# GIproject

July 13, 2020

[1]:
```python
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

[2]:
```python
data = pd.read_csv('Indicators.csv')
```

## 1 Getting a feel for the data

[3]:
```python
data.head()
```

[3]:
```
  CountryName CountryCode                                      IndicatorName  \
0  Arab World         ARB  Adolescent fertility rate (births per 1,000 wo…
1  Arab World         ARB  Age dependency ratio (% of working-age populat…
2  Arab World         ARB  Age dependency ratio, old (% of working-age po…
3  Arab World         ARB  Age dependency ratio, young (% of working-age …
4  Arab World         ARB        Arms exports (SIPRI trend indicator values)

     IndicatorCode  Year         Value
0      SP.ADO.TFRT  1960  1.335609e+02
1      SP.POP.DPND  1960  8.779760e+01
2   SP.POP.DPND.OL  1960  6.634579e+00
3   SP.POP.DPND.YG  1960  8.102333e+01
4   MS.MIL.XPRT.KD  1960  3.000000e+06
```

[4]:
```python
data.describe()
```

[4]:
```
               Year         Value
count  5.656458e+06  5.656458e+06
mean   1.994464e+03  1.070501e+12
std    1.387895e+01  4.842469e+13
min    1.960000e+03 -9.824821e+15
25%    1.984000e+03  5.566242e+00
50%    1.997000e+03  6.357450e+01
75%    2.006000e+03  1.346722e+07
max    2.015000e+03  1.103367e+16
```

[5]:
```
data.shape
```

[5]: (5656458, 6)

[6]:
```
#checking for any null values - returns False, so no need to remove any null␣
↪values
data.isnull().values.any()
```

[6]: False

## 2 First, we will investigate EU urban population vs GDP

This is over the range of data available - from 1960 to 2015

[7]:
```
urban_pop = data[data['IndicatorName'] == 'Urban population']
```

[8]:
```
urban_pop.shape
```

[8]: (13374, 6)

[9]:
```
gdp = data[data['IndicatorName'] == 'GDP per capita (constant 2005 US$)']
```

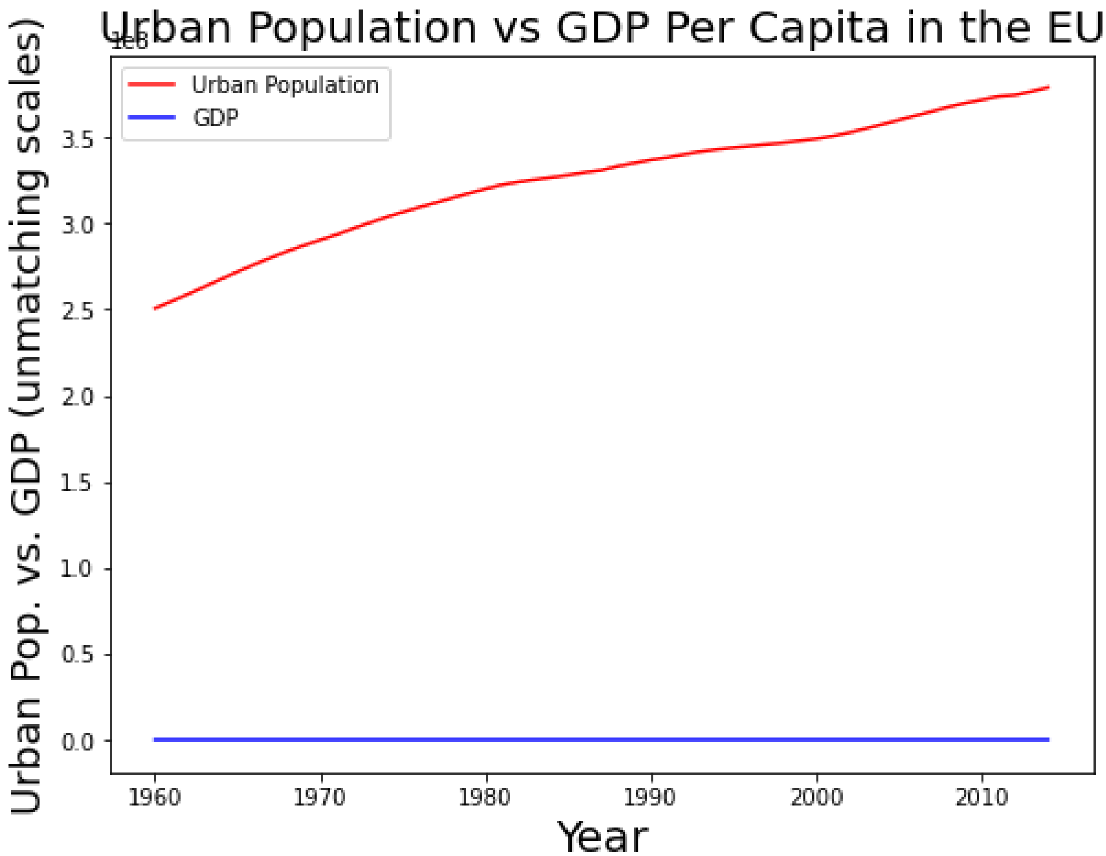
[10]:
```
gdp.shape
```

[10]: (9951, 6)

[11]:
```
eu_urban_pop = urban_pop[urban_pop['CountryName'] == 'European Union']
eu_gdp = gdp[gdp['CountryName'] == 'European Union']
```

## 3 Plotting the EU urban population data vs GDP

[12]:
```
plt.figure(figsize=(8, 6))

plt.plot(eu_urban_pop['Year'], eu_urban_pop['Value'], color= 'Red')
plt.plot(eu_gdp['Year'], eu_gdp['Value'], color = 'Blue')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('Urban Pop. vs. GDP (unmatching scales)', fontsize = 18)
plt.legend(['Urban Population', 'GDP'])
plt.title('Urban Population vs GDP Per Capita in the EU', fontsize = 20)

plt.show()
```




So, we cannot get any useful insight as these data are on very different scales.

The data will have to be normalised to better show how each has changed in the specified time

# 4 Creating a normalisation function

Creating a function that normalises the input data to be scaled between 0 and 1

The function uses the column values to be normalised, and the minimum and maximum values for that column

It can be used further in the analysis for any column

[13]:
```
def feature_normaliser(a, min, max):

    normalised_features = []

    for record in a:

        feature_normed = (record - min) / (max - min)

```

```
        normalised_features.append(feature_normed)

    return normalised_features
```

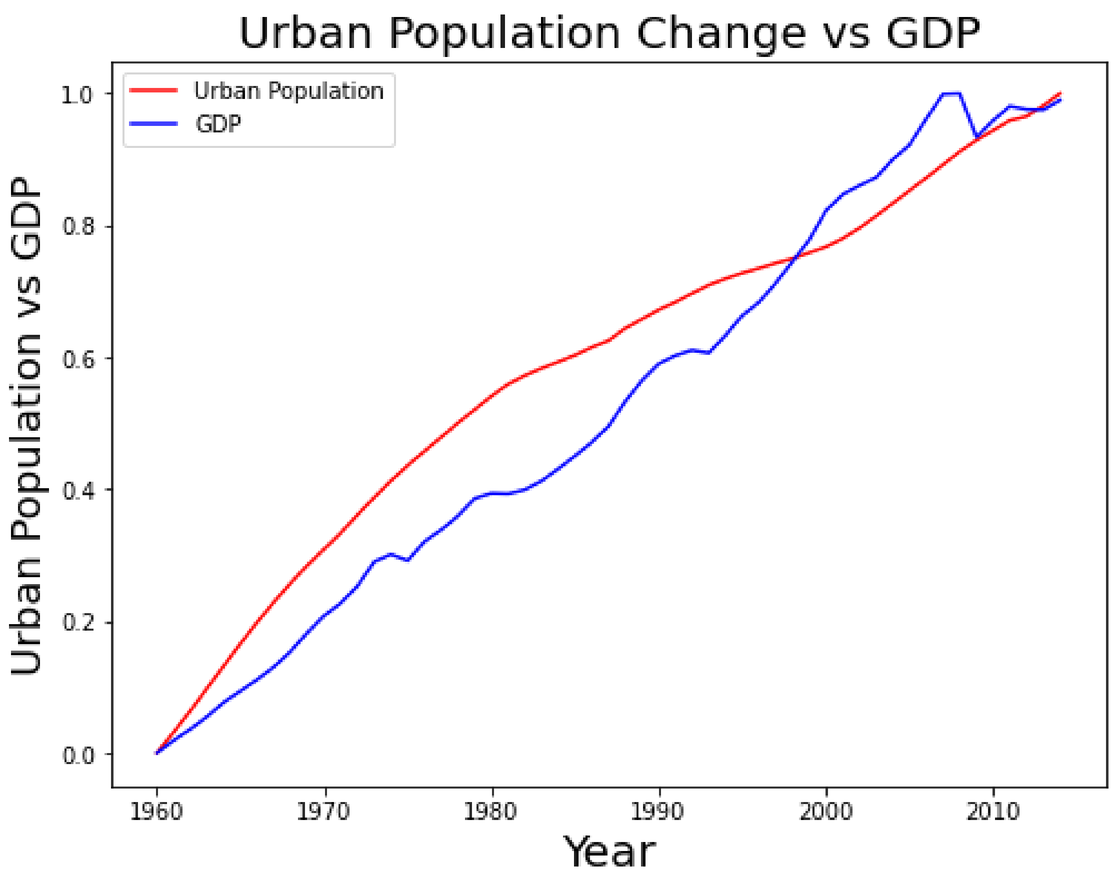
[14]:
```
eu_urb_min = eu_urban_pop['Value'].min()
eu_urb_max = eu_urban_pop['Value'].max()

eu_gdp_min = eu_gdp['Value'].min()
eu_gdp_max = eu_gdp['Value'].max()

print('The change in EU urban population is %s' % (eu_urb_max - eu_urb_min))
```

```
The change in EU urban population is 128373991.0
```

[15]:
```
# normalising the urban population and GDP data using the function above:

normed_pop = feature_normaliser(eu_urban_pop['Value'], eu_urb_min, eu_urb_max)

normed_gdp = feature_normaliser(eu_gdp['Value'], eu_gdp_min, eu_gdp_max)
```

## 5 Plotting the normalised data

[16]:
```
plt.figure(figsize=(8, 6))

plt.plot(eu_urban_pop['Year'], normed_pop, color = 'Red')
plt.plot(eu_gdp['Year'], normed_gdp, color = 'Blue')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('Urban Population vs GDP', fontsize = 18)
plt.legend(['Urban Population', 'GDP'])
plt.title('Urban Population Change vs GDP', fontsize = 20)

plt.show()
```




So clearly there has been a significant increase in the proportion of people living in urban areas, as well as a significant upward trend in GDP.

To examine this relative to rural areas, we will plot the data for rural populations in the EU against the data above

## 6 Including the rural population data in the plot

[17]:
```
# finding the rural population data

rural_pop = data[data['IndicatorName'] == 'Rural population']

eu_rural_pop = rural_pop[rural_pop['CountryName'] == 'European Union']
eu_rural_pop.head()
```

[17]:
```
          CountryName CountryCode      IndicatorName IndicatorCode  Year  \
896    European Union         EUU  Rural population   SP.RUR.TOTL  1960
24212  European Union         EUU  Rural population   SP.RUR.TOTL  1961
50889  European Union         EUU  Rural population   SP.RUR.TOTL  1962
```

```
79323    European Union          EUU   Rural population   SP.RUR.TOTL  1963
107957   European Union          EUU   Rural population   SP.RUR.TOTL  1964

               Value
896      158832585.0
24212    158290687.0
50889    157809721.0
79323    157145688.0
107957   156441716.0
```

[18]:
```
# finding the maximum and minimum values for rural population so we can␣
↪normalise the data:

eu_rural_min = eu_rural_pop['Value'].min()
eu_rural_max = eu_rural_pop['Value'].max()

# use the normalisation function to scale the values between 0 and 1

eu_rural_normed = feature_normaliser(eu_rural_pop['Value'], eu_rural_min,␣
↪eu_rural_max)

print('The change in EU rural population is %s' % (eu_rural_max - eu_rural_min))

print('The difference between EU urban and rural populations in 2015 is %s' %␣
↪(eu_urb_max - eu_rural_max))
```

```
The change in EU rural population is 29557934.0
The difference between EU urban and rural populations in 2015 is 220207283.0
```

# 7 Plotting EU Urban, Rural, and GDP data

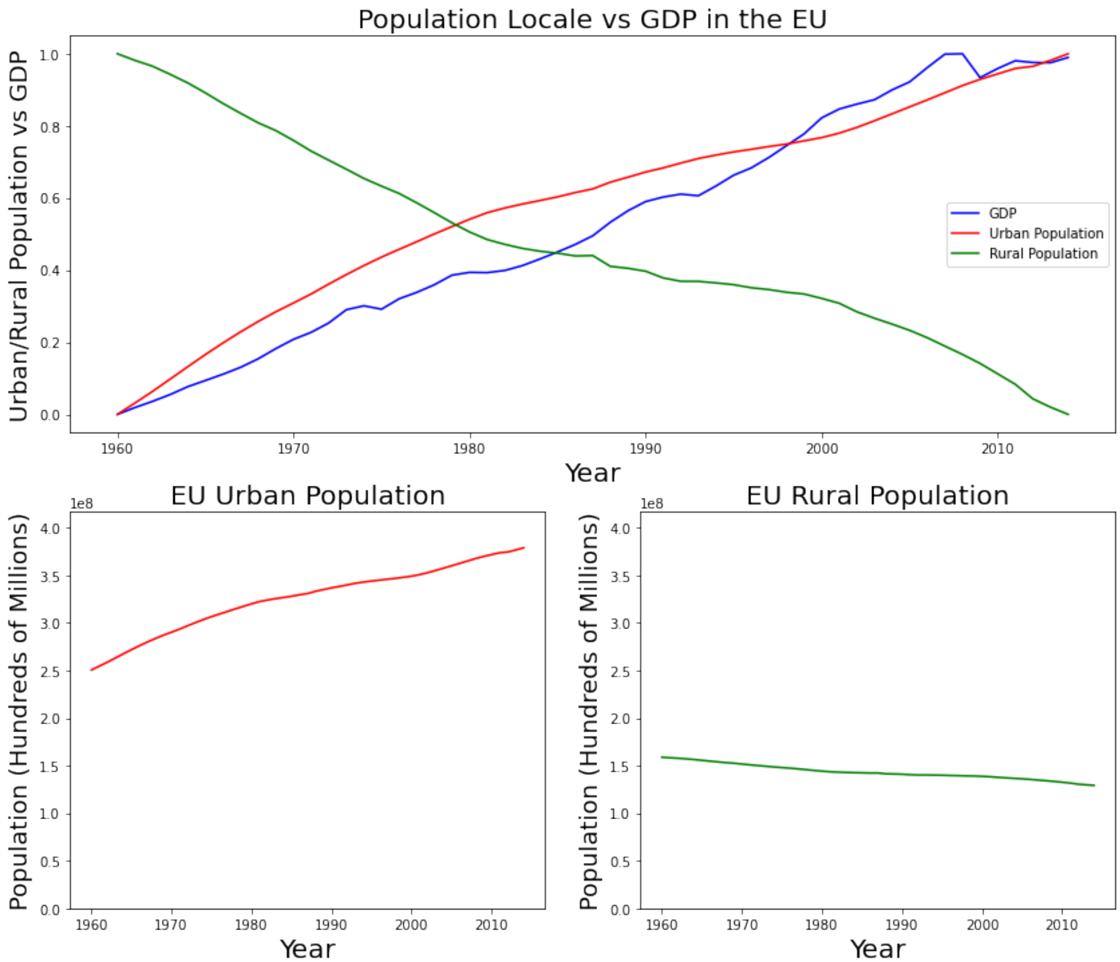
We will plot the normalised data for urban population, rural population, and GDP

This will give insight into how much each changes relative to one another in this time frame

We will also plot the non-normalised population data to see how it is changing on the whole

[19]:
```
# creating a figure to contain the three plots

plt.figure(figsize=(14, 12))

# creating the normalised plot with rural, urban, and GDP data

plt.subplot(2, 1, 1)

plt.plot(eu_gdp['Year'],
         normed_gdp,
         color = 'Blue')
```

```
plt.plot(eu_urban_pop['Year'],
         normed_pop,
         color = 'Red')

plt.plot(eu_rural_pop['Year'],
         eu_rural_normed,
         color = 'Green')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('Urban/Rural Population vs GDP', fontsize = 18)
plt.legend(['GDP', 'Urban Population', 'Rural Population'])
plt.title('Population Locale vs GDP in the EU', fontsize = 20)

# creating the non-normalised plot for urban population

plt.subplot(2, 2, 3, ylim = (0, (eu_urb_max)*1.1)) # the y-scale will be the␣
↪same for urban and rural pops

plt.plot(eu_urban_pop['Year'], eu_urban_pop['Value'], color = 'Red')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('Population (Hundreds of Millions)', fontsize = 18)
plt.title('EU Urban Population', fontsize = 20)

# creating the non-normalised plot for rural population

plt.subplot(2, 2, 4, ylim = (0, (eu_urb_max) * 1.1))

plt.plot(eu_rural_pop['Year'], eu_rural_pop['Value'], color = 'Green')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('Population (Hundreds of Millions)', fontsize = 18)
plt.title('EU Rural Population', fontsize = 20)

plt.show()
```




[20]:
```
# Finding exactly the difference between urban population growth and rural population decline

# This will give some insight as to how much urban growth cannot be attributed to rural decline

eu_urban_change = (eu_urb_max - eu_urb_min)

eu_rural_change = (eu_rural_max - eu_rural_min)

print('The difference between urban population growth and rural population decline is %s' % (eu_urban_change - eu_rural_change))
```

```
The difference between urban population growth and rural population decline is
98816057.0
```

# 8 Investigating the plots above

So, for developed nations in the EU, increased GDP is strongly correlated with urban population growth and rural population decline

However, the change in population is greater for urban populations than for rural populations. This means that not all urban population growth can be attributed to rural population flight.

Urban populations have increased by 98,816,057 more people than the rural populations have declined.

So, it is likely immigration that accounts for much of this urban population growth.

How much of this can be attributed to natural growth i.e. births?

For a deeper analysis it would be interesting to look into the relationship between GDP, emigration, immigration, birth rates, and the proportion of workers who are immigrants. However, this is outside the scope of this anaylsis.

Would this make a strong case for immigration?

# 9 Comparison - plotting the same data for Sub-Saharan Africa

The European picture is fairly clear - more and more people are living in urban areas while fewer people are living in rural areas. This relates to steadily increasing GDP

Is this trend seen in a continent with significantly poorer countries?

Firstly, for the sake of comparison we will look at the same data over the same time period for all of Sub-Saharan Africa

[21]:
```
# selecting out data for all sub-Saharan African countries

Africa = data[data['CountryName'] == 'Sub-Saharan Africa (all income levels)']

Africa_gdp = Africa[Africa['IndicatorName'] == 'GDP per capita (constant 2005␣
↪US$)']

Africa_urb_pop = Africa[Africa['IndicatorName'] == 'Urban population']

Africa_rur_pop = Africa[Africa['IndicatorName'] == 'Rural population']
```

[22]:
```
# selecting the minimum and maximum values for normalisation

Africa_gdp_min = Africa_gdp['Value'].min()
Africa_gdp_max = Africa_gdp['Value'].max()

Africa_urb_min = Africa_urb_pop['Value'].min()
Africa_urb_max = Africa_urb_pop['Value'].max()

Africa_rur_min = Africa_rur_pop['Value'].min()
```

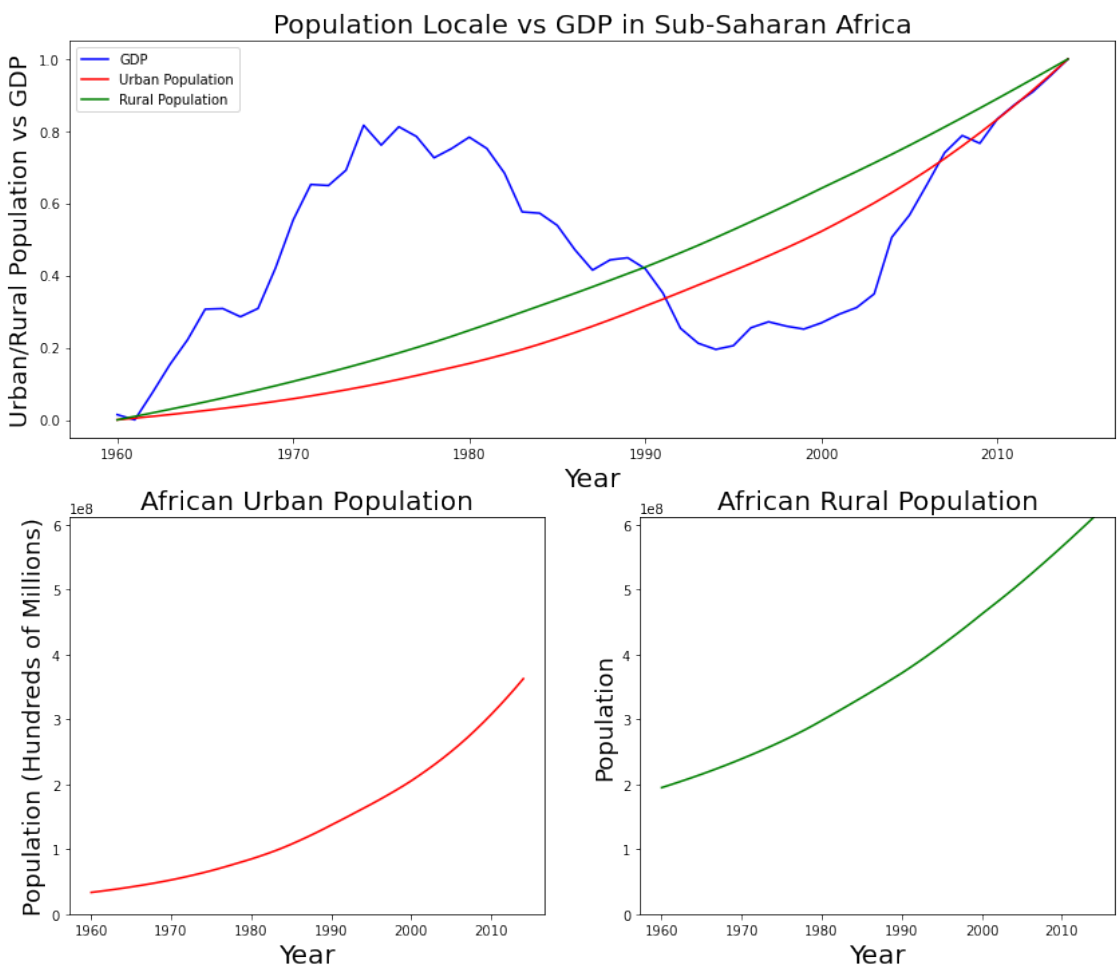
```
Africa_rur_max = Africa_rur_pop['Value'].max()
```

[23]:
```
# normalising the Africa data using the function created earlier

Africa_gdp_normed = feature_normaliser(Africa_gdp['Value'],
                                       Africa_gdp_min,
                                       Africa_gdp_max)

Africa_urb_normed = feature_normaliser(Africa_urb_pop['Value'],
                                       Africa_urb_min,
                                       Africa_urb_max)

Africa_rur_normed = feature_normaliser(Africa_rur_pop['Value'],
                                       Africa_rur_min,
                                       Africa_rur_max)
```

## 10 Plotting African Urban, Rural and GDP data

[24]:
```
# plotting the data in the same way as the EU data

plt.figure(figsize=(14, 12))

plt.subplot(2, 1, 1)

plt.plot(Africa_gdp['Year'],
         Africa_gdp_normed,
         color = 'Blue')

plt.plot(Africa_urb_pop['Year'],
         Africa_urb_normed,
         color = 'Red')

plt.plot(Africa_rur_pop['Year'],
         Africa_rur_normed,
         color = 'Green')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('Urban/Rural Population vs GDP', fontsize = 18)
plt.legend(['GDP', 'Urban Population', 'Rural Population'])
plt.title('Population Locale vs GDP in Sub-Saharan Africa', fontsize = 20)

plt.subplot(2, 2, 3, ylim = (0, Africa_rur_max))

plt.plot(Africa_urb_pop['Year'], Africa_urb_pop['Value'], color = 'Red')

plt.xlabel('Year', fontsize = 20)
```

```
plt.ylabel('Population (Hundreds of Millions)', fontsize = 18)
plt.title('African Urban Population', fontsize = 20)

plt.subplot(2, 2, 4, ylim = (0, Africa_rur_max))

plt.plot(Africa_rur_pop['Year'], Africa_rur_pop['Value'], color = 'Green')
plt.xlabel('Year', fontsize = 20)
plt.ylabel('Population', fontsize = 18)
plt.title('African Rural Population', fontsize = 20)

plt.show()
```




There are a number of features of this graphic that are interesting

Firstly, GDP was rising up until around 1975, and then for some reason the continent entered a recession that has only recently come to an end in the early 2000s. The continent is now in a growth phase.

What could be the cause(s) of this?

Also, both urban AND rural populations have been growing steadily from 1960, with marginally more growth in the rural populations.

Interestingly, more people live in rural areas than urban areas in African countries and this trend looks set to continue due to the rural population growth.

According to just these features and this graphic, however, it is not **strictly** necessary for GDP growth to be tied to urban population growth.

To investigate this, it will be useful to look at the numbers to get a sense of the **scale** of the growth

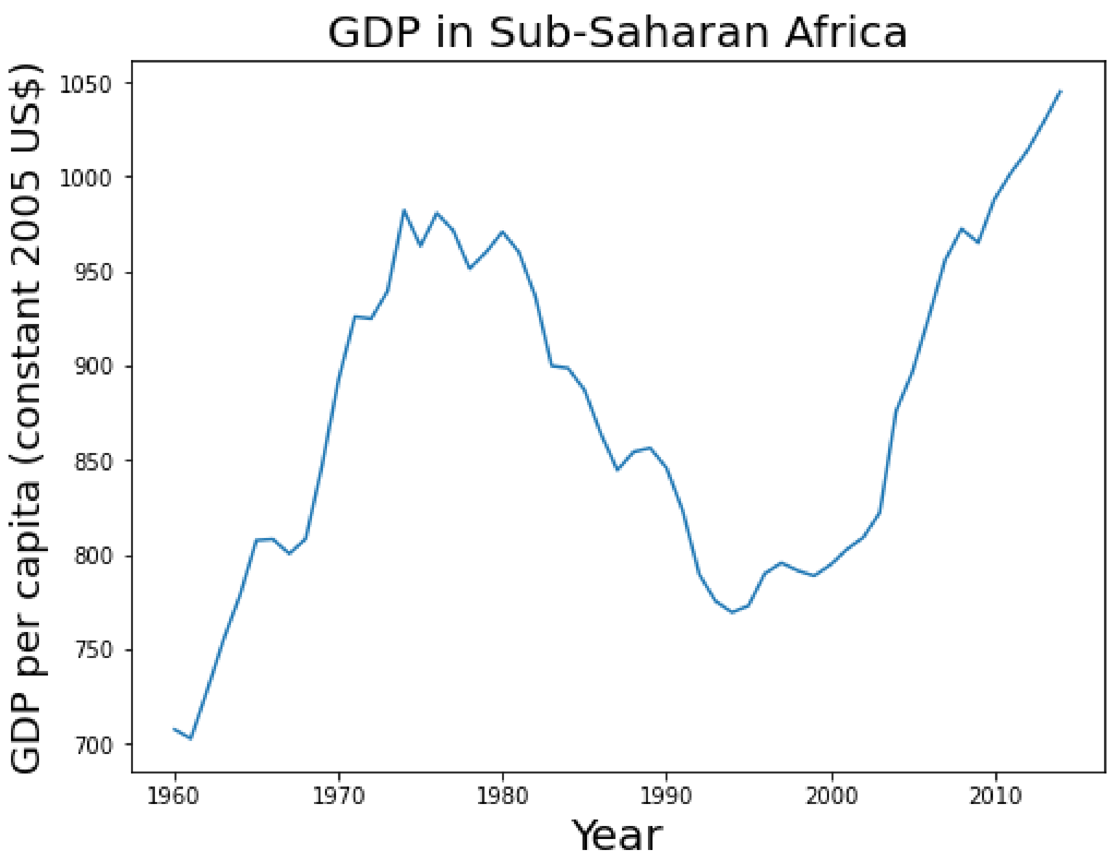
[25]:
```
x = Africa_urb_max - Africa_urb_min
y = Africa_rur_max - Africa_rur_min

difference = (y - x)
print(x, y, difference)
```

```
329355420.0 416691151.0 87335731.0
```

## 11 Investigating the magnitude of African GDP growth

To show this we will plot only the Sub-Saharan Africa GDP data without normalisation

[27]:
```
# creating a new figure that will contain just the GDP data without␣
↪normalisation

plt.figure(figsize=(8, 6))

plt.plot(Africa_gdp['Year'],
         Africa_gdp['Value'])

plt.xlabel('Year', fontsize = 20)
plt.ylabel('GDP per capita (constant 2005 US$)', fontsize = 18)
plt.title('GDP in Sub-Saharan Africa', fontsize = 20)

plt.show()
```




So, despite there seeming to be very large fluctuations in GDP (growth and decline), the absolute GDP change is in fact constrained to within around **350 constant 2005 US Dollars per capita**

[28]:
```
print('The African GDP per capita is %s' % (Africa_gdp_max))
print('The EU GDP per capita is %s' % (eu_gdp_max))
```

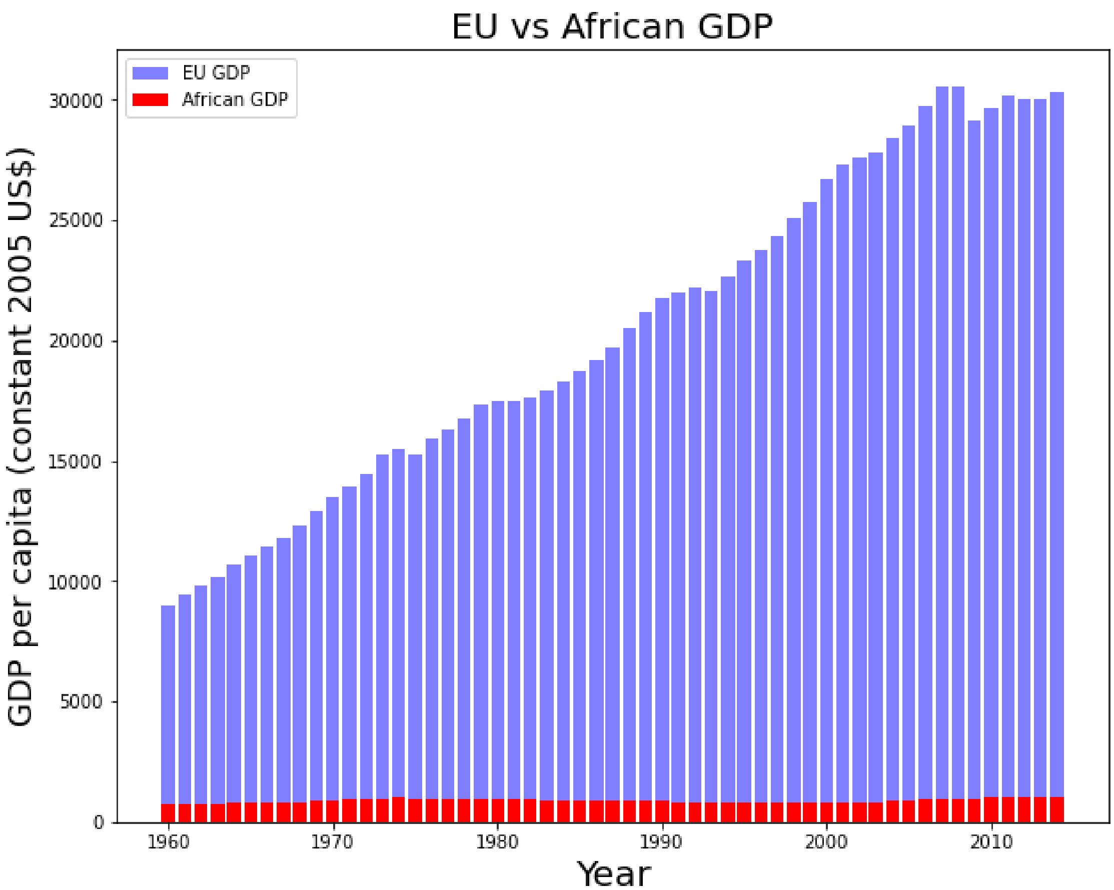
```
The African GDP per capita is 1044.87095463383
The EU GDP per capita is 30562.714986320298
```

## 12 Comparing the EU and African GDP data

This may support the idea that significant GDP growth is tied to urban population growth

[30]:
```
plt.figure(figsize=(10, 8))

plt.bar(eu_gdp['Year'], eu_gdp['Value'], color = 'Blue', alpha = 0.5)

plt.bar(Africa_gdp['Year'], Africa_gdp['Value'], color = 'Red')

plt.xlabel('Year', fontsize = 20)
```

```python
plt.ylabel('GDP per capita (constant 2005 US$)', fontsize = 18)
plt.title('EU vs African GDP', fontsize = 20)
plt.legend(['EU GDP', 'African GDP'])

plt.show()
```




As can be clearly seen, African GDP has shown no significant growth in the years in question, but the EU GDP growth has been significant and fairly steady.

So, a clear positive effect has been seen along with the population growth in urban areas. However, the fact remains that rural populations are also strongly treding downwards, and this will not come without having some kind of effect on the countries in question.

For example, what is the effect on fuel consumption? What is the effect on CO2 emmissions? What is the effect on population health?

What other effects can be seen in this time scale? GDP growth is positive, but are there any negative indicators over this time frame?

# 13 How does this relate to CO2 emissions?

With increasing GDP and increasing urban populations, what do we see with CO2 emissions?

The data includes a measure of the CO2 emissions per constant 2005 US$ of GDP, so this will be investigated first.

Ideally we will see emissions reduce per dollar of GDP over time - this would be an indicator of progress.

We will also plot this data for Sub-Saharan Africa to see how the CO2 emissions are changing per dollar of GDP over time.

Secondly, we will plot absolute CO2 emissions to see the trend over time.

[68]:
```
#CO2 emissions (kg per 2005 US$ of GDP)

co2_gdp_data = data[data['IndicatorName'] == 'CO2 emissions (kg per 2005 US$ of␣
↪GDP)']

eu_co2_gdp = co2_gdp_data[co2_gdp_data['CountryName'] == 'European Union']

africa_co2_gdp = co2_gdp_data[co2_gdp_data['CountryName'] == 'Sub-Saharan␣
↪Africa (all income levels)']

eu_co2.head()
```

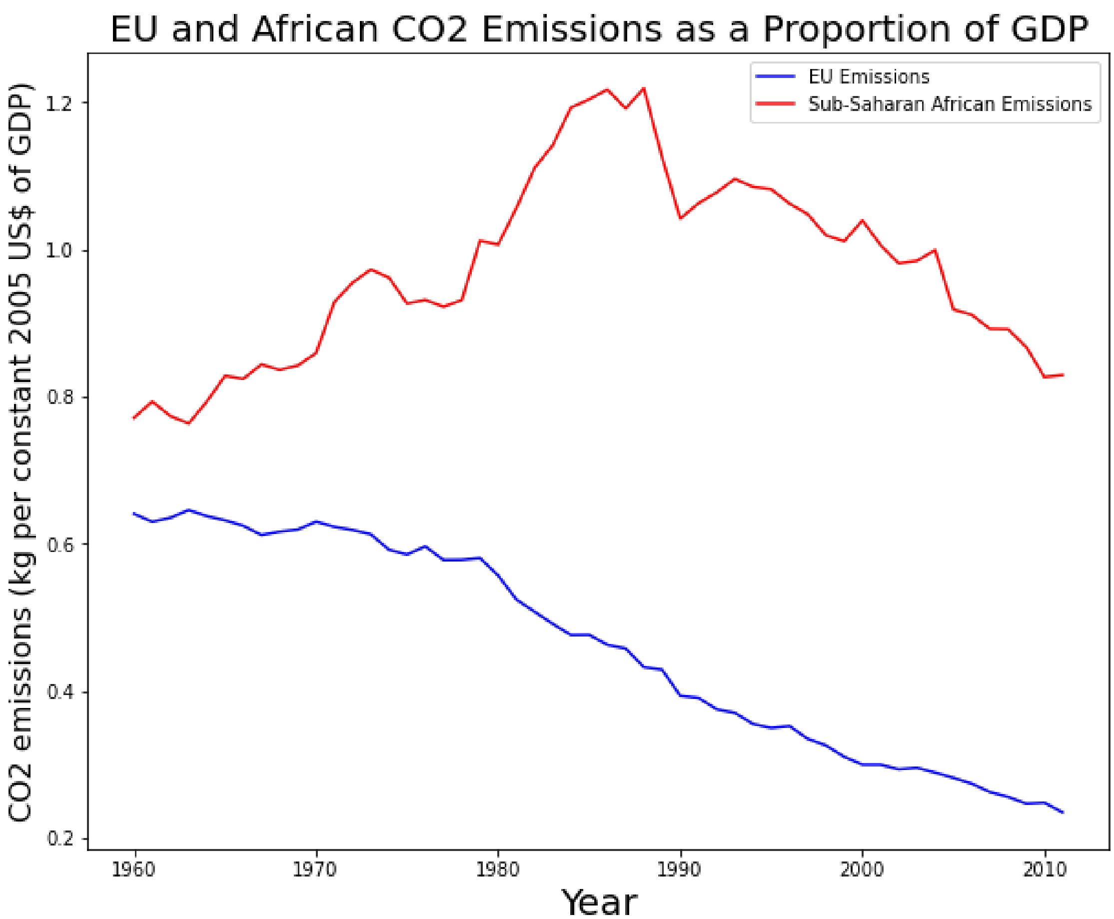
[68]:
```
          CountryName CountryCode       IndicatorName   IndicatorCode  Year  \
799     European Union         EUU  CO2 emissions (kt)  EN.ATM.CO2E.KT  1960
24108   European Union         EUU  CO2 emissions (kt)  EN.ATM.CO2E.KT  1961
50772   European Union         EUU  CO2 emissions (kt)  EN.ATM.CO2E.KT  1962
79205   European Union         EUU  CO2 emissions (kt)  EN.ATM.CO2E.KT  1963
107839  European Union         EUU  CO2 emissions (kt)  EN.ATM.CO2E.KT  1964

               Value
799     2.359870e+06
24108   2.446231e+06
50772   2.587557e+06
79205   2.763983e+06
107839  2.879926e+06
```

[69]:
```
africa_co2.head()
```

[69]:
```
                                   CountryName CountryCode  \
3022    Sub-Saharan Africa (all income levels)         SSF
26697   Sub-Saharan Africa (all income levels)         SSF
53518   Sub-Saharan Africa (all income levels)         SSF
81935   Sub-Saharan Africa (all income levels)         SSF
110602  Sub-Saharan Africa (all income levels)         SSF
```

```
              IndicatorName    IndicatorCode  Year          Value
3022    CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  124501.721969
26697   CO2 emissions (kt)  EN.ATM.CO2E.KT  1961  130220.854922
53518   CO2 emissions (kt)  EN.ATM.CO2E.KT  1962  134787.761338
81935   CO2 emissions (kt)  EN.ATM.CO2E.KT  1963  141307.998081
110602  CO2 emissions (kt)  EN.ATM.CO2E.KT  1964  155107.955734
```

[75]:
```
plt.figure(figsize=(10, 8))

plt.plot(eu_co2_gdp['Year'], eu_co2_gdp['Value'], color = 'Blue')

plt.plot(africa_co2_gdp['Year'], africa_co2_gdp['Value'], color = 'Red')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('CO2 emissions (kg per constant 2005 US$ of GDP)', fontsize = 16)
plt.title('EU and African CO2 Emissions as a Proportion of GDP', fontsize = 20)
plt.legend(['EU Emissions', 'Sub-Saharan African Emissions'])

plt.show()
```



So, for the EU there is an encouraging trend in CO2 emissions per US$ of GDP decreasing in the years between 1960 and 2015.

It is perhaps unsurprising that the CO2 emissions in Sub-Saharan African countries are higher per dollar of GDP.

## 14 Now looking at absolute CO2 emissions

[71]:
```
abs_co2 = data[data['IndicatorName'] == 'CO2 emissions (kt)']

eu_co2 = abs_co2[abs_co2['CountryName'] == 'European Union']

africa_co2 = abs_co2[abs_co2['CountryName'] == 'Sub-Saharan Africa (all income␣
↪levels)']

eu_co2.head()
```

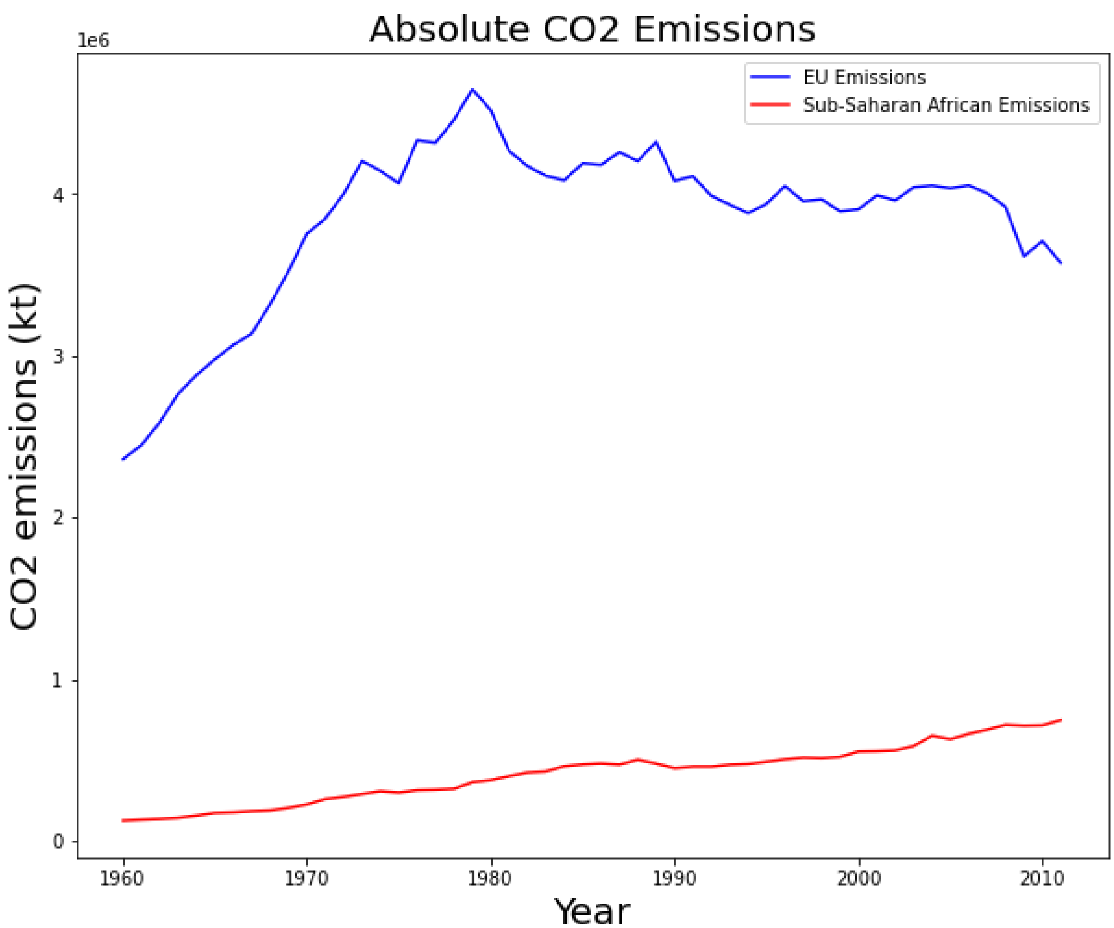
[71]:
```
            CountryName CountryCode        IndicatorName   IndicatorCode  Year  \
799      European Union         EUU   CO2 emissions (kt)  EN.ATM.CO2E.KT  1960
24108    European Union         EUU   CO2 emissions (kt)  EN.ATM.CO2E.KT  1961
50772    European Union         EUU   CO2 emissions (kt)  EN.ATM.CO2E.KT  1962
79205    European Union         EUU   CO2 emissions (kt)  EN.ATM.CO2E.KT  1963
107839   European Union         EUU   CO2 emissions (kt)  EN.ATM.CO2E.KT  1964

                Value
799      2.359870e+06
24108    2.446231e+06
50772    2.587557e+06
79205    2.763983e+06
107839   2.879926e+06
```

[72]:
```
africa_co2.head()
```

[72]:
```
                                    CountryName CountryCode  \
3022     Sub-Saharan Africa (all income levels)         SSF
26697    Sub-Saharan Africa (all income levels)         SSF
53518    Sub-Saharan Africa (all income levels)         SSF
81935    Sub-Saharan Africa (all income levels)         SSF
110602   Sub-Saharan Africa (all income levels)         SSF

              IndicatorName   IndicatorCode  Year          Value
3022     CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  124501.721969
26697    CO2 emissions (kt)  EN.ATM.CO2E.KT  1961  130220.854922
53518    CO2 emissions (kt)  EN.ATM.CO2E.KT  1962  134787.761338
81935    CO2 emissions (kt)  EN.ATM.CO2E.KT  1963  141307.998081
110602   CO2 emissions (kt)  EN.ATM.CO2E.KT  1964  155107.955734
```

[73]:

```
plt.figure(figsize=(10, 8))

plt.plot(eu_co2['Year'], eu_co2['Value'], color = 'Blue')

plt.plot(africa_co2['Year'], africa_co2['Value'], color = 'Red')

plt.xlabel('Year', fontsize = 20)
plt.ylabel('CO2 emissions (kt)', fontsize = 20)
plt.title('Absolute CO2 Emissions', fontsize = 20)
plt.legend(['EU Emissions', 'Sub-Saharan African Emissions'])

plt.show()
```




So, despite the downward trend in the EU for less CO2 emissions per dollar of GDP, the emissions are still extremely high with respect to all nations in Sub-Saharan Africa.

Will this downward slope seen in 2010 - 2015 continue as it needs to for the environment?

## 15 Final remarks

There is still much to be investigated with regard to the relationship between population change, urbanisation, GDP, and CO2 emissions. However, there are some interesting observations to be made with the data used in this investigation.

Clearly, rural populations are in decline in the EU as a whole, but this is not seen in developing countries like many of those in Sub-Saharan Africa. In fact, the opposite is true in these developing countries - rural populations are growing.

Also, EU GDP has been rising significantly and steadily, and again this is not true for Sub-Saharan Africa as a whole.

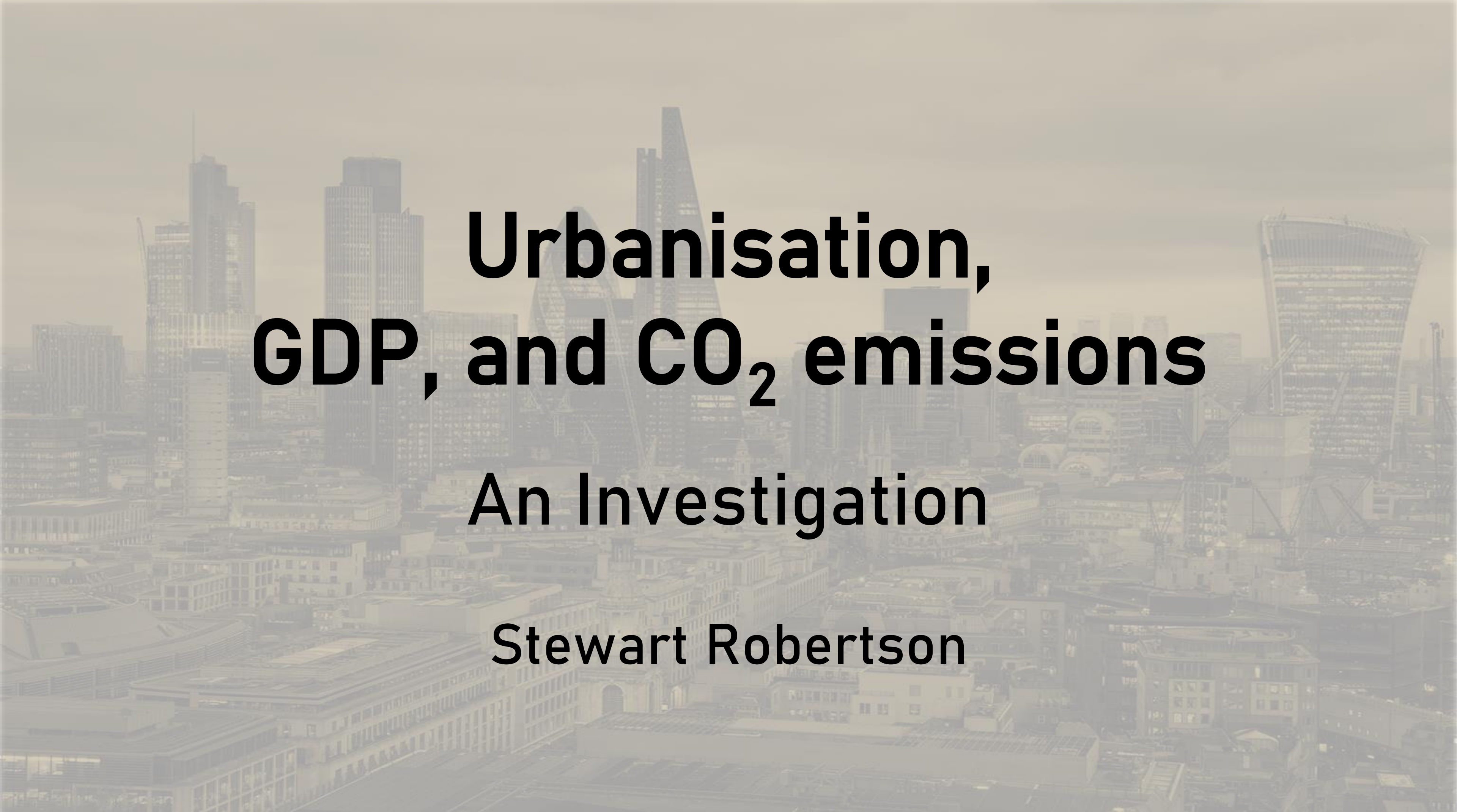
It is likely that this urbanisation and transition from more rural living to urbanised living and economic growth are mututally inclusive, at least with the way the economy has functioned since 1960.

Will this apply to Africa as the countries continue to develop? Or can a new, more digital economy allow for the maintenance of rural living?

The fact that GDP growth is outpacing CO2 emissions in the EU is encouraging. But is the decline in emissions happening quickly enough?

Should there be more global focus on integrating sustainability into developing countries such as those in Africa, so that their CO2 emissions never reach the levels reached in the EU and elsewhere?

What will become of the more rural parts of the countries in question as their populations decline? Will this happen at an increasing rate as we advance ever more technologically? Or will the reverse happen - where people can work together more remotely via technology?



# Urbanisation, GDP, and $CO_2$ emissions

## An Investigation

Stewart Robertson

# The Dataset

This project uses the “World Development Indicators” dataset, made available by the World Bank.

This is a fairly large dataset, with approximately 5.7 million records.

The data is collected to provide a measure of how the countries of the world are developing according to numerous criteria over time.

The time frame over which the data spans is 55 years – from 1960 to 2015.

# Motivation

In most developed countries, rural populations are declining while urban populations are growing.

This has many implications – the allocation of resources by government to rural and urban areas may be disproportionate, strain may be placed on the environment, cities and towns may struggle to adequately house those moving into urban areas, etc.

However, with this urbanisation also comes many positive effects. For example, improved access to facilities and services for those in urban areas, greater economic opportunity, more access to health care, etc.

Whether this population change has a net positive effect is a complicated question, but the investigation can begin with a look at its relationship with GDP and $CO_2$ emissions.

A comparison between developed countries and developing countries can add further insight – do we see common or conflicting trends?

The data can then be used to inform further reasearch and highlight points of interest. E.g. what should change in developed countries to maximise GDP growth? Can the $CO_2$ emissions projections inform government policy? Local governments could use the data and investigate which regions specifically are experiencing the most growth, and use this to aid in planning for the future.

# Research Questions

This investigation will focus on one continent of mostly highly developed countries (the EU), and contrast this with one continent of mostly developing countries (Sub-Saharan Africa).

It will attempt to answer the following questions:

- Approximately how much of the EU's urban population growth can be attributed to rural population decline?
- What is the relationship between rural and urban population change and GDP in the EU?
- What is the relationship between rural and urban population change and GDP in Sub-Saharan Africa?
- What is the relationship between rural and urban population change and $CO_2$ emissions in the EU?
- What is the relationship between rural and urban population change and $CO_2$ in Sub-Saharan Africa?

Note – all GDP data used is in constant 2005 US $

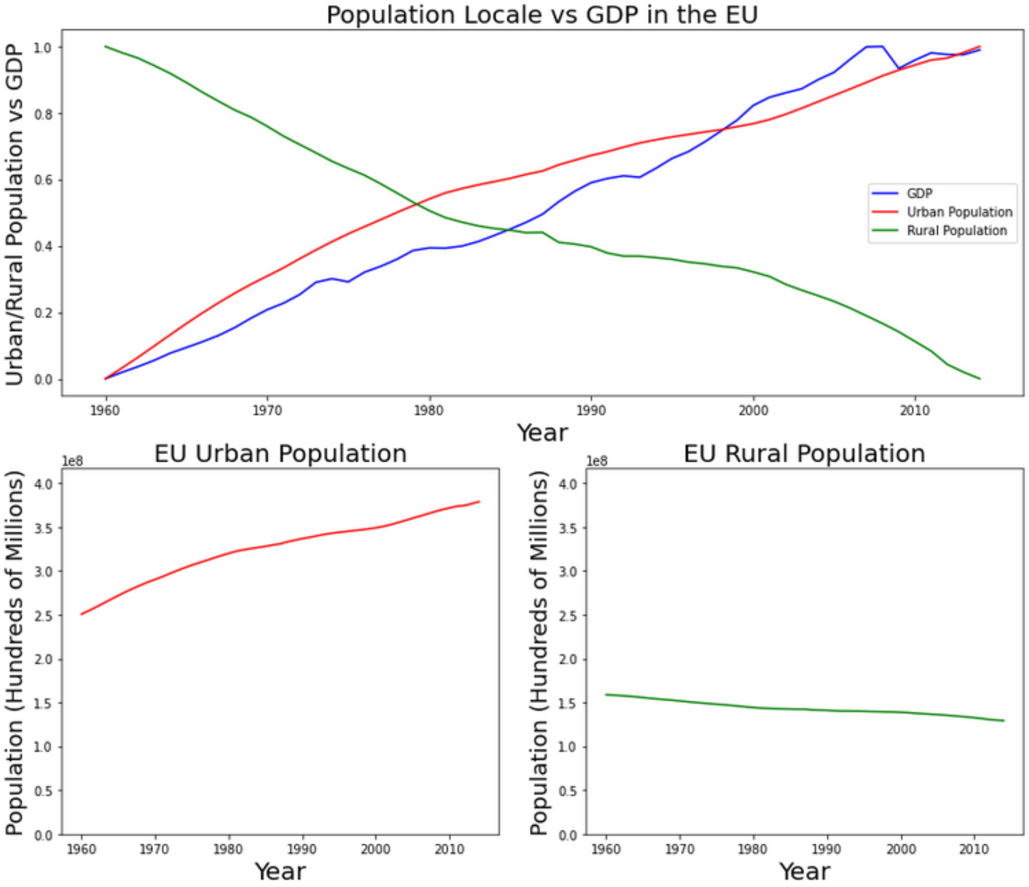
# Findings - EU

The plot at the top has been normalised to greater show the relationship between the features.

Urban populations have grown by over 128 million people, while rural populations have declined by over 29 million.

Assuming 100% of the rural decline was due to migration to urban areas, this still leaves growth of over 98 million people in urban areas unaccounted for. Some of this will be due to “natural” growth i.e. births, but much will be due to immigration. This is an interesting point for further analysis.

Clearly, GDP growth and urban population growth are strongly correlated. This relationship is complicated and bi-directional, meaning one influences the other. This is due to environmental and socioeconomic factors, culture and values, etc. The extent to which they influence the other, and which comes first, would require extensive investigation.



# Findings – Sub-Saharan Africa

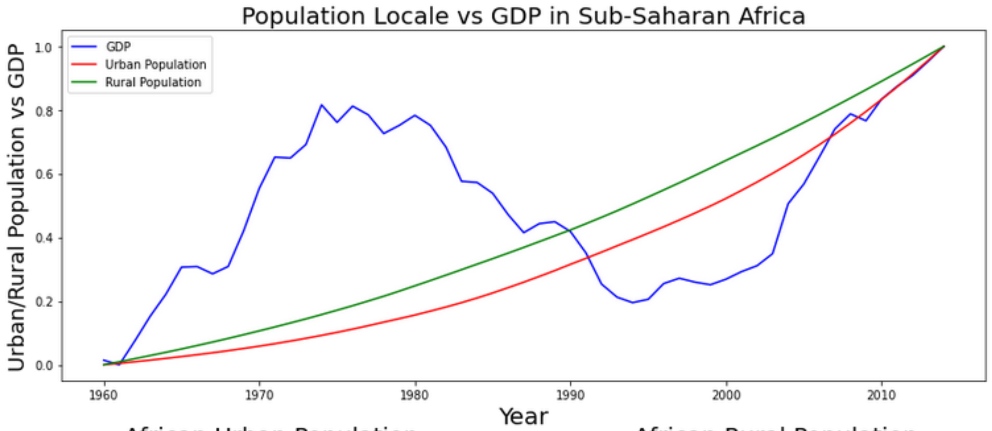
Again, the top plot is normalised to greater show the relationship between the features.

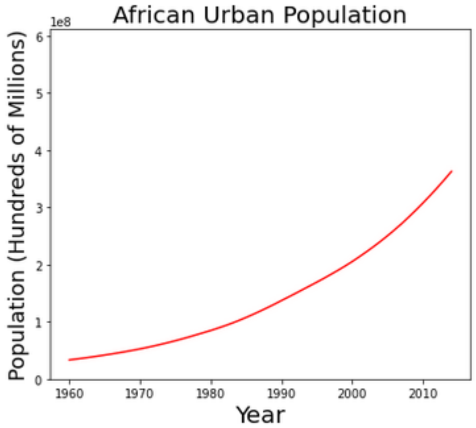
Urban populations grew by over 329 million people, while rural populations grew by over 416 million. So, this contrasts the EU, as rural populations are greater than urban populations in Africa and are growing faster too.

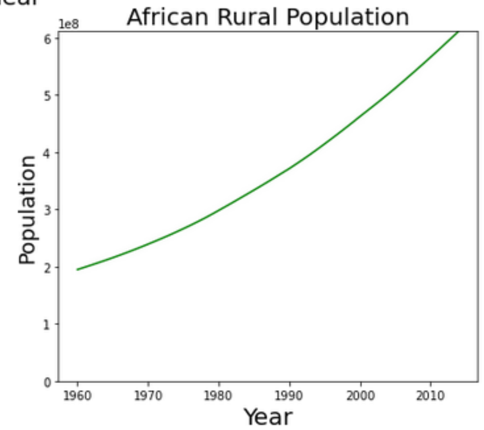
For both populations, this is immense growth in the given time frame and is interesting in itself, and highlights one of the main issues facing these countries – excessive child birth. Countries cannot provide enough high quality jobs for the growing population, and it is hindering development.

There is no correlation with population change and GDP also, which is, as with the EU,  due to socioeconomic, cultural and environmental factors.

To get a deeper understanding of the how these findings relate to each other, it will be useful to compare the magnitude of the GDP change.



# Magnitude of EU vs Sub-Saharan Africa GDP Change

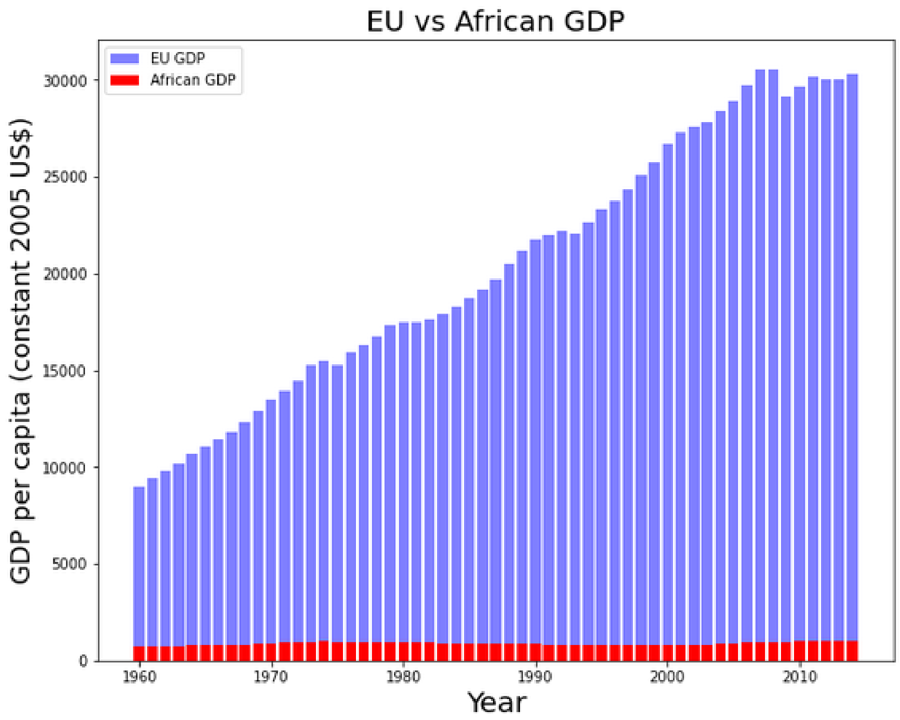
So, despite the normalised plot on the previous slide showing large fluctuations in African GDP, it is in fact constrained to within around $350 per capita.

This supports urbanisation insofar as it correlates with GDP growth.

The increase in EU GDP is significant and consistent, with some exceptions such as after the 2008 financial crisis.

This increase in GDP could offset the potential decline of rural areas which have seen their populations dwindle, as money generated in the urban environment can be invested in infrastructure and maintenance of rural areas, as well as further development of urban areas.

This also highlights the work that must be done to develop the nations in Africa if their GDP is to be in any way comparable to that of EU countries.



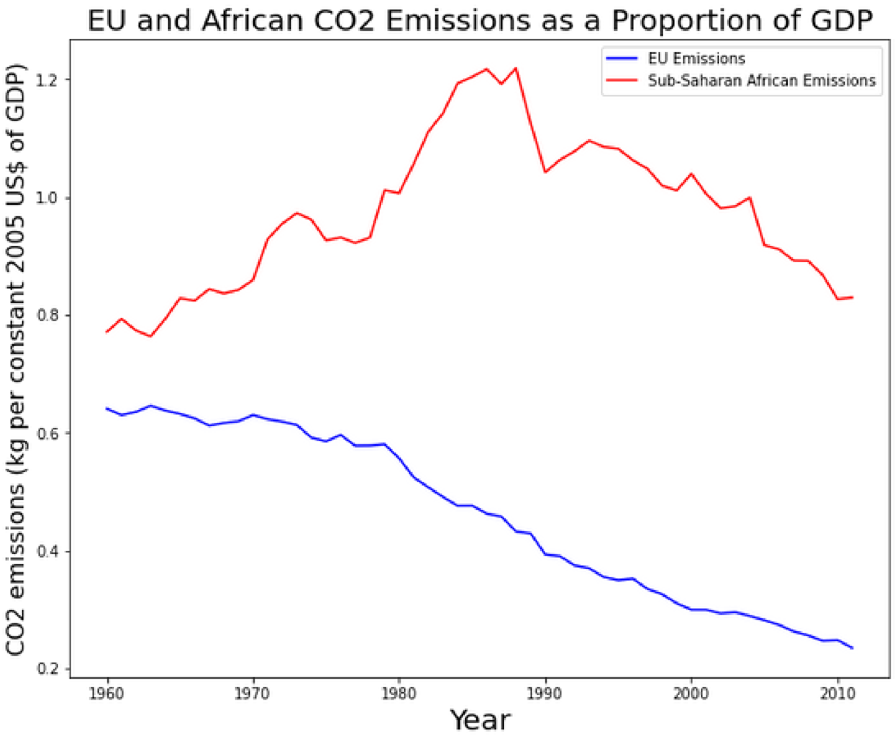
# $CO_2$ Emissions As a Proportion of GDP

Encouragingly, with the growth in EU GDP we are seeing a strong decline in the $CO_2$ emissions per dollar of GDP.

There is also a downward trend in Africa as of the 1990s, and hopefully this will continue as the economies grow. Despite this, the emissions per dollar for Africa are much higher than the EU, suggesting low efficiency. This may be due to poor infrastructure, "dirty" energy sources such as coal, lack of investment, industrial structure, etc.

This plot only gives an insight into energy efficiency per dollar – to get a more detailed picture we will look at the magnitude of these emissions.

Are the emissions being reduced enough?



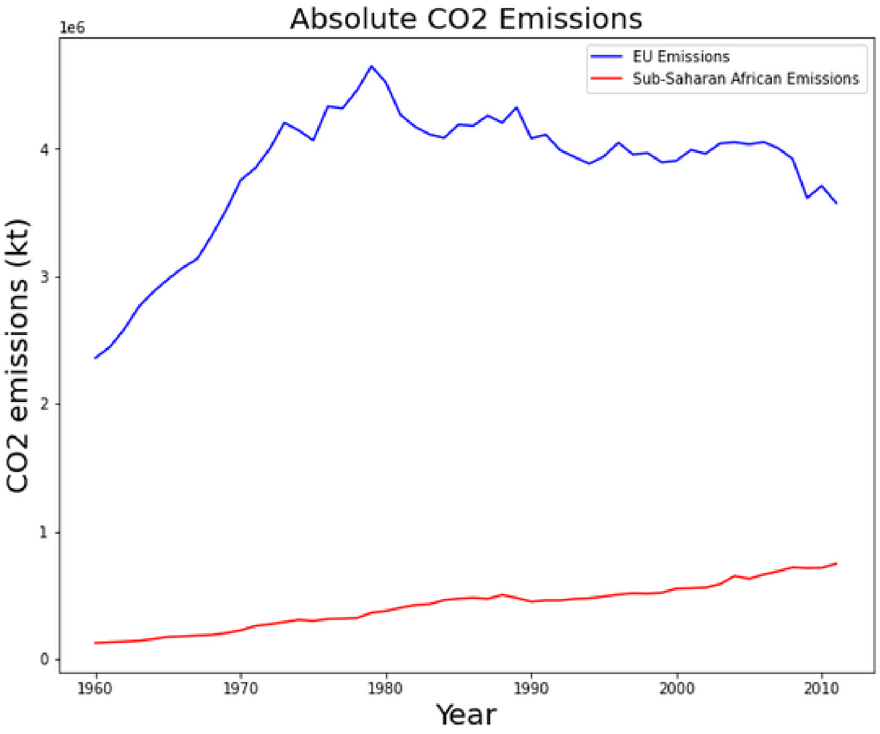
# Absolute $CO_2$ emissions

So, Sub-Saharan Africa still only produces a fraction of the $CO_2$ emissions of the EU despite its inefficiency

GDP growth is outpacing $CO_2$ emissions in the EU, which is a strong positive sign.

However, massive $CO_2$ emissions are still a by-product of this urban way of living and economic set up.

It is also quite concerning that African $CO_2$ emissions are steadily rising. As these countries develop, sustainability should be built-in to the economies so emissions do not reach the levels of developed countries like those in the EU.



# Final Remarks

For the EU, urban population growth and GDP growth are strongly correlated and the relationship is bi-directional. This is due to socioeconomic, cultural, political and environmental factors.

Contrastingly, urban population growth in Sub-Saharan Africa does not correlate with GDP growth, and rural populations in Sub-Saharan Africa are greater than urban populations. This will be due to many of the same factors as those which influence the EU, plus factors which apply only to developing countries such as excessively high birth rate, lack of jobs, etc.

$CO_2$ emissions, however, have been very high over this time period for the EU, so this GDP growth has not come without cost. The focus for Europe must be to even more aggressively reduce $CO_2$ emissions per dollar of GDP. The Paris climate agreement has EU countries pledging to emit at least 40% below 1990 levels by 2030 – whether this will be met is yet to be seen.

The need for in-built sustainable development within the African continent is illustrated in the visualisations. With the population figures in Africa, if those countries were at a comparable level of development as Europe without “greener” development and infrastructure, global $CO_2$ emissions would rise again and the progress made with things such as the Paris Climate Agreement would be undone.

# Acknowledgements

I would like to acknowledge the World Bank for making such complete and high quality data available.

I would also like to acknowledge my peers who will review this work, and the staff at UCSD and EdX for providing this Data Science Micromasters.

Lastly, I'd like to acknowledge Marinko Škare and Sanja Blažević of Juraj Dobrila University of Pula, Faculty of Economics, for their essay which helped form the base of this investigation.

# References

"Population and Economic Growth: A Review Essay" by Marinko Škare and Sanja Blažević. This aided in selecting the research area.

All investigation and data retrieval once the research area was chosen was my own.