In [6]:
```
import pandas as pd
import numpy as np
world_bank = pd.read_excel ("WorldBank.xlsx").rename ({"Country Name": "Country"}, axi
world_bank ["Population (M)"] = (world_bank ["GDP (USD)"]/world_bank ["GDP per capita
world_bank.head ()
```

Out[6]:

	Country	Country Code	Region	IncomeGroup	Year	Birth rate, crude (per 1,000 people)	Death rate, crude (per 1,000 people)	Electric power consumption (kWh per capita)	GDP (USD)
0	Afghanistan	AFG	South Asia	Low income	2018	NaN	NaN	NaN	1.936300e+10
1	Afghanistan	AFG	South Asia	Low income	2017	33.211	6.575	NaN	2.019180e+10
2	Afghanistan	AFG	South Asia	Low income	2016	33.981	6.742	NaN	1.936260e+10
3	Afghanistan	AFG	South Asia	Low income	2015	34.809	6.929	NaN	1.990710e+10
4	Afghanistan	AFG	South Asia	Low income	2014	35.706	7.141	NaN	2.048490e+10

In [7]:
```
world_bank.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 12449 entries, 0 to 12448
Data columns (total 16 columns):
 #   Column                                                Non-Null Count  Dt
ype
---  ------                                                --------------  --
---
 0   Country                                               12449 non-null  ob
ject
 1   Country Code                                          12449 non-null  ob
ject
 2   Region                                                12449 non-null  ob
ject
 3   IncomeGroup                                           12449 non-null  ob
ject
 4   Year                                                  12449 non-null  in
t64
 5   Birth rate, crude (per 1,000 people)                  11440 non-null  fl
oat64
 6   Death rate, crude (per 1,000 people)                  11416 non-null  fl
oat64
 7   Electric power consumption (kWh per capita)           5848 non-null   fl
oat64
 8   GDP (USD)                                             9578 non-null   fl
oat64
 9   GDP per capita (USD)                                  9575 non-null   fl
oat64
 10  Individuals using the Internet (% of population)      5064 non-null   fl
oat64
 11  Infant mortality rate (per 1,000 live births)         9984 non-null   fl
oat64
 12  Life expectancy at birth (years)                      11176 non-null  fl
oat64
 13  Population density (people per sq. km of land area)   11845 non-null  fl
oat64
 14  Unemployment (% of total labor force) (modeled ILO estimate)  5208 non-null   fl
oat64
 15  Population (M)                                        9575 non-null   fl
oat64
dtypes: float64(11), int64(1), object(4)
memory usage: 1.5+ MB
```

In [8]:
```
world_bank.describe()
```

Out[8]:

	Year	Birth rate, crude (per 1,000 people)	Death rate, crude (per 1,000 people)	Electric power consumption (kWh per capita)	GDP (USD)	GDP per capita (USD)	Individu using t Internet populatic
count	12449.00000	11440.000000	11416.000000	5848.000000	9.578000e+03	9575.000000	5064.0000
mean	1989.00000	28.643276	10.588539	3175.294686	1.700740e+11	8231.812259	23.3344
std	17.03007	13.131893	5.489382	4467.139298	8.979866e+11	16173.539954	28.3193
min	1960.00000	6.900000	1.127000	0.000000	8.824450e+06	34.790600	0.0000
25%	1974.00000	16.600000	6.863750	390.385750	1.393010e+09	513.145500	0.5949
50%	1989.00000	27.545500	9.200000	1541.895000	7.275305e+09	1852.810000	8.4062
75%	2004.00000	40.881250	12.687000	4313.767500	4.857782e+10	7774.565000	41.2959
max	2018.00000	58.227000	54.444000	54799.200000	2.050000e+13	189171.000000	100.0000

In [9]:
```
hdi=pd.read_csv("hdi.csv")
wb_hdi_2014 = world_bank.query("Year == 2014"). merge (hdi [["iso3", "hdi_2014"]], how
wb_hdi_2014.head()
```

Out[9]:

	Country	Country Code	Region	IncomeGroup	Year	Birth rate, crude (per 1,000 people)	Death rate, crude (per 1,000 people)	Electric power consumption (kWh per capita)	GDP (USD)
0	Afghanistan	AFG	South Asia	Low income	2014	35.706	7.141	NaN	2.048490e+10
1	Albania	ALB	Europe & Central Asia	Upper middle income	2014	12.259	7.219	2309.37	1.322820e+10
2	Algeria	DZA	Middle East & North Africa	Upper middle income	2014	25.538	4.709	1362.87	2.140000e+11
3	American Samoa	ASM	East Asia & Pacific	Upper middle income	2014	17.500	4.200	NaN	6.430000e+08
4	Andorra	AND	Europe & Central Asia	High income: nonOECD	2014	NaN	NaN	NaN	3.350740e+09

In [10]:
```
wb_hdi_2014.describe()
```

Out[10]:

	Year	Birth rate, crude (per 1,000 people)	Death rate, crude (per 1,000 people)	Electric power consumption (kWh per capita)	GDP (USD)	GDP per capita (USD)	Individuals using the Internet (% of population)	I mor rate 1,00 bi
count	211.0	204.000000	204.000000	139.000000	2.010000e+02	201.000000	199.000000	190.0
mean	2014.0	20.859378	7.693044	4270.600563	3.907779e+11	17779.147856	45.738525	23.84
std	0.0	10.280302	2.706626	5981.464101	1.546117e+12	27113.375249	28.934518	21.75
min	2014.0	7.900000	1.127000	39.055800	3.729180e+07	248.845000	0.990000	1.80
25%	2014.0	11.975000	5.836250	858.174500	6.047810e+09	2163.160000	17.730000	6.80
50%	2014.0	17.799000	7.501500	2588.300000	3.133500e+10	6684.800000	46.160000	15.05
75%	2014.0	28.663750	9.219250	5478.100000	2.060000e+11	20258.000000	69.890000	36.77
max	2014.0	47.988000	16.433000	53832.500000	1.750000e+13	189171.000000	98.160000	93.00

In [11]:
```
gdp_pivot = world_bank.pivot_table(index="Year", columns="Region", values="GDP (USD)",
gdp_pivot.head()
```

Out[11]:

Region	East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	North America	South Asia	Sub-S
Year							
1960	1.464117e+11	2.834935e+11	5.739654e+10	1.410025e+10	5.430845e+11	4.746783e+10	2.64887
1961	1.466530e+11	3.017402e+11	6.076987e+10	1.480286e+10	6.046452e+11	5.062976e+10	2.75348
1962	1.499092e+11	3.299221e+11	9.396297e+10	1.464985e+10	6.479622e+11	5.403093e+10	2.98442
1963	1.672525e+11	3.638093e+11	9.482963e+10	1.646011e+10	6.848096e+11	6.077062e+10	3.41977
1964	1.921179e+11	4.028910e+11	1.056766e+11	1.817288e+10	7.362342e+11	6.960261e+10	3.33803

In [12]:
```
pop_pivot = world_bank. pivot_table(index="Year", columns="Region", values="Populatior
pop_pivot.head()
```

Out[12]:

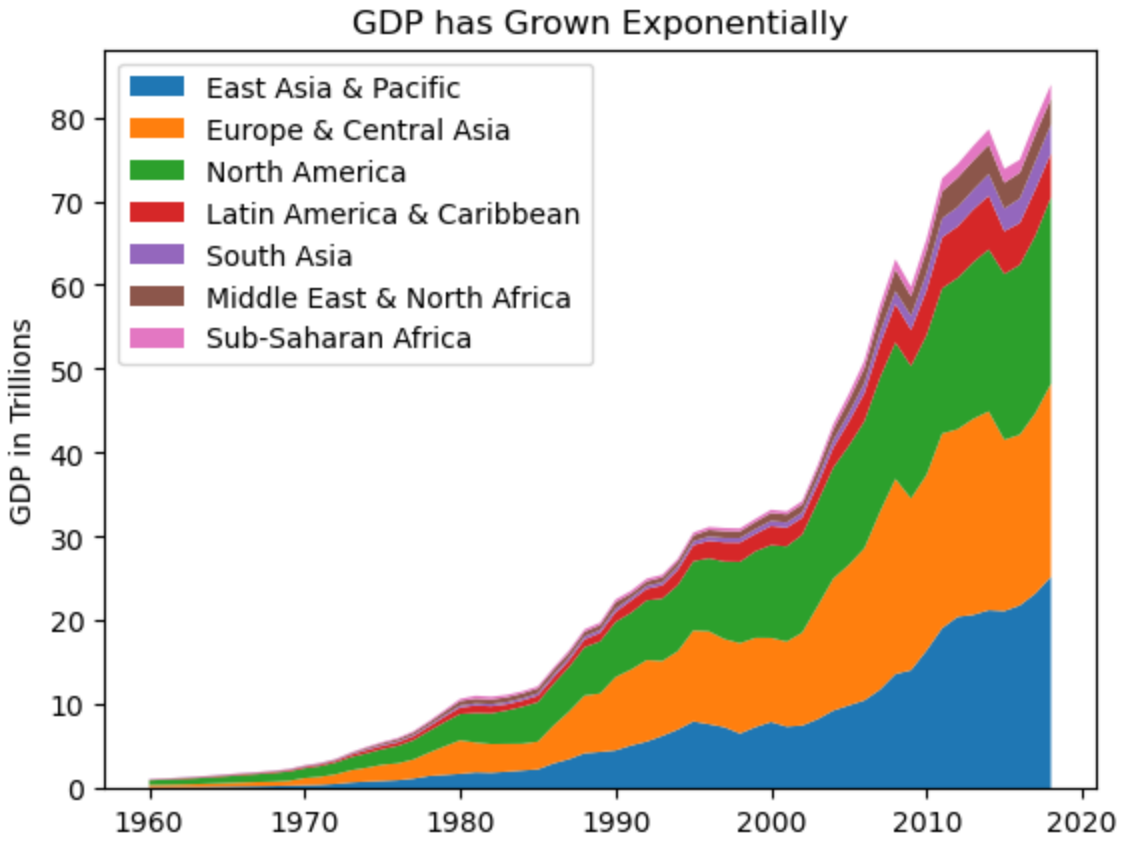
Region	East Asia & Pacific	Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	North America	South Asia	Sub-Saharan Africa
Year							
1960	872.149068	280.746688	186.852628	59.303286	180.615844	572.526561	168.518542
1961	871.192827	283.656647	192.312725	60.945298	201.909874	584.618194	172.577530
1962	881.036241	286.741448	219.093142	62.667579	205.167974	597.165027	177.066655
1963	902.065879	289.789614	225.199250	64.416446	208.372005	610.162283	181.468995
1964	922.572657	292.833206	231.411905	66.231025	211.318874	623.598506	186.049084

In [13]:
```
wb_hdi_by_region = wb_hdi_2014.groupby("Region").agg ({"hdi_2014": "mean"}).sort_value
wb_hdi_by_region.head()
```

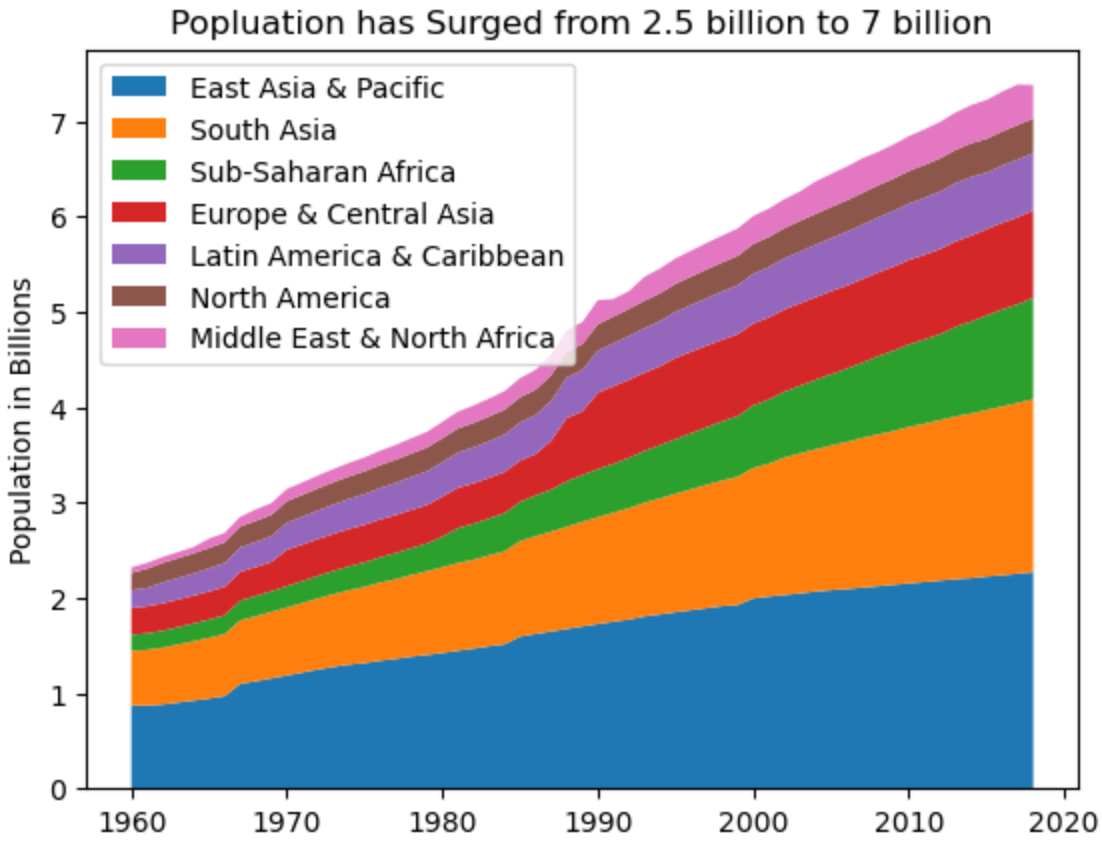
Out[13]:

Region	hdi_2014
North America	0.922000
Europe & Central Asia	0.845320
Latin America & Caribbean	0.739697
Middle East & North Africa	0.736619
East Asia & Pacific	0.729963

In [14]:
```
import matplotlib.pyplot as plt
import seaborn as sns

# Assuming gdp_pivot is a pre-defined DataFrame with a proper structure
fig, ax = plt.subplots()
ax.stackplot(
    gdp_pivot.index,
    [gdp_pivot[region]/1_000_000_000_000 for region in gdp_pivot.iloc[-1].sort_values(
    labels=[region for region in gdp_pivot.iloc[-1].sort_values(ascending=False).index
)
ax.legend(loc="upper left")
ax.set_title("GDP has Grown Exponentially")
ax.set_ylabel("GDP in Trillions")
plt.show()
```

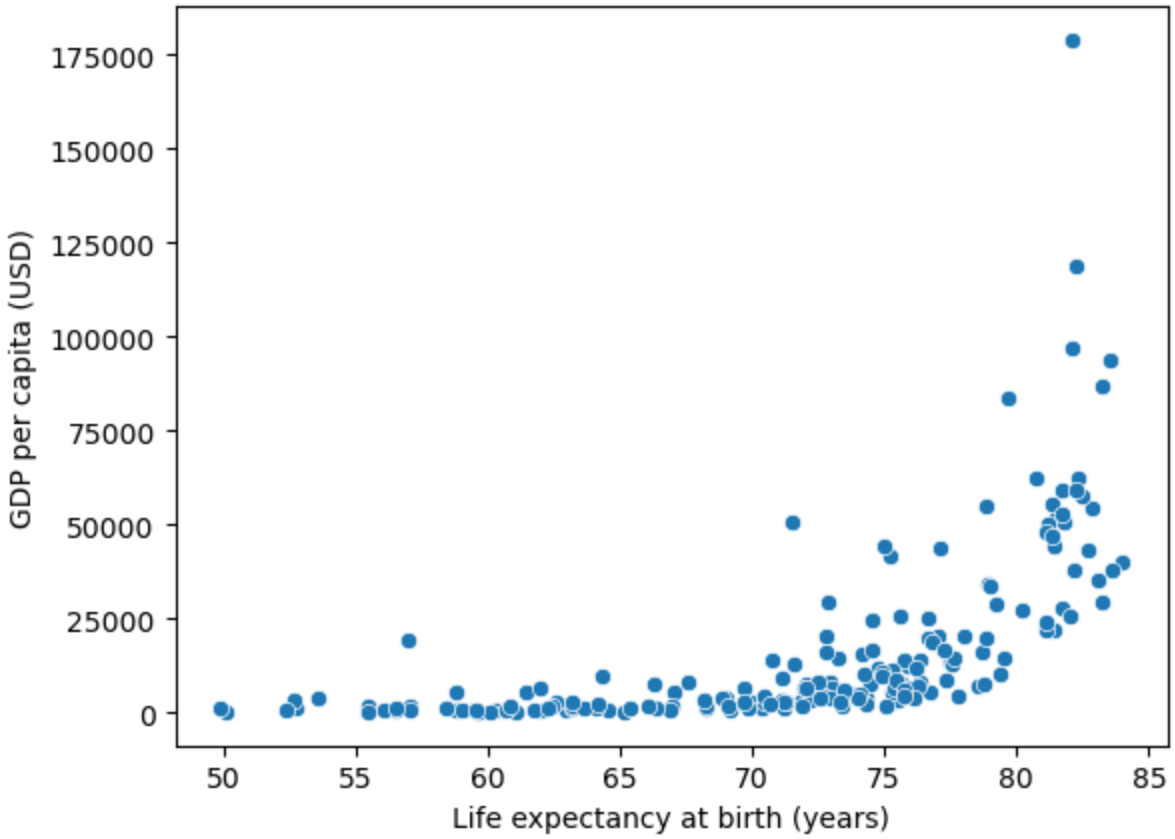


In [15]:
```
import matplotlib.pyplot as plt
import seaborn as sns

# Assuming pop_pivot is a pre-defined DataFrame
fig, ax = plt.subplots()
ax.stackplot(
    pop_pivot.index,
    [pop_pivot[region]/1000 for region in pop_pivot.iloc[-1].sort_values(ascending=Fal
    labels=[region for region in pop_pivot.iloc[-1].sort_values(ascending=False).index
)
ax.legend(loc="upper left")
ax.set_title("Popluation has Surged from 2.5 billion to 7 billion")
ax.set_ylabel("Population in Billions")
plt.show()
```



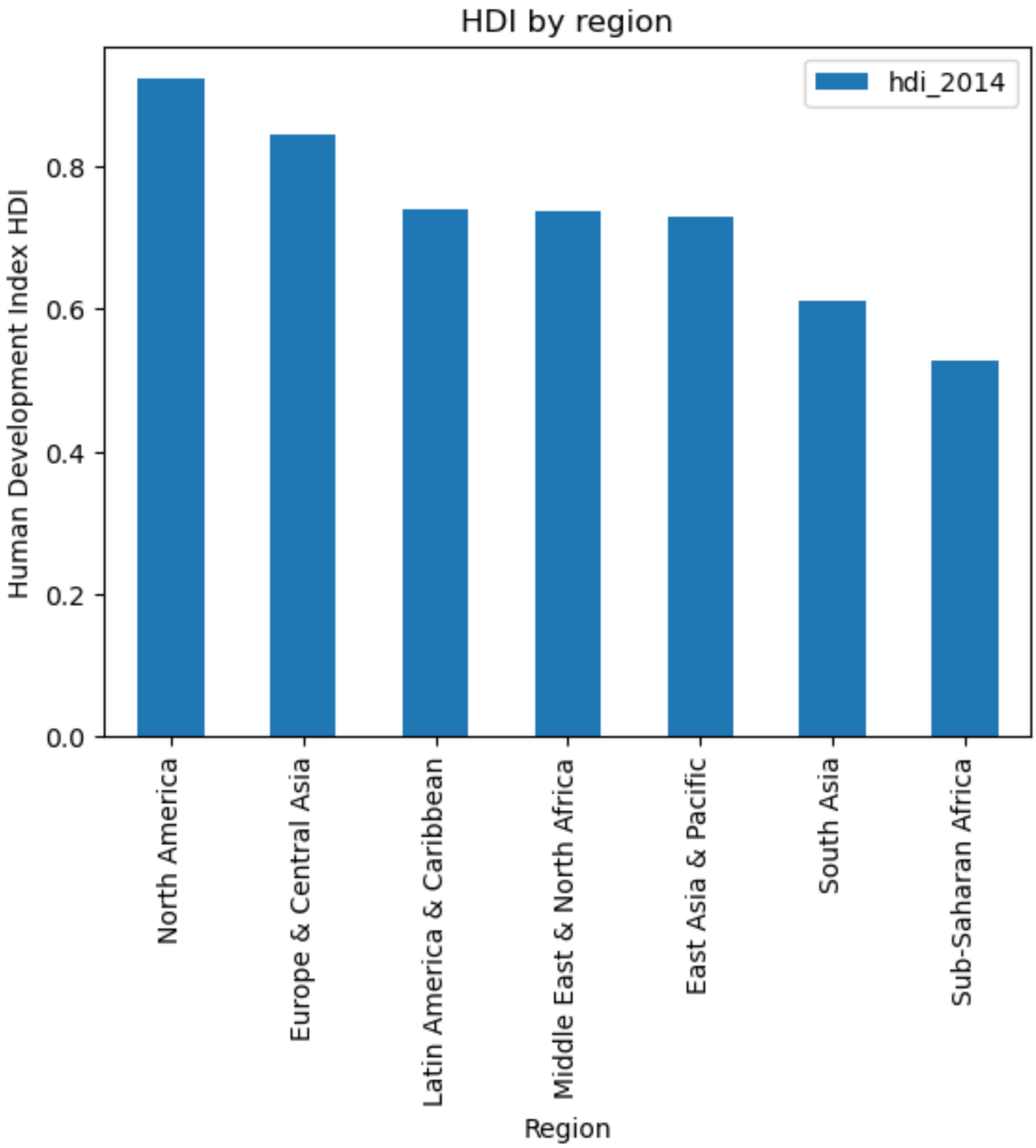
In [16]:
```
sns.scatterplot(
data=wb_hdi_2014,
x="Life expectancy at birth (years)",
y="GDP per capita (USD)",
)
```

Out[16]:
```
<Axes: xlabel='Life expectancy at birth (years)', ylabel='GDP per capita (USD)'>
```



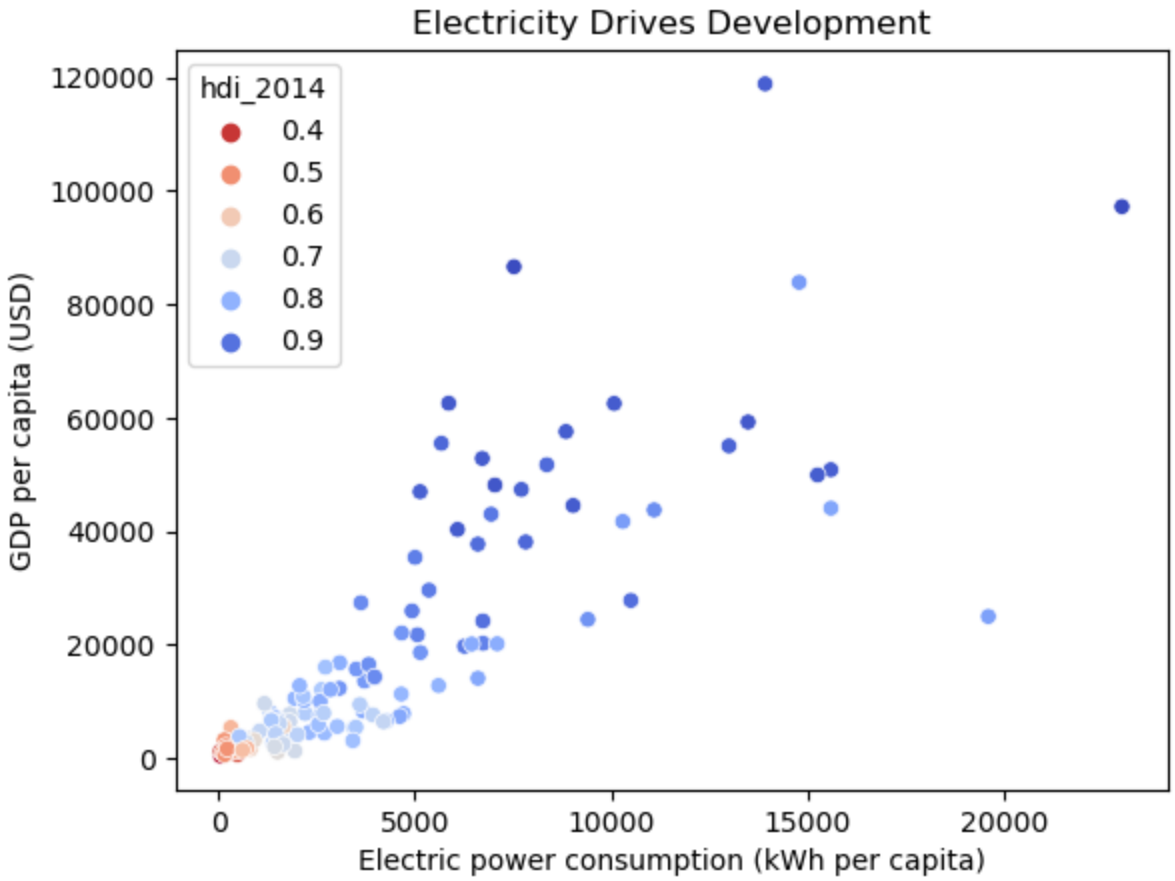
In [18]:
```
wb_hdi_by_region.plot.bar(title="HDI by region",ylabel="Human Development Index HDI")
```

Out[18]:
```
<Axes: title={'center': 'HDI by region'}, xlabel='Region', ylabel='Human Development
Index HDI'>
```

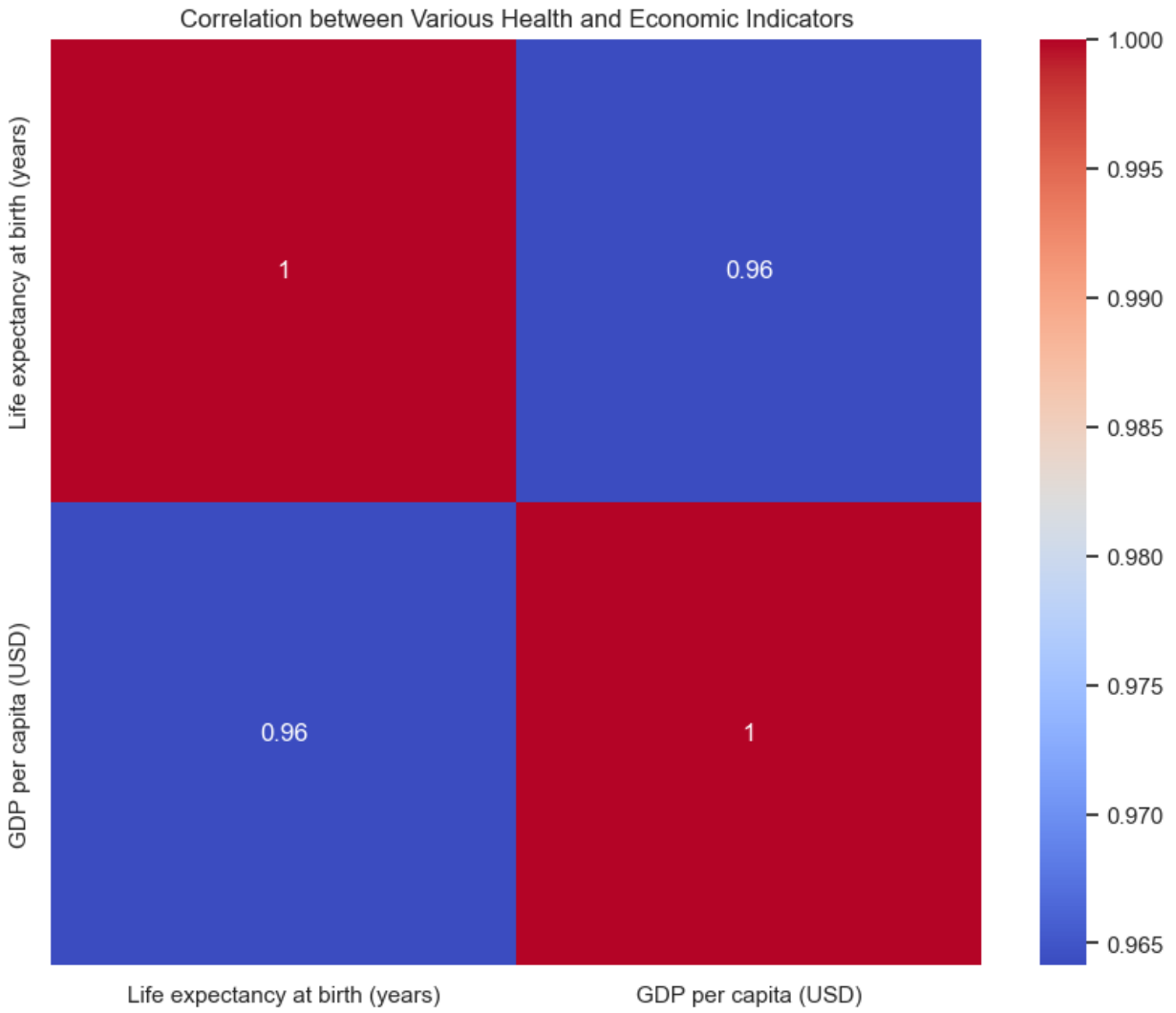


In [19]:
```
sns.scatterplot(
    data=wb_hdi_2014.query("Country != 'Iceland'"),
    x="Electric power consumption (kWh per capita)",
    y="GDP per capita (USD)",
    hue="hdi_2014",
    palette="coolwarm_r"
).set(title="Electricity Drives Development")
```

Out[19]:
```
[Text(0.5, 1.0, 'Electricity Drives Development')]
```



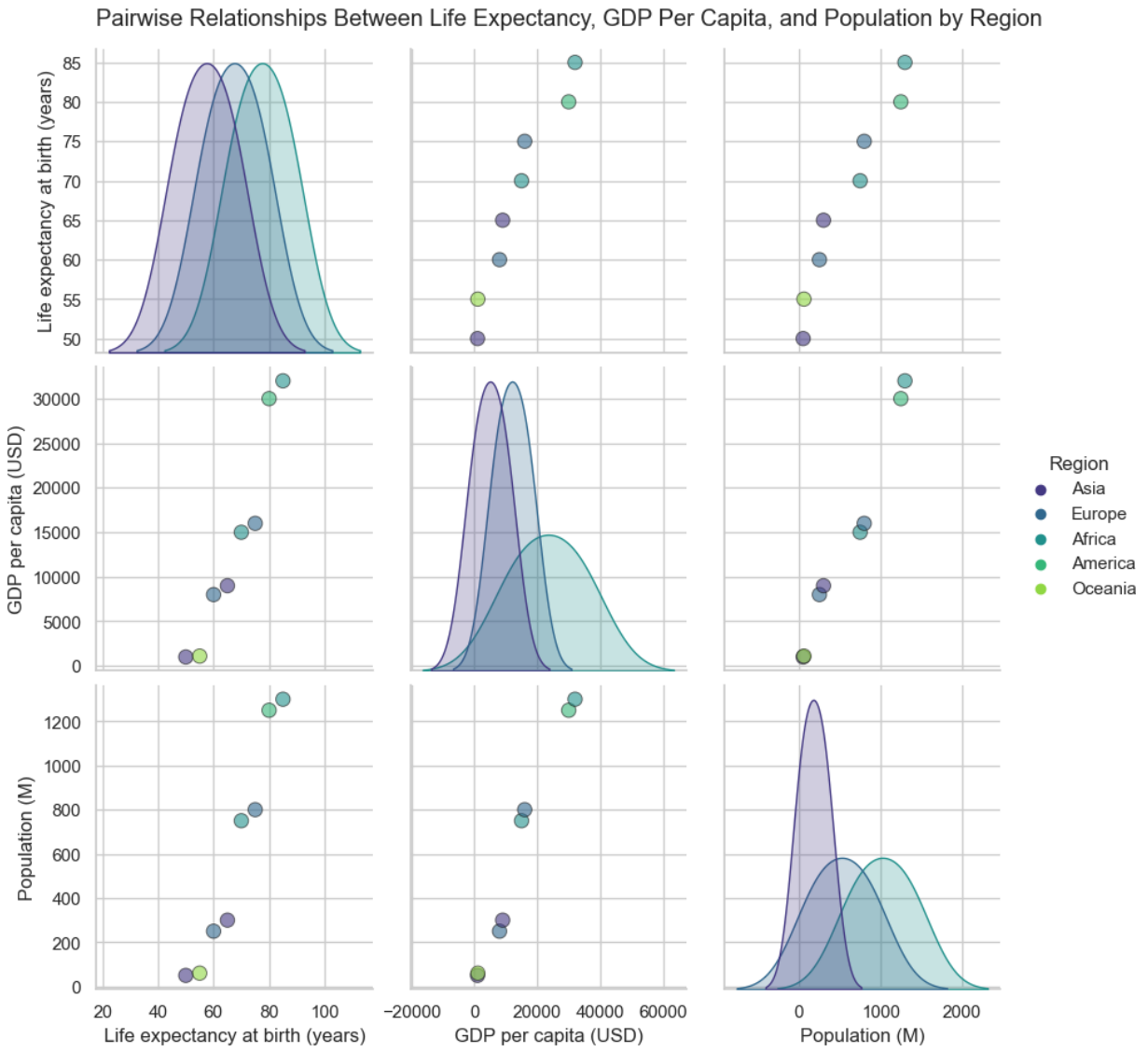
In [72]:
```
import seaborn as sns
import matplotlib.pyplot as plt

# Calculating correlation matrix
corr = wb_hdi_2014[['Life expectancy at birth (years)', 'GDP per capita (USD)', ]].cor

# Plotting heatma
plt.figure(figsize=(10, 8))
sns.heatmap(corr, annot=True, cmap='coolwarm')
plt.title('Correlation between Various Health and Economic Indicators')
plt.show()
```



In [76]:
```
data_for_plotting = wb_hdi_2014[['Life expectancy at birth (years)', 'GDP per capita (
```

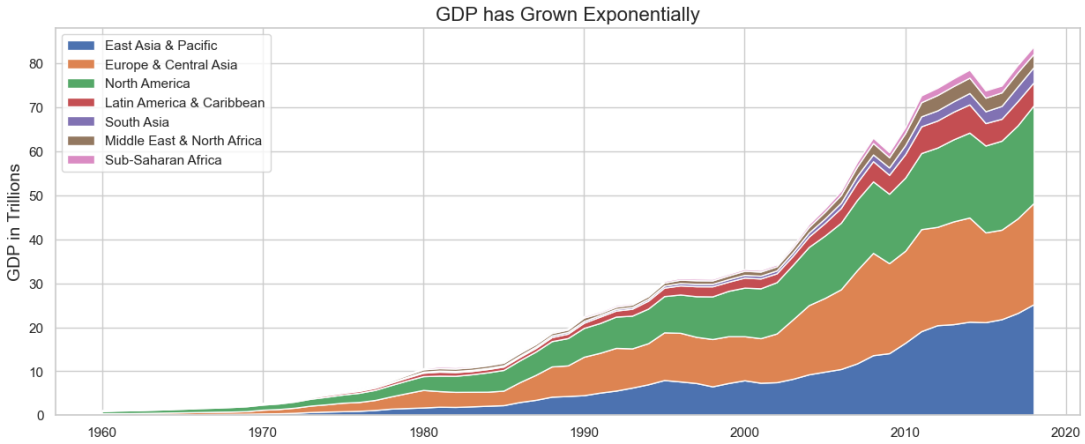
In [77]:
```
# Create the pair plot
pair_plot = sns.pairplot(data=data_for_plotting, hue='Region', height=3,
                         vars=['Life expectancy at birth (years)', 'GDP per capita (US
                         diag_kind='kde',  # KDE plots on the diagonal
                         plot_kws={'alpha': 0.6, 's': 80, 'edgecolor': 'k'},  # Adjust
                         palette='viridis')  # Color palette for different regions

# Set titles and labels for better readability (optional)
pair_plot.fig.suptitle('Pairwise Relationships Between Life Expectancy, GDP Per Capita

# Display the plot
plt.show()
```

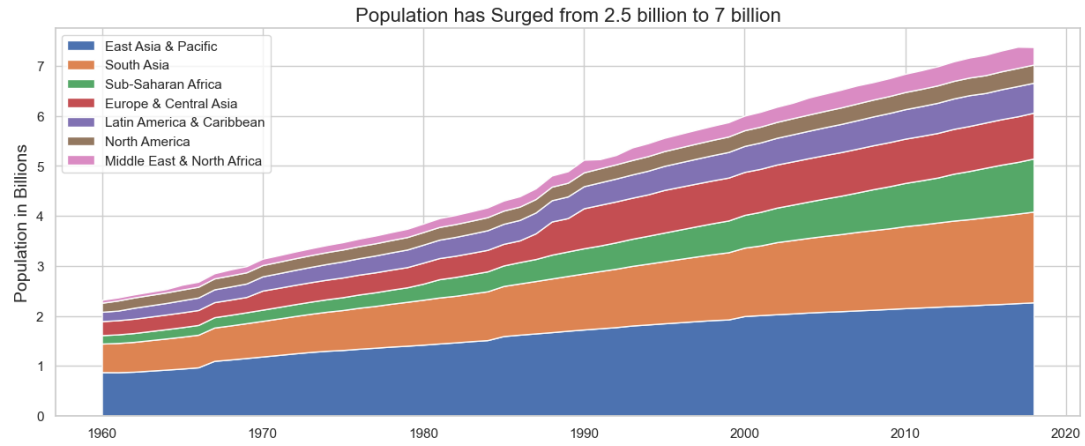
```
C:\Users\manju\anaconda3\Lib\site-packages\seaborn\axisgrid.py:118: UserWarning: The
figure layout has changed to tight
  self._figure.tight_layout(*args, **kwargs)
```

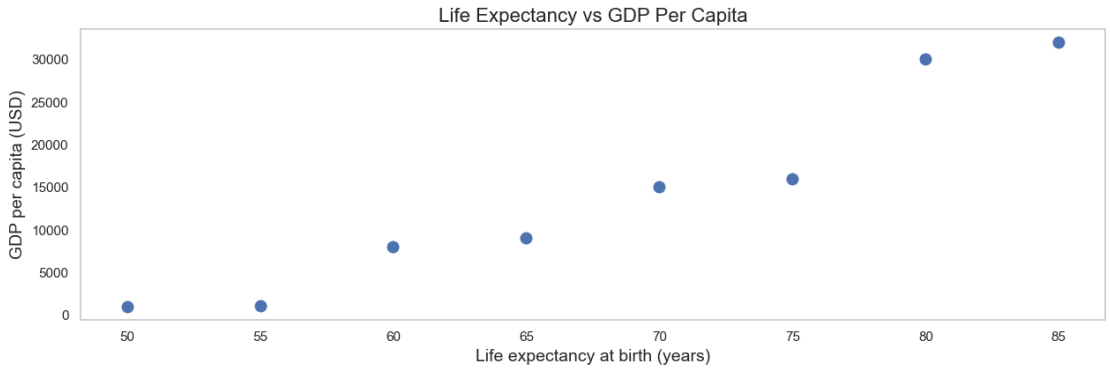


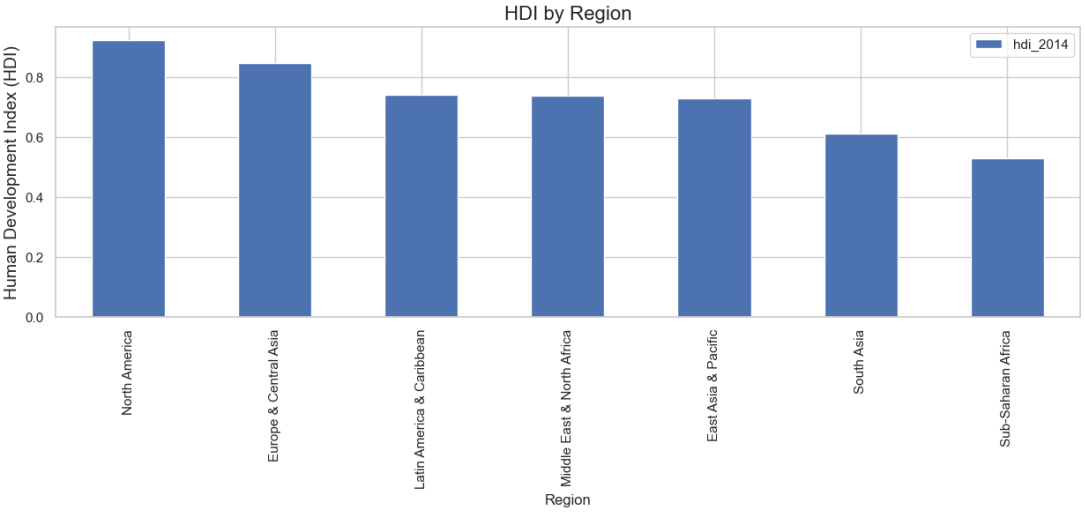
In [81]:
```
import matplotlib.pyplot as plt
import seaborn as sns
from matplotlib.gridspec import GridSpec

# Creating the figure and specifying the grid layout
fig = plt.figure(figsize=(15, 30))
gs = GridSpec(6, 1, figure=fig, height_ratios=[2, 2, 1.5, 1.5, 0.1, 1.5])  # Allocate

# Adjust the spacing more precisely
plt.subplots_adjust(hspace=0.4, top=0.95)

# GDP Stack Plot
ax1 = fig.add_subplot(gs[0, 0])
ax1.stackplot(
    gdp_pivot.index,
    [gdp_pivot[region]/1_000_000_000_000 for region in gdp_pivot.iloc[-1].sort_values(
    labels=[region for region in gdp_pivot.iloc[-1].sort_values(ascending=False).index
)
ax1.legend(loc="upper left")
ax1.set_title("GDP has Grown Exponentially", fontsize=16)
ax1.set_ylabel("GDP in Trillions", fontsize=14)

# Population Stack Plot
ax2 = fig.add_subplot(gs[1, 0])
```

```
ax2.stackplot(
    pop_pivot.index,
    [pop_pivot[region]/1000 for region in pop_pivot.iloc[-1].sort_values(ascending=Fal
    labels=[region for region in pop_pivot.iloc[-1].sort_values(ascending=False).index
)
ax2.legend(loc="upper left")
ax2.set_title("Population has Surged from 2.5 billion to 7 billion", fontsize=16)
ax2.set_ylabel("Population in Billions", fontsize=14)

# Scatter Plot for Life Expectancy vs GDP Per Capita
ax3 = fig.add_subplot(gs[2, 0])
scatter_plot = sns.scatterplot(
    data=wb_hdi_2014,
    x="Life expectancy at birth (years)",
    y="GDP per capita (USD)",
    ax=ax3,
    s=120  # Increased marker size
)
ax3.set_title("Life Expectancy vs GDP Per Capita", fontsize=16)
ax3.set_xlabel("Life expectancy at birth (years)", fontsize=14)
ax3.set_ylabel("GDP per capita (USD)", fontsize=14)
scatter_plot.grid(False)  # Remove grid

# Bar Chart for HDI by Region
ax4 = fig.add_subplot(gs[3, 0])
wb_hdi_by_region.plot.bar(ax=ax4)
ax4.set_title("HDI by Region", fontsize=16)
ax4.set_ylabel("Human Development Index (HDI)", fontsize=14)

# Add an empty space row for better spacing
ax5 = fig.add_subplot(gs[4, 0])
ax5.axis('off')

# Heatmap for Correlation
ax6 = fig.add_subplot(gs[5, 0])
corr = wb_hdi_2014[['Life expectancy at birth (years)', 'GDP per capita (USD)']].corr(
sns.heatmap(corr, annot=True, cmap='coolwarm', ax=ax6)
ax6.set_title('Correlation between Various Health and Economic Indicators', fontsize=1

# Show the full figure with all plots
plt.show()
```

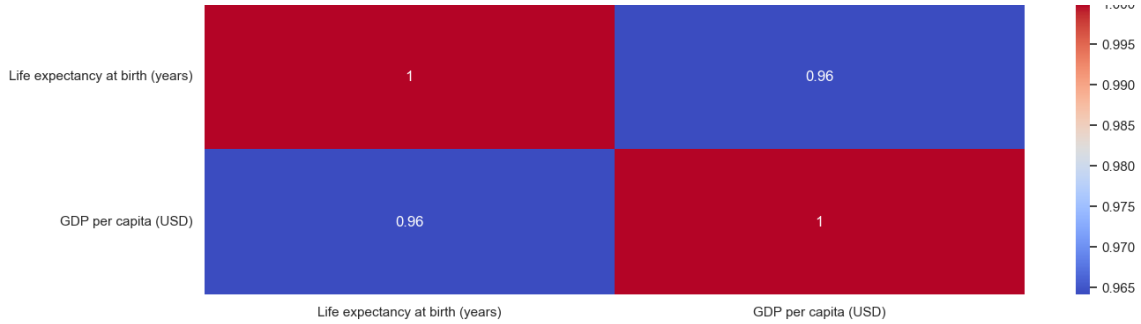

GDP has Grown Exponentially
East Asia & Pacific
Europe & Central Asia
North America
Latin America & Caribbean
South Asia
Middle East & North Africa
Sub-Saharan Africa
GDP in Trillions
0
10
20
30
40
50
60
70
80
1960
1970
1980
1990
2000
2010
2020


Population has Surged from 2.5 billion to 7 billion
East Asia & Pacific
South Asia
Sub-Saharan Africa
Europe & Central Asia
Latin America & Caribbean
North America
Middle East & North Africa
Population in Billions
0
1
2
3
4
5
6
7
1960
1970
1980
1990
2000
2010
2020


Life Expectancy vs GDP Per Capita
GDP per capita (USD)
0
5000
10000
15000
20000
25000
30000
50
55
60
65
70
75
80
85
Life expectancy at birth (years)


HDI by Region
hdi_2014
Human Development Index (HDI)
0.0
0.2
0.4
0.6
0.8
North America
Europe & Central Asia
Latin America & Caribbean
Middle East & North Africa
East Asia & Pacific
South Asia
Sub-Saharan Africa
Region


Correlation between Various Health and Economic Indicators


Life expectancy at birth (years)
GDP per capita (USD)
1
0.96
0.96
1
Life expectancy at birth (years)
GDP per capita (USD)
0.995
0.990
0.985
0.980
0.975
0.970
0.965

In [ ]: