Understanding the life expectancy of Indian population using the world development indicators

In [61]:

```
import pandas as pd
import numpy as np
import random
import matplotlib.pyplot as plt
```

In [62]:

```
data = pd.read_csv('./world-development-indicators/Indicators.csv')
data.shape
```

Out[62]:

```
(5656458, 6)
```

In [63]:

```
data.head(10)
```

Out[63]:

	CountryName	CountryCode	IndicatorName	IndicatorCode	Year	Value
0	Arab World	ARB	Adolescent fertility rate (births per 1,000 wo...	SP.ADO.TFRT	1960	1.335609e+02
1	Arab World	ARB	Age dependency ratio (% of working-age populat...	SP.POP.DPND	1960	8.779760e+01
2	Arab World	ARB	Age dependency ratio, old (% of working-age po...	SP.POP.DPND.OL	1960	6.634579e+00
3	Arab World	ARB	Age dependency ratio, young (% of working-age ...	SP.POP.DPND.YG	1960	8.102333e+01
4	Arab World	ARB	Arms exports (SIPRI trend indicator values)	MS.MIL.XPRT.KD	1960	3.000000e+06
5	Arab World	ARB	Arms imports (SIPRI trend indicator values)	MS.MIL.MPRT.KD	1960	5.380000e+08
6	Arab World	ARB	Birth rate, crude (per 1,000 people)	SP.DYN.CBRT.IN	1960	4.769789e+01
7	Arab World	ARB	CO2 emissions (kt)	EN.ATM.CO2E.KT	1960	5.956399e+04
8	Arab World	ARB	CO2 emissions (metric tons per capita)	EN.ATM.CO2E.PC	1960	6.439635e-01
9	Arab World	ARB	CO2 emissions from gaseous fuel consumption (%...	EN.ATM.CO2E.GF.ZS	1960	5.041292e+00

In [66]:

```
# How many unique indicators are there ? (should be the same #)
indicators = data['IndicatorName'].unique().tolist()
```

In [38]:

```
indicators
```

```
 'Prevalence of anemia among children (% of children under 5)',
 'Prevalence of anemia among non-pregnant women (% of women ages 15-49)',
 'Prevalence of anemia among pregnant women (%)',
 'Prevalence of HIV, female (% ages 15-24)',
 'Prevalence of HIV, male (% ages 15-24)',
 'Prevalence of HIV, total (% of population ages 15-49)',
 'Proportion of seats held by women in national parliaments (%)',
 'Ratio of female to male labor force participation rate (%) (modeled ILO estimate)',
 'Refugee population by country or territory of asylum',
 'Refugee population by country or territory of origin',
 'Renewable electricity output (% of total electricity output)',
 'Renewable energy consumption (% of total final energy consumption)',
 'SF6 gas emissions (thousand metric tons of CO2 equivalent)',
 'Terrestrial and marine protected areas (% of total territorial area)',
 'Terrestrial protected areas (% of total land area)',
 "Women's share of population ages 15+ living with HIV (%)",
 'Completeness of death registration with cause-of-death information (%)',
 'Share of women in wage employment in the nonagricultural sector (% of total nonagricultural employment)',
 'HFC gas emissions (thousand metric tons of CO2 equivalent)',
 'Household final consumption expenditure, PPP (constant 2011 international $)',
 'Net official aid received (constant 2012 US$)',
 'Net official aid received (current US$)',
 'PFC gas emissions (thousand metric tons of CO2 equivalent)',
 'GNI per capita, PPP (constant 2011 international $)',
 'GNI, PPP (constant 2011 international $)',
 'Incidence of tuberculosis (per 100,000 people)',
 'Tuberculosis case detection rate (%, all forms)',
 'Unmet need for contraception (% of married women ages 15-49)',
 'Community health workers (per 1,000 people)',
 'Investment in transport with private participation (current US$)',
 'Number of neonatal deaths',
 'Nurses and midwives (per 1,000 people)',
 'Cash surplus/deficit (% of GDP)',
 'Expense (% of GDP)',
 'Military expenditure (% of central government expenditure)',
 'Revenue, excluding grants (% of GDP)',
 'Tax revenue (% of GDP)',
 'Grants and other revenue (% of revenue)',
 'Other taxes (% of revenue)',
 'Taxes on goods and services (% of revenue)',
 'Taxes on income, profits and capital gains (% of revenue)',
 'Taxes on international trade (% of revenue)',
 'Interest payments (% of revenue)',
 'Central government debt, total (% of GDP)',
 'Interest payments (% of expense)',
 'Net incurrence of liabilities, domestic (% of GDP)',
 'Net incurrence of liabilities, foreign (% of GDP)',
 'Children (0-14) living with HIV',
 'Adjusted net savings, including particulate emission damage (current US$)',
 'Adjusted savings: particulate emission damage (current US$)',
 'Agricultural methane emissions (thousand metric tons of CO2 equivalent)',
 'Agricultural nitrous oxide emissions (thousand metric tons of CO2 equivalent)',
 'Industrial nitrous oxide emissions (thousand metric tons of CO2 equivalent)',
 'PPP conversion factor, GDP (LCU per international $)',
 'Price level ratio of PPP conversion factor (GDP) to market exchange rate',
 'PPP conversion factor, private consumption (LCU per international $)',
 'GHG net emissions/removals by LUCF (Mt of CO2 equivalent)',
 'Investment in telecoms with private participation (current US$)',
 'S&P Global Equity Indices (annual % change)',
 'Cash surplus/deficit (current LCU)',
 'Central government debt, total (current LCU)',
 'Compensation of employees (% of expense)',
 'Compensation of employees (current LCU)',
 'Customs and other import duties (% of tax revenue)',
 'Customs and other import duties (current LCU)',
 'Expense (current LCU)',
 'Goods and services expense (% of expense)',
 'Goods and services expense (current LCU)',
 'Grants and other revenue (current LCU)',
 'Interest payments (current LCU)',
 'Net incurrence of liabilities, domestic (current LCU)',
 'Net incurrence of liabilities, foreign (current LCU)',
 'Net ODA received (% of central government expense)',
 'Other expense (% of expense)',
 'Other expense (current LCU)',
 'Other taxes (current LCU)',
 'Revenue, excluding grants (current LCU)',
 'Subsidies and other transfers (% of expense)',
 'Subsidies and other transfers (current LCU)',
 'Tax revenue (current LCU)',
 'Taxes on exports (% of tax revenue)',
 'Taxes on exports (current LCU)',
 'Taxes on goods and services (% value added of industry and services)',
 'Taxes on goods and services (current LCU)',
 'Taxes on international trade (current LCU)',
 'Taxes on income, profits and capital gains (% of total taxes)',
 'Taxes on income, profits and capital gains (current LCU)',
 'Social contributions (% of revenue)',
 'Social contributions (current LCU)',
 'Investment in energy with private participation (current US$)',
 'ARI treatment (% of children under 5 taken to a health provider)',
 'Diarrhea treatment (% of children under 5 receiving oral rehydration and continued feeding)',
 'Female headed households (% of households with a female head)',
 'Low-birthweight babies (% of births)',
 'Employment to population ratio, 15+, female (%) (modeled ILO estimate)',
 'Employment to population ratio, 15+, male (%) (modeled ILO estimate)',
 'Employment to population ratio, 15+, total (%) (modeled ILO estimate)',
 'Employment to population ratio, ages 15-24, female (%) (modeled ILO estimate)',
 'Employment to population ratio, ages 15-24, male (%) (modeled ILO estimate)',
 'Employment to population ratio, ages 15-24, total (%) (modeled ILO estimate)',
 'Prevalence of undernourishment (% of population)',
 'Unemployment, female (% of female labor force)',
 'Unemployment, male (% of male labor force)',
 'Unemployment, total (% of total labor force)',
 ...]
```

In [67]:

```
data[data['CountryName'].str.contains('India')]
```

Out[67]:

	CountryName	CountryCode	IndicatorName	IndicatorCode	Year	Value
11561	India	IND	Adolescent fertility rate (births per 1,000 wo...	SP.ADO.TFRT	1960	103.938000
11562	India	IND	Age dependency ratio (% of working-age populat...	SP.POP.DPND	1960	76.559538
11563	India	IND	Age dependency ratio, old (% of working-age po...	SP.POP.DPND.OL	1960	5.403730
11564	India	IND	Age dependency ratio, young (% of working-age ...	SP.POP.DPND.YG	1960	71.155808
11565	India	IND	Agriculture, value added (% of GDP)	NV.AGR.TOTL.ZS	1960	42.561131
...	...	...	...	...	...	...
5648838	India	IND	Time required to register property (days)	IC.PRP.DURS	2015	47.000000
5648839	India	IND	Time required to start a business (days)	IC.REG.DURS	2015	29.000000
5648840	India	IND	Time to prepare and pay taxes (hours)	IC.TAX.DURS	2015	243.000000
5648841	India	IND	Time to resolve insolvency (years)	IC.ISV.DURS	2015	4.300000
5648842	India	IND	Total tax rate (% of commercial profits)	IC.TAX.TOTL.CP.ZS	2015	60.600000

35721 rows × 6 columns

In [83]:

```
Men_indicator = 'Life expectancy at birth, male \(years'
Women_indicator = 'Life expectancy at birth, female \(years'
Total_indicator = 'Life expectancy at birth, total \(years'

Country = 'India'

Indicator_mask_men = data['IndicatorName'].str.contains(Men_indicator)
Indicator_mask_women = data['IndicatorName'].str.contains(Women_indicator)
Indicator_mask_total = data['IndicatorName'].str.contains(Total_indicator)

Country_mask = data['CountryName'].str.contains(Country)

Filtered_data_men = data[Indicator_mask_men & Country_mask]
Filtered_data_women = data[Indicator_mask_women & Country_mask]
Filtered_data_total = data[Indicator_mask_total & Country_mask]

Filtered_data_total.head()
```

Out[83]:

	CountryName	CountryCode	IndicatorName	IndicatorCode	Year	Value
11684	India	IND	Life expectancy at birth, total (years)	SP.DYN.LE00.IN	1960	41.171951
36635	India	IND	Life expectancy at birth, total (years)	SP.DYN.LE00.IN	1961	41.790488
64177	India	IND	Life expectancy at birth, total (years)	SP.DYN.LE00.IN	1962	42.417415
92622	India	IND	Life expectancy at birth, total (years)	SP.DYN.LE00.IN	1963	43.052732
121419	India	IND	Life expectancy at birth, total (years)	SP.DYN.LE00.IN	1964	43.698415

In [97]:

```
Filtered_data_total.shape
```

Out[97]:

```
(54, 6)
```

In [98]:

```
Filtered_years = Filtered_data_total['Year'].unique().tolist()
len(Filtered_years)
```

Out[98]:

```
54
```

In [99]:

```
print('The Filtered data is From' , min(Filtered_years) , 'To' , max(Filtered_years) )
```

```
The Filtered data is From 1960 To 2013
```

In [100...

```
Filtered_values = Filtered_data_total['Value'].unique().tolist()
len(Filtered_values)
```

Out[100...

```
54
```

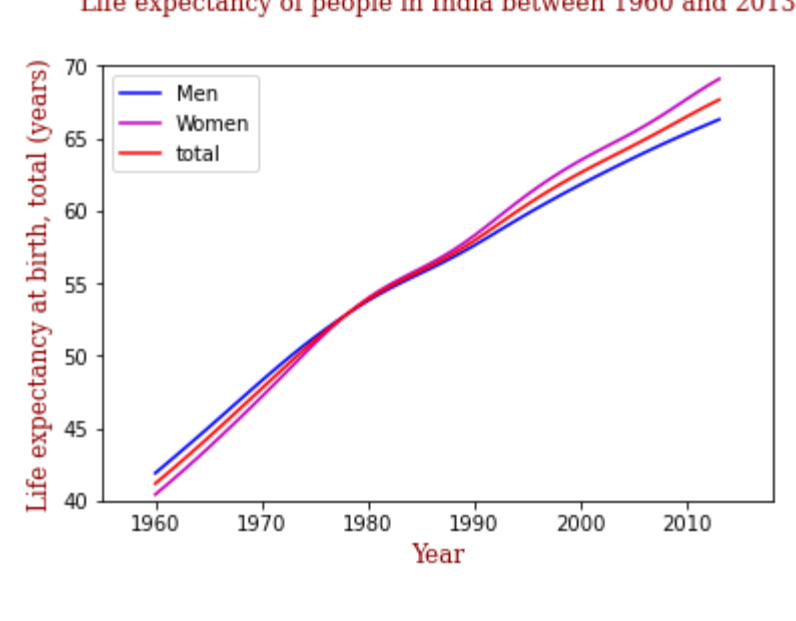
In [102...

```
print('The Filtered values are From' , min(Filtered_values) , 'To' , max(Filtered_values) )
```

```
The Filtered values are From 41.1719512195122 To 67.6604146341463
```

In [133...

```
Years = Filtered_data_total['Year'].values

Life_expectancy_men = Filtered_data_men['Value'].values
Life_expectancy_women = Filtered_data_women['Value'].values
Life_expectancy_total = Filtered_data_total['Value'].values


font = {'family': 'serif',
        'color':  'darkred',
        'weight': 'normal',
        'size': 12,
        }

plt.plot(Years, Life_expectancy_men, label = 'Men', color = 'b')
plt.plot(Years, Life_expectancy_women,label = 'Women', color = 'm')
plt.plot(Years, Life_expectancy_total, label = 'total', color = 'r')

plt.xlabel('Year' ,  fontdict = font )
plt.ylabel(Filtered_data_total['IndicatorName'].iloc[0] ,  fontdict=font)

plt.title('Life expectancy of people in India between 1960 and 2013' , x = 0.5 , y = 1.1 ,  fontdict=font)

plt.axis([1955, 2018,40,70])

plt.legend()
plt.show()
```



1. Looks like the women were the ones with least life expectancy at the beginning of the data, but eventually had the most life expectancy by 2013

2. Inversely, men had higher life expetancy and the trend was reversed by having lower life expectancy than women by 2013

3. 1980 was the time when both women and men were living to the same age (No more lonely elders! ;))

In [138...

```
Men_emp_indicator = 'Employment to population ratio, ages 15-24, male'
Women_emp_indicator = 'Employment to population ratio, ages 15-24, female'

Emp_indicator_mask_men = data['IndicatorName'].str.contains(Men_emp_indicator)
Emp_indicator_mask_women = data['IndicatorName'].str.contains(Women_emp_indicator)

Filtered_emp_data_men = data[Emp_indicator_mask_men & Country_mask]
Filtered_emp_data_women = data[Emp_indicator_mask_women & Country_mask]

Filtered_emp_data_women.head()
```

Out[138...

	CountryName	CountryCode	IndicatorName	IndicatorCode	Year	Value
2054814	India	IND	Employment to population ratio, ages 15-24, fe...	SL.EMP.1524.SP.FE.ZS	1991	27.000000
2172930	India	IND	Employment to population ratio, ages 15-24, fe...	SL.EMP.1524.SP.FE.ZS	1992	26.700001
2295904	India	IND	Employment to population ratio, ages 15-24, fe...	SL.EMP.1524.SP.FE.ZS	1993	26.400000
2420692	India	IND	Employment to population ratio, ages 15-24, fe...	SL.EMP.1524.SP.FE.ZS	1994	27.200001
2550554	India	IND	Employment to population ratio, ages 15-24, fe...	SL.EMP.1524.SP.FE.ZS	1995	26.500000

In [143...

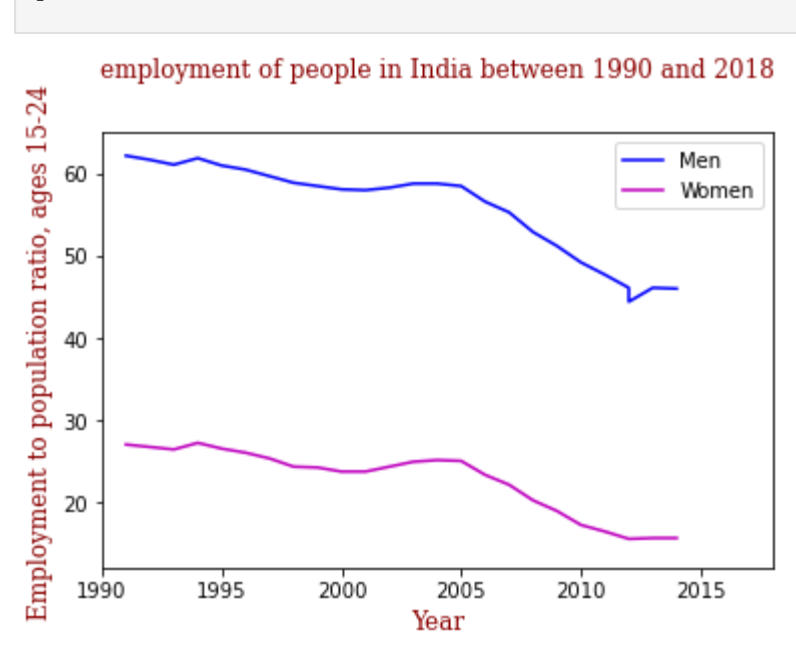
```
Filtered_emp_data_women['Year'].unique
```

Out[143...

```
<bound method Series.unique of 2054814    1991
2172930    1992
2295904    1993
2420692    1994
2550554    1995
2685578    1996
2821009    1997
2957994    1998
3097276    1999
3245630    2000
3397970    2001
3550595    2002
3704827    2003
3860195    2004
4029702    2005
4208244    2006
4387098    2007
4567413    2008
4746539    2009
4929353    2010
5111156    2011
5286839    2012
5286840    2012
5453040    2013
5585256    2014
Name: Year, dtype: int64>
```

In [156...

```
Years = Filtered_emp_data_women['Year']

Emp_men = Filtered_emp_data_men['Value'].values
Emp_women = Filtered_emp_data_women['Value'].values


font = {'family': 'serif',
        'color':  'darkred',
        'weight': 'normal',
        'size': 12,
        }

plt.plot(Years, Emp_men, label = 'Men', color = 'b')
plt.plot(Years, Emp_women,label = 'Women', color = 'm')

plt.xlabel('Year' ,  fontdict = font )
plt.ylabel('Employment to population ratio, ages 15-24' ,  fontdict=font)

plt.title('employment of people in India between 1990 and 2018' , x = 0.5 , y = 1.1 ,  fontdict=font)

plt.axis([1990, 2018,12,65])

plt.legend()
plt.show()
```



1. Contrary to the popular belief that the employment rates of people around the world have improved in the 21st century, the chart above states otherwise. The employment ratio of both men and women has declined simultaneously between 1990 and 2015.

2. This could also mean that more and more young people aged between 15 and 24 went to school, which is a good sign.

3. If we looked at the data of employment of people aged 25-35 alongside their income levels between 1990 and 2015, we might be able to understand if there was a positive impact of lower employment between age 15-24 on their overall quality of life.

4. We wanted to understand the reason for improvement in life expectancy in women between 1969 and 2010 and decline of the same for men, in relation to the employment levels.

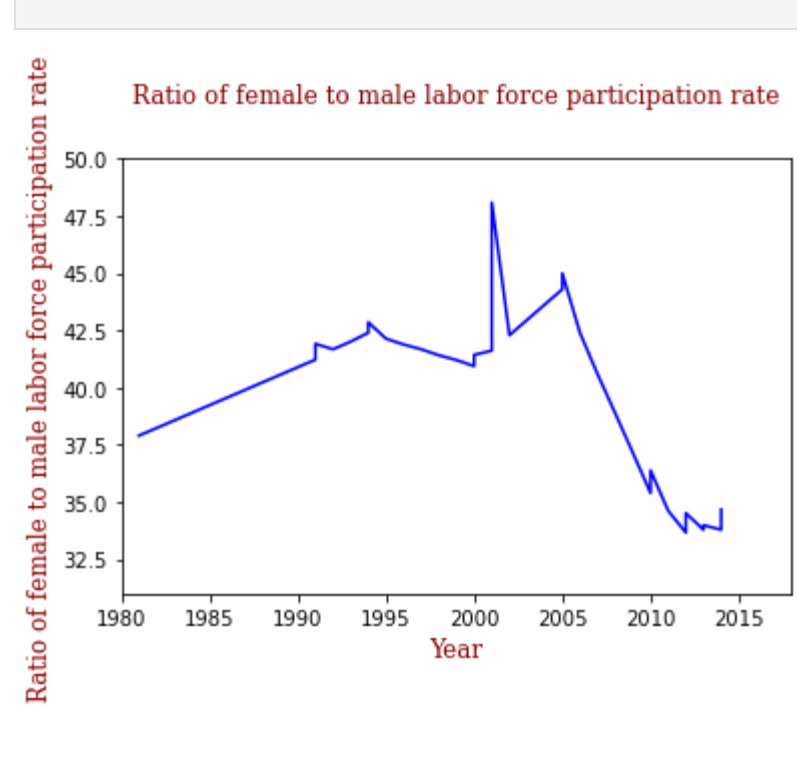
Unfortunately due to lack of data from 1960 to 1990 (especially during 1980's when the life expectancy trend has reversed), we were not able to find any correlation between the employment levels and life expectancy.

In [169...

```
labor_ratio_indicator = 'Ratio of female to male labor force participation rate'
labor_ratio_indicator_mask = data['IndicatorName'].str.contains(labor_ratio_indicator)

Filtered_labor_data = data[labor_ratio_indicator_mask & Country_mask]
```

In [170...

```
Years = Filtered_labor_data['Year']

labor_data = Filtered_labor_data['Value'].values

font = {'family': 'serif',
        'color':  'darkred',
        'weight': 'normal',
        'size': 12,
        }

plt.plot(Years, labor_data, color = 'b')

plt.xlabel('Year' ,  fontdict = font )
plt.ylabel('Ratio of female to male labor force participation rate' ,  fontdict=font)

plt.title('Ratio of female to male labor force participation rate' , x = 0.5 , y = 1.1 ,  fontdict=font)

plt.axis([1960, 2018,31,50])

plt.show()
```



Labor force participation rate is the proportion of the population ages 15 and older that is economically active: all people who supply labor for the production of goods and services during a specified period. Ratio of female to male labor force participation rate is calculated by dividing female labor force participation rate by male labor force participation rate and multiplying by 100.

We are mainly interested in the data / trend between 1970 and 1990 to be able to understand the decline men's life expectancy. As we see from the graph above, the number women working during the time period only went up. Though we can see a decline in this ratio around 1995 and there was a sudden spike of the ratio between 2000 and 2005, this did not translate into any meaningful trend on the life expectancy chart.

2005 seem to be a direction changing year to this ratio. We can see a sudden and continuous decline in the ratio, which again did not show any effect on the life expectancy.

In [ ]: