## Mini-project

June 29, 2018

```
In [305]: import pandas as pd
          import numpy as np
          import random
          import matplotlib.pyplot as plt
          from matplotlib.offsetbox import (TextArea, DrawingArea, OffsetImage,
                                             AnnotationBbox)
In [306]: data = pd.read_csv('./world-development-indicators/Indicators.csv')
          data.shape
Out[306]: (5656458, 6)
  Lets explore the health expenditure of India compared to other developing countries like
China.
In [308]: health_filter='Health expenditure, total \(\% of GDP\)'
          india_code = 'IND'
          china_code = 'CHN'
          india_mask = data['CountryCode'].str.contains(india_code)
          china_mask = data['CountryCode'].str.contains(china_code)
          health_mask = data['IndicatorName'].str.contains(health_filter)
          india_health_exp = data[india_mask & health_mask]
          china_health_exp = data[china_mask & health_mask]
In [310]: india_health_exp.head(20)
Out [310]:
                  CountryName CountryCode
                                                                   IndicatorName
          2550719
                        India
                                       IND
                                           Health expenditure, total (% of GDP)
                                      IND Health expenditure, total (% of GDP)
          2685740
                        India
                                           Health expenditure, total (% of GDP)
          2821173
                        India
                                      IND
          2958149
                        India
                                      IND Health expenditure, total (% of GDP)
          3097445
                        India
                                           Health expenditure, total (% of GDP)
                                      IND
          3245806
                        India
                                      IND
                                           Health expenditure, total (% of GDP)
          3398139
                        India
                                      IND
                                           Health expenditure, total (% of GDP)
          3550764
                                      IND Health expenditure, total (% of GDP)
                        India
```

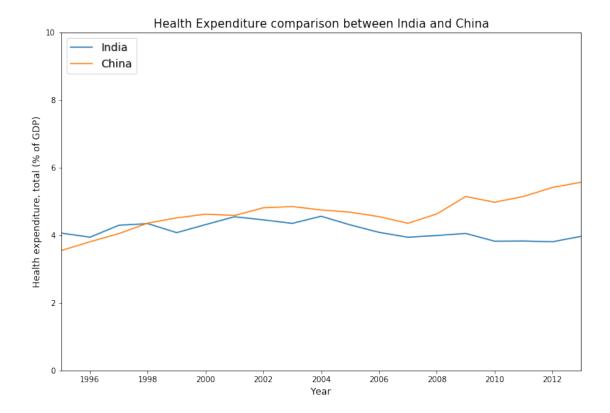
```
3705004
              India
                                  Health expenditure, total (% of GDP)
                             IND
                                  Health expenditure, total (% of GDP)
3860366
              India
                             IND
                                  Health expenditure, total (% of GDP)
              India
                             IND
4029881
                                  Health expenditure, total (% of GDP)
4208428
              India
                             IND
                                  Health expenditure, total (% of GDP)
4387273
              India
                             IND
                             IND
                                  Health expenditure, total (% of GDP)
4567586
              India
4746720
              India
                             IND
                                  Health expenditure, total (% of GDP)
                                  Health expenditure, total (% of GDP)
4929535
              India
                             IND
                                  Health expenditure, total (% of GDP)
5111336
              India
                             IND
                                  Health expenditure, total (% of GDP)
5287013
              India
                             IND
                                  Health expenditure, total (% of GDP)
5453181
              India
                             IND
          {\tt IndicatorCode}
                         Year
                                   Value
2550719
         SH.XPD.TOTL.ZS
                          1995
                                4.064824
2685740
         SH.XPD.TOTL.ZS
                          1996
                                3.942175
         SH.XPD.TOTL.ZS
2821173
                          1997
                                4.294396
2958149
         SH.XPD.TOTL.ZS
                          1998
                                4.343979
                          1999
3097445
         SH.XPD.TOTL.ZS
                                4.075596
         SH.XPD.TOTL.ZS
                          2000
                                4.314108
3245806
3398139
         SH.XPD.TOTL.ZS
                          2001
                                4.546835
3550764
         SH.XPD.TOTL.ZS
                          2002
                                4.454388
         SH.XPD.TOTL.ZS
3705004
                          2003
                                4.350801
3860366
         SH.XPD.TOTL.ZS
                          2004
                               4.561793
         SH.XPD.TOTL.ZS
                          2005
4029881
                                4.307127
4208428
         SH.XPD.TOTL.ZS
                          2006
                                4.085378
4387273
         SH.XPD.TOTL.ZS
                          2007
                                3.940753
4567586
         SH.XPD.TOTL.ZS
                          2008
                                3.991911
4746720
         SH.XPD.TOTL.ZS
                          2009
                                4.053157
         SH.XPD.TOTL.ZS
4929535
                          2010
                                3.824004
5111336
         SH.XPD.TOTL.ZS
                          2011
                                3.829788
5287013
         SH.XPD.TOTL.ZS
                          2012
                                3.806117
5453181
         SH.XPD.TOTL.ZS
                          2013 3.967993
```

In [311]: china\_health\_exp.head(20)

Out[311]:	CountryName	CountryCode		IndicatorName \		
2523687	China	CHN	Health expenditure,	total (	% of	GDP)
2658660	China	CHN	Health expenditure,	total (	% of	GDP)
2794266	China	CHN	Health expenditure,	total (	% of	GDP)
2930960	China	CHN	Health expenditure,	total (	% of	GDP)
3069205	China	CHN	Health expenditure,	total (	% of	GDP)
3215461	China	CHN	Health expenditure,	total (	% of	GDP)
3368744	China	CHN	Health expenditure,	total (	% of	GDP)
3520209	China	CHN	Health expenditure,	total (	% of	GDP)
3675129	China	CHN	Health expenditure,	total (	% of	GDP)
3829429	China	CHN	Health expenditure,	total (	% of	GDP)
3994434	China	CHN	Health expenditure,	total (	% of	GDP)
4174223	China	CHN	Health expenditure,	total (	% of	GDP)

```
Health expenditure, total (% of GDP)
4352416
              China
                            CHN
                            CHN Health expenditure, total (% of GDP)
4532834
              China
                                Health expenditure, total (% of GDP)
4712383
              China
                            CHN
              China
                            CHN
                                Health expenditure, total (% of GDP)
4893608
                                Health expenditure, total (% of GDP)
5077193
              China
                            CHN
                                Health expenditure, total (% of GDP)
              China
                            CHN
5252940
                                Health expenditure, total (% of GDP)
5422999
              China
                            CHN
         IndicatorCode Year
                                  Value
2523687
         SH.XPD.TOTL.ZS
                        1995
                               3.544976
         SH.XPD.TOTL.ZS
                         1996
                               3.806619
2658660
         SH.XPD.TOTL.ZS
2794266
                         1997
                               4.047860
2930960
         SH.XPD.TOTL.ZS
                         1998
                              4.358571
3069205
         SH.XPD.TOTL.ZS
                         1999 4.513426
3215461
         SH.XPD.TOTL.ZS
                         2000 4.622937
3368744
        SH.XPD.TOTL.ZS
                         2001 4.583395
3520209
        SH.XPD.TOTL.ZS
                         2002 4.811688
3675129
        SH.XPD.TOTL.ZS
                         2003 4.847566
3829429
         SH.XPD.TOTL.ZS
                        2004 4.747541
3994434
        SH.XPD.TOTL.ZS
                         2005 4.682612
4174223
         SH.XPD.TOTL.ZS
                         2006 4.550478
4352416
        SH.XPD.TOTL.ZS
                         2007 4.354222
4532834
        SH.XPD.TOTL.ZS
                         2008 4.628598
4712383
        SH.XPD.TOTL.ZS
                         2009 5.145722
4893608
        SH.XPD.TOTL.ZS
                         2010 4.976277
5077193
        SH.XPD.TOTL.ZS
                         2011
                               5.145995
        SH.XPD.TOTL.ZS
                         2012
5252940
                              5.413016
5422999
        SH.XPD.TOTL.ZS
                        2013 5.567502
```

Lets plot these dataframe together on line plot to visualize it better



In 1995, India's health expenditure (% of GDP) was 4.06% compared to China's 3.5% but India's health expenditure have rather decreased over the years with 3.9% of GDP in 2013 compared to China's 5.56%.

Lets explore the Health expenditure per capita and it's correlation with the Life expectancy of polulation of India.

```
In [313]: health_exp_filter = 'Health expenditure per capita \((current US\$\))'
          country_filter = 'IND'
          mask1 = data['CountryCode'].str.contains(country_filter)
          mask2 = data['IndicatorName'].str.contains(health_exp_filter)
          health_exp_data = data[mask1 & mask2]
In [314]: health_exp_data.head()
Out[314]:
                  CountryName CountryCode
                                                                           IndicatorName
          2550713
                        India
                                           Health expenditure per capita (current US$)
                                       IND
          2685734
                        India
                                       IND
                                            Health expenditure per capita (current US$)
                                           Health expenditure per capita (current US$)
                        India
          2821167
                                       IND
                                            Health expenditure per capita (current US$)
          2958143
                        India
                                       IND
          3097439
                        India
                                       IND
                                            Health expenditure per capita (current US$)
                  IndicatorCode Year
                                            Value
```

```
1995 16.088374
          2550713
                   SH.XPD.PCAP
         2685734
                   SH.XPD.PCAP
                                1996 16.226104
         2821167
                   SH.XPD.PCAP
                                1997
                                      18.774171
         2958143
                   SH.XPD.PCAP
                                1998 18.840850
         3097439
                   SH.XPD.PCAP
                                1999 18.683437
In [315]: life_exp_filter = 'Life expectancy at birth, total \((years\)'
         mask3 = data['IndicatorName'].str.contains(life_exp_filter)
         life_exp_data = data[mask1 & mask3]
In [316]: life_exp_data.head()
Out [316]:
                CountryName CountryCode
                                                                   IndicatorName \
          11684
                      India
                                    IND Life expectancy at birth, total (years)
                                         Life expectancy at birth, total (years)
                      India
         36635
                                    IND
                                    IND Life expectancy at birth, total (years)
         64177
                      India
                                    IND Life expectancy at birth, total (years)
         92622
                      India
                      India
                                    IND Life expectancy at birth, total (years)
          121419
                  IndicatorCode Year
                                           Value
         11684
                 SP.DYN.LEOO.IN 1960 41.171951
                 SP.DYN.LEOO.IN 1961 41.790488
         36635
         64177
                 SP.DYN.LEOO.IN 1962 42.417415
         92622
                 SP.DYN.LEOO.IN 1963 43.052732
          121419 SP.DYN.LEOO.IN 1964 43.698415
```

Before calculating the correlation between the life expectancy and health expenditure we need to slice the data further since health expenditure data is only available between 1995 to 2013.

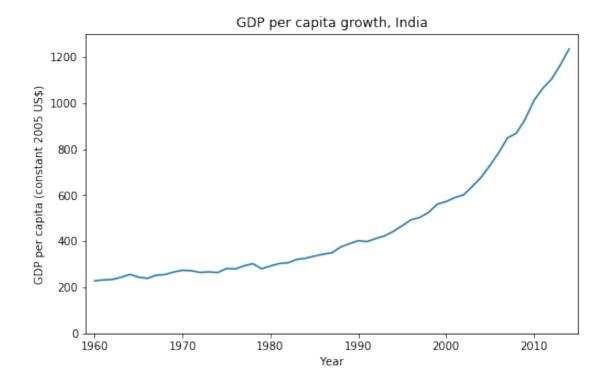
```
In [317]: filtered_life_exp_data = life_exp_data[life_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'].isin(health_exp_data['Year'
```

Finding: Correlation value of 0.95 between the Health expenditure and Life expectancy clearly suggests growth in health expenditure have lead to proportional growth in life expectancy in India.

Lets explore the GDP per capita data for India

```
In [321]: gdp_filter = 'GDP per capita \((constant 2005'))
          mask4 = data['IndicatorName'].str.contains(gdp_filter)
          gdp_data = data[mask1 & mask4]
In [323]: gdp_data.describe()
Out [323]:
                       Year
                                   Value
                   55.00000
                               55.000000
          count
                1987.00000
                              472.921846
          mean
          std
                   16.02082
                              272.215645
          min
                 1960.00000
                              228.304470
          25%
                 1973.50000
                              273.019022
          50%
                 1987.00000
                              349.639894
          75%
                 2000.50000
                              580.708296
                 2014.00000 1233.949344
          max
```

Exploring the changes in GDP per capita of India by plotting it on line plot



We can see there has been a substantial growth in GDP since 1990.

In [328]: country\_data.tail()

Lets see if we can find a correlation between the Health expenditure, Life expectancy and GDP per capita by plotting it on the scatterplot.

First, we will merge the three dataframes based on years into new dataframe 'country\_data'

```
In [325]: country_data = pd.merge(gdp_data, health_exp_data[['IndicatorName','Year','Value']],
In [326]: country_data.head()
Out[326]:
            CountryName CountryCode
                                                           IndicatorName_x
                                                                              IndicatorCode
          0
                   India
                                       GDP per capita (constant 2005 US$)
                                                                             NY.GDP.PCAP.KD
                                 IND
          1
                   India
                                       GDP per capita (constant 2005 US$)
                                 IND
                                                                             NY.GDP.PCAP.KD
          2
                   India
                                 IND
                                       GDP per capita (constant 2005 US$)
                                                                             NY.GDP.PCAP.KD
                                       GDP per capita (constant 2005 US$)
          3
                   India
                                 IND
                                                                             NY.GDP.PCAP.KD
          4
                   India
                                 IND
                                       GDP per capita (constant 2005 US$)
                                                                             NY.GDP.PCAP.KD
                                                 Value_y
             Year
                       Value_x IndicatorName_y
                   228.304470
          0
             1960
                                            NaN
                                                     NaN
          1
             1961
                    232.142053
                                            NaN
                                                     NaN
             1962
                    234.166685
                                            NaN
                                                     NaN
          3
             1963
                    243.176418
                                            NaN
                                                     NaN
             1964
                    255.963668
                                            NaN
                                                     NaN
```

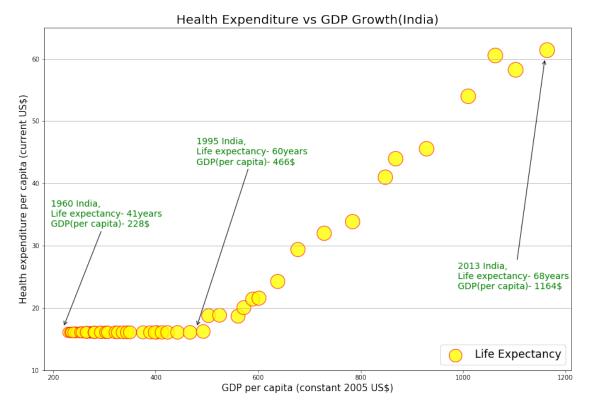
```
Out [328]:
             CountryName CountryCode
                                                          IndicatorName_x \
          50
                   India
                                      GDP per capita (constant 2005 US$)
                                 IND
          51
                   India
                                 IND
                                      GDP per capita (constant 2005 US$)
          52
                   India
                                      GDP per capita (constant 2005 US$)
                                 IND
                                      GDP per capita (constant 2005 US$)
          53
                   India
                                 IND
                                 IND
                                      GDP per capita (constant 2005 US$)
          54
                   India
               IndicatorCode Year
                                        Value_x \
          50 NY.GDP.PCAP.KD
                              2010
                                    1010.309221
          51 NY.GDP.PCAP.KD
                              2011
                                    1063.159868
          52 NY.GDP.PCAP.KD
                                    1102.910323
                              2012
          53 NY.GDP.PCAP.KD
                              2013
                                    1164.342834
          54 NY.GDP.PCAP.KD
                              2014
                                    1233.949344
                                          IndicatorName_y
                                                              Value_y
          50 Health expenditure per capita (current US$)
                                                            53.995063
          51 Health expenditure per capita (current US$)
                                                            60.544216
          52 Health expenditure per capita (current US$)
                                                            58.246599
          53 Health expenditure per capita (current US$)
                                                            61.408213
          54
                                                       NaN
                                                                  NaN
In [329]: country_data.isnull().any()
Out [329]: CountryName
                             False
          CountryCode
                             False
          IndicatorName x
                             False
          IndicatorCode
                             False
          Year
                             False
          Value_x
                             False
          IndicatorName_y
                              True
          Value_y
                              True
          dtype: bool
```

We can see there are NaN values for IndicatorName\_y and Value\_y since the health expenditure per capita is not available for years 1960 to 1994 and 2014. We can fill the NaN value with last known values using fillna method assuming health expenditure (1960 to 1994) less than equal to health expenditure (1995).

```
False
          Value_x
          IndicatorName_y
                             False
                             False
          Value_y
          dtype: bool
In [333]: country_data.tail()
             CountryName CountryCode
Out [333]:
                                                         IndicatorName_x \
                                      GDP per capita (constant 2005 US$)
          50
                   India
                                 IND GDP per capita (constant 2005 US$)
          51
                   India
                   India
                                      GDP per capita (constant 2005 US$)
          52
                                 IND
          53
                   India
                                 IND
                                      GDP per capita (constant 2005 US$)
                                 IND GDP per capita (constant 2005 US$)
          54
                   India
               IndicatorCode Year
                                        Value_x \
          50 NY.GDP.PCAP.KD
                              2010
                                    1010.309221
          51 NY.GDP.PCAP.KD
                                    1063.159868
                              2011
          52 NY.GDP.PCAP.KD
                              2012
                                    1102.910323
          53 NY.GDP.PCAP.KD
                              2013
                                    1164.342834
          54 NY.GDP.PCAP.KD
                              2014 1233.949344
                                          IndicatorName_y
                                                             Value_y
          50 Health expenditure per capita (current US$)
                                                            53.995063
          51 Health expenditure per capita (current US$)
                                                            60.544216
          52 Health expenditure per capita (current US$)
                                                            58.246599
          53 Health expenditure per capita (current US$)
                                                            61.408213
          54 Health expenditure per capita (current US$)
                                                            61.408213
In [334]: country_data = pd.merge(country_data, life_exp_data[['IndicatorName','Year','Value']]
In [335]: country_data.isnull().any()
Out[335]: CountryName
                             False
          CountryCode
                             False
          IndicatorName_x
                             False
          IndicatorCode
                             False
          Year
                             False
          Value_x
                             False
                             False
          IndicatorName_y
          Value_y
                             False
          IndicatorName
                             False
          Value
                             False
          dtype: bool
In [336]: country_data.tail()
             CountryName CountryCode
Out [336]:
                                                         IndicatorName_x \
          49
                   India
                                      GDP per capita (constant 2005 US$)
                                 IND
          50
                                 IND
                                      GDP per capita (constant 2005 US$)
                   India
```

```
51
                  India
                                IND GDP per capita (constant 2005 US$)
         52
                                IND GDP per capita (constant 2005 US$)
                  India
         53
                   India
                                IND GDP per capita (constant 2005 US$)
              IndicatorCode Year
                                       Value x \
         49 NY.GDP.PCAP.KD
                             2009
                                    928.977520
         50 NY.GDP.PCAP.KD
                             2010 1010.309221
         51 NY.GDP.PCAP.KD
                             2011 1063.159868
         52 NY.GDP.PCAP.KD
                             2012 1102.910323
         53 NY.GDP.PCAP.KD 2013 1164.342834
                                         IndicatorName_y
                                                            Value_y \
         49 Health expenditure per capita (current US$)
                                                          45.575644
         50 Health expenditure per capita (current US$)
                                                          53.995063
         51 Health expenditure per capita (current US$)
                                                          60.544216
         52 Health expenditure per capita (current US$)
                                                          58.246599
         53 Health expenditure per capita (current US$)
                                                          61.408213
                                       IndicatorName
                                                          Value
         49 Life expectancy at birth, total (years) 66.102634
         50 Life expectancy at birth, total (years)
                                                      66.506146
         51 Life expectancy at birth, total (years) 66.904171
         52 Life expectancy at birth, total (years) 67.289878
         53 Life expectancy at birth, total (years) 67.660415
In [337]: fig, axis = plt.subplots(figsize=(15,10))
          # Grid lines, Xticks, Xlabel, Ylabel
         axis.yaxis.grid(True)
         axis.set_title('Health Expenditure vs GDP Growth(India)',fontsize=20)
         axis.set_xlabel(country_data['IndicatorName_x'].iloc[0],fontsize=15)
         axis.set_ylabel(country_data['IndicatorName_y'].iloc[0],fontsize=15)
         axis.set_ylim(10,65)
         axis.annotate("1960 India, \nLife expectancy- 41years \nGDP(per capita)- 228$",
                          xy=(220, 17),
                          xycoords='data',
                          xytext=(195, 33),
                          textcoords='data',
                          arrowprops=dict(arrowstyle="->",
                          connectionstyle="arc3"),fontsize=14, color='green'
                     )
         axis.annotate("1995 India, \nLife expectancy- 60years \nGDP(per capita)- 466$",
                          xy=(480, 17),
                          xycoords='data',
                          xytext=(480, 43),
                          textcoords='data',
```

```
arrowprops=dict(arrowstyle="->",
                connectionstyle="arc3"),fontsize=14, color='green'
            )
axis.annotate("2013 India, \nLife expectancy- 68years \nGDP(per capita)- 1164$",
                xy=(1160, 60),
                xycoords='data',
                xytext=(990, 23),
                textcoords='data',
                arrowprops=dict(arrowstyle="->",
                connectionstyle="arc3"),fontsize=14, color='green'
            )
X = country_data['Value_x']
Y = country_data['Value_y']
S = country_data['Value']**1.5
axis.scatter(X, Y, s=S,label='Life Expectancy',
               alpha=0.8, color='yellow', edgecolors='Red')
axis.legend(loc='lower right',prop={'weight':'roman','size':'xx-large'})
plt.show()
```



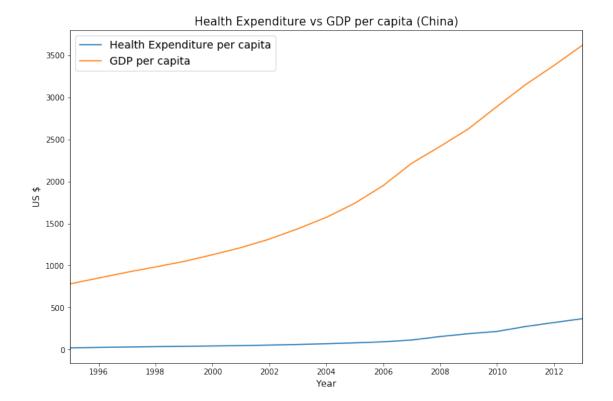
Scatterplot and correlation value of 0.93 clearly shows there is strong relationship between the Health expenditure and GDP growth. Now lets verify these findings by analysing the data for similar developing economies.

```
similar developing economies.
In [339]: china_health_exp = data[mask2 & china_mask]
          china_gdp_data = data[mask4 & china_mask]
In [340]: china_health_exp.head()
Out [340]:
                  CountryName CountryCode
                                                                           IndicatorName
          2523681
                        China
                                            Health expenditure per capita (current US$)
                                       CHN
          2658654
                        China
                                            Health expenditure per capita (current US$)
                                       CHN
          2794260
                        China
                                       CHN Health expenditure per capita (current US$)
                        China
                                       CHN
                                           Health expenditure per capita (current US$)
          2930954
                                            Health expenditure per capita (current US$)
          3069199
                        China
                                       CHN
                  IndicatorCode
                                 Year
                                            Value
                                        20.742711
          2523681
                    SH.XPD.PCAP
                                  1995
          2658654
                    SH.XPD.PCAP
                                  1996
                                        25.975256
                    SH.XPD.PCAP
          2794260
                                  1997
                                        30.511171
          2930954
                    SH.XPD.PCAP
                                  1998
                                        34.926572
          3069199
                    SH.XPD.PCAP
                                  1999
                                        38.195679
In [341]: china_gdp_data.head()
Out [341]:
                 CountryName CountryCode
                                                                 IndicatorName \
          7187
                       China
                                      CHN
                                           GDP per capita (constant 2005 US$)
                                           GDP per capita (constant 2005 US$)
          31484
                       China
                                      CHN
                       China
                                           GDP per capita (constant 2005 US$)
          58684
                                      CHN
                                           GDP per capita (constant 2005 US$)
          87069
                       China
                                      CHN
                                           GDP per capita (constant 2005 US$)
          115795
                       China
                                      CHN
                   IndicatorCode Year
                                              Value
                  NY.GDP.PCAP.KD 1960
                                         121.191585
          7187
          31484
                  NY.GDP.PCAP.KD 1961
                                          89.005585
                  NY.GDP.PCAP.KD
                                  1962
                                          83.334735
          58684
          87069
                  NY.GDP.PCAP.KD
                                  1963
                                          89.605409
          115795 NY.GDP.PCAP.KD
                                  1964
                                        103.571525
```

In [342]: china\_data = pd.merge(china\_health\_exp, china\_gdp\_data[['IndicatorName','Year','Value

In [343]: china\_data.head()

```
Out [343]:
            CountryName CountryCode
                                                                  IndicatorName_x \
          0
                  China
                                CHN Health expenditure per capita (current US$)
                                     Health expenditure per capita (current US$)
          1
                  China
                                CHN
          2
                  China
                                CHN
                                     Health expenditure per capita (current US$)
                                CHN Health expenditure per capita (current US$)
          3
                  China
          4
                                CHN
                                     Health expenditure per capita (current US$)
                  China
            IndicatorCode Year
                                   Value_x
                                                                IndicatorName_y \
              SH.XPD.PCAP
                           1995
                                 20.742711 GDP per capita (constant 2005 US$)
                                            GDP per capita (constant 2005 US$)
          1
              SH.XPD.PCAP
                           1996
                                 25.975256
          2
              SH.XPD.PCAP
                                            GDP per capita (constant 2005 US$)
                           1997
                                 30.511171
          3
              SH.XPD.PCAP
                                 34.926572 GDP per capita (constant 2005 US$)
                           1998
                                 38.195679 GDP per capita (constant 2005 US$)
              SH.XPD.PCAP
                           1999
                 Value_y
          0
              782.093690
          1
              850.750357
          2
              919.786248
          3
              982.548142
          4 1048.284370
In [344]: fig,ax = plt.subplots(figsize=(12,8))
          line1, = ax.plot(china_data['Year'].values, china_data['Value_x'].values, label="Hea
          line2, = ax.plot(china_data['Year'].values, china_data['Value_y'].values, label="GDP"
          ax.set_xlabel('Year', fontsize='12')
          ax.set_ylabel('US $', fontsize='12')
          ax.set_title('Health Expenditure vs GDP per capita (China)',fontsize='15')
          ax.legend(handles=[line1, line2], loc='upper left', fontsize='14')
          #ax.set_ylim([0,10])
          ax.set_xlim([china_data['Year'].min(),china_data['Year'].max()])
          plt.show()
```



With correlation value of 0.97 for China it confirms our previous findings that increase in Health expenditure inturn leads to economic growth.

Now that we know the importance of Health care, let explore it further to find the public contribution to health expenditure and what effects it will have on the Indian population?

```
In [347]: mask5 = data['IndicatorName'].str.contains('Health expenditure, public \(\\\\\\\\\\\) of total
          mask6 = data['IndicatorName'].str.contains('Poverty headcount ratio at national poverty)
          public_health_exp_data = data[mask1 & mask5]
          poverty_data = data[mask1 & mask6]
In [348]: public_health_exp_data.tail()
Out [348]:
                   CountryName CountryCode
          4746719
                         India
                                        IND
          4929534
                         India
                                        IND
          5111335
                         India
                                        IND
```

IND

IND

5287012

5453180

India

India

```
IndicatorName IndicatorCode \
          4746719
                   Health expenditure, public (% of total health ...
                                                                         SH.XPD.PUBL
          4929534
                   Health expenditure, public (% of total health ...
                                                                         SH.XPD.PUBL
                   Health expenditure, public (% of total health ...
          5111335
                                                                         SH.XPD.PUBL
                   Health expenditure, public (% of total health ...
          5287012
                                                                         SH.XPD.PUBL
                   Health expenditure, public (% of total health ...
          5453180
                                                                         SH.XPD.PUBL
                   Year
                             Value
          4746719
                   2009
                         30.034031
                   2010
                         30.195803
          4929534
                         29.826705
          5111335
                   2011
          5287012
                   2012
                         30.455623
                   2013
                         32.215077
          5453180
In [349]: public_health_exp_data.describe()
Out [349]:
                        Year
                                   Value
                              19.000000
                   19.000000
          count
                 2004.000000
                              26.878575
          mean
                               2.809142
                    5.627314
          std
                              21.912007
          min
                 1995.000000
          25%
                 1999.500000
                              24.691476
          50%
                 2004.000000
                              26.682532
          75%
                 2008.500000
                              29.390584
                 2013.000000
                              32.215077
          max
In [350]: poverty_data.tail()
                  CountryName CountryCode \
Out [350]:
          2296334
                        India
                                       IND
          3860676
                        India
                                       IND
                        India
                                       IND
          4747049
          5111668
                        India
                                       IND
                                                        IndicatorName IndicatorCode
                   Poverty headcount ratio at national poverty li...
          2296334
                                                                         SI.POV.NAHC
          3860676
                   Poverty headcount ratio at national poverty li...
                                                                         SI.POV.NAHC
                   Poverty headcount ratio at national poverty li...
          4747049
                                                                         SI.POV.NAHC
                   Poverty headcount ratio at national poverty li...
                                                                         SI.POV.NAHC
          5111668
                        Value
                   Year
          2296334
                   1993
                          45.3
                          37.2
          3860676
                   2004
                   2009
                          29.8
          4747049
                          21.9
          5111668 2011
```