SHRI VAISHNAV VIDYAPEETH VISHWAVIDYALAYA, INDORE

Project Report

On

Life Expectancy Data

Submitted By

1. **Divyam Pithawa [19100BTCSAII05259]**
2. **Mradul Agrawal [19100BTCSAII05269]**
3. **Rudraksh Thorat [19100BTCSAII05280]**
4. **Sarthak Nahar [19100BTCSAII05283]**
5. **Shivam Sharma [19100BTCSAII05286]**

Under the Guidance of

Mr. Sachin Garg

SHRI VAISHNAV INSTITUTE OF INFORMATION TECHNOLOGY

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

JULY 2021-JANUARY 2021

INTRODUCTION

Although there have been lot of studies undertaken in the past on factors affecting life expectancy considering demographic variables, income composition and mortality rates. It was found that affect of immunization and human development index was not taken into account in the past. Also, some of the past research was done considering multiple linear regression based on data set of one year for all the countries. Hence, this gives motivation to resolve both the factors stated previously by formulating a regression model based on mixed effects model and multiple linear regression while considering data from a period of 2000 to 2015 for all the countries. Important immunization like Hepatitis B, Polio and Diphtheria will also be considered. In a nutshell, this study will focus on immunization factors, mortality factors, economic factors, social factors and other health related factors as well. Since the observations this dataset are based on different countries, it will be easier for a country to determine the predicting factor which is contributing to lower value of life expectancy. This will help in suggesting a country which area should be given importance in order to efficiently improve the life expectancy of its population.

Content

The project relies on accuracy of data. The Global Health Observatory (GHO) data repository under World Health Organization (WHO) keeps track of the health status as well as many other related factors for all countries The data-sets are made available to public for the purpose of health data analysis. The data-set related to life expectancy, health factors for 193 countries has been collected from the same WHO data repository website and its corresponding economic data was collected from United Nation website. Among all categories of health-related factors only those critical factors were chosen which are more representative. It has been observed that in the past 15 years , there has been a huge development in health sector resulting in improvement of human mortality rates especially in the developing nations in comparison to the past 30 years. Therefore, in this project we have considered data from year 2000-2015 for 193 countries for further analysis. The individual data files have been merged together into a single data-set. On initial visual inspection of the data showed some missing values. As the data-sets were from WHO, we found no evident errors. Missing data was handled in R software by using Missmap command. The result indicated that most

of the missing data was for population, Hepatitis B and GDP. The missing data were from less known countries like Vanuatu, Tonga, Togo, Cabo Verde etc. Finding all data for these countries was difficult and hence, it was decided that we exclude these countries from the final model data-set. The final merged file(final dataset) consists of 22 Columns and 2938 rows which meant 20 predicting variables. All predicting variables was then divided into several broad categories:Immunization related factors, Mortality factors, Economical factors and Social factors.

Acknowledgements

The data was collected from WHO and United Nations website with the help of Deeksha Russell and Duan Wang.

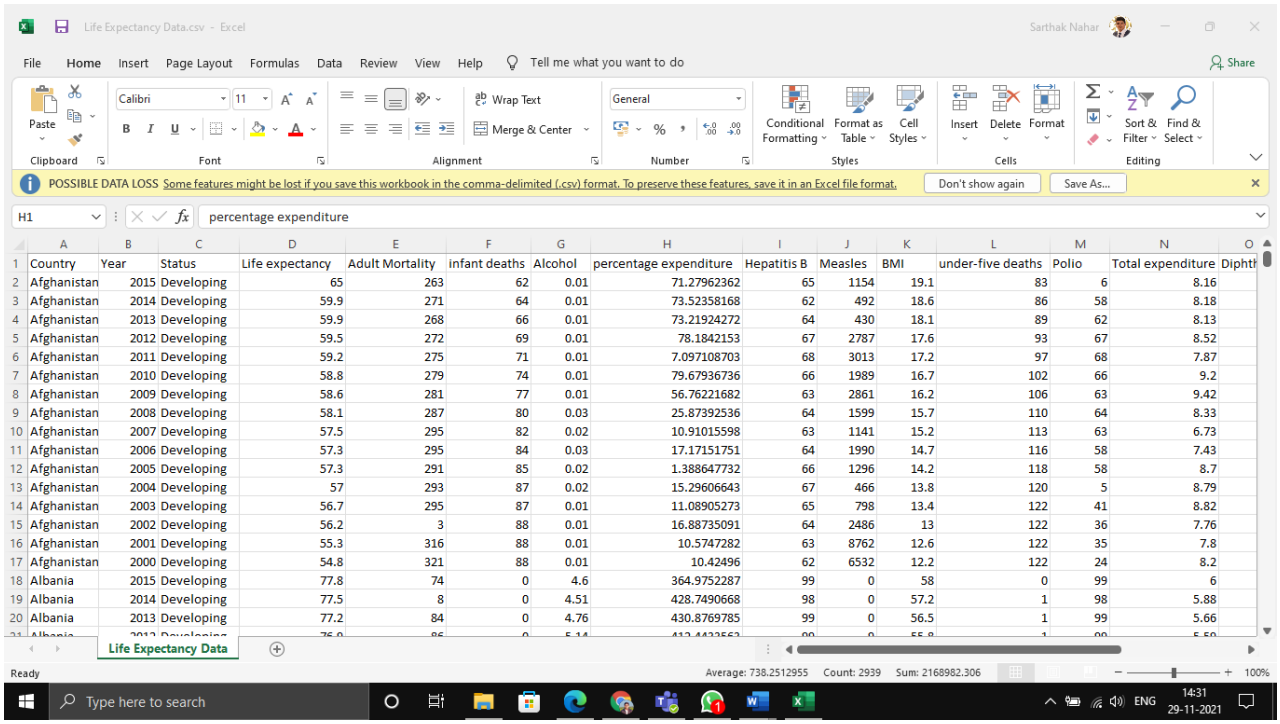
Public Dataset available at below website

https://www.kaggle.com/kumarajarshi/life-expectancy-who

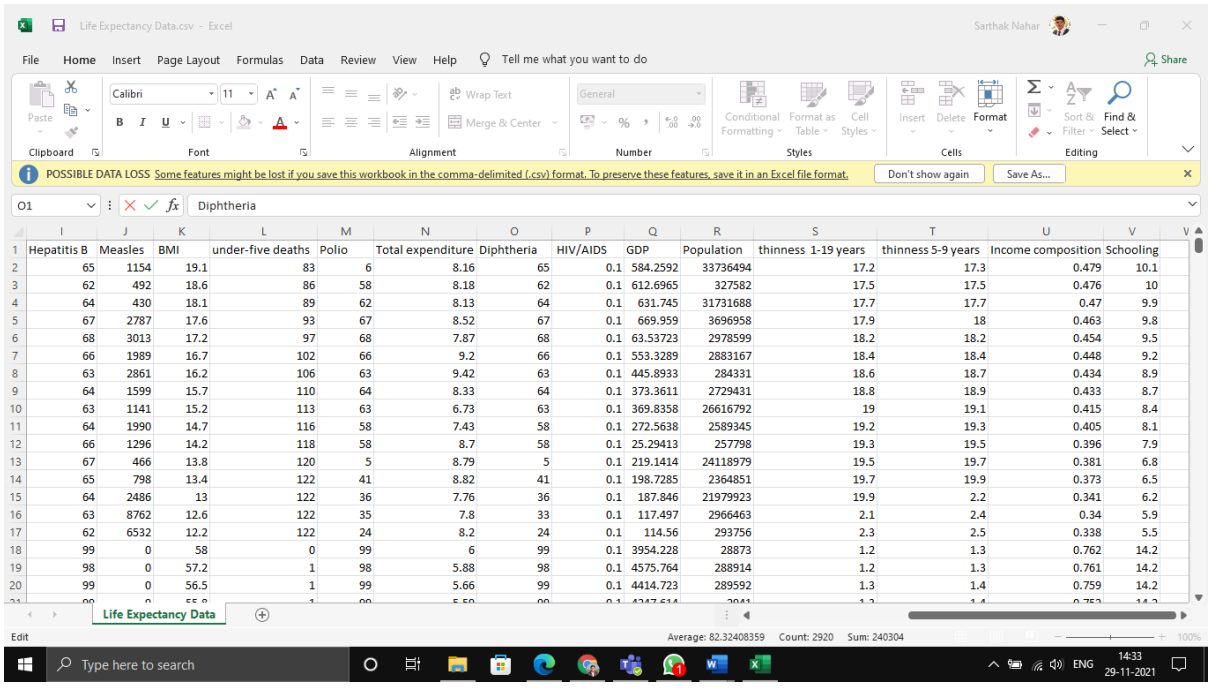
Tools Used

Jupyter Notebook.

DataSet (Screenshot) : -



Country	Year	Status	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria
Afghanistan	2015	Developing	65	263	62	0.01	71.27962362	65	1154	19.1	83	6	8.16	
Afghanistan	2014	Developing	59.9	271	64	0.01	73.52358168	62	492	18.6	86	58	8.18	
Afghanistan	2013	Developing	59.9	268	66	0.01	73.21924272	64	430	18.1	89	62	8.13	
Afghanistan	2012	Developing	59.5	272	69	0.01	78.1842153	67	2787	17.6	93	67	8.52	
Afghanistan	2011	Developing	59.2	275	71	0.01	7.097108703	68	3013	17.2	97	68	7.87	
Afghanistan	2010	Developing	58.8	279	74	0.01	79.67936736	66	1989	16.7	102	66	9.2	
Afghanistan	2009	Developing	58.6	281	77	0.01	56.76221682	63	2861	16.2	106	63	9.42	
Afghanistan	2008	Developing	58.1	287	80	0.03	25.87392536	64	1599	15.7	110	64	8.33	
Afghanistan	2007	Developing	57.5	295	82	0.02	10.91015598	63	1141	15.2	113	63	6.73	
Afghanistan	2006	Developing	57.3	295	84	0.03	17.17151751	64	1990	14.7	116	58	7.43	
Afghanistan	2005	Developing	57.3	291	85	0.02	1.388647732	66	1296	14.2	118	58	8.7	
Afghanistan	2004	Developing	57	293	87	0.02	15.29606643	67	466	13.8	120	5	8.79	
Afghanistan	2003	Developing	56.7	295	87	0.01	11.08905273	65	798	13.4	122	41	8.82	
Afghanistan	2002	Developing	56.2	3	88	0.01	16.88735091	64	2486	13	122	36	7.76	
Afghanistan	2001	Developing	55.3	316	88	0.01	10.5747282	63	8762	12.6	122	35	7.8	
Afghanistan	2000	Developing	54.8	321	88	0.01	10.42496	62	6532	12.2	122	24	8.2	
Albania	2015	Developing	77.8	74	0	4.6	364.9752287	99	0	58	0	99	6	
Albania	2014	Developing	77.5	8	0	4.51	428.7490668	98	0	57.2	1	98	5.88	
Albania	2013	Developing	77.2	84	0	4.76	430.8769785	99	0	56.5	1	99	5.66	



Hepatitis B	Measles	BMI	under-five deaths	Polio	Total expenditure	Diphtheria	HIV/AIDS	GDP	Population	thinness 1-19 years	thinness 5-9 years	Income composition	Schooling
65	1154	19.1	83	6	8.16	65	0.1	584.2592	33736494	17.2	17.3	0.479	10.1
62	492	18.6	86	58	8.18	62	0.1	612.6965	327582	17.5	17.5	0.476	10
64	430	18.1	89	62	8.13	64	0.1	631.745	31731688	17.7	17.7	0.47	9.9
67	2787	17.6	93	67	8.52	67	0.1	669.959	3696958	17.9	18	0.463	9.8
68	3013	17.2	97	68	7.87	68	0.1	63.53723	2978599	18.2	18.2	0.454	9.5
66	1989	16.7	102	66	9.2	66	0.1	553.3289	2883167	18.4	18.4	0.448	9.2
63	2861	16.2	106	63	9.42	63	0.1	445.8933	284331	18.6	18.7	0.434	8.9
64	1599	15.7	110	64	8.33	64	0.1	373.3611	2729431	18.8	18.9	0.433	8.7
63	1141	15.2	113	63	6.73	63	0.1	369.8358	26616792	19	19.1	0.415	8.4
64	1990	14.7	116	58	7.43	58	0.1	272.5638	2589345	19.2	19.3	0.405	8.1
66	1296	14.2	118	58	8.7	58	0.1	25.29413	257798	19.3	19.5	0.396	7.9
67	466	13.8	120	5	8.79	5	0.1	219.1414	24118979	19.5	19.7	0.381	6.8
65	798	13.4	122	41	8.82	41	0.1	198.7285	2364851	19.7	19.9	0.373	6.5
64	2486	13	122	36	7.76	36	0.1	187.846	21979923	19.9	2.2	0.341	6.2
63	8762	12.6	122	35	7.8	33	0.1	117.497	2966463	2.1	2.4	0.34	5.9
62	6532	12.2	122	24	8.2	24	0.1	114.56	293756	2.3	2.5	0.338	5.5
99	0	58	0	99	6	99	0.1	3954.228	28873	1.2	1.3	0.762	14.2
98	0	57.2	1	98	5.88	98	0.1	4575.764	288914	1.2	1.3	0.761	14.2
99	0	56.5	1	99	5.66	99	0.1	4414.723	289592	1.3	1.4	0.759	14.2

About Columns of Dataset: -

1. Country - Country
2. Year - Year
3. Status - Developed or Developing status
4. Life expectancy - Life Expectancy in age
5. Adult Mortality - Adult Mortality Rates of both sexes (probability of dying between 15 and 60 years per 1000 population)
6. Infant Deaths - Number of Infant Deaths per 1000 population
7. Alcohol - Alcohol, recorded per capita (15+) consumption (in litres of pure alcohol)
8. Percentage Expenditure - Expenditure on health as a percentage of Gross Domestic Product per capita(%)
9. Hepatitis B - Hepatitis B (HepB) immunization coverage among 1-year-olds (%)
10. Measles - Number of reported cases per 1000 population
11. BMI - Average Body Mass Index of entire population
12. Under-Five Deaths - Number of under-five deaths per 1000 population
13. Polio - Polio (Pol3) immunization coverage among 1-year-olds (%)
14. Total Expenditure - General government expenditure on health as a percentage of total government expenditure (%)
15. Diphtheria - Diphtheria tetanus toxoid and pertussis (DTP3) immunization coverage among 1-year-olds (%)
16. HIV/AIDS - Deaths per 1 000 live births HIV/AIDS (0-4 years)
17. GDP - Gross Domestic Product per capita (in USD)
18. Population - Population of the country
19. Thinness 1-19 Years - Prevalence of thinness among children and adolescents for Age 10 to 19 (%)
20. Thinness 5-9 Years - Prevalence of thinness among children for Age 5 to 9(%)
21. Income Composition of Resources - Human Development Index in terms of income composition of resources (index ranging from 0 to 1)
22. Schooling - Number of years of Schooling(years)

Problem Statement are: -

Que 1. Finding top 10 Countries with maximum number of Life Expectancy.

Que 2. Find the Correlation between GDP and Life Expectancy.

Que 3. Which Disease affect Life Expectancy the most.

Que 4. Which Country's Life Expectancy increased the most in 15 years.

Que 5. Top Five Country with Highest Adult Mortality.

Que 6. Draw a Graph b/w Income Composition and Population.

Que 7. In which Year the Total Expenditure is Highest in India.

Que 8. What are the predicting variables that actually affecting the Life Expectancy?

Que 9. Should a country having a lower life expectancy value (<65) increase its healthcare expenditure in order to improve its average lifespan?

Que 10. How does Infant and Adult mortality rates affect Life Expectancy?

Que 11. Does Life Expectancy have positive or negative correlation with Status of the Country, drinking alcohol, BMI, GDP etc.

Que 12. What is the impact of schooling on the lifespan of humans?

Que 13. Do densely populated countries tend to have lower life expectancy?

Que 14. Draw a Line Graph of Year wise GDP of India.

Que 15. Which Country infant die more often.

Que 16. In which Year the Under Five Death is Highest.

Que 17. Draw a Graph between Year and Schooling.

Que 18. Country which consumed most alcohol.

Que 19. Countries whose government shows awareness about their health condition (more expenditure on health issues).

Que 20. In which year India shows healthiest Body Mass Index (BMI).

Que 21. Top 5 countries whose infant mortality rate is lowest.

Que 22. Countries whose literacy rate is best.

Que 23. Countries with lowest literacy rate.

Que 24. In which Year the Total Expenditure is Lowest in India.

Que 25. In which year the Income Composition of Resources is Highest in India.

Que 26. How much the GDP has increased in the last 15 years in India.

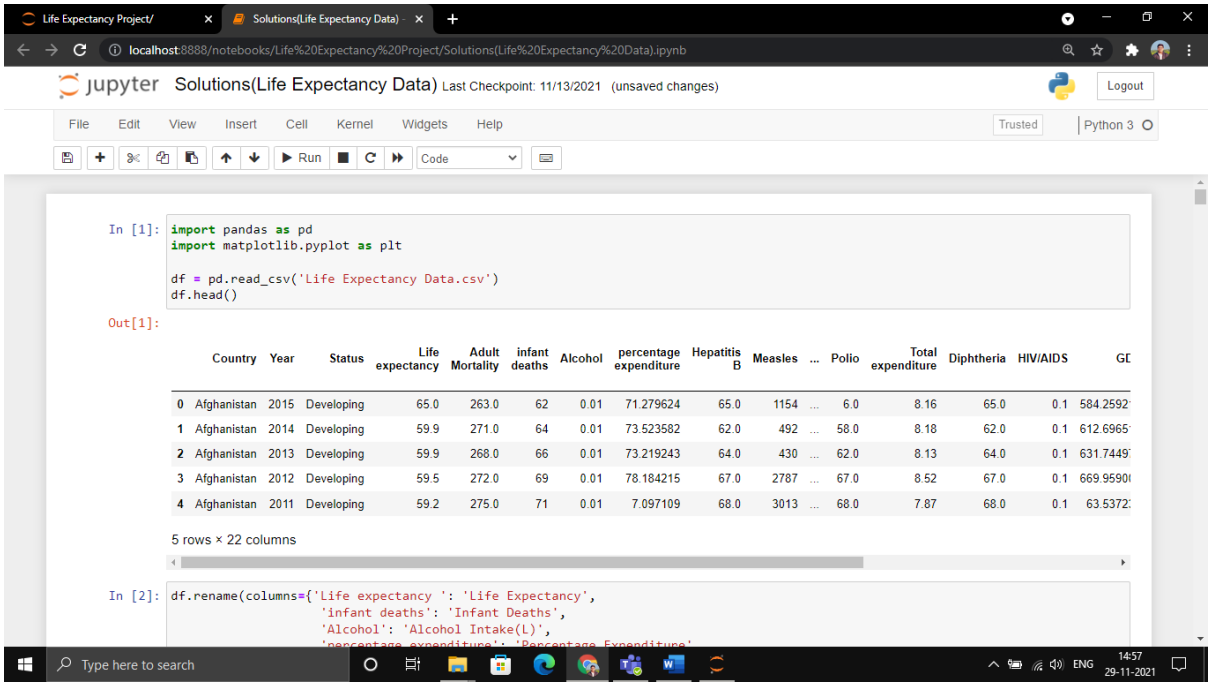
Que 27. How many countries are Developed till the year 2015.

Que 28. List of Countries which has Developed in the last 15 Years.

Que 29. Draw a Graph of Year wise Schooling in India.

Que 30. Draw a Graph of Life Expectancy of different Countries.

Step 1. Importing Data set to a Python Code



Life Expectancy Project/ | Solutions(Life Expectancy Data) - | localhost:8888/notebooks/Life%20Expectancy%20Project/Solutions(Life%20Expectancy%20Data).ipynb

jupyter Solutions(Life Expectancy Data) Last Checkpoint: 11/13/2021 (unsaved changes) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

In [1]:
```
import pandas as pd
import matplotlib.pyplot as plt

df = pd.read_csv('Life Expectancy Data.csv')
df.head()
```

Out[1]:

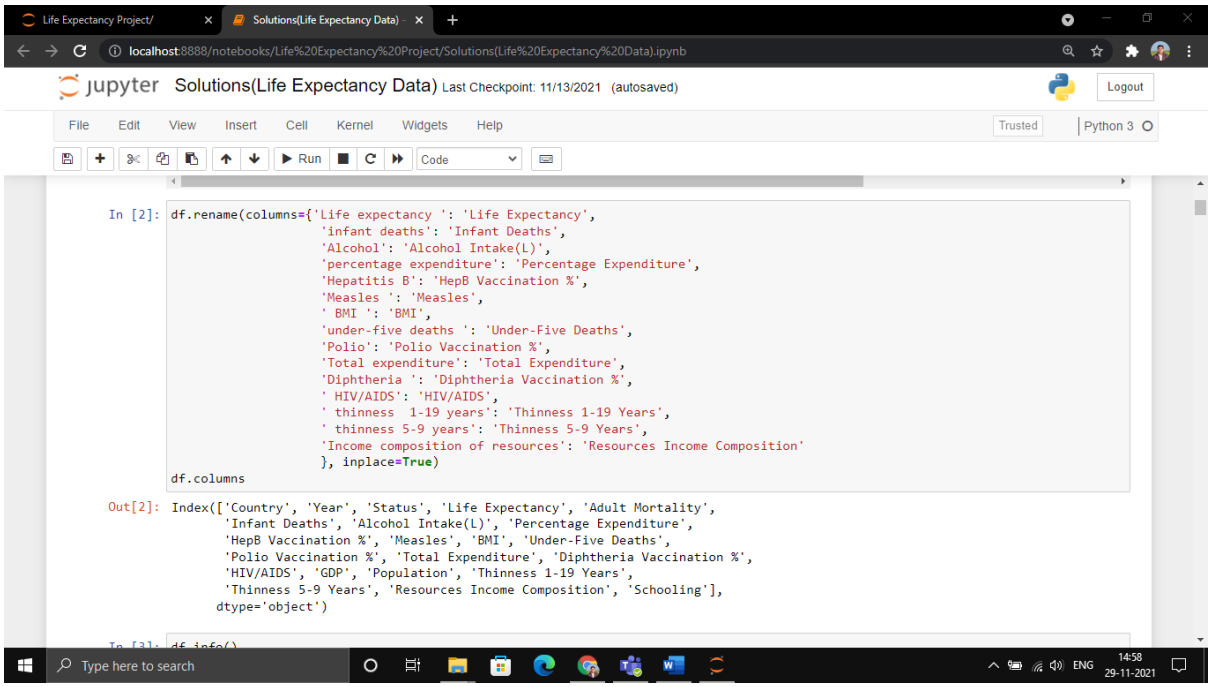
	Country	Year	Status	Life expectancy	Adult Mortality	infant deaths	Alcohol	percentage expenditure	Hepatitis B	Measles	...	Polio	Total expenditure	Diphtheria	HIV/AIDS	GD
0	Afghanistan	2015	Developing	65.0	263.0	62	0.01	71.279624	65.0	1154	...	6.0	8.16	65.0	0.1	584.2592
1	Afghanistan	2014	Developing	59.9	271.0	64	0.01	73.523582	62.0	492	...	58.0	8.18	62.0	0.1	612.6965
2	Afghanistan	2013	Developing	59.9	268.0	66	0.01	73.219243	64.0	430	...	62.0	8.13	64.0	0.1	631.7449
3	Afghanistan	2012	Developing	59.5	272.0	69	0.01	78.184215	67.0	2787	...	67.0	8.52	67.0	0.1	669.9590
4	Afghanistan	2011	Developing	59.2	275.0	71	0.01	7.097109	68.0	3013	...	68.0	7.87	68.0	0.1	63.5372

5 rows × 22 columns

In [2]:
```
df.rename(columns={'Life expectancy ': 'Life Expectancy',
                   'infant deaths': 'Infant Deaths',
                   'Alcohol': 'Alcohol Intake(L)',
                   'percentage expenditure': 'Percentage Expenditure',
```

Type here to search ENG 14:57 29-11-2021

Step 2. Renaming Columns of a Data Set



Life Expectancy Project/ | Solutions(Life Expectancy Data) - | localhost:8888/notebooks/Life%20Expectancy%20Project/Solutions(Life%20Expectancy%20Data).ipynb

jupyter Solutions(Life Expectancy Data) Last Checkpoint: 11/13/2021 (autosaved) Logout

File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

Run Code

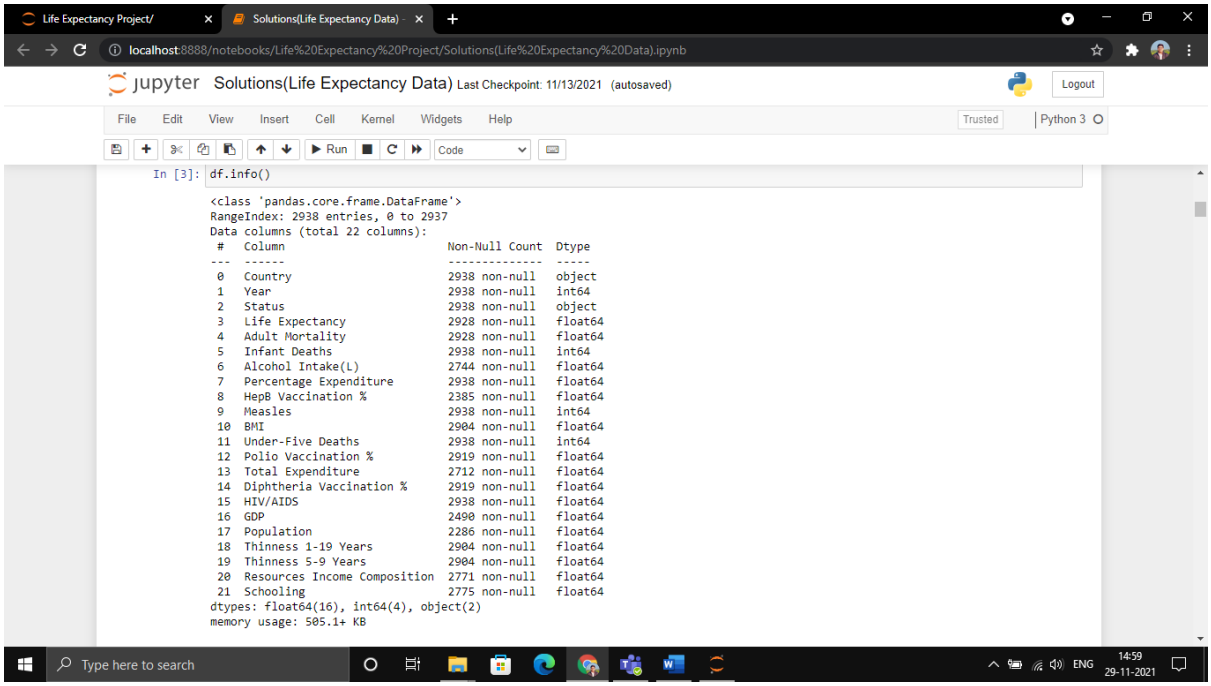
In [2]:
```
df.rename(columns={'Life expectancy ': 'Life Expectancy',
                   'infant deaths': 'Infant Deaths',
                   'Alcohol': 'Alcohol Intake(L)',
                   'percentage expenditure': 'Percentage Expenditure',
                   'Hepatitis B': 'HepB Vaccination %',
                   'Measles ': 'Measles',
                   ' BMI ': 'BMI',
                   'under-five deaths ': 'Under-Five Deaths',
                   'Polio': 'Polio Vaccination %',
                   'Total expenditure': 'Total Expenditure',
                   'Diphtheria ': 'Diphtheria Vaccination %',
                   ' HIV/AIDS': 'HIV/AIDS',
                   ' thinness  1-19 years': 'Thinness 1-19 Years',
                   ' thinness 5-9 years': 'Thinness 5-9 Years',
                   'Income composition of resources': 'Resources Income Composition'
                   }, inplace=True)
df.columns
```

Out[2]:
```
Index(['Country', 'Year', 'Status', 'Life Expectancy', 'Adult Mortality',
       'Infant Deaths', 'Alcohol Intake(L)', 'Percentage Expenditure',
       'HepB Vaccination %', 'Measles', 'BMI', 'Under-Five Deaths',
       'Polio Vaccination %', 'Total Expenditure', 'Diphtheria Vaccination %',
       'HIV/AIDS', 'GDP', 'Population', 'Thinness 1-19 Years',
       'Thinness 5-9 Years', 'Resources Income Composition', 'Schooling'],
      dtype='object')
```

In [3]: df.info()

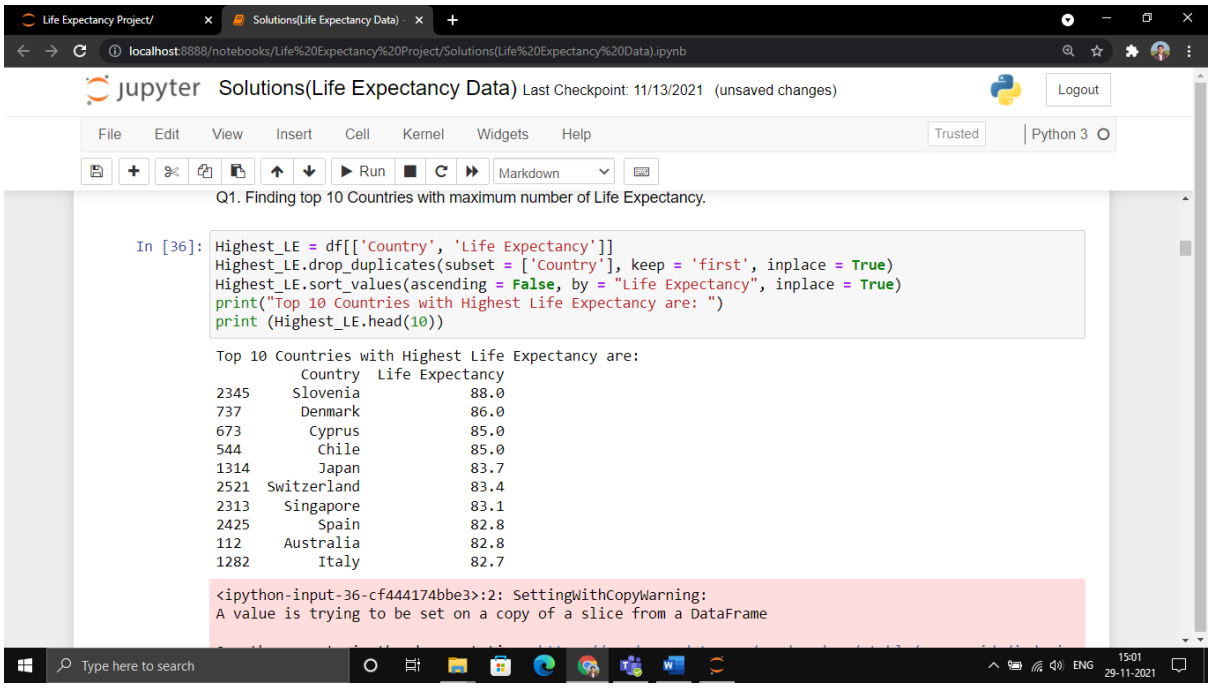
Type here to search ENG 14:58 29-11-2021

Step 3. Finding info about Data Set

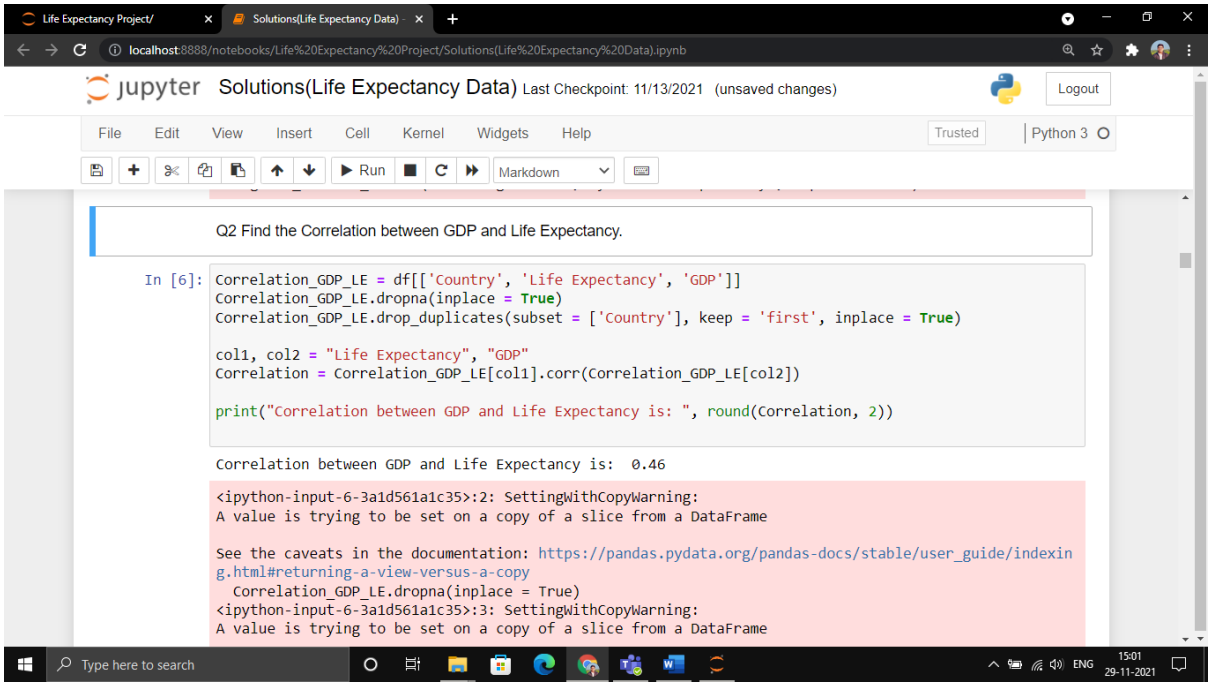


Solutions of Problem Statement: -

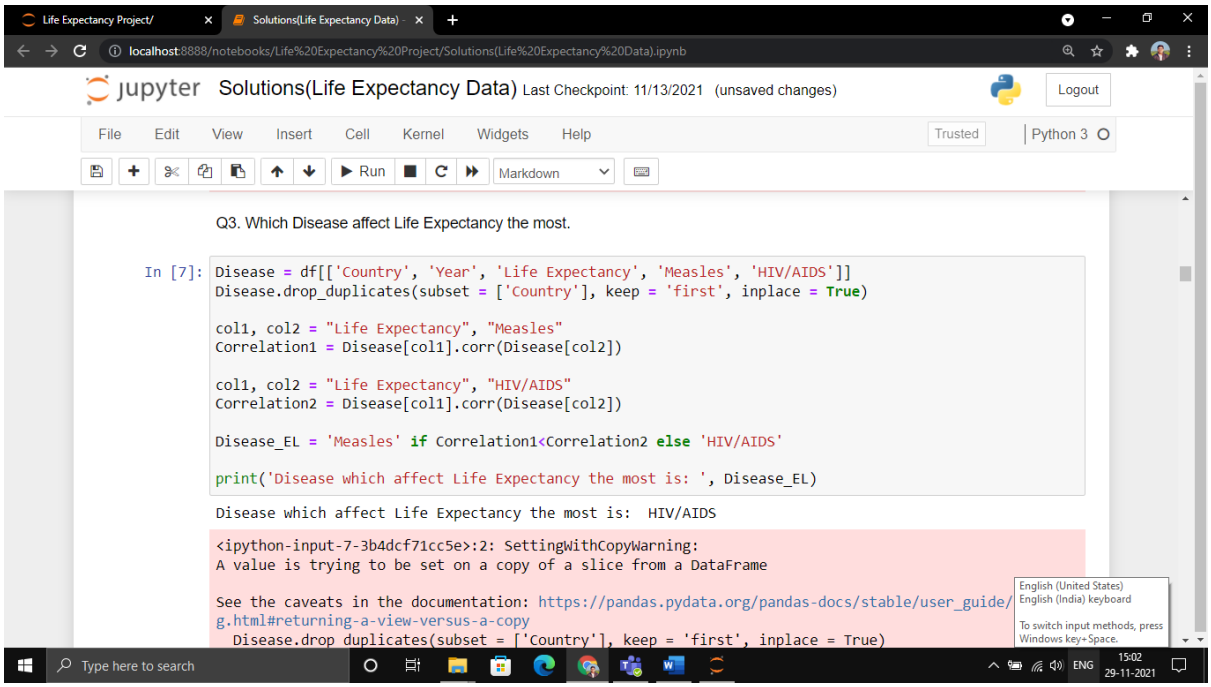
Que 1. Finding top 10 Countries with maximum number of Life Expectancy.



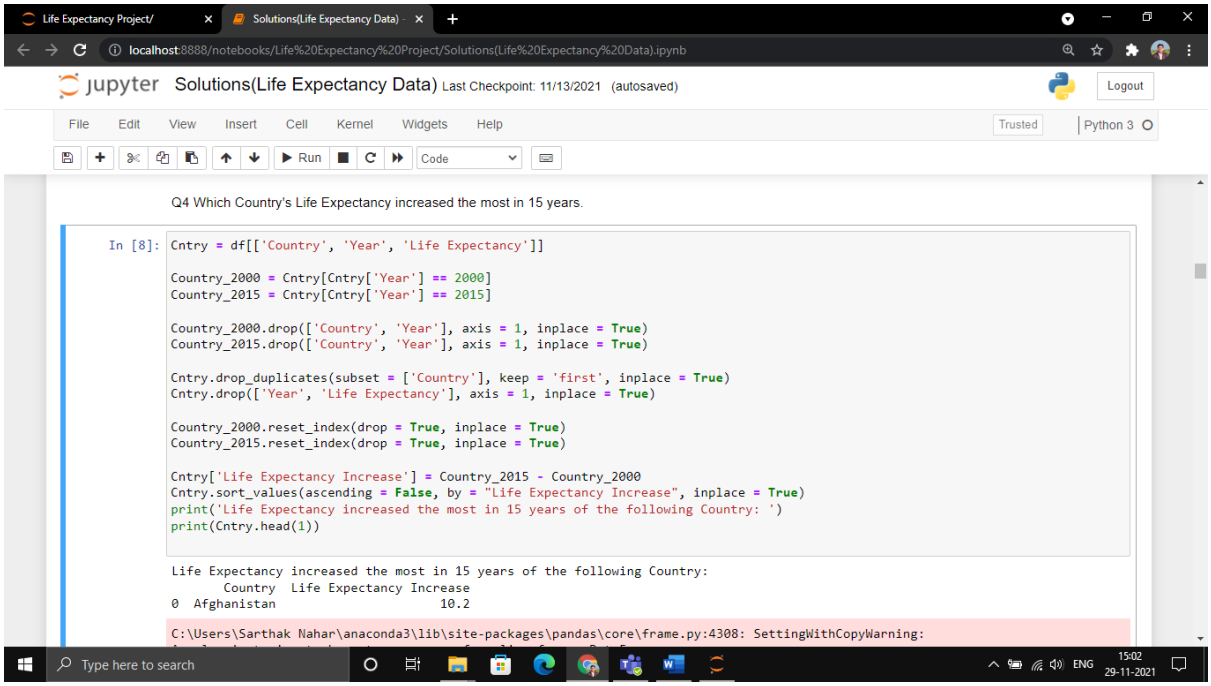
Que 2. Find the Correlation between GDP and Life Expectancy.



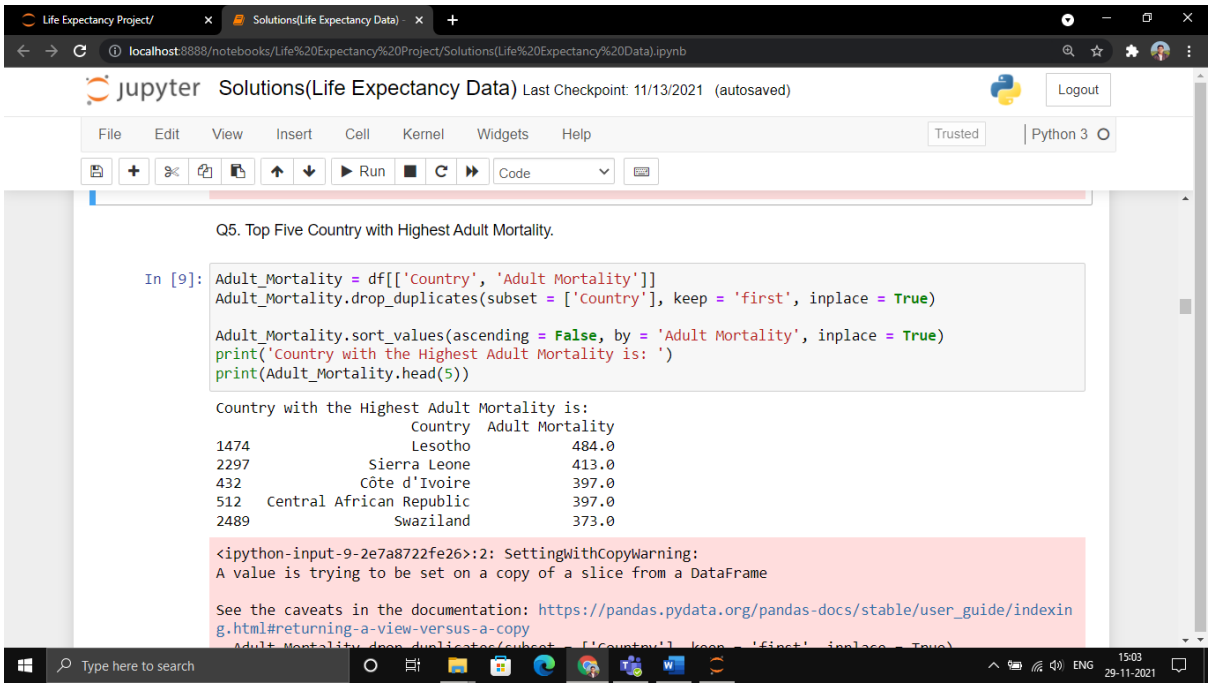
Que 3. Which Disease affect Life Expectancy the most.



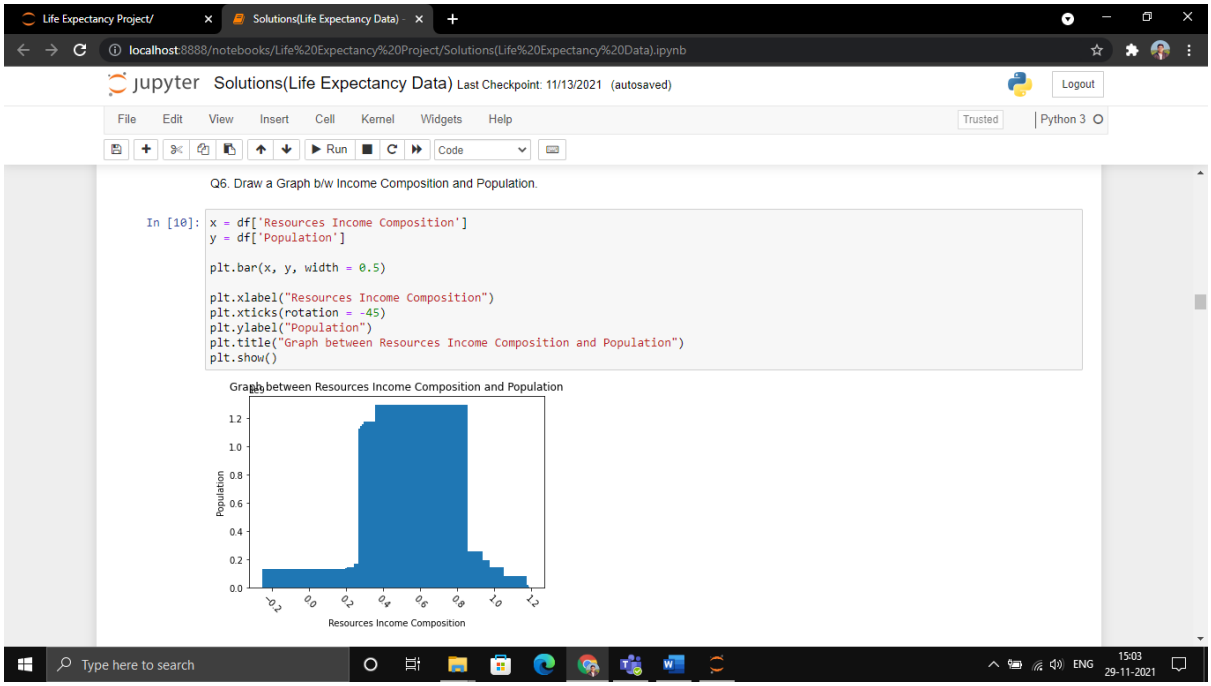
Que 4. Which Country's Life Expectancy increased the most in 15 years.



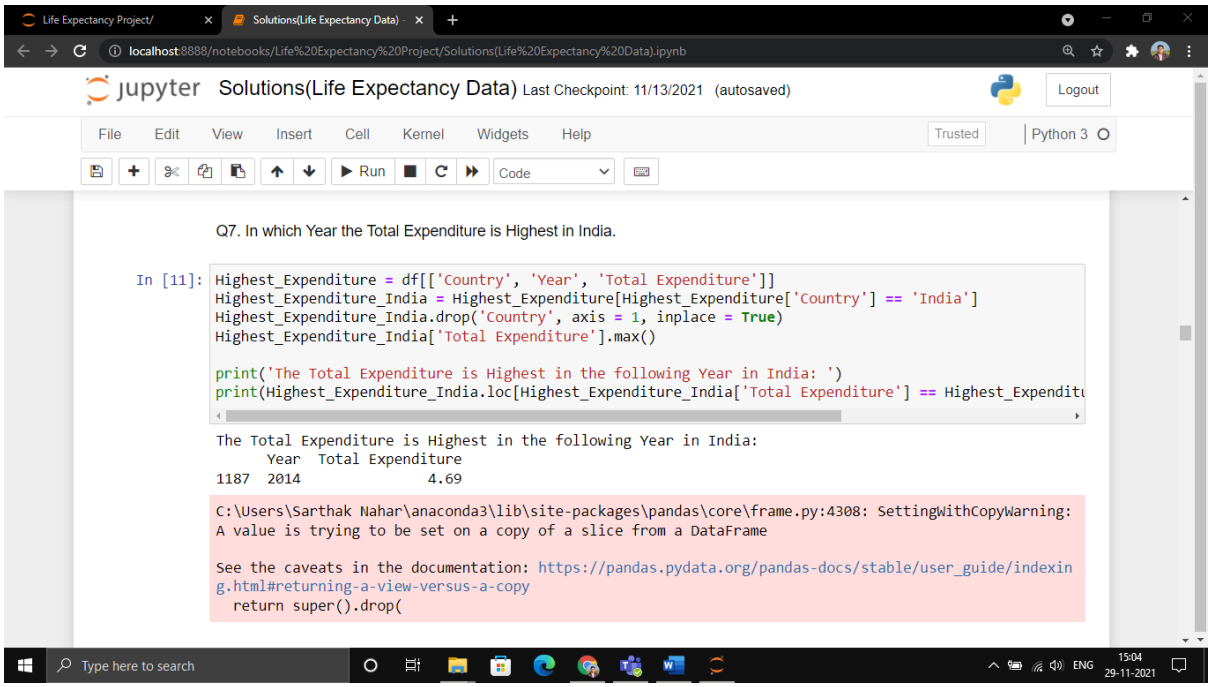
Que 5. Top Five Country with Highest Adult Mortality.



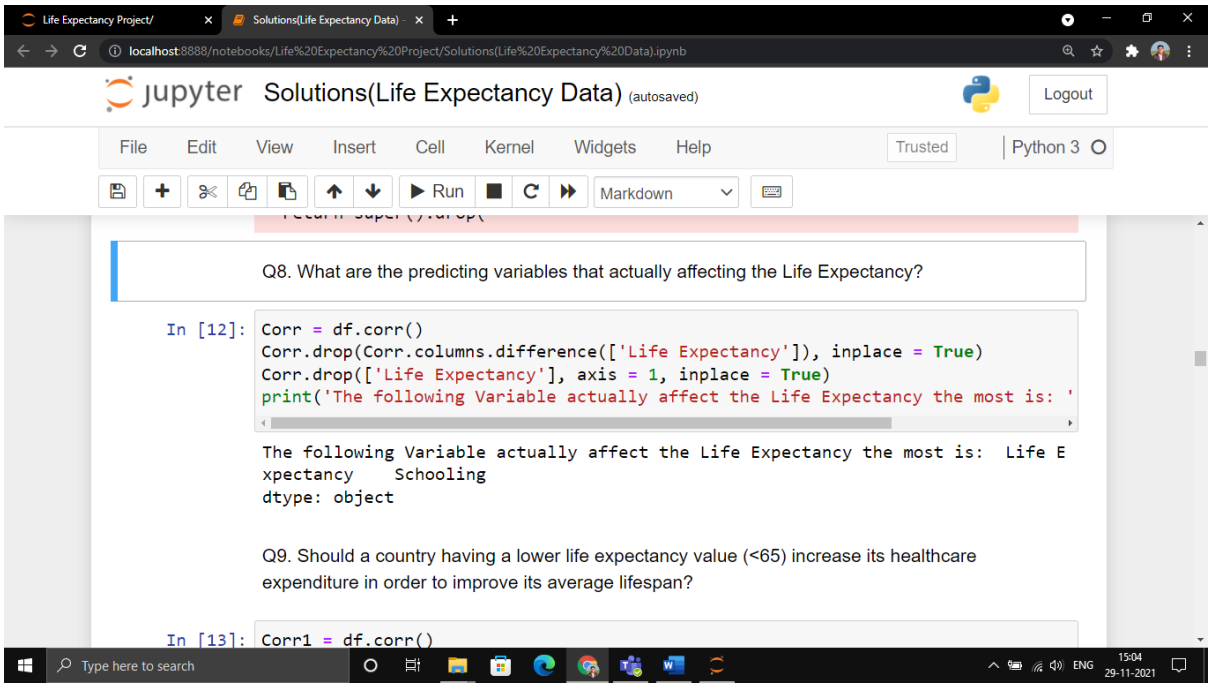
Que 6. Draw a Graph b/w Income Composition and Population.



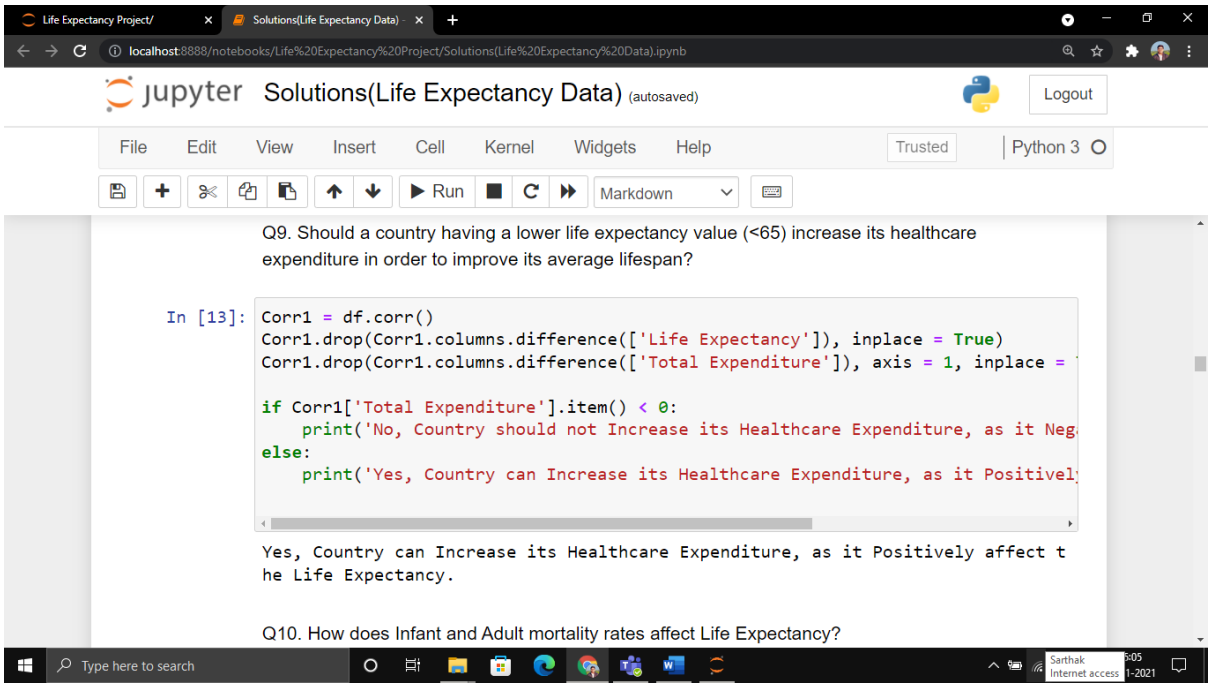
Que 7. In which Year the Total Expenditure is Highest in India.



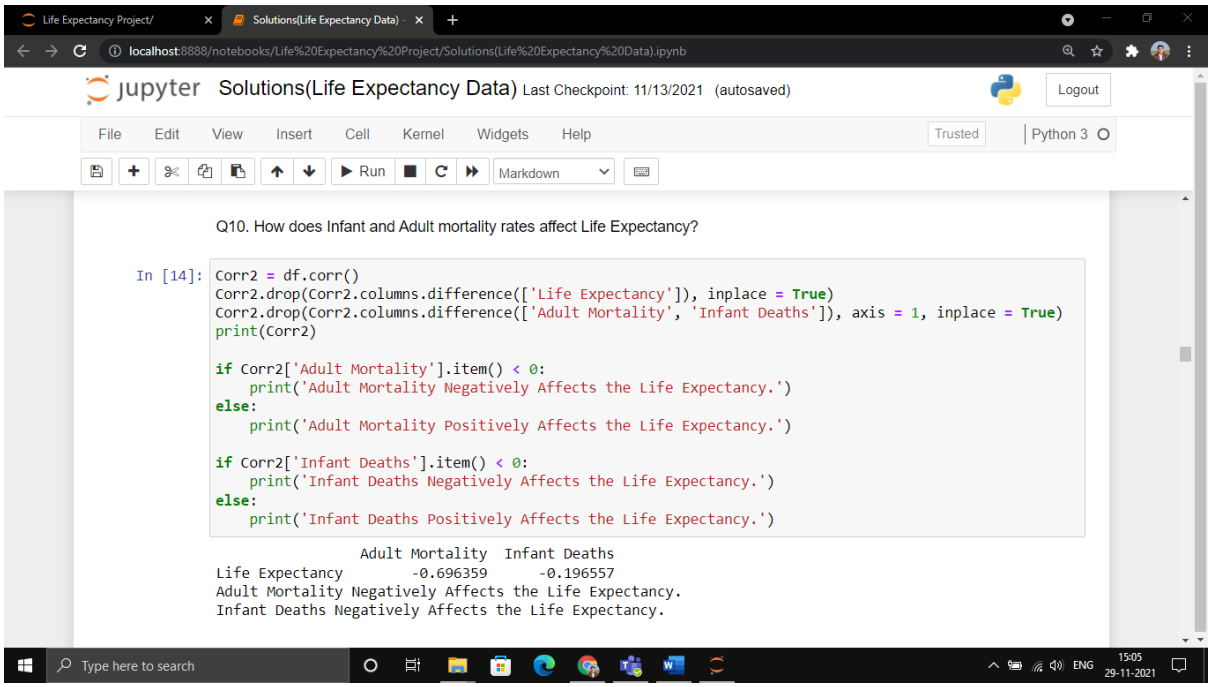
Que 8. What are the predicting variables that actually affecting the Life Expectancy?



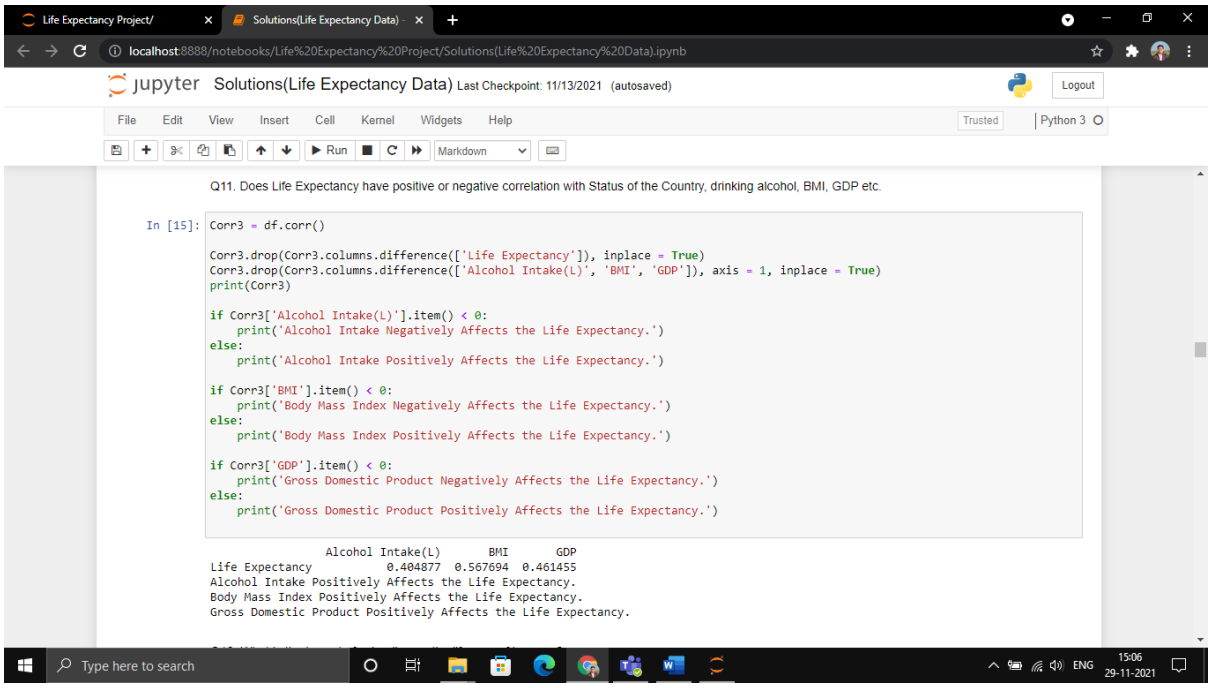
Que 9. Should a country having a lower life expectancy value (<65) increase its healthcare expenditure in order to improve its average lifespan?



Que 10. How does Infant and Adult mortality rates affect Life Expectancy?



Que 11. Does Life Expectancy have positive or negative correlation with Status of the Country, drinking alcohol, BMI, GDP etc.

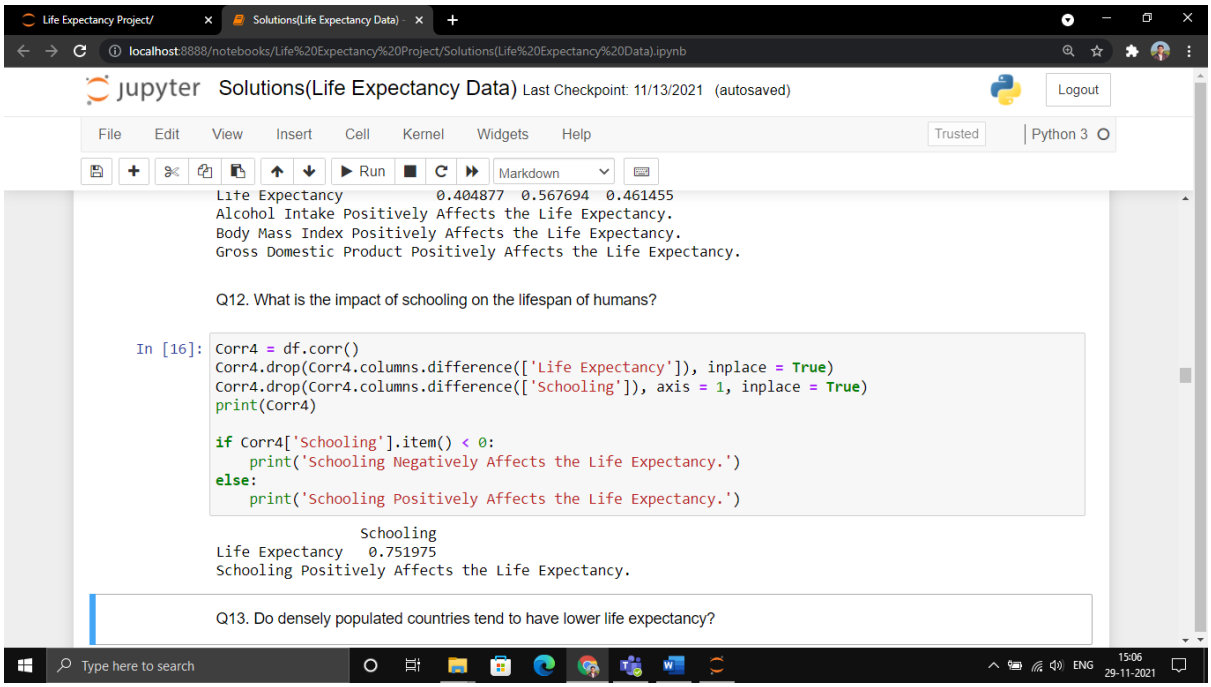


Q11. Does Life Expectancy have positive or negative correlation with Status of the Country, drinking alcohol, BMI, GDP etc.

In [15]:
```
Corr3 = df.corr()

Corr3.drop(Corr3.columns.difference(['Life Expectancy']), inplace = True)
Corr3.drop(Corr3.columns.difference(['Alcohol Intake(L)', 'BMI', 'GDP']), axis = 1, inplace = True)
print(Corr3)

if Corr3['Alcohol Intake(L)'].item() < 0:
    print('Alcohol Intake Negatively Affects the Life Expectancy.')
else:
    print('Alcohol Intake Positively Affects the Life Expectancy.')

if Corr3['BMI'].item() < 0:
    print('Body Mass Index Negatively Affects the Life Expectancy.')
else:
    print('Body Mass Index Positively Affects the Life Expectancy.')

if Corr3['GDP'].item() < 0:
    print('Gross Domestic Product Negatively Affects the Life Expectancy.')
else:
    print('Gross Domestic Product Positively Affects the Life Expectancy.')
```

```
                 Alcohol Intake(L)       BMI       GDP
Life Expectancy           0.404877  0.567694  0.461455
Alcohol Intake Positively Affects the Life Expectancy.
Body Mass Index Positively Affects the Life Expectancy.
Gross Domestic Product Positively Affects the Life Expectancy.
```

Que 12. What is the impact of schooling on the lifespan of humans?



```
Life Expectancy           0.404877  0.567694  0.461455
Alcohol Intake Positively Affects the Life Expectancy.
Body Mass Index Positively Affects the Life Expectancy.
Gross Domestic Product Positively Affects the Life Expectancy.
```

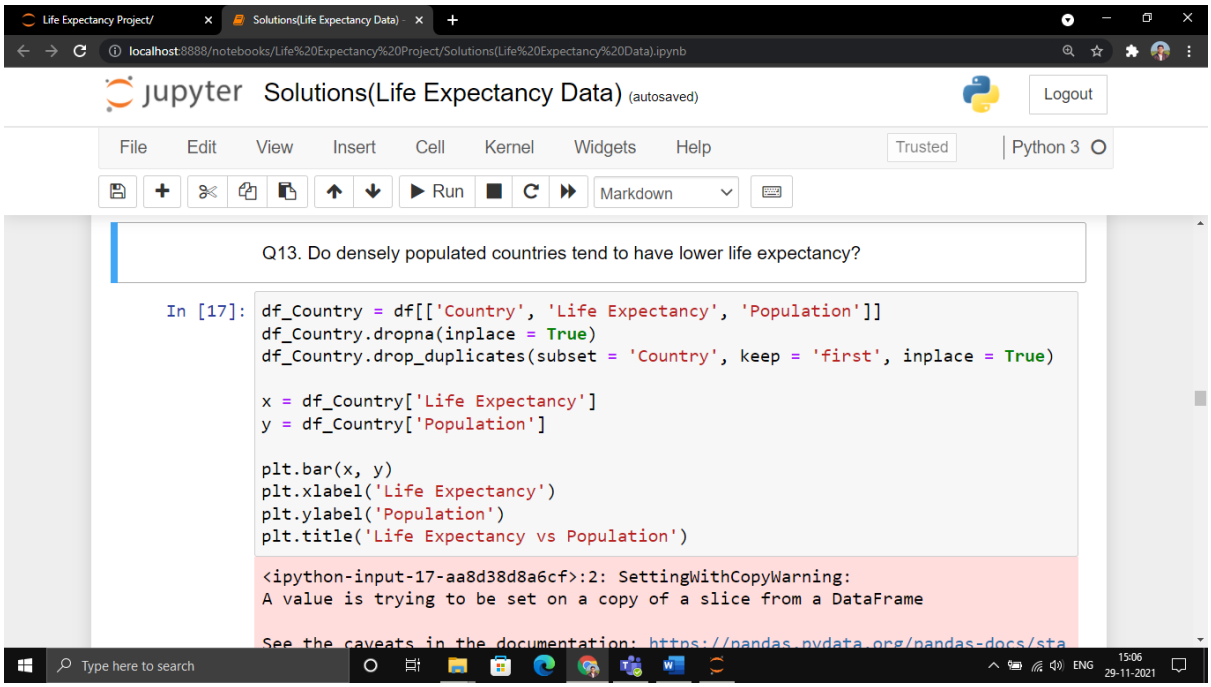
Q12. What is the impact of schooling on the lifespan of humans?

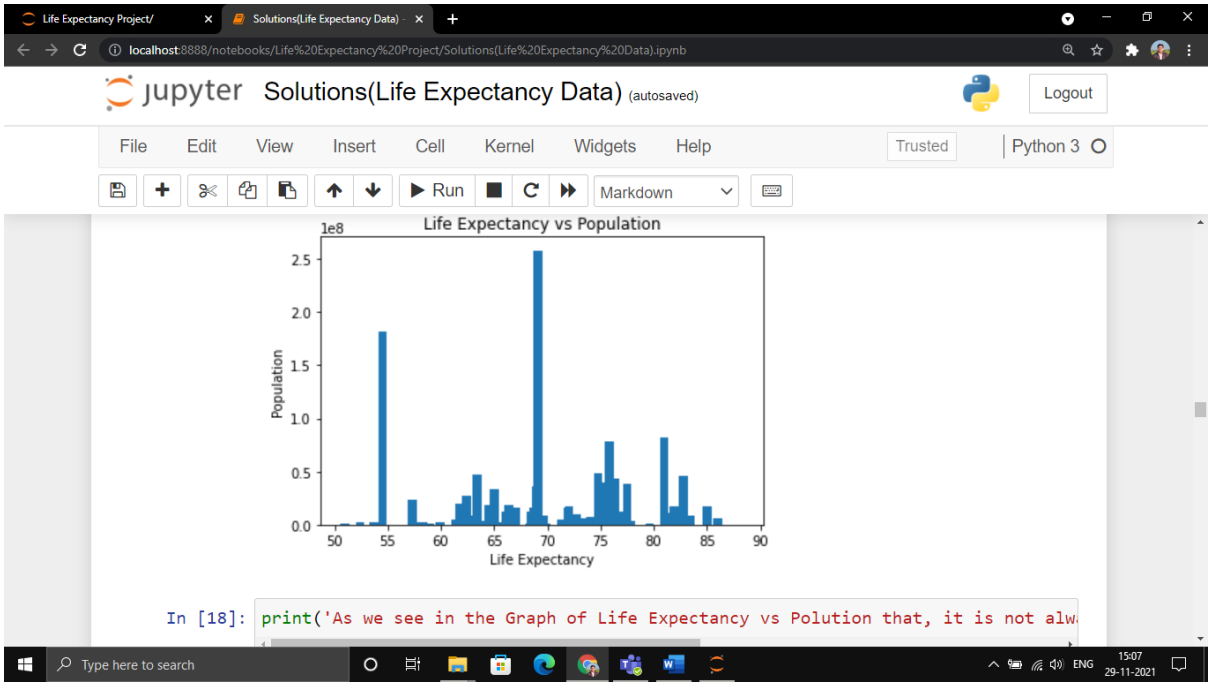
In [16]:
```
Corr4 = df.corr()
Corr4.drop(Corr4.columns.difference(['Life Expectancy']), inplace = True)
Corr4.drop(Corr4.columns.difference(['Schooling']), axis = 1, inplace = True)
print(Corr4)

if Corr4['Schooling'].item() < 0:
    print('Schooling Negatively Affects the Life Expectancy.')
else:
    print('Schooling Positively Affects the Life Expectancy.')
```

```
                Schooling
Life Expectancy   0.751975
Schooling Positively Affects the Life Expectancy.
```

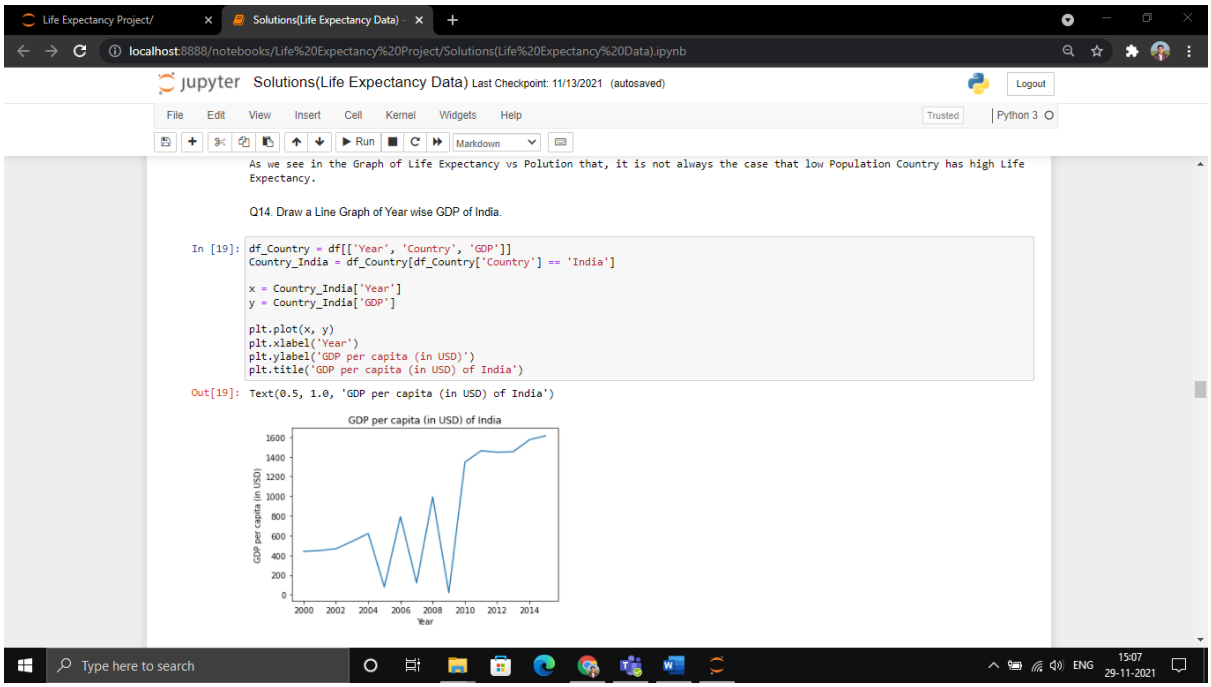
Q13. Do densely populated countries tend to have lower life expectancy?

Que 13. Do densely populated countries tend to have lower life expectancy?

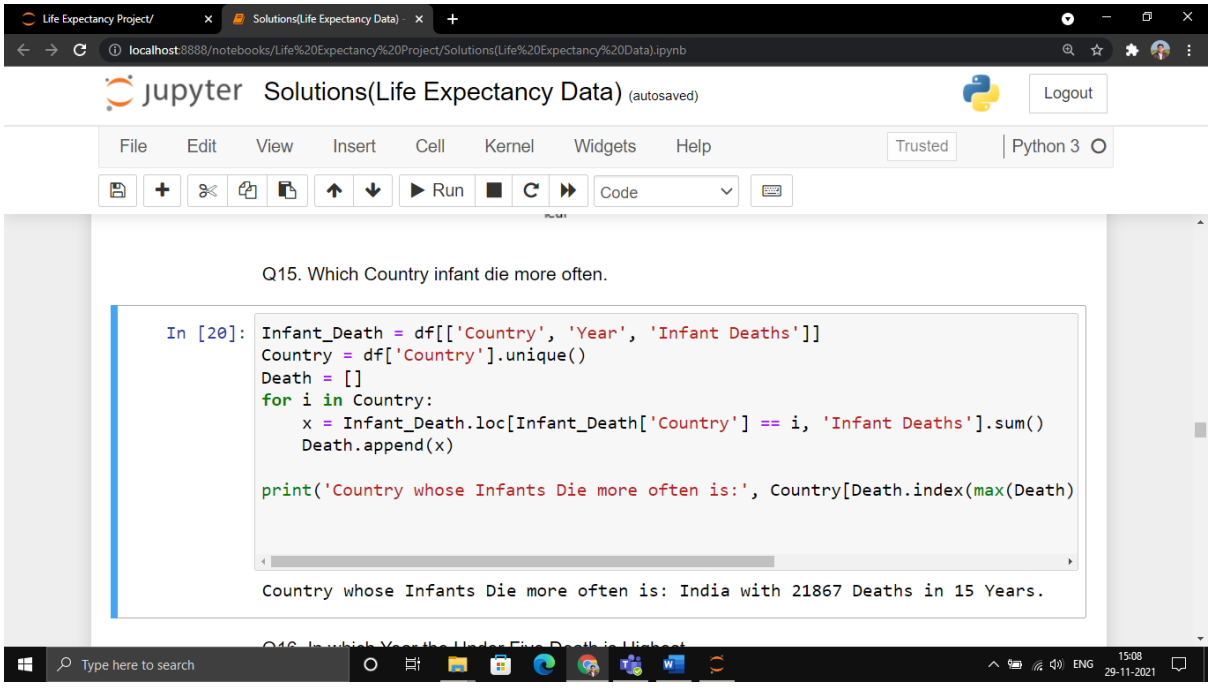




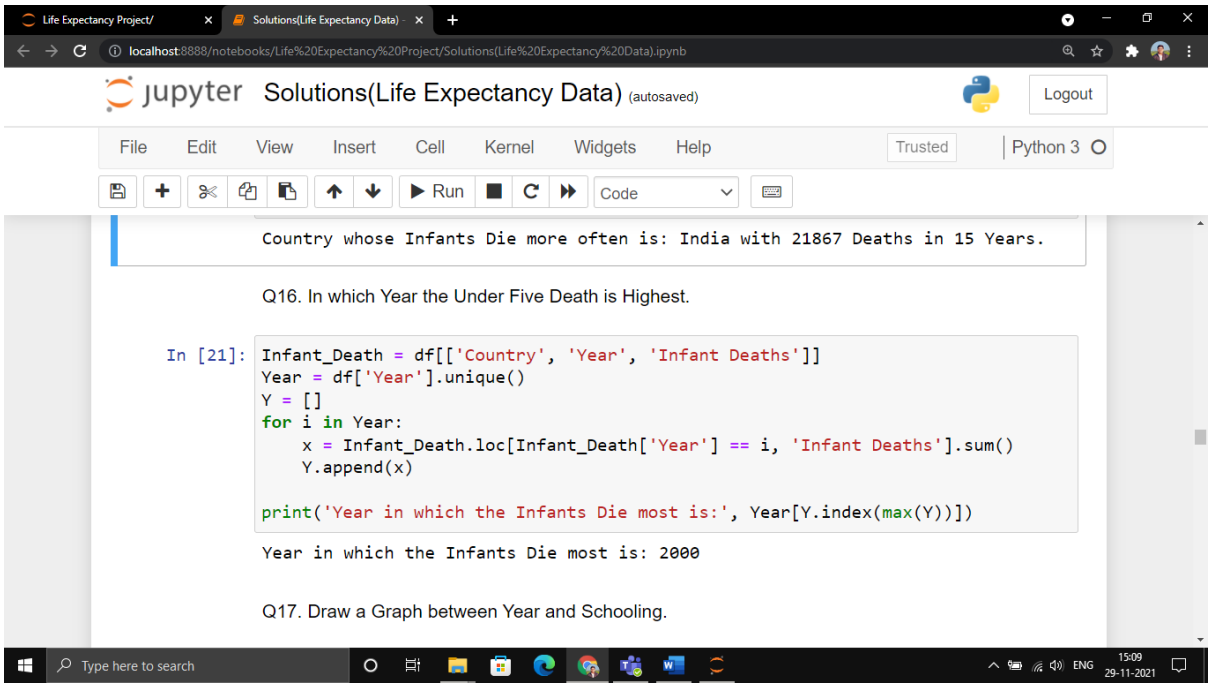
Que 14. Draw a Line Graph of Year wise GDP of India.



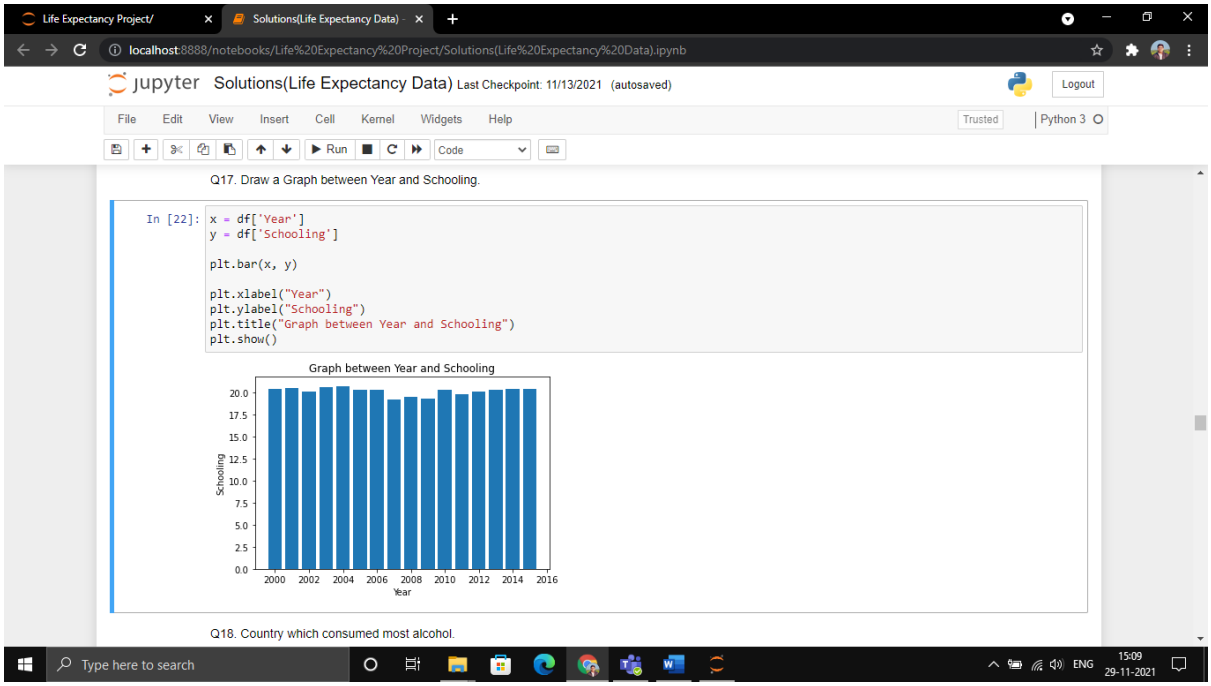
Que 15. Which Country infant die more often.



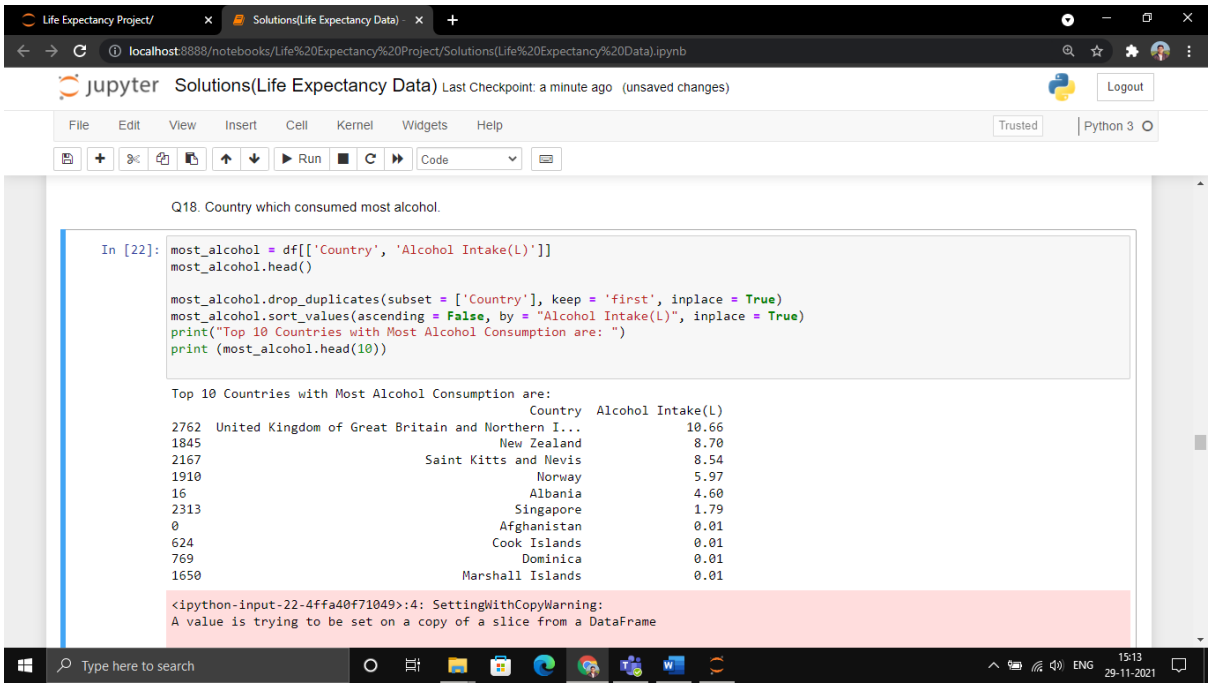
Que 16. In which Year the Under Five Death is Highest.



Que 17. Draw a Graph between Year and Schooling.



Que 18. Country which consumed most alcohol.



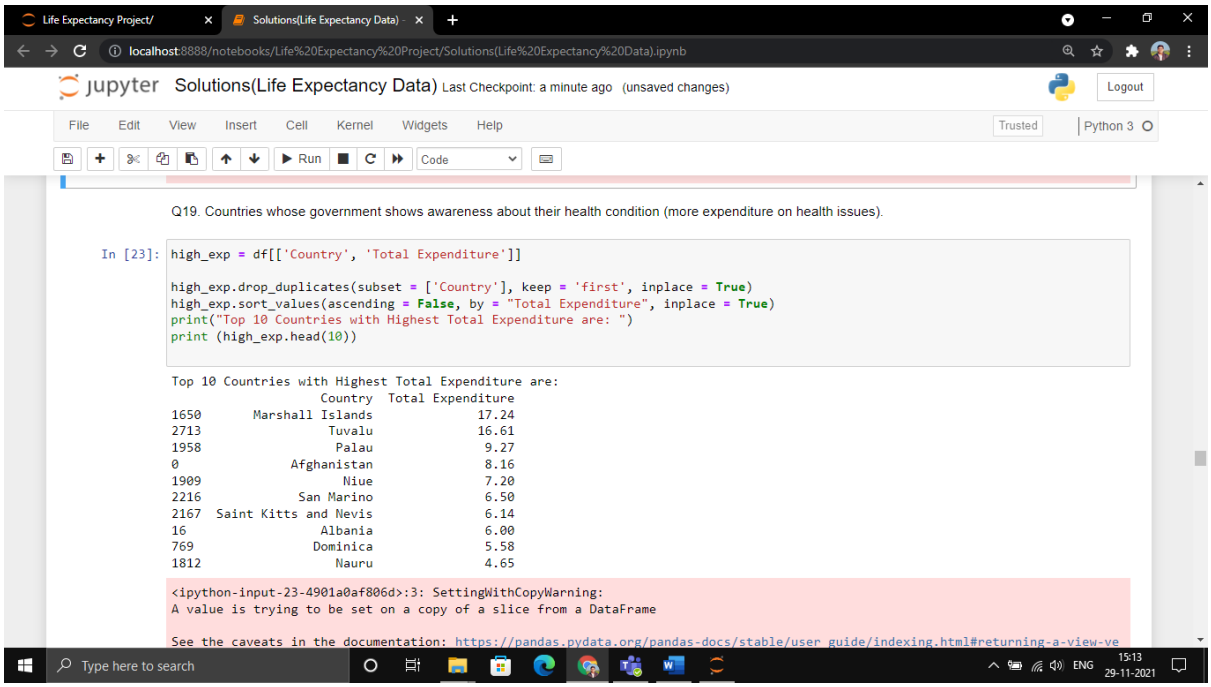
Q18. Country which consumed most alcohol.

```
In [22]: most_alcohol = df[['Country', 'Alcohol Intake(L)']]
most_alcohol.head()

most_alcohol.drop_duplicates(subset = ['Country'], keep = 'first', inplace = True)
most_alcohol.sort_values(ascending = False, by = "Alcohol Intake(L)", inplace = True)
print("Top 10 Countries with Most Alcohol Consumption are: ")
print (most_alcohol.head(10))
```

```
Top 10 Countries with Most Alcohol Consumption are:
                                          Country  Alcohol Intake(L)
2762  United Kingdom of Great Britain and Northern I...              10.66
1845                                   New Zealand               8.70
2167                            Saint Kitts and Nevis               8.54
1910                                         Norway               5.97
16                                          Albania               4.60
2313                                      Singapore               1.79
0                                       Afghanistan               0.01
624                                    Cook Islands               0.01
769                                        Dominica               0.01
1650                                Marshall Islands               0.01
```

```
<ipython-input-22-4ffa40f71049>:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

Que 19. Countries whose government shows awareness about their health condition (more expenditure on health issues).



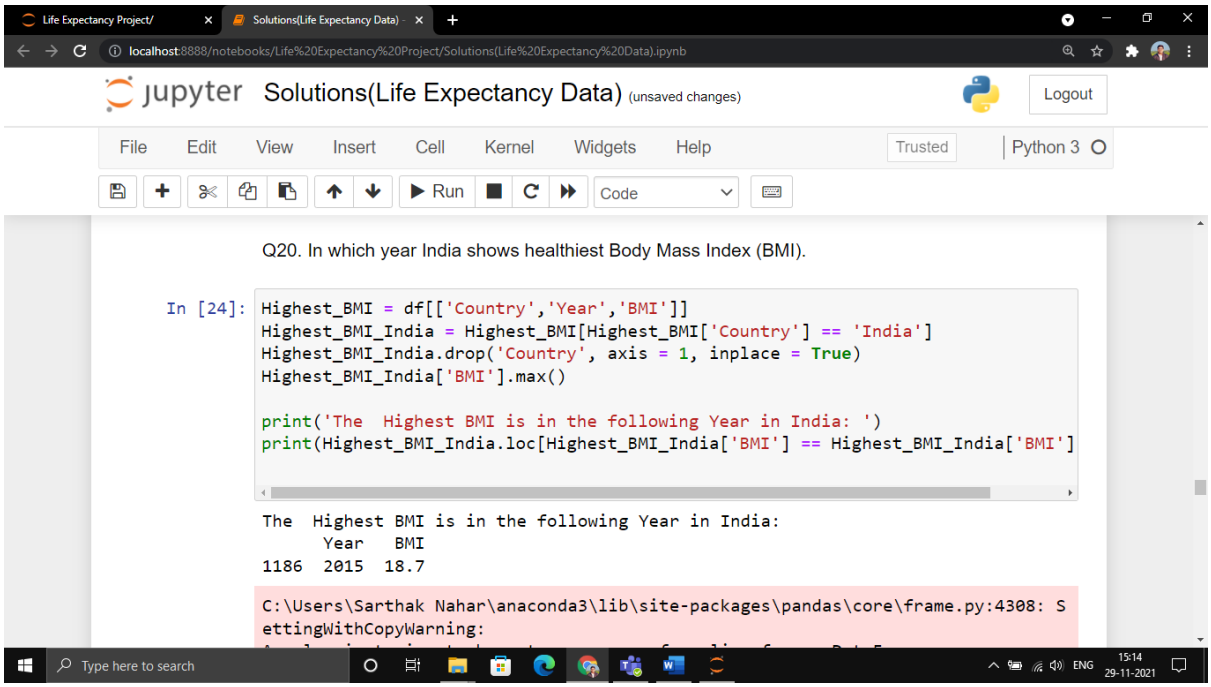
Q19. Countries whose government shows awareness about their health condition (more expenditure on health issues).

```
In [23]: high_exp = df[['Country', 'Total Expenditure']]

high_exp.drop_duplicates(subset = ['Country'], keep = 'first', inplace = True)
high_exp.sort_values(ascending = False, by = "Total Expenditure", inplace = True)
print("Top 10 Countries with Highest Total Expenditure are: ")
print (high_exp.head(10))
```

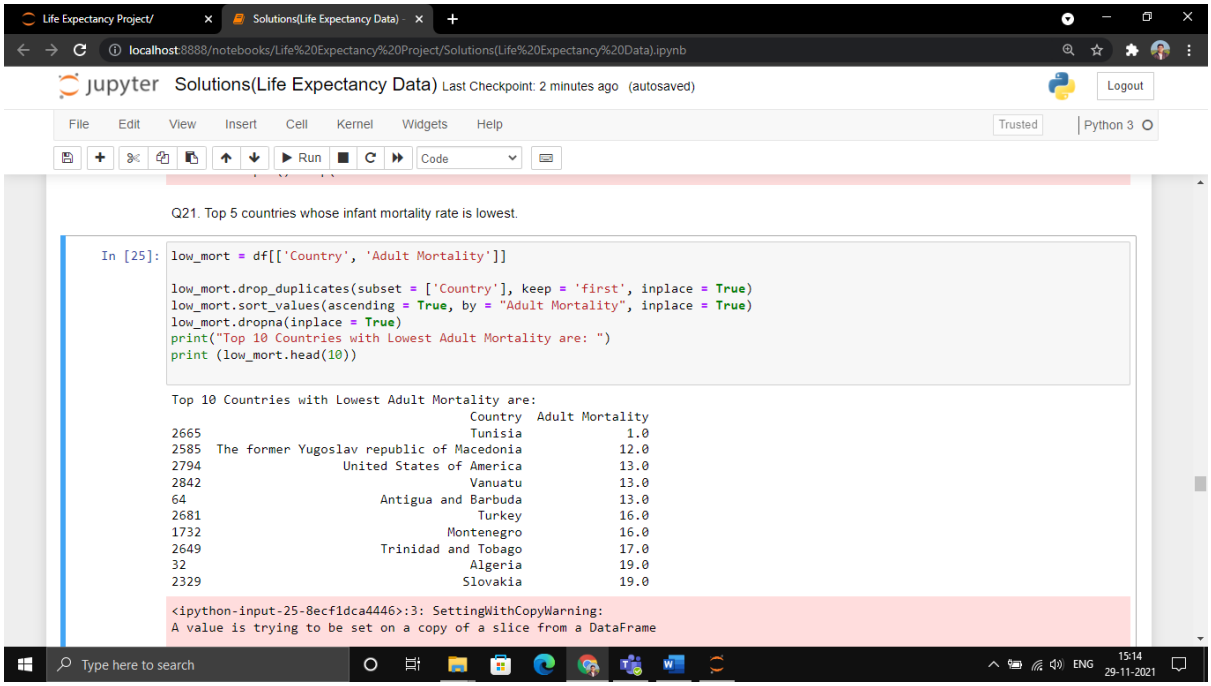
```
Top 10 Countries with Highest Total Expenditure are:
                      Country  Total Expenditure
1650       Marshall Islands              17.24
2713                 Tuvalu              16.61
1958                  Palau               9.27
0               Afghanistan               8.16
1909                   Niue               7.20
2216             San Marino               6.50
2167  Saint Kitts and Nevis               6.14
16                   Albania               6.00
769                 Dominica               5.58
1812                   Nauru               4.65
```

```
<ipython-input-23-4901a0af806d>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-ve
```

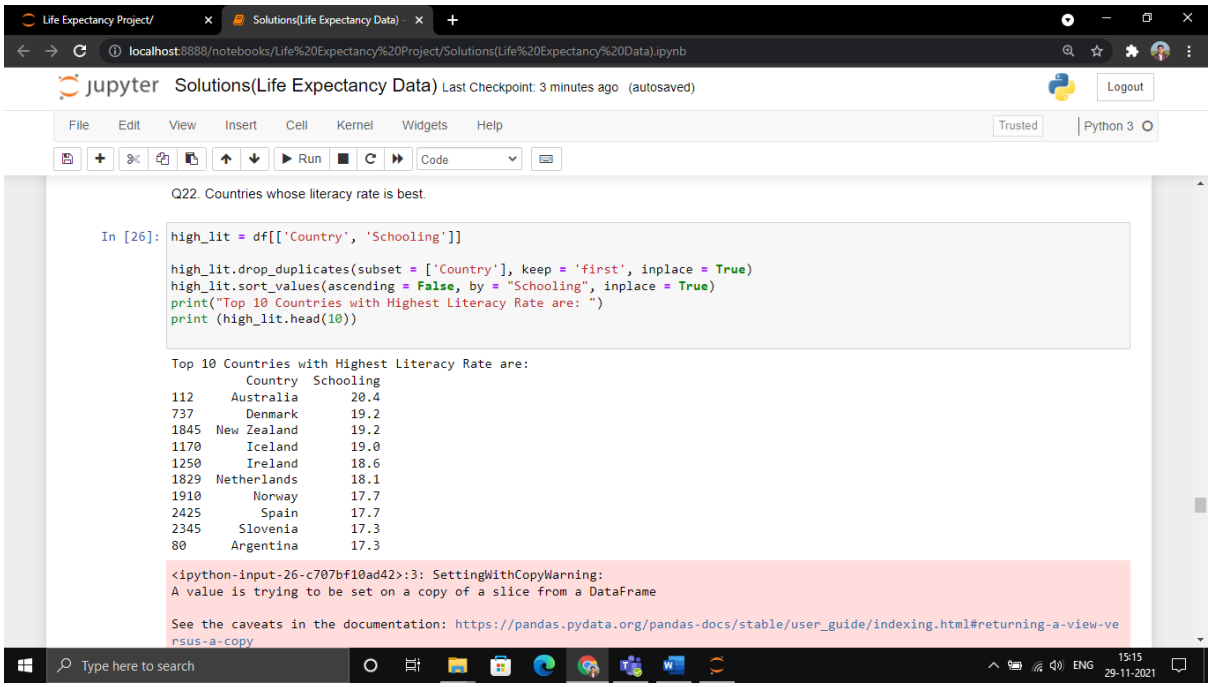
Que 20. In which year India shows healthiest Body Mass Index (BMI).



Que 21. Top 5 countries whose infant mortality rate is lowest.



Que 22. Countries whose literacy rate is best.



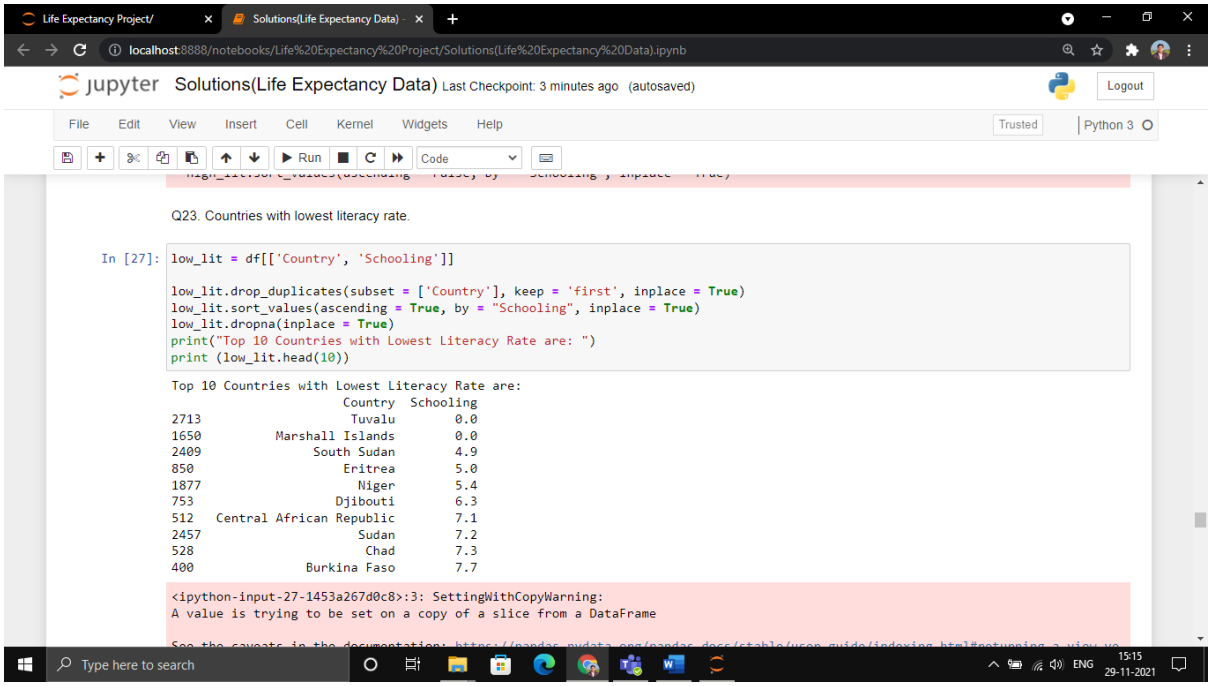
Q22. Countries whose literacy rate is best.

```
In [26]: high_lit = df[['Country', 'Schooling']]

high_lit.drop_duplicates(subset = ['Country'], keep = 'first', inplace = True)
high_lit.sort_values(ascending = False, by = "Schooling", inplace = True)
print("Top 10 Countries with Highest Literacy Rate are: ")
print (high_lit.head(10))
```

```
Top 10 Countries with Highest Literacy Rate are:
          Country  Schooling
112     Australia       20.4
737       Denmark       19.2
1845  New Zealand       19.2
1170      Iceland       19.0
1250      Ireland       18.6
1829  Netherlands       18.1
1910       Norway       17.7
2425        Spain       17.7
2345     Slovenia       17.3
80      Argentina       17.3
```

```
<ipython-input-26-c707bf10ad42>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
```

Que 23. Countries with lowest literacy rate.



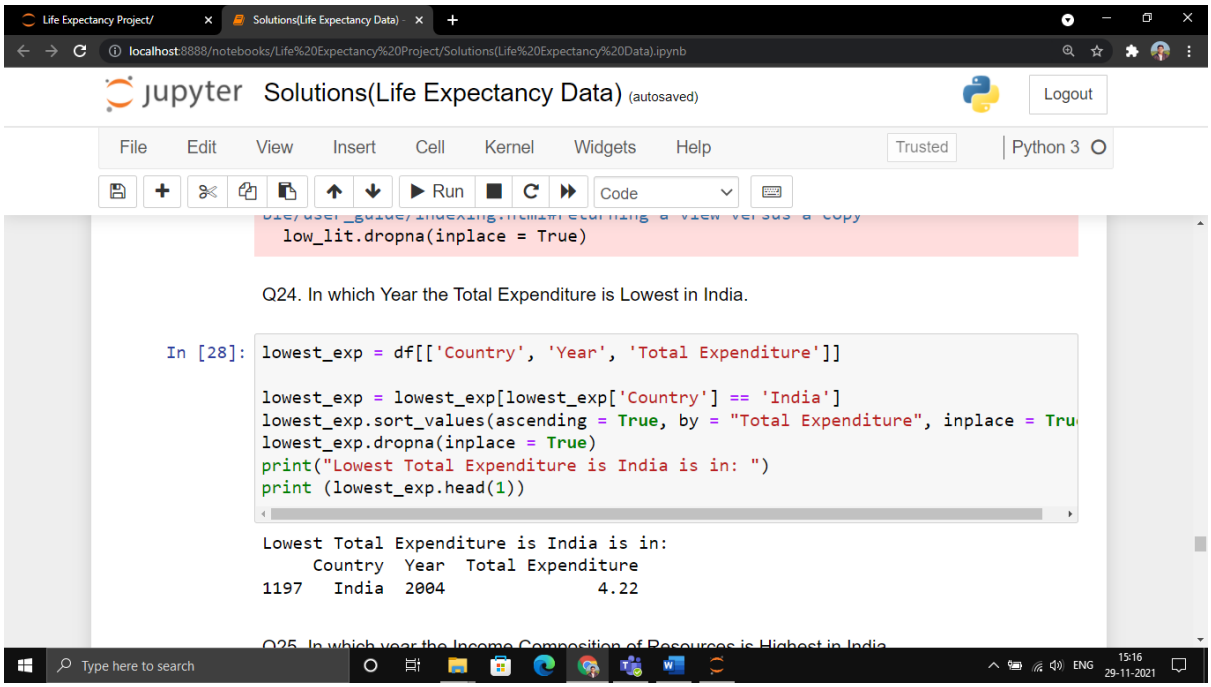
Q23. Countries with lowest literacy rate.

```
In [27]: low_lit = df[['Country', 'Schooling']]

low_lit.drop_duplicates(subset = ['Country'], keep = 'first', inplace = True)
low_lit.sort_values(ascending = True, by = "Schooling", inplace = True)
low_lit.dropna(inplace = True)
print("Top 10 Countries with Lowest Literacy Rate are: ")
print (low_lit.head(10))
```

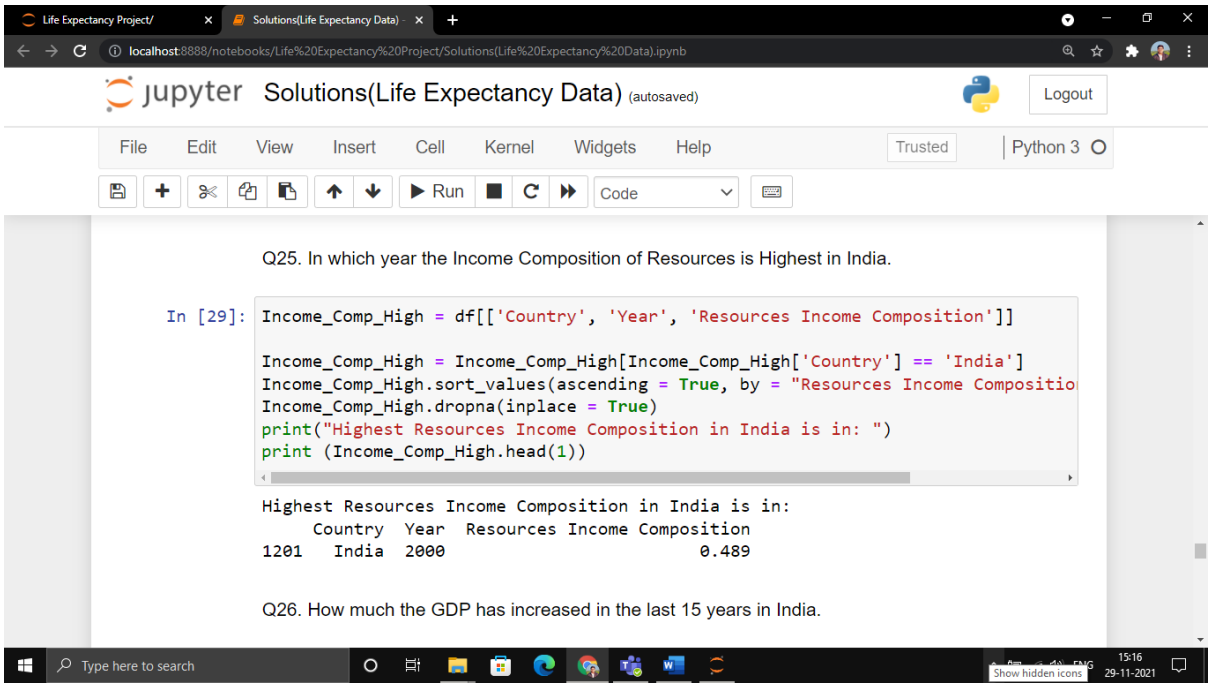
```
Top 10 Countries with Lowest Literacy Rate are:
                        Country  Schooling
2713                      Tuvalu        0.0
1650             Marshall Islands       0.0
2409                  South Sudan       4.9
850                       Eritrea       5.0
1877                        Niger       5.4
753                      Djibouti       6.3
512   Central African Republic        7.1
2457                        Sudan       7.2
528                          Chad       7.3
400                  Burkina Faso       7.7
```

```
<ipython-input-27-1453a267d0c8>:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

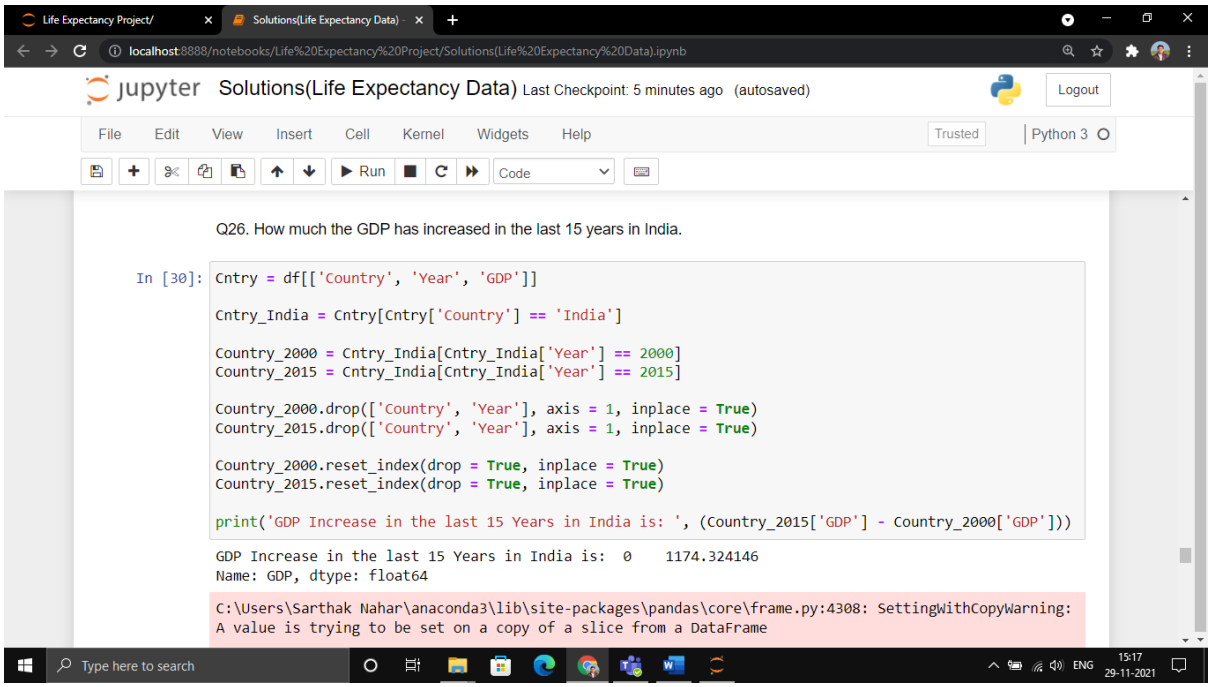
Que 24. In which Year the Total Expenditure is Lowest in India.



Que 25. In which year the Income Composition of Resources is Highest in India.



Que 26. How much the GDP has increased in the last 15 years in India.



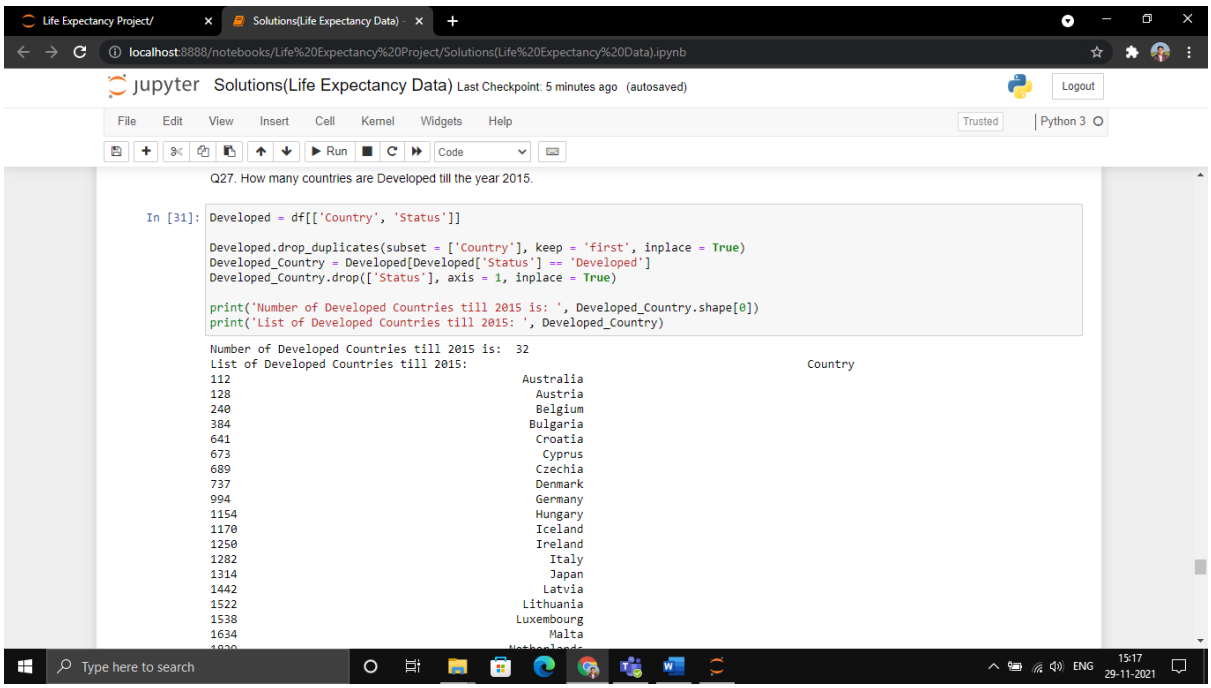
Q26. How much the GDP has increased in the last 15 years in India.

```
In [30]: Cntry = df[['Country', 'Year', 'GDP']]

         Cntry_India = Cntry[Cntry['Country'] == 'India']

         Country_2000 = Cntry_India[Cntry_India['Year'] == 2000]
         Country_2015 = Cntry_India[Cntry_India['Year'] == 2015]

         Country_2000.drop(['Country', 'Year'], axis = 1, inplace = True)
         Country_2015.drop(['Country', 'Year'], axis = 1, inplace = True)

         Country_2000.reset_index(drop = True, inplace = True)
         Country_2015.reset_index(drop = True, inplace = True)

         print('GDP Increase in the last 15 Years in India is: ', (Country_2015['GDP'] - Country_2000['GDP']))
```

```
GDP Increase in the last 15 Years in India is:  0    1174.324146
Name: GDP, dtype: float64
```

```
C:\Users\Sarthak Nahar\anaconda3\lib\site-packages\pandas\core\frame.py:4308: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

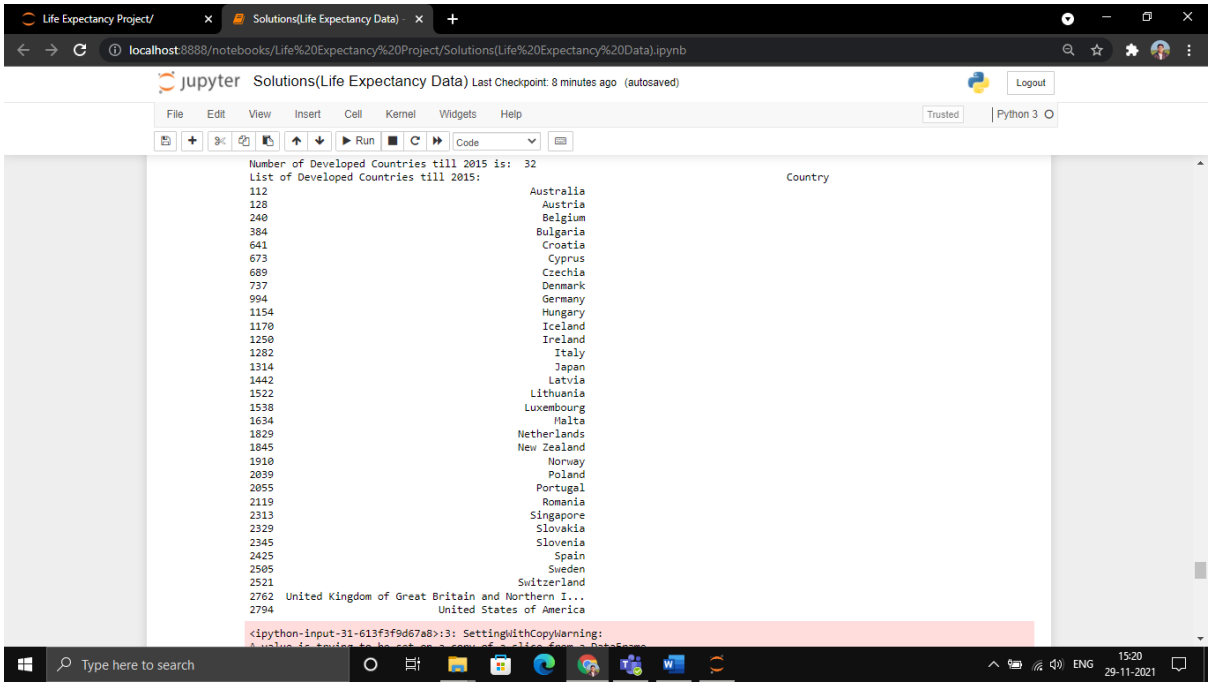
Que 27. How many countries are Developed till the year 2015.



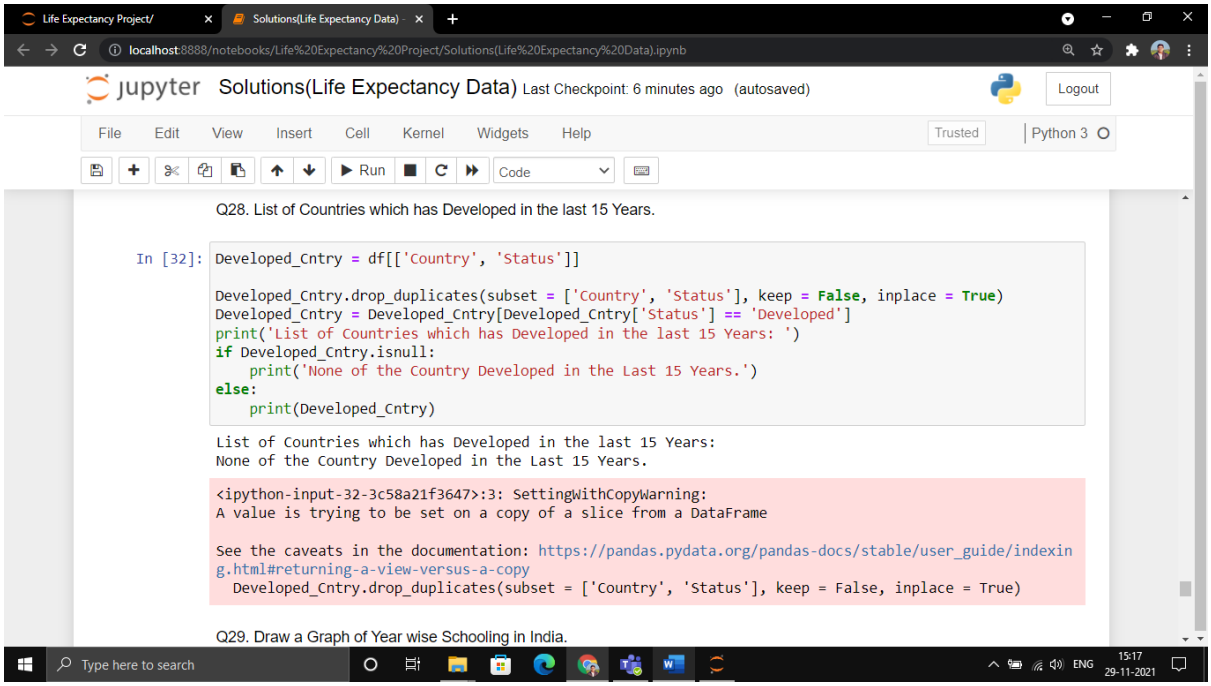
Q27. How many countries are Developed till the year 2015.

```
In [31]: Developed = df[['Country', 'Status']]

         Developed.drop_duplicates(subset = ['Country'], keep = 'first', inplace = True)
         Developed_Country = Developed[Developed['Status'] == 'Developed']
         Developed_Country.drop(['Status'], axis = 1, inplace = True)

         print('Number of Developed Countries till 2015 is: ', Developed_Country.shape[0])
         print('List of Developed Countries till 2015: ', Developed_Country)
```

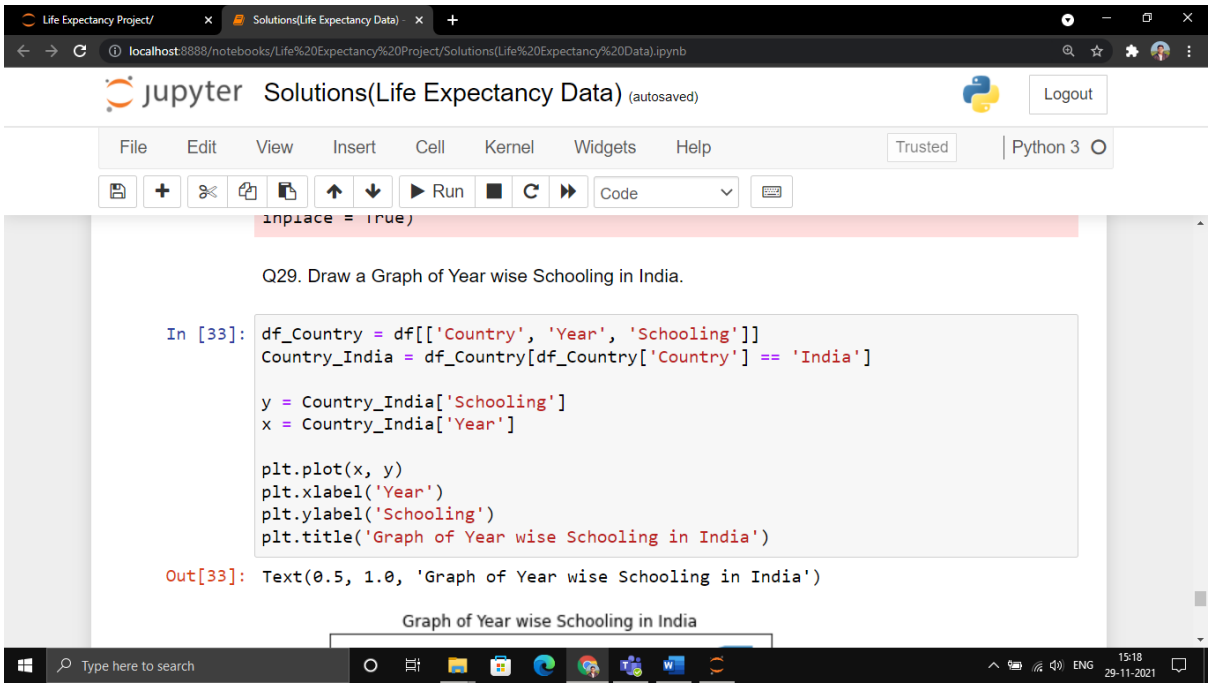
```
Number of Developed Countries till 2015 is:  32
List of Developed Countries till 2015:                                        Country
112                                        Australia
128                                          Austria
240                                          Belgium
384                                         Bulgaria
641                                          Croatia
673                                           Cyprus
689                                          Czechia
737                                          Denmark
994                                          Germany
1154                                         Hungary
1170                                         Iceland
1250                                         Ireland
1282                                           Italy
1314                                           Japan
1442                                          Latvia
1522                                       Lithuania
1538                                      Luxembourg
1634                                           Malta
```

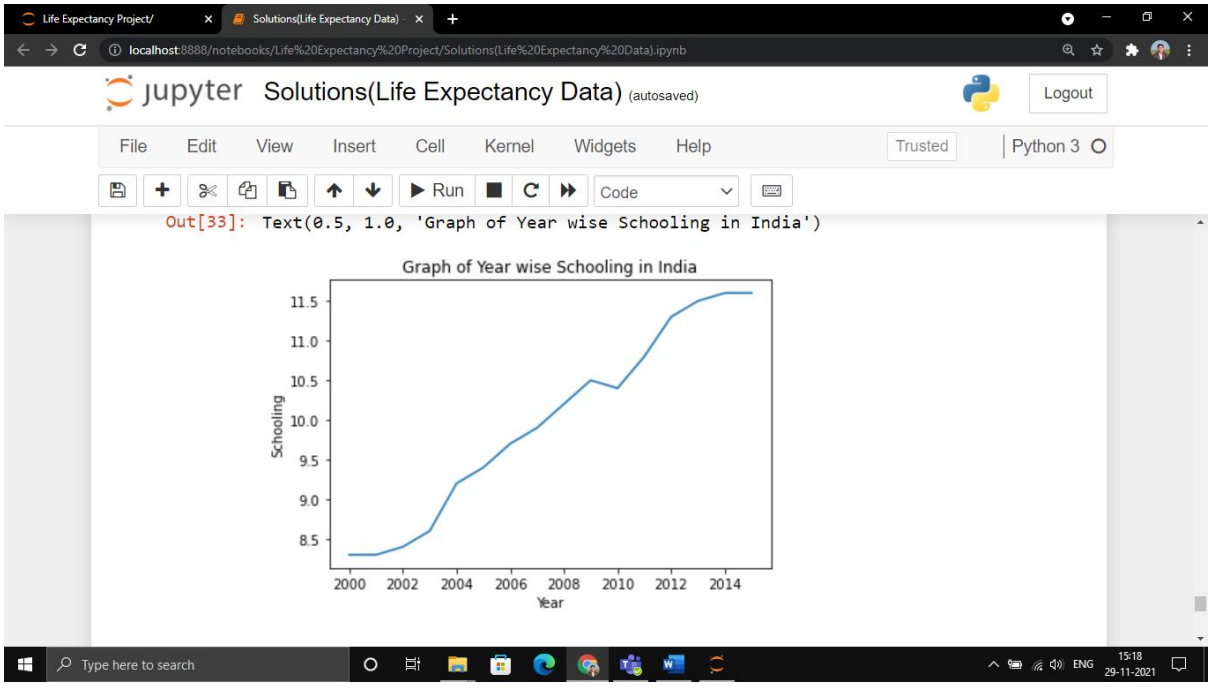


Que 28. List of Countries which has Developed in the last 15 Years.



Que 29. Draw a Graph of Year wise Schooling in India.





Que 30. Draw a Graph of Life Expectancy of different Countries.