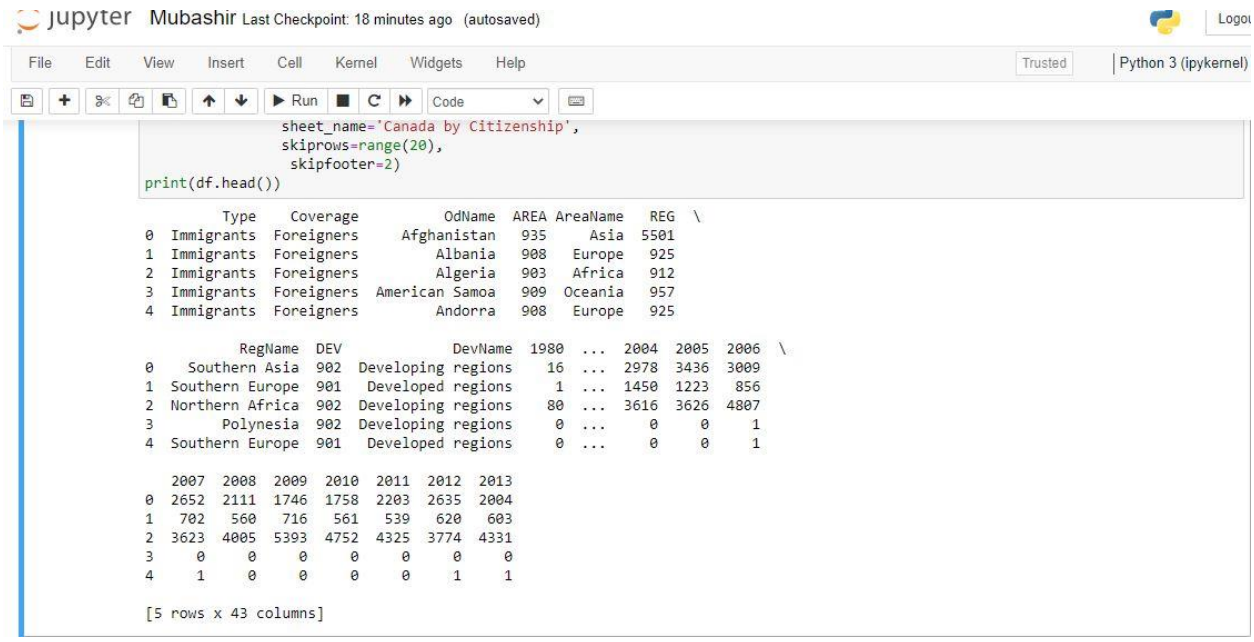
Step 1

Output



Step 2

Output

Step 3

```
In [19]: df.rename(columns={'OdName':'Country','AreaName':'Continent', 'RegName':'Region'}, inplace=True)
         print(df.columns)
```

Output

```
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],
      dtype='object')
```

Step 4

```
In [27]: df['Total']=df.sum(axis=1)
         print(df.head)
```

Output

```
<bound method NDFrame.head of               Continent             Region            DevName  1980  1981  \
OdName
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
...                  ...                ...                ...   ...   ...
Viet Nam            Asia  South-Eastern Asia  Developing regions  1191  1829
Western Sahara    Africa     Northern Africa  Developing regions     0     0
Yemen               Asia        Western Asia  Developing regions     1     2
Zambia            Africa      Eastern Africa  Developing regions    11    17
Zimbabwe          Africa      Eastern Africa  Developing regions    72   114

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

Step 5

```
In [14]: print (df.isnull().sum())
```

Output

```
OdName      0
AreaName    0
RegName     0
DevName     0
1980        0
1981        0
1982        0
1983        0
1984        0
1985        0
```

Step 6

```
In [16]: df=df.set_index('OdName')
         print(df.head)
```

Output

```
<bound method NDFrame.head of                AreaName            RegName            DevName  1980  1981  \
OdName
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
...                 ...                ...                ...   ...   ...
Viet Nam           Asia  South-Eastern Asia  Developing regions  1191  1829
Western Sahara   Africa     Northern Africa  Developing regions     0     0
Yemen              Asia        Western Asia  Developing regions     1     2
Zambia           Africa      Eastern Africa  Developing regions    11    17
Zimbabwe         Africa      Eastern Africa  Developing regions    72   114

                1982  1983  1984  1985  1986  ...  2004  2005  2006  2007  \
OdName                                        ...
Afghanistan       39    47    71   340   496  ...  2978  3436  3009  2652
Albania            0     0     0     0     1  ...  1450  1223   856   702
Algeria           71    69    63    44    69  ...  3616  3626  4807  3623
American Samoa     0     0     0     0     0  ...     0     0     1     0
Andorra            0     0     0     0     2  ...     0     0     1     1
```

Step 7

```
In [17]: print (plt.style.available)
         mpl.style.use(['ggplot'])
```

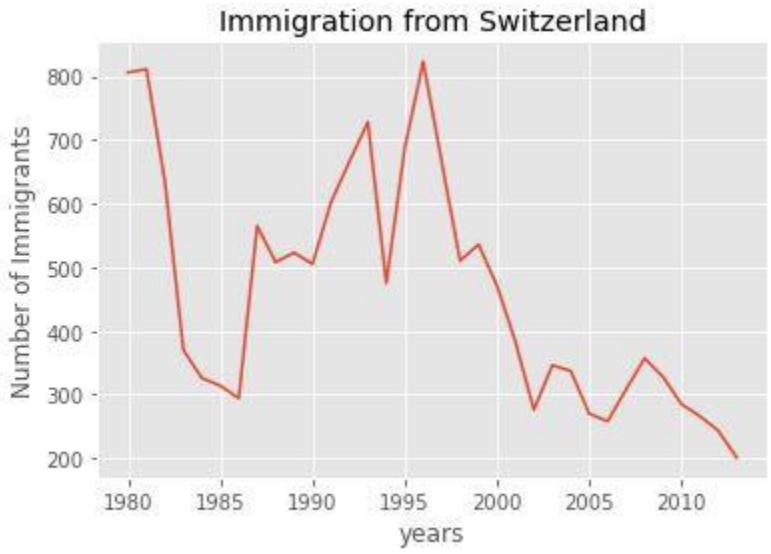
Output

```
['Solarize_Light2', '_classic_test_patch', 'bmh', 'classic', 'dark_background', 'fast', 'fivethirtyeight', 'ggplot', 'grayscal
e', 'seaborn', 'seaborn-bright', 'seaborn-colorblind', 'seaborn-dark', 'seaborn-dark-palette', 'seaborn-darkgrid', 'seaborn-de
p', 'seaborn-muted', 'seaborn-notebook', 'seaborn-paper', 'seaborn-pastel', 'seaborn-poster', 'seaborn-talk', 'seaborn-ticks',
'seaborn-white', 'seaborn-whitegrid', 'tableau-colorblind10']
```

Step 8

```
In [18]: years=list(map(int,range(1980,2014)))
         df.loc['Switzerland', years]
         df.loc['Switzerland', years].plot()
         plt.title('Immigration from Switzerland')
         plt.ylabel('Number of Immigrants')
         plt.xlabel('years')
         plt.show()
```

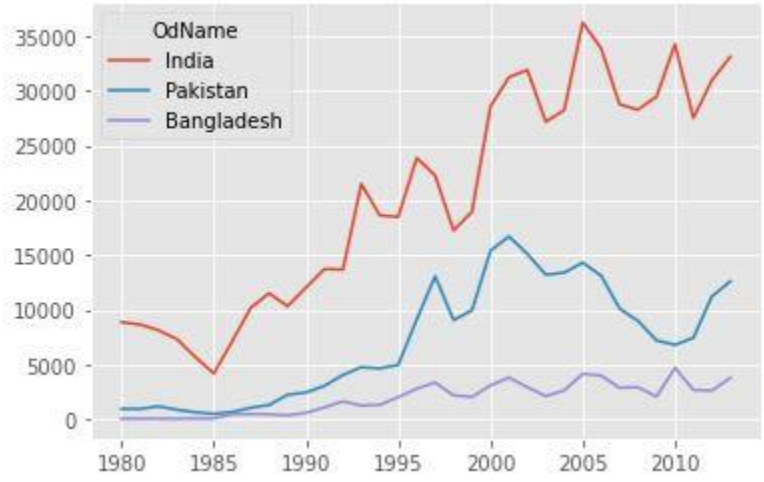
Output



Step 9

```
In [33]: ind_pak_ban=df.loc[['India','Pakistan','Bangladesh'],years]
         ind_pak_ban.head()
         ind_pak_ban.T
         ind_pak_ban.T.plot()
         plt.show()
```

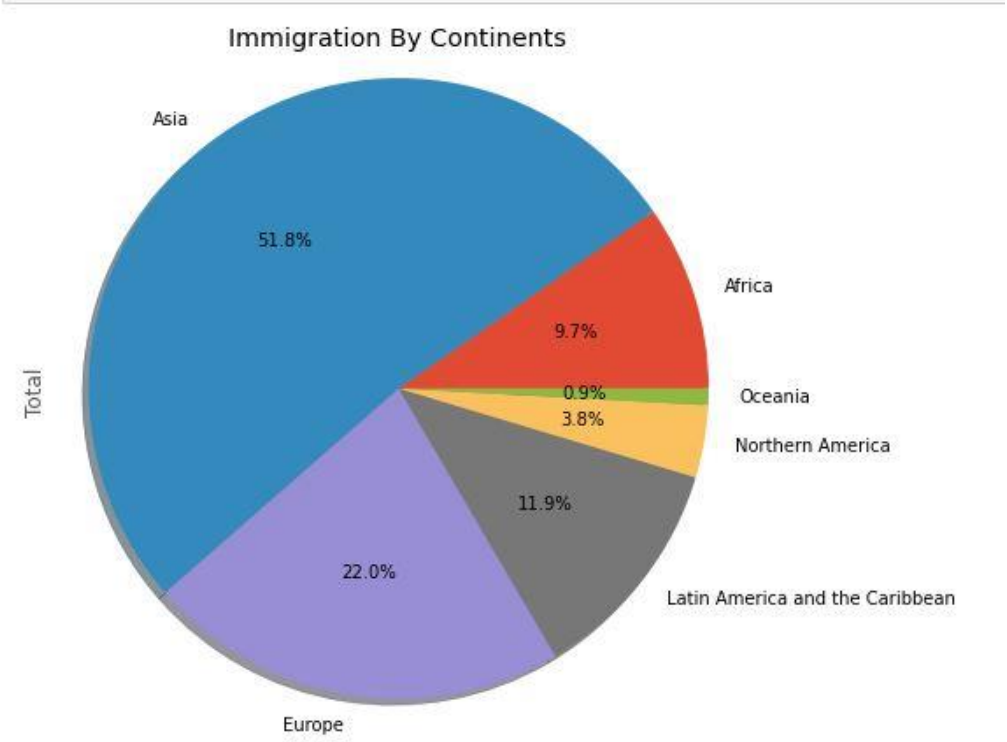
Output



Step 10

```
cont=df.groupby('Continent',axis=0).sum()
cont['Total'].plot(kind='pie',figsize=(7,7),
                   autopct='%1.1f%%',
                   shadow=True)
plt.title('Immigration By Continents')
plt.axis('equal')
plt.show()
```

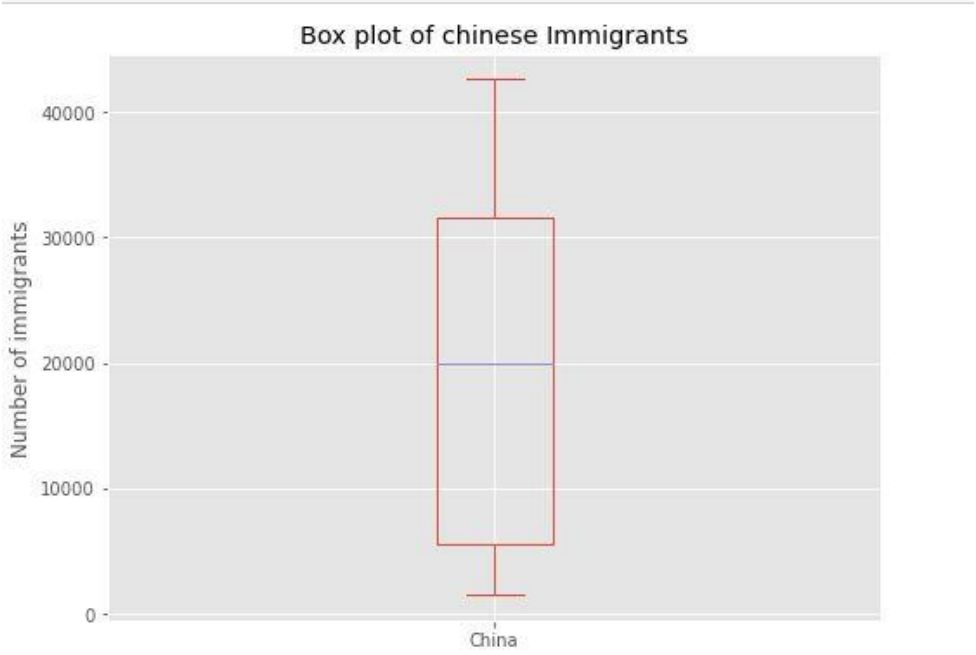
Output



Step 11

```
In [38]: china=df.loc[['China'], years].T
         china.plot(kind='box', figsize=(8, 6))
         plt.title('Box plot of chinese Immigrants')
         plt.ylabel('Number of immigrants')
         plt.show()
```

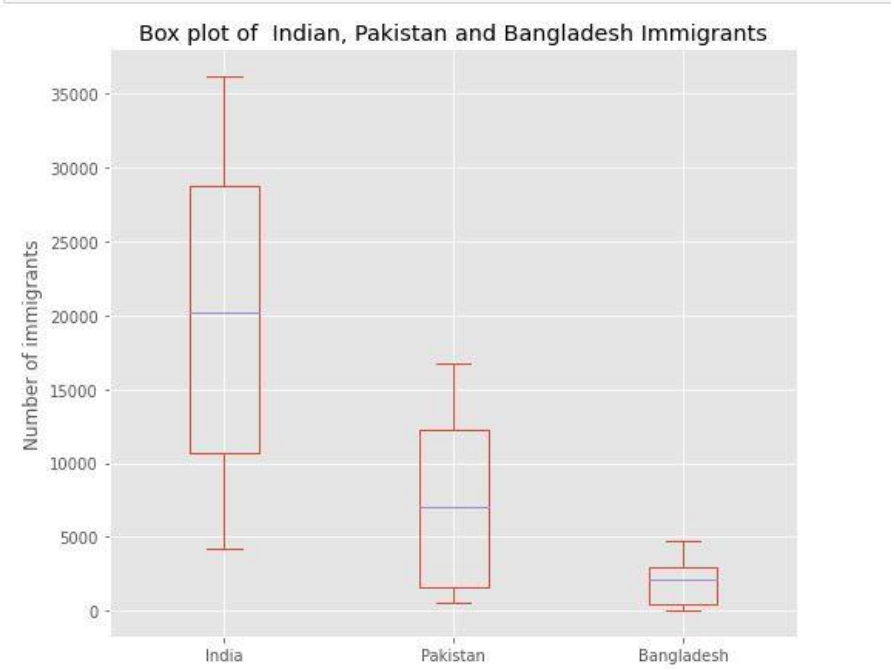
Output



Step 12

```
In [39]: ind_pak_ban.T.plot(kind='box',figsize=(8, 7))
plt.title('Box plot of  Indian, Pakistan and Bangladesh Immigrants')
plt.ylabel('Number of immigrants')
plt.show()
```

Output



Step 13

```
In [43]: totalPerYear=pd.DataFrame(df[years].sum(axis=0))
         totalPerYear.head()
         print(totalPerYear.head())
         totalPerYear.index = map(int, totalPerYear.index)
         totalPerYear.reset_index(inplace=True)
         totalPerYear.head()
         print(totalPerYear.head())
```

Output

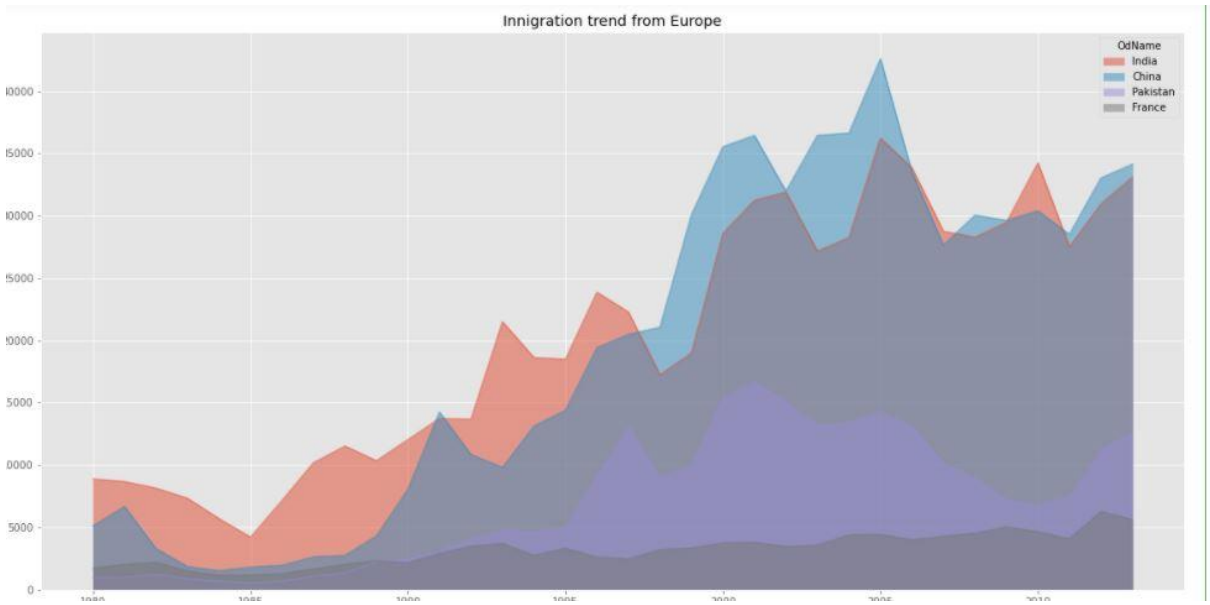
```
           0
1980   99137
1981  110563
1982  104271
1983   75550
1984   73417
   index       0
0   1980   99137
1   1981  110563
2   1982  104271
3   1983   75550
4   1984   73417
```

Step 14

```
top = df.loc[['India', 'China', 'Pakistan', 'France'], years]
top = top.T
print(top)
colors = ['Black', 'Green', 'Blue', 'Red']
top.plot( kind='area', stacked=False, figsize=(20, 10))
plt.title('Innigration trend from Europe')
plt.ylabel('Number of Immigrants')
plt.xlabel('Years')
plt.show()
```

Output

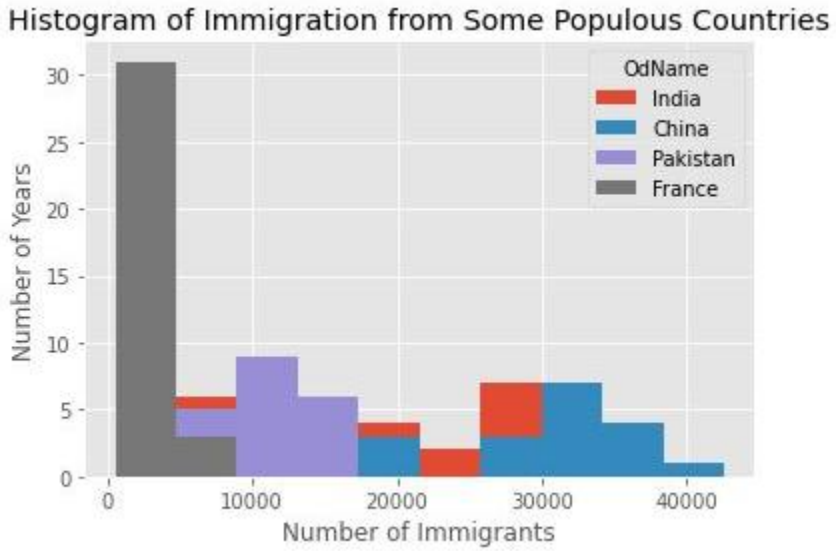
OdName	India	China	Pakistan	France
1980	8880	5123	978	1729
1981	8670	6682	972	2027
1982	8147	3308	1201	2219
1983	7338	1863	900	1490
1984	5704	1527	668	1169
1985	4211	1816	514	1177
1986	7150	1960	691	1298
1987	10189	2643	1072	1658
1988	11522	2758	1334	2038
1989	10343	4323	2261	2332
1990	12041	8076	2470	2177
1991	13734	14255	3079	2911



Step 15


```
In [47]: top.plot.hist()
         plt.title('Histogram of Immigration from Some Populous Countries')
         plt.ylabel('Number of Years')
         plt.xlabel('Number of Immigrants')
```

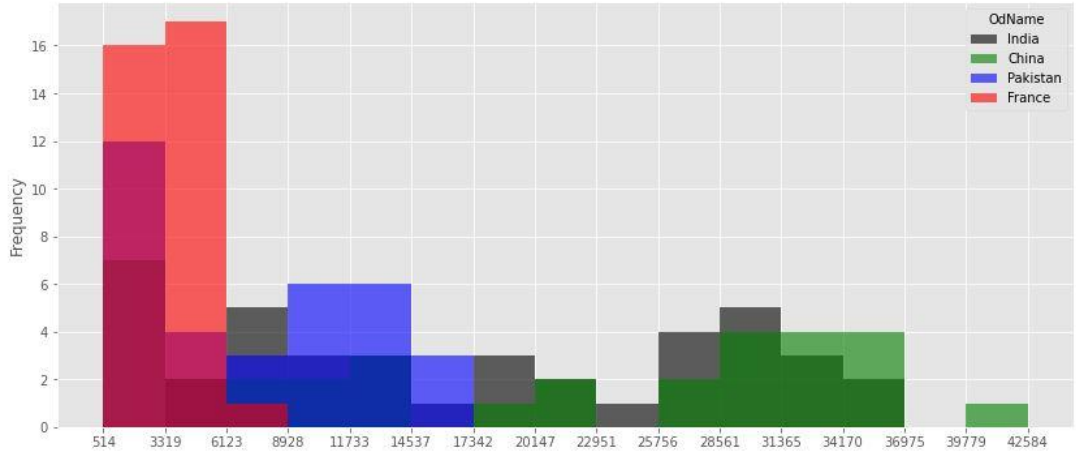
Output



Step 16

```
In [50]: count, bin_edges = np.histogram(top, 15)
         top.plot(kind='hist', figsize=(14, 6), bins=15, alpha=0.6, xticks=bin_edges, color=colors)
         plt.show()
```

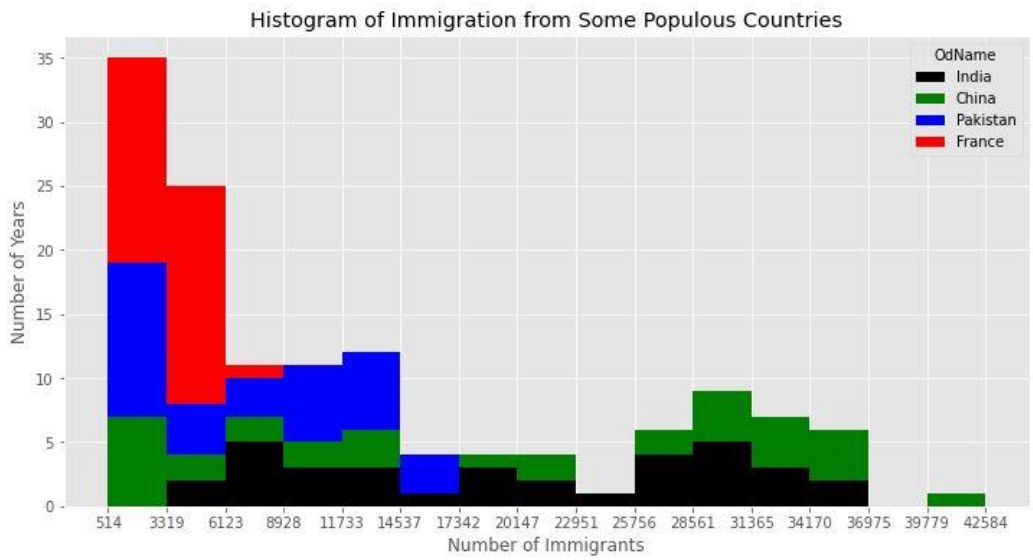
Output



Step 17

```
In [52]: top.plot(kind='hist',
                  figsize=(12, 6),
                  bins=15,
                  xticks=bin_edges,
                  color=colors,
                  stacked=True,
                 )
         plt.title('Histogram of Immigration from Some Populous Countries')
         plt.ylabel('Number of Years')
         plt.xlabel('Number of Immigrants')
         plt.show()
```

Output

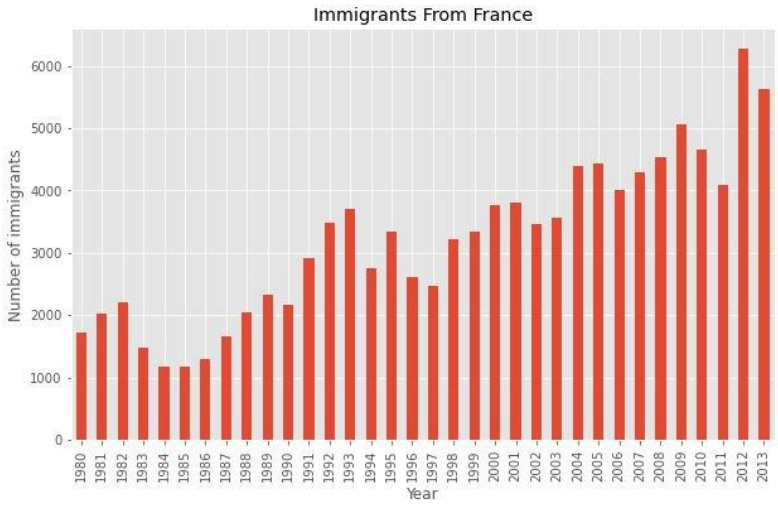


Step 18

```
In [54]: france = df.loc['France', years]
         france.plot(kind='bar', figsize=(10, 6))
         plt.xlabel('Year')
         plt.ylabel('Number of immigrants')
         plt.title('Immigrants From France')
```

Output

```
Out[54]: Text(0.5, 1.0, 'Immigrants From France')
```

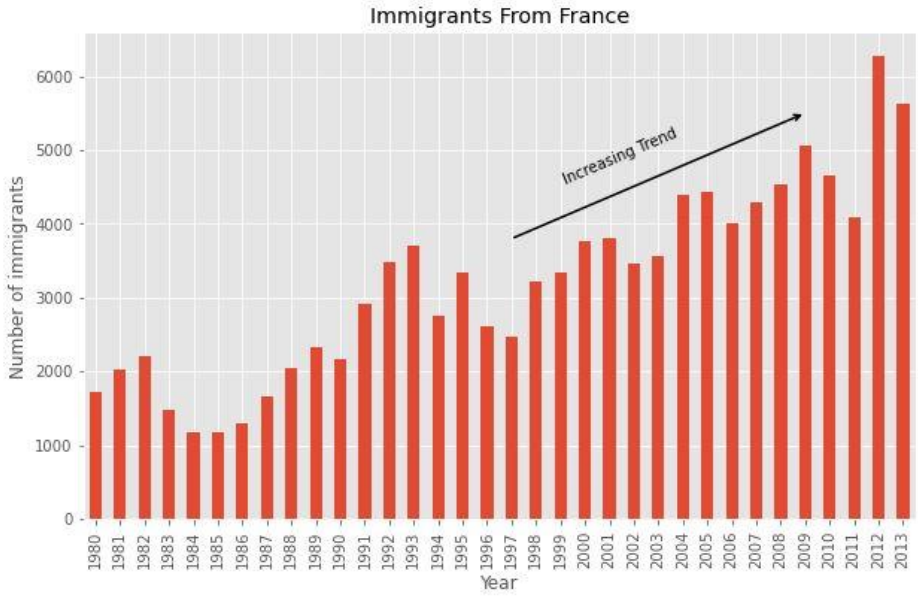


Step 19

```
In [63]: france.plot(kind='bar', figsize=(10, 6))
         plt.xlabel('Year')
         plt.ylabel('Number of immigrants')
         plt.title('Immigrants From France')
         plt.annotate('Increasing Trend',
                      xy=(19, 4500),
                      rotation=23,
                      va='bottom',
                      ha='left')
         plt.annotate('',
                      xy=(29, 5500),
                      xytext=(17, 3800),
                      xycoords='data',
                      arrowprops=dict(arrowstyle='->', connectionstyle='arc3', color='black', lw=1.5))
```

Output

```
Out[63]: Text(17, 3800, '')
```

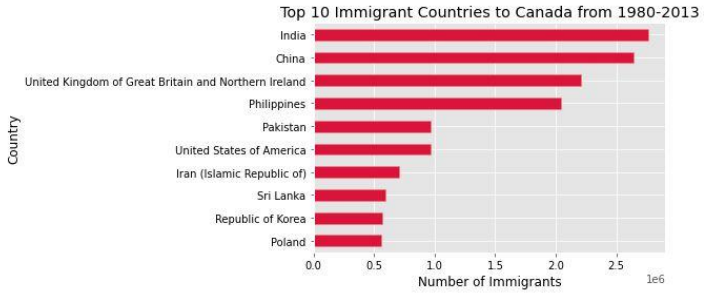


Step 20

```
In [67]: df_top10 = pd.DataFrame(df.nlargest(10, 'Total')['Total'].sort_values( ascending=True))
         df_top10.plot.barh(legend=False, color='crimson', edgecolor ='LightCoral')
         plt.title('Top 10 Immigrant Countries to Canada from 1980-2013', color='black')
         plt.xlabel('Number of Immigrants', color='black')
         plt.ylabel('Country', color='black')
         plt.xticks(color='black')
         plt.yticks(color='black')
```

Output

```
Text(0, 6, 'Philippines'),
Text(0, 7, 'United Kingdom of Great Britain and Northern Ireland'),
Text(0, 8, 'China'),
Text(0, 9, 'India')])
```

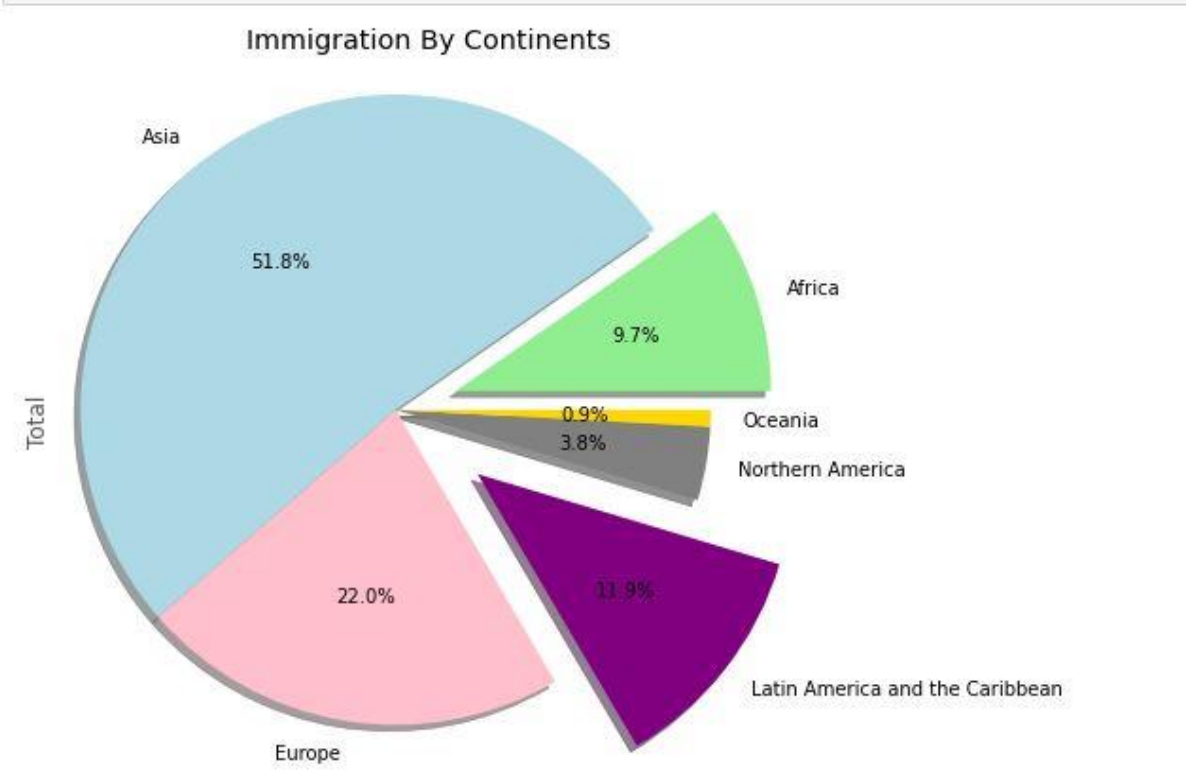


5 Part#3: Challenging Questions

Q1. Improve the appearance of the pie plot we did in step 4 of the plotting exercise subsection. First, choose the colors as light green, light blue, pink, purple, grey, and gold. Second, set the explode attribute. Finally, set the starting angle of the pie plot to any value from 0 to 360 accordingly.

Ans:-

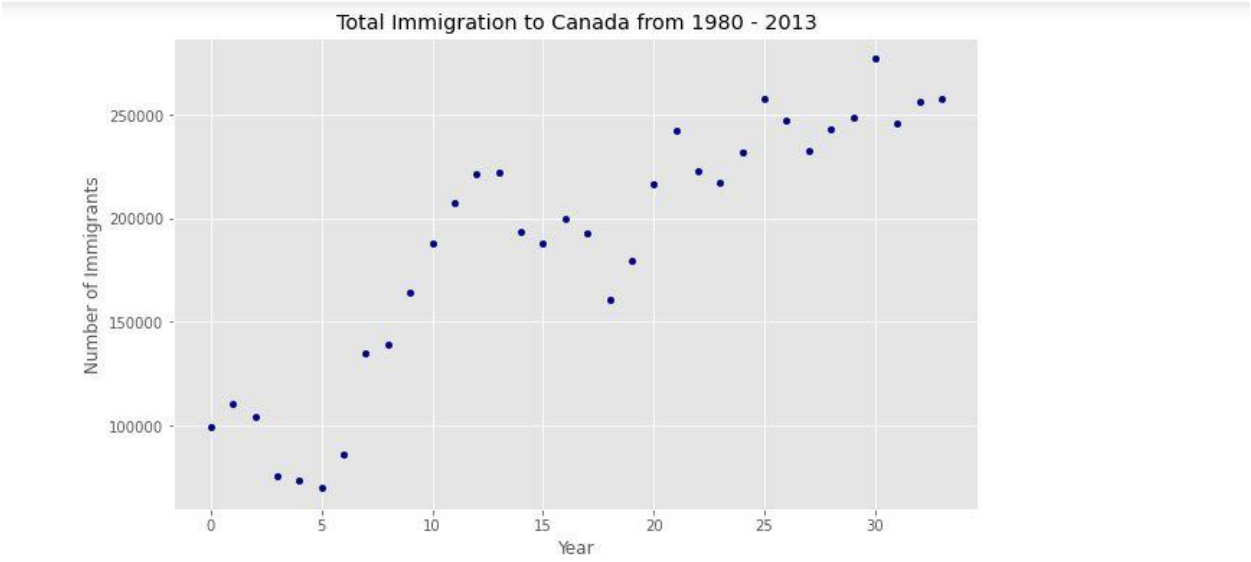
In [117]:
```
cont=df.groupby('Continent',axis=0).sum()
myexplode = [0.2, 0, 0, 0.33,0 ,0]
mycolors = ["lightgreen", "lightblue","pink","purple", "grey", "gold"]
cont['Total'].plot(kind='pie',figsize=(7,7),
                   autopct='%1.1f%%', colors=mycolors,explode = myexplode,

                   shadow=True
    )

plt.title('Immigration By Continents')
plt.axis('equal')
plt.show()
```



Q2. Produce a scatter plot using the DataFrame in step 7 of the plotting exercise subsection to see the trend in the number of immigrants to Canada over the years

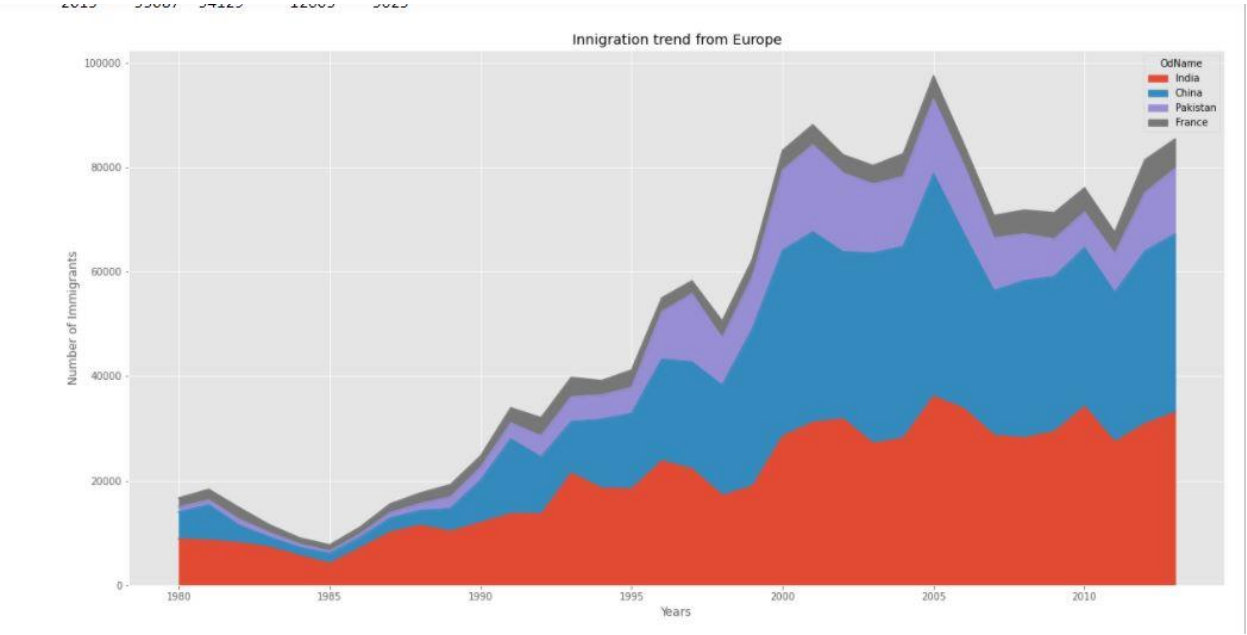
Ans:-

In [123]:
```
df_t = pd.DataFrame(df[years].sum(axis=0))
df_t.index = map(int, df_tot.index)

df_t.reset_index(inplace = True)

df_t.columns = ['year', 'total']


df_t.plot(kind='scatter', x='year', y='total', figsize=(10, 6), color='darkblue')

plt.title('Total Immigration to Canada from 1980 - 2013')
plt.xlabel('Year')
plt.ylabel('Number of Immigrants')
plt.show()
```



Q3. Produce an area plot using the "top" DataFrame in step 8 of the plotting exercise subsection to see the individual countries' area plot. This can be done using the "*stacked*"feature enabled i.e. set the value for the "*stacked*" feature as "*True*".

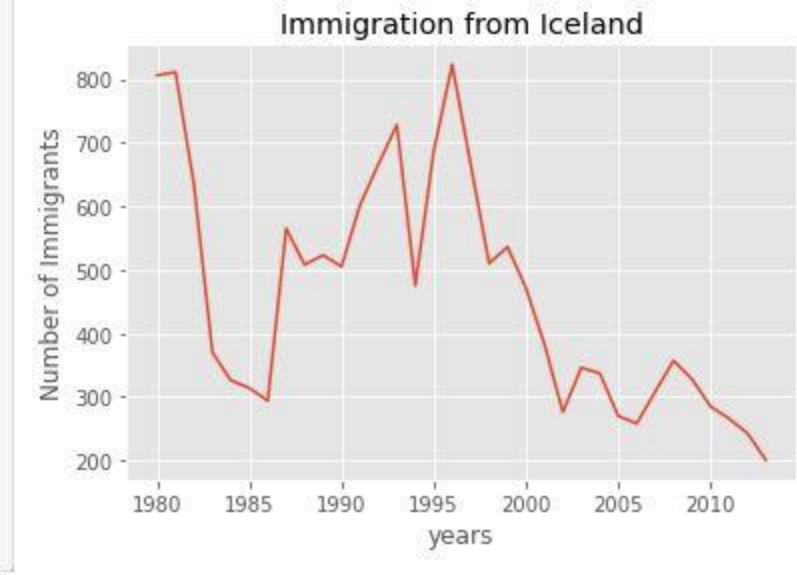
In [126]:
```
top = df.loc[['India', 'China', 'Pakistan', 'France'], years]
top = top.T
print(top)
colors = ['Black', 'Green', 'Blue', 'Red']
top.plot( kind='area', stacked=True, figsize=(20, 10))
plt.title('Innigration trend from Europe')
plt.ylabel('Number of Immigrants')
plt.xlabel('Years')
plt.show()
```



Q4. Compare the number of Icelandic immigrants (country = 'Iceland') to Canada from the year 1980 to 2013 using a vertical bar plot. Explain in a maximum of five-line how you did it and what do you understand from looking at the output.

Ans:- Immigrants of Iceland is decreased in 1980-1985, 800 number of immigrant's are in 1980 and in 1985 the number of immigrant is decreased to 300. The graph show that in 1995 again number of immigrant is increased up to 800, later on it decrease till 2014.

```
In [128]: years=list(map(int,range(1980,2014)))
          df.loc['Iceland', years]
          df.loc['Switzerland', years].plot()
          plt.title('Immigration from Iceland')
          plt.ylabel('Number of Immigrants')
          plt.xlabel('years')
          plt.show()
```



Q5. Produce a horizontal bar plot showing the total number of immigrants to Canada from the top 15 countries, for the period 1980 - 2013. Label each country with the total immigrant count using the scripting layer. Explain in a maximum of five-line how you did it and what do you understand from looking at the output.

Ans:- We draw the graph of top 15 15 countries, for the period 1980 – 2013, the graph show that india has highest number of immigrents and Romania has lowest number of immigrents.

```
In [130]: df_top15 = pd.DataFrame(df.nlargest(15, 'Total')['Total'].sort_values( ascending=True))
          df_top15.plot.barh(legend=False, color='crimson', edgecolor ='LightCoral')
          plt.title('Top 15 Immigrant Countries to Canada from 1980-2013', color='black')
          plt.xlabel('Number of Immigrants', color='black')
          plt.ylabel('Country', color='black')
          plt.xticks(color='black')
          plt.yticks(color='black')
```