
1980         0
1981         0
1982         0
1983         0
1984         0
1985         0
1986         0
1987         0
1988         0
1989         0
1990         0
1991         0
1992         0
1993         0
1994         0
1995         0
1996         0
1997         0
1998         0
1999         0
2000         0
2001         0
2002         0
2003         0
2004         0
2005         0
2006         0
2007         0
2008         0
2009         0
2010         0
2011         0
2012         0
2013         0
Total        0
```

```
dtype: int64
```

[17]:
```
# Summary statistik data

df_can.describe()
```

[17]:
```
              1980          1981          1982          1983          1984  \
count   195.000000    195.000000    195.000000    195.000000    195.000000
mean    508.394872    566.989744    534.723077    387.435897    376.497436
std    1949.588546   2152.643752   1866.997511   1204.333597   1198.246371
min       0.000000      0.000000      0.000000      0.000000      0.000000
25%       0.000000      0.000000      0.000000      0.000000      0.000000
50%      13.000000     10.000000     11.000000     12.000000     13.000000
75%     251.500000    295.500000    275.000000    173.000000    181.000000
max   22045.000000  24796.000000  20620.000000  10015.000000  10170.000000

             1985          1986          1987          1988          1989  \
count   195.000000    195.000000    195.000000    195.000000    195.000000
mean    358.861538    441.271795    691.133333    714.389744    843.241026
std    1079.309600   1225.576630   2109.205607   2443.606788   2555.048874
min       0.000000      0.000000      0.000000      0.000000      0.000000
25%       0.000000      0.500000      0.500000      1.000000      1.000000
50%      17.000000     18.000000     26.000000     34.000000     44.000000
75%     197.000000    254.000000    434.000000    409.000000    508.500000
max    9564.000000   9470.000000  21337.000000  27359.000000  23795.000000

       ...          2005          2006          2007          2008  \
count  ...    195.000000    195.000000    195.000000    195.000000
mean   ...   1320.292308   1266.958974   1191.820513   1246.394872
std    ...   4425.957828   3926.717747   3443.542409   3694.573544
min    ...      0.000000      0.000000      0.000000      0.000000
25%    ...     28.500000     25.000000     31.000000     31.000000
50%    ...    210.000000    218.000000    198.000000    205.000000
75%    ...    832.000000    842.000000    899.000000    934.500000
max    ...  42584.000000  33848.000000  28742.000000  30037.000000

              2009          2010          2011          2012          2013  \
count   195.000000    195.000000    195.000000    195.000000    195.000000
mean   1275.733333   1420.287179   1262.533333   1313.958974   1320.702564
std    3829.630424   4462.946328   4030.084313   4247.555161   4237.951988
min       0.000000      0.000000      0.000000      0.000000      0.000000
25%      36.000000     40.500000     37.500000     42.500000     45.000000
50%     214.000000    211.000000    179.000000    233.000000    213.000000
75%     888.000000    932.000000    772.000000    783.000000    796.000000
max   29622.000000  38617.000000  36765.000000  34315.000000  34129.000000

              Total
count  1.950000e+02
```

```
mean    6.573490e+04
std     1.835710e+05
min     2.000000e+00
25%     1.904000e+03
50%     1.003600e+04
75%     4.447900e+04
max     1.383808e+06

[8 rows x 35 columns]
```

1.1.3 Pandas Intermediate: Indexing and Selection

[18]:
```
# Filter berdasarkan negara (Country)

df_can.Country
```

[18]:
```
0          Afghanistan
1              Albania
2              Algeria
3       American Samoa
4              Andorra
              ...
190           Viet Nam
191     Western Sahara
192              Yemen
193             Zambia
194           Zimbabwe
Name: Country, Length: 195, dtype: object
```

[19]:
```
# Filter countries ('OdName') and the data for years: 1980 - 1985.

df_can[['Country', 1980, 1981, 1982, 1983, 1984, 1985]]
```

[19]:
```
              Country  1980  1981  1982  1983  1984  1985
0         Afghanistan    16    39    39    47    71   340
1             Albania     1     0     0     0     0     0
2             Algeria    80    67    71    69    63    44
3      American Samoa     0     1     0     0     0     0
4             Andorra     0     0     0     0     0     0
..                ...   ...   ...   ...   ...   ...   ...
190         Viet Nam  1191  1829  2162  3404  7583  5907
191   Western Sahara     0     0     0     0     0     0
192            Yemen     1     2     1     6     0    18
193           Zambia    11    17    11     7    16     9
194         Zimbabwe    72   114   102    44    32    29

[195 rows x 7 columns]
```

[20]:
```
df_can.set_index('Country', inplace=True)
```

[21]:
```
df_can.head(5)
```

[21]:
```
               Continent           Region             DevName  1980  1981  \
Country
Afghanistan         Asia    Southern Asia  Developing regions    16    39
Albania           Europe  Southern Europe   Developed regions     1     0
Algeria           Africa  Northern Africa  Developing regions    80    67
American Samoa   Oceania        Polynesia  Developing regions     0     1
Andorra           Europe  Southern Europe   Developed regions     0     0

                1982  1983  1984  1985  1986  ...  2005  2006  2007  2008  \
Country                                       ...
Afghanistan       39    47    71   340   496  ...  3436  3009  2652  2111
Albania            0     0     0     0     1  ...  1223   856   702   560
Algeria           71    69    63    44    69  ...  3626  4807  3623  4005
American Samoa     0     0     0     0     0  ...     0     1     0     0
Andorra            0     0     0     0     2  ...     0     1     1     0

                2009  2010  2011  2012  2013   Total
Country
Afghanistan     1746  1758  2203  2635  2004  117278
Albania          716   561   539   620   603   31398
Algeria         5393  4752  4325  3774  4331  138878
American Samoa     0     0     0     0     0      12
Andorra            0     0     0     1     1      30

[5 rows x 38 columns]
```

[22]:
```
df_can.index.name = None
```

[23]:
```
print(df_can.loc['Japan'])
```

```
Continent                Asia
Region            Eastern Asia
DevName      Developed regions
1980                       701
1981                       756
1982                       598
1983                       309
1984                       246
1985                       198
1986                       248
1987                       422
1988                       324
1989                       494
1990                       379
1991                       506
```

```
1992                      605
1993                      907
1994                      956
1995                      826
1996                      994
1997                      924
1998                      897
1999                     1083
2000                     1010
2001                     1092
2002                      806
2003                      817
2004                      973
2005                     1067
2006                     1212
2007                     1250
2008                     1284
2009                     1194
2010                     1168
2011                     1265
2012                     1214
2013                      982
Total                   55414
Name: Japan, dtype: object
```

[24]:
```
print(df_can.loc['Japan', 2013])
```

```
982
```

[25]:
```
print(df_can.loc['Japan', [1980, 1981, 1982, 1983, 1984, 1984]])
```

```
1980    701
1981    756
1982    598
1983    309
1984    246
1984    246
Name: Japan, dtype: object
```

[26]:
```
df_can.columns = list(map(str, df_can.columns))
[print (type(x)) for x in df_can.columns.values]
```

```
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
```

```
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
<class 'str'>
```

[26]:

```
[None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
```

```
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None,
 None]
```

[27]:
```
years = list(map(str, range(1980, 2014)))
years
```

[27]:
```
['1980',
 '1981',
 '1982',
 '1983',
 '1984',
 '1985',
 '1986',
 '1987',
 '1988',
 '1989',
 '1990',
 '1991',
 '1992',
 '1993',
 '1994',
 '1995',
 '1996',
 '1997',
```

```
 '1998',
 '1999',
 '2000',
 '2001',
 '2002',
 '2003',
 '2004',
 '2005',
 '2006',
 '2007',
 '2008',
 '2009',
 '2010',
 '2011',
 '2012',
 '2013']
```

[28]:
```
condition = df_can['Continent'] == 'Asia'
print(condition)
```

```
Afghanistan         True
Albania            False
Algeria            False
American Samoa     False
Andorra            False
                   ...
Viet Nam            True
Western Sahara     False
Yemen               True
Zambia             False
Zimbabwe           False
Name: Continent, Length: 195, dtype: bool
```

[29]:
```
df_can[condition]
```

[29]:

	Continent	Region	\
Afghanistan	Asia	Southern Asia	
Armenia	Asia	Western Asia	
Azerbaijan	Asia	Western Asia	
Bahrain	Asia	Western Asia	
Bangladesh	Asia	Southern Asia	
Bhutan	Asia	Southern Asia	
Brunei Darussalam	Asia	South-Eastern Asia	
Cambodia	Asia	South-Eastern Asia	
China	Asia	Eastern Asia	
China, Hong Kong Special Administrative Region	Asia	Eastern Asia	
China, Macao Special Administrative Region	Asia	Eastern Asia	
Cyprus	Asia	Western Asia	

Democratic People's Republic of Korea	Asia	Eastern Asia
Georgia	Asia	Western Asia
India	Asia	Southern Asia
Indonesia	Asia	South-Eastern Asia
Iran (Islamic Republic of)	Asia	Southern Asia
Iraq	Asia	Western Asia
Israel	Asia	Western Asia
Japan	Asia	Eastern Asia
Jordan	Asia	Western Asia
Kazakhstan	Asia	Central Asia
Kuwait	Asia	Western Asia
Kyrgyzstan	Asia	Central Asia
Lao People's Democratic Republic	Asia	South-Eastern Asia
Lebanon	Asia	Western Asia
Malaysia	Asia	South-Eastern Asia
Maldives	Asia	Southern Asia
Mongolia	Asia	Eastern Asia
Myanmar	Asia	South-Eastern Asia
Nepal	Asia	Southern Asia
Oman	Asia	Western Asia
Pakistan	Asia	Southern Asia
Philippines	Asia	South-Eastern Asia
Qatar	Asia	Western Asia
Republic of Korea	Asia	Eastern Asia
Saudi Arabia	Asia	Western Asia
Singapore	Asia	South-Eastern Asia
Sri Lanka	Asia	Southern Asia
State of Palestine	Asia	Western Asia
Syrian Arab Republic	Asia	Western Asia
Tajikistan	Asia	Central Asia
Thailand	Asia	South-Eastern Asia
Turkey	Asia	Western Asia
Turkmenistan	Asia	Central Asia
United Arab Emirates	Asia	Western Asia
Uzbekistan	Asia	Central Asia
Viet Nam	Asia	South-Eastern Asia
Yemen	Asia	Western Asia

	DevName	1980	\
Afghanistan	Developing regions	16	
Armenia	Developing regions	0	
Azerbaijan	Developing regions	0	
Bahrain	Developing regions	0	
Bangladesh	Developing regions	83	
Bhutan	Developing regions	0	
Brunei Darussalam	Developing regions	79	
Cambodia	Developing regions	12	

							\
China	Developing regions	5123					
China, Hong Kong Special Administrative Region	Developing regions	0					
China, Macao Special Administrative Region	Developing regions	0					
Cyprus	Developing regions	132					
Democratic People's Republic of Korea	Developing regions	1					
Georgia	Developing regions	0					
India	Developing regions	8880					
Indonesia	Developing regions	186					
Iran (Islamic Republic of)	Developing regions	1172					
Iraq	Developing regions	262					
Israel	Developing regions	1403					
Japan	Developed regions	701					
Jordan	Developing regions	177					
Kazakhstan	Developing regions	0					
Kuwait	Developing regions	1					
Kyrgyzstan	Developing regions	0					
Lao People's Democratic Republic	Developing regions	11					
Lebanon	Developing regions	1409					
Malaysia	Developing regions	786					
Maldives	Developing regions	0					
Mongolia	Developing regions	0					
Myanmar	Developing regions	80					
Nepal	Developing regions	1					
Oman	Developing regions	0					
Pakistan	Developing regions	978					
Philippines	Developing regions	6051					
Qatar	Developing regions	0					
Republic of Korea	Developing regions	1011					
Saudi Arabia	Developing regions	0					
Singapore	Developing regions	241					
Sri Lanka	Developing regions	185					
State of Palestine	Developing regions	0					
Syrian Arab Republic	Developing regions	315					
Tajikistan	Developing regions	0					
Thailand	Developing regions	56					
Turkey	Developing regions	481					
Turkmenistan	Developing regions	0					
United Arab Emirates	Developing regions	0					
Uzbekistan	Developing regions	0					
Viet Nam	Developing regions	1191					
Yemen	Developing regions	1					

	1981	1982	1983	1984	1985	\
Afghanistan	39	39	47	71	340	
Armenia	0	0	0	0	0	
Azerbaijan	0	0	0	0	0	
Bahrain	2	1	1	1	3	

Bangladesh	84	86	81	98	92
Bhutan	0	0	0	1	0
Brunei Darussalam	6	8	2	2	4
Cambodia	19	26	33	10	7
China	6682	3308	1863	1527	1816
China, Hong Kong Special Administrative Region	0	0	0	0	0
China, Macao Special Administrative Region	0	0	0	0	0
Cyprus	128	84	46	46	43
Democratic People's Republic of Korea	1	3	1	4	3
Georgia	0	0	0	0	0
India	8670	8147	7338	5704	4211
Indonesia	178	252	115	123	100
Iran (Islamic Republic of)	1429	1822	1592	1977	1648
Iraq	245	260	380	428	231
Israel	1711	1334	541	446	680
Japan	756	598	309	246	198
Jordan	160	155	113	102	179
Kazakhstan	0	0	0	0	0
Kuwait	0	8	2	1	4
Kyrgyzstan	0	0	0	0	0
Lao People's Democratic Republic	6	16	16	7	17
Lebanon	1119	1159	789	1253	1683
Malaysia	816	813	448	384	374
Maldives	0	0	1	0	0
Mongolia	0	0	0	0	0
Myanmar	62	46	31	41	23
Nepal	1	6	1	2	4
Oman	0	0	8	0	0
Pakistan	972	1201	900	668	514
Philippines	5921	5249	4562	3801	3150
Qatar	0	0	0	0	0
Republic of Korea	1456	1572	1081	847	962
Saudi Arabia	0	1	4	1	2
Singapore	301	337	169	128	139
Sri Lanka	371	290	197	1086	845
State of Palestine	0	0	0	0	0
Syrian Arab Republic	419	409	269	264	385
Tajikistan	0	0	0	0	0
Thailand	53	113	65	82	66
Turkey	874	706	280	338	202
Turkmenistan	0	0	0	0	0
United Arab Emirates	2	2	1	2	0
Uzbekistan	0	0	0	0	0
Viet Nam	1829	2162	3404	7583	5907
Yemen	2	1	6	0	18

	1986	...	2005	2006	\

Afghanistan	496	...	3436	3009
Armenia	0	...	224	218
Azerbaijan	0	...	359	236
Bahrain	0	...	12	12
Bangladesh	486	...	4171	4014
Bhutan	0	...	5	10
Brunei Darussalam	12	...	4	5
Cambodia	8	...	370	529
China	1960	...	42584	33518
China, Hong Kong Special Administrative Region	0	...	729	712
China, Macao Special Administrative Region	0	...	21	32
Cyprus	48	...	7	9
Democratic People's Republic of Korea	0	...	14	10
Georgia	0	...	114	125
India	7150	...	36210	33848
Indonesia	127	...	632	613
Iran (Islamic Republic of)	1794	...	5837	7480
Iraq	265	...	2226	1788
Israel	1212	...	2446	2625
Japan	248	...	1067	1212
Jordan	181	...	1940	1827
Kazakhstan	0	...	506	408
Kuwait	4	...	66	35
Kyrgyzstan	0	...	173	161
Lao People's Democratic Republic	21	...	42	74
Lebanon	2576	...	3709	3802
Malaysia	425	...	593	580
Maldives	0	...	0	0
Mongolia	0	...	59	64
Myanmar	18	...	210	953
Nepal	13	...	607	540
Oman	0	...	14	18
Pakistan	691	...	14314	13127
Philippines	4166	...	18139	18400
Qatar	1	...	11	2
Republic of Korea	1208	...	5832	6215
Saudi Arabia	5	...	198	252
Singapore	205	...	392	298
Sri Lanka	1838	...	4930	4714
State of Palestine	0	...	453	627
Syrian Arab Republic	493	...	1458	1145
Tajikistan	0	...	85	46
Thailand	78	...	575	500
Turkey	257	...	2065	1638
Turkmenistan	0	...	40	26
United Arab Emirates	5	...	31	42
Uzbekistan	0	...	330	262

Viet Nam	2741	...	1852	3153	
Yemen	7	...	161	140	

	2007	2008	2009	2010	\
Afghanistan	2652	2111	1746	1758	
Armenia	198	205	267	252	
Azerbaijan	203	125	165	209	
Bahrain	22	9	35	28	
Bangladesh	2897	2939	2104	4721	
Bhutan	7	36	865	1464	
Brunei Darussalam	11	10	5	12	
Cambodia	460	354	203	200	
China	27642	30037	29622	30391	
China, Hong Kong Special Administrative Region	674	897	657	623	
China, Macao Special Administrative Region	16	12	21	21	
Cyprus	4	7	6	18	
Democratic People's Republic of Korea	7	19	11	45	
Georgia	132	112	128	126	
India	28742	28261	29456	34235	
Indonesia	657	661	504	712	
Iran (Islamic Republic of)	6974	6475	6580	7477	
Iraq	2406	3543	5450	5941	
Israel	2401	2562	2316	2755	
Japan	1250	1284	1194	1168	
Jordan	1421	1581	1235	1831	
Kazakhstan	436	394	431	377	
Kuwait	62	53	68	67	
Kyrgyzstan	135	168	173	157	
Lao People's Democratic Republic	53	32	39	54	
Lebanon	3467	3566	3077	3432	
Malaysia	600	658	640	802	
Maldives	2	1	7	4	
Mongolia	82	59	118	169	
Myanmar	1887	975	1153	556	
Nepal	511	581	561	1392	
Oman	16	10	7	14	
Pakistan	10124	8994	7217	6811	
Philippines	19837	24887	28573	38617	
Qatar	5	9	6	18	
Republic of Korea	5920	7294	5874	5537	
Saudi Arabia	188	249	246	330	
Singapore	690	734	366	805	
Sri Lanka	4123	4756	4547	4422	
State of Palestine	441	481	400	654	
Syrian Arab Republic	1056	919	917	1039	
Tajikistan	44	15	50	52	
Thailand	487	519	512	499	

Turkey	1463	1122	1238	1492
Turkmenistan	37	13	20	30
United Arab Emirates	37	33	37	86
Uzbekistan	284	215	288	289
Viet Nam	2574	1784	2171	1942
Yemen	122	133	128	211

	2011	2012	2013	Total
Afghanistan	2203	2635	2004	117278
Armenia	236	258	207	6620
Azerbaijan	138	161	57	5298
Bahrain	21	39	32	950
Bangladesh	2694	2640	3789	131136
Bhutan	1879	1075	487	11752
Brunei Darussalam	6	3	6	1200
Cambodia	196	233	288	13076
China	28502	33024	34129	1319924
China, Hong Kong Special Administrative Region	591	728	774	18654
China, Macao Special Administrative Region	13	33	29	568
Cyprus	6	12	16	2252
Democratic People's Republic of Korea	97	66	17	776
Georgia	139	147	125	4136
India	27509	30933	33087	1383808
Indonesia	390	395	387	26300
Iran (Islamic Republic of)	7479	7534	11291	351846
Iraq	6196	4041	4918	139578
Israel	1970	2134	1945	133016
Japan	1265	1214	982	55414
Jordan	1635	1206	1255	70812
Kazakhstan	381	462	348	16980
Kuwait	58	73	48	4050
Kyrgyzstan	159	278	123	4706
Lao People's Democratic Republic	22	25	15	2178
Lebanon	3072	1614	2172	230718
Malaysia	409	358	204	48834
Maldives	3	1	1	60
Mongolia	103	68	99	1904
Myanmar	368	193	262	18490
Nepal	1129	1185	1308	20444
Oman	10	13	11	448
Pakistan	7468	11227	12603	483200
Philippines	36765	34315	29544	1022782
Qatar	3	14	6	314
Republic of Korea	4588	5316	4509	285162
Saudi Arabia	278	286	267	6850
Singapore	219	146	141	29158
Sri Lanka	3309	3338	2394	296716

```
State of Palestine                                555    533    462    13024
Syrian Arab Republic                             1005    650   1009    62970
Tajikistan                                         47     34     39     1006
Thailand                                          396    296    400    18348
Turkey                                           1257   1068    729    63562
Turkmenistan                                       20     20     14      620
United Arab Emirates                               60     54     46     1672
Uzbekistan                                        162    235    167     6736
Viet Nam                                         1723   1731   2112   194292
Yemen                                             160    174    217     5970

[49 rows x 38 columns]
```

[30]:
```
df_can[(df_can['Continent']=='Asia') & (df_can['Region']=='South-Eastern Asia')]
```

[30]:
```
                                Continent             Region  \
Brunei Darussalam                    Asia  South-Eastern Asia
Cambodia                             Asia  South-Eastern Asia
Indonesia                            Asia  South-Eastern Asia
Lao People's Democratic Republic     Asia  South-Eastern Asia
Malaysia                             Asia  South-Eastern Asia
Myanmar                              Asia  South-Eastern Asia
Philippines                          Asia  South-Eastern Asia
Singapore                            Asia  South-Eastern Asia
Thailand                             Asia  South-Eastern Asia
Viet Nam                             Asia  South-Eastern Asia

                                          DevName  1980  1981  1982  1983  \
Brunei Darussalam                Developing regions    79     6     8     2
Cambodia                         Developing regions    12    19    26    33
Indonesia                        Developing regions   186   178   252   115
Lao People's Democratic Republic Developing regions    11     6    16    16
Malaysia                         Developing regions   786   816   813   448
Myanmar                          Developing regions    80    62    46    31
Philippines                      Developing regions  6051  5921  5249  4562
Singapore                        Developing regions   241   301   337   169
Thailand                         Developing regions    56    53   113    65
Viet Nam                         Developing regions  1191  1829  2162  3404

                                 1984  1985  1986  ...   2005   2006   2007  \
Brunei Darussalam                   2     4    12  ...      4      5     11
Cambodia                           10     7     8  ...    370    529    460
Indonesia                         123   100   127  ...    632    613    657
Lao People's Democratic Republic    7    17    21  ...     42     74     53
Malaysia                          384   374   425  ...    593    580    600
Myanmar                            41    23    18  ...    210    953   1887
Philippines                      3801  3150  4166  ...  18139  18400  19837
Singapore                         128   139   205  ...    392    298    690
```

```
Thailand                          82    66    78  ...    575    500    487
Viet Nam                        7583  5907  2741  ...   1852   3153   2574

                                2008  2009  2010  2011  2012  2013  \
Brunei Darussalam                 10     5    12     6     3     6
Cambodia                         354   203   200   196   233   288
Indonesia                        661   504   712   390   395   387
Lao People's Democratic Republic  32    39    54    22    25    15
Malaysia                         658   640   802   409   358   204
Myanmar                          975  1153   556   368   193   262
Philippines                    24887 28573 38617 36765 34315 29544
Singapore                        734   366   805   219   146   141
Thailand                         519   512   499   396   296   400
Viet Nam                        1784  2171  1942  1723  1731  2112

                                Total
Brunei Darussalam                1200
Cambodia                        13076
Indonesia                       26300
Lao People's Democratic Republic 2178
Malaysia                        48834
Myanmar                         18490
Philippines                   1022782
Singapore                       29158
Thailand                        18348
Viet Nam                       194292

[10 rows x 38 columns]
```

[32]:
```
print('data dimensions:', df_can.shape)
print(df_can.columns)
df_can.head(4)
```

```
data dimensions: (195, 38)
Index(['Continent', 'Region', 'DevName', '1980', '1981', '1982', '1983',
       '1984', '1985', '1986', '1987', '1988', '1989', '1990', '1991', '1992',
       '1993', '1994', '1995', '1996', '1997', '1998', '1999', '2000', '2001',
       '2002', '2003', '2004', '2005', '2006', '2007', '2008', '2009', '2010',
       '2011', '2012', '2013', 'Total'],
      dtype='object')
```

[32]:
```
               Continent           Region             DevName  1980  1981  \
Afghanistan         Asia     Southern Asia  Developing regions    16    39
Albania           Europe   Southern Europe   Developed regions     1     0
Algeria           Africa   Northern Africa  Developing regions    80    67
American Samoa   Oceania         Polynesia  Developing regions     0     1

                 1982  1983  1984  1985  1986  ...  2005  2006  2007  2008  \
```

```
Afghanistan       39    47    71   340   496  ...  3436  3009  2652  2111
Albania            0     0     0     0     1  ...  1223   856   702   560
Algeria           71    69    63    44    69  ...  3626  4807  3623  4005
American Samoa     0     0     0     0     0  ...     0     1     0     0

                2009  2010  2011  2012  2013   Total
Afghanistan     1746  1758  2203  2635  2004  117278
Albania          716   561   539   620   603   31398
Algeria         5393  4752  4325  3774  4331  138878
American Samoa     0     0     0     0     0      12

[4 rows x 38 columns]
```

1.1.4 Visualisasi Data dengan Matplotlib

Matplotlib: Standard Python Visualization Library

[33]:
```
%matplotlib inline

import matplotlib as mpl
import matplotlib.pyplot as plt
```

[34]:
```
print(plt.style.available)
mpl.style.use(['ggplot'])
```

```
['seaborn-bright', 'bmh', 'seaborn-white', 'seaborn-paper', 'fast', 'seaborn-
ticks', 'seaborn-poster', 'seaborn-deep', 'classic', 'seaborn-talk', 'seaborn-
muted', 'seaborn-colorblind', 'Solarize_Light2', 'seaborn', 'seaborn-pastel',
'dark_background', 'seaborn-darkgrid', 'tableau-colorblind10', 'seaborn-
whitegrid', 'ggplot', 'seaborn-notebook', '_classic_test', 'fivethirtyeight',
'grayscale', 'seaborn-dark', 'seaborn-dark-palette']
```

1.1.5 Line Pots

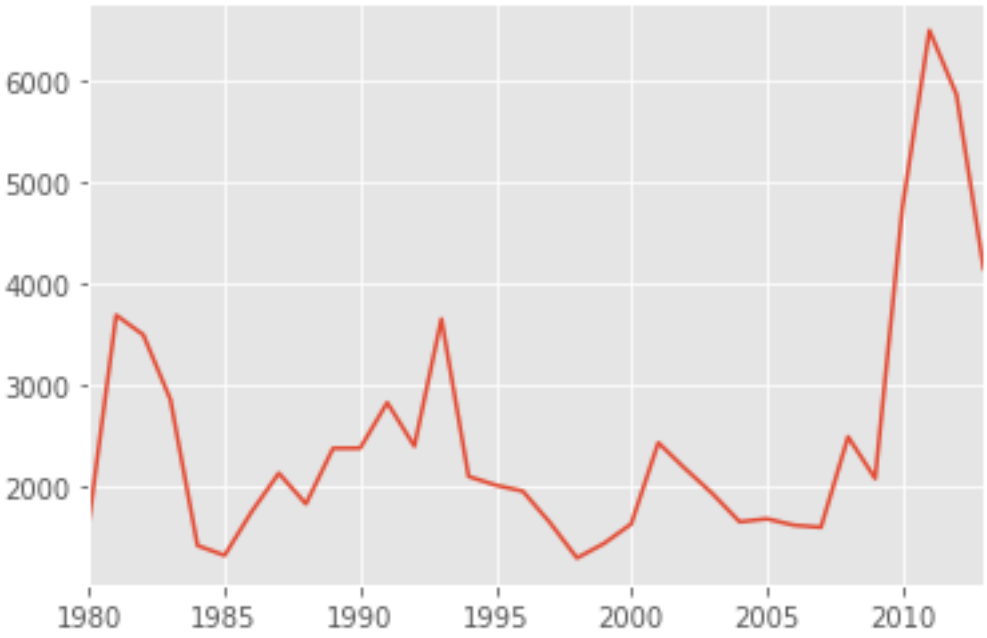
Pada tahun 2010, terjadi gempa di haiti dan Pemerintah Canada menerima pengungsi dari haiti. Kita akan melihat grafik jumlah imigran dari haiti.

[35]:
```
haiti = df_can.loc['Haiti', years] # passing in years 1980 - 2013 to exclude␣
 ↪the 'total' column
haiti.head()
```

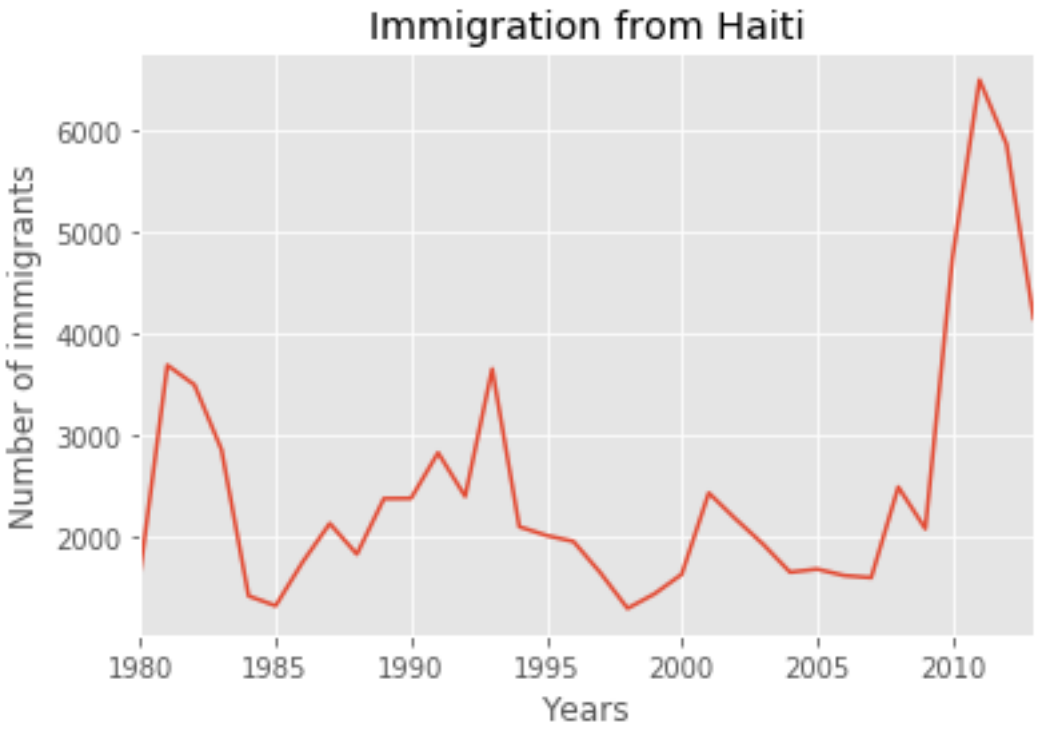
[35]:
```
1980    1666
1981    3692
1982    3498
1983    2860
1984    1418
Name: Haiti, dtype: object
```

[36]:
```
haiti.plot()
```

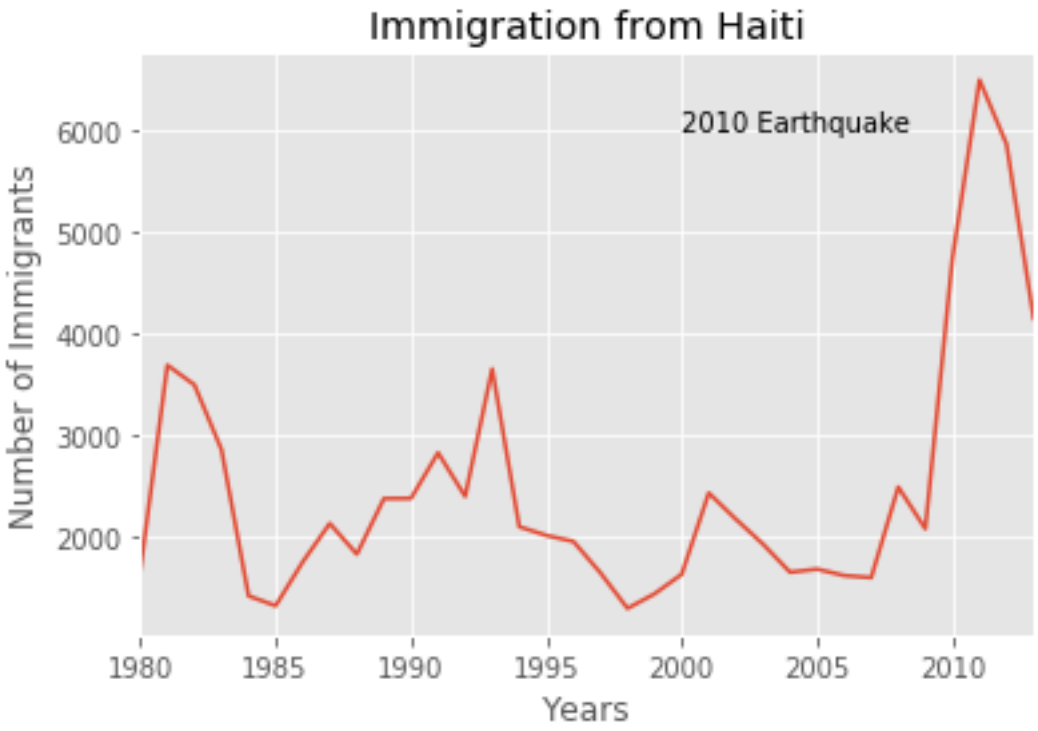
[36]: `<matplotlib.axes._subplots.AxesSubplot at 0x7f467b13d630>`



[37]:
```
# Mengubah tipe index menjadi integer

haiti.index = haiti.index.map(int)
haiti.plot(kind='line')

plt.title('Immigration from Haiti')
plt.ylabel('Number of immigrants')
plt.xlabel('Years')

plt.show()
```



[38]:
```
# Menambahkan title dan label

haiti.plot(kind='line')

plt.title('Immigration from Haiti')
plt.ylabel('Number of Immigrants')
plt.xlabel('Years')

# annotate the 2010 Earthquake.
plt.text(2000, 6000, '2010 Earthquake') # see note below

plt.show()
```



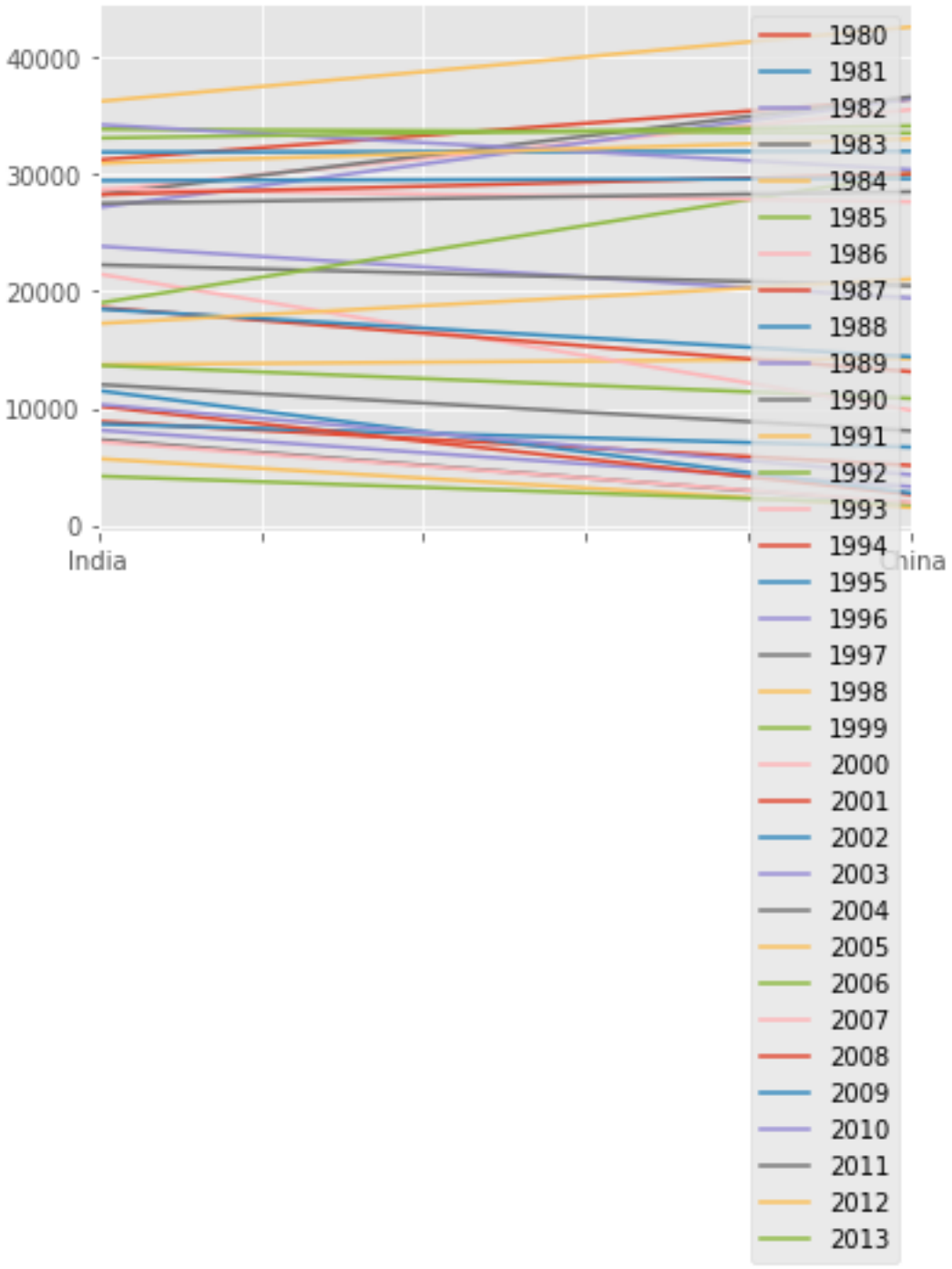
Menampilkan Imigran dari India dan cina

[43]:
```
# Menampilkan data imigran dari cina dan india

df_CI = df_can.loc[['India', 'China'], years]
df_CI.head()
```

[43]:
```
         1980  1981  1982  1983  1984  1985  1986   1987   1988   1989  ...  \
India    8880  8670  8147  7338  5704  4211  7150  10189  11522  10343  ...
China    5123  6682  3308  1863  1527  1816  1960   2643   2758   4323  ...

          2004   2005   2006   2007   2008   2009   2010   2011   2012   2013
India    28235  36210  33848  28742  28261  29456  34235  27509  30933  33087
China    36619  42584  33518  27642  30037  29622  30391  28502  33024  34129

[2 rows x 34 columns]
```

[44]:
```
# Menampilkan grafik plot line

df_CI.plot(kind='line')
```

[44]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7f467ca24c88>
```

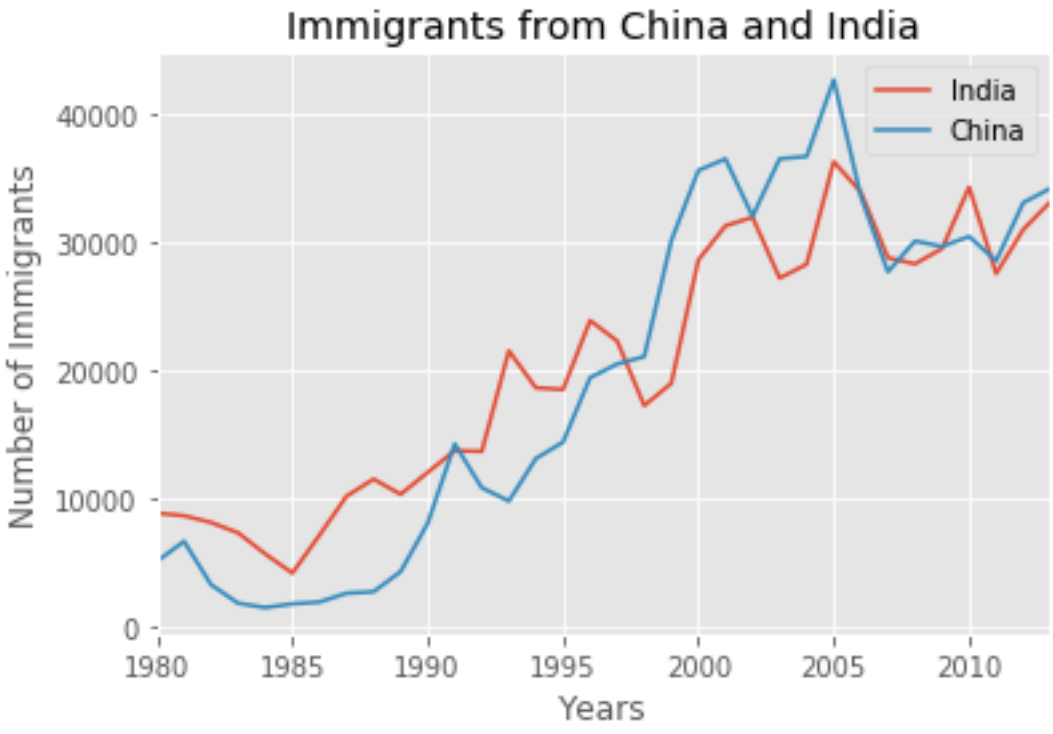


[45]:
```
# Melakukan transpose data

df_CI = df_CI.transpose()
df_CI.head()
```

[45]:

	India	China
1980	8880	5123
1981	8670	6682
1982	8147	3308
1983	7338	1863
1984	5704	1527

[46]:
```
# Mengubah data df_CI ke integer untuk plotting

df_CI.index = df_CI.index.map(int)
df_CI.plot(kind='line')

plt.title('Immigrants from China and India')
plt.ylabel('Number of Immigrants')
plt.xlabel('Years')

plt.show()
```



Membandingkan trend dari 5 terbanyak negara yang menjadi imigran di Kanada

[47]:
```
# Melakukan Sorting nilai terbanyak ke terendah

df_can.sort_values(by='Total', ascending=False, axis=0, inplace=True)
```

[48]:
```
# Menampilkan 5 negara terbanyak

df_top5 = df_can.head(5)
```

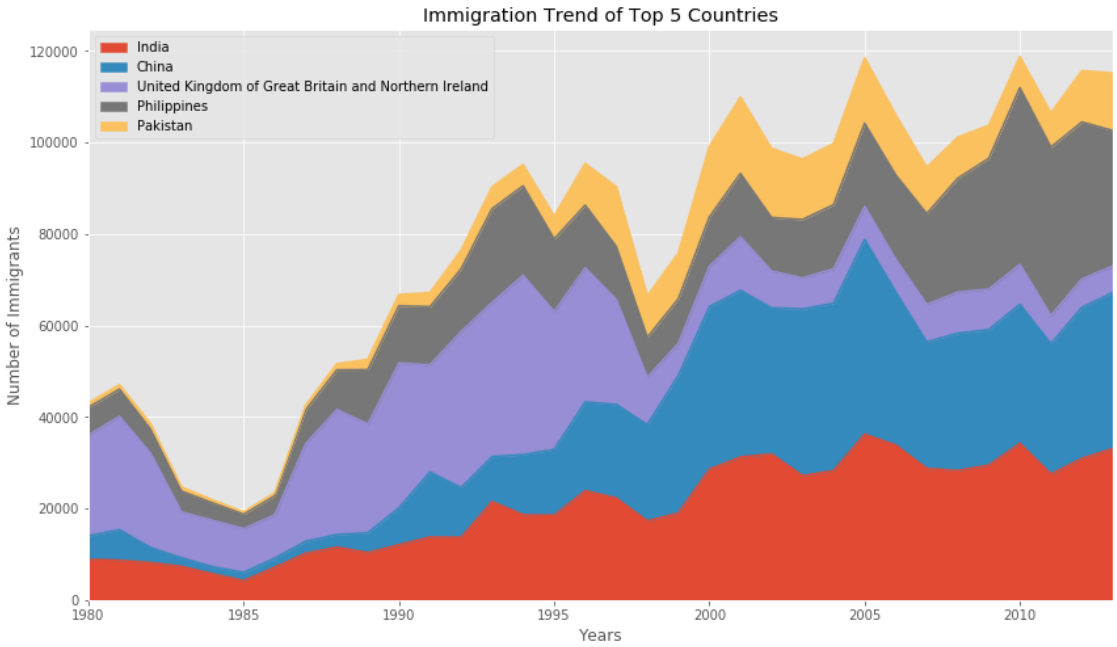
[49]:
```
df_top5 = df_top5[years].transpose()
```

[50]:
```
print(df_top5)
```

```
      India  China  United Kingdom of Great Britain and Northern Ireland  \
1980   8880   5123                                                22045
1981   8670   6682                                                24796
1982   8147   3308                                                20620
1983   7338   1863                                                10015
1984   5704   1527                                                10170
1985   4211   1816                                                 9564
1986   7150   1960                                                 9470
1987  10189   2643                                                21337
1988  11522   2758                                                27359
1989  10343   4323                                                23795
1990  12041   8076                                                31668
1991  13734  14255                                                23380
1992  13673  10846                                                34123
1993  21496   9817                                                33720
1994  18620  13128                                                39231
1995  18489  14398                                                30145
1996  23859  19415                                                29322
1997  22268  20475                                                22965
1998  17241  21049                                                10367
1999  18974  30069                                                 7045
2000  28572  35529                                                 8840
2001  31223  36434                                                11728
2002  31889  31961                                                 8046
2003  27155  36439                                                 6797
2004  28235  36619                                                 7533
2005  36210  42584                                                 7258
2006  33848  33518                                                 7140
2007  28742  27642                                                 8216
2008  28261  30037                                                 8979
2009  29456  29622                                                 8876
2010  34235  30391                                                 8724
2011  27509  28502                                                 6204
2012  30933  33024                                                 6195
2013  33087  34129                                                 5827

      Philippines  Pakistan
1980         6051       978
1981         5921       972
1982         5249      1201
```

```
1983          4562        900
1984          3801        668
1985          3150        514
1986          4166        691
1987          7360       1072
1988          8639       1334
1989         11865       2261
1990         12509       2470
1991         12718       3079
1992         13670       4071
1993         20479       4777
1994         19532       4666
1995         15864       4994
1996         13692       9125
1997         11549      13073
1998          8735       9068
1999          9734       9979
2000         10763      15400
2001         13836      16708
2002         11707      15110
2003         12758      13205
2004         14004      13399
2005         18139      14314
2006         18400      13127
2007         19837      10124
2008         24887       8994
2009         28573       7217
2010         38617       6811
2011         36765       7468
2012         34315      11227
2013         29544      12603
```

[64]:
```
# Menampilkan grafik Top 5 countries

df_top5.index = df_top5.index.map(int)
df_top5.plot(kind='area', figsize=(14, 8))

plt.title('Immigration Trend of Top 5 Countries')
plt.ylabel('Number of Immigrants')
plt.xlabel('Years')

plt.show()
```



[65]:
```
df_can.to_csv('Data Imigran Kanada.csv', index=False)
```

[ ]:
```

```