GHG Emissions and Livestocks

by Dung (Yung) Nguyen

Dataset(s)

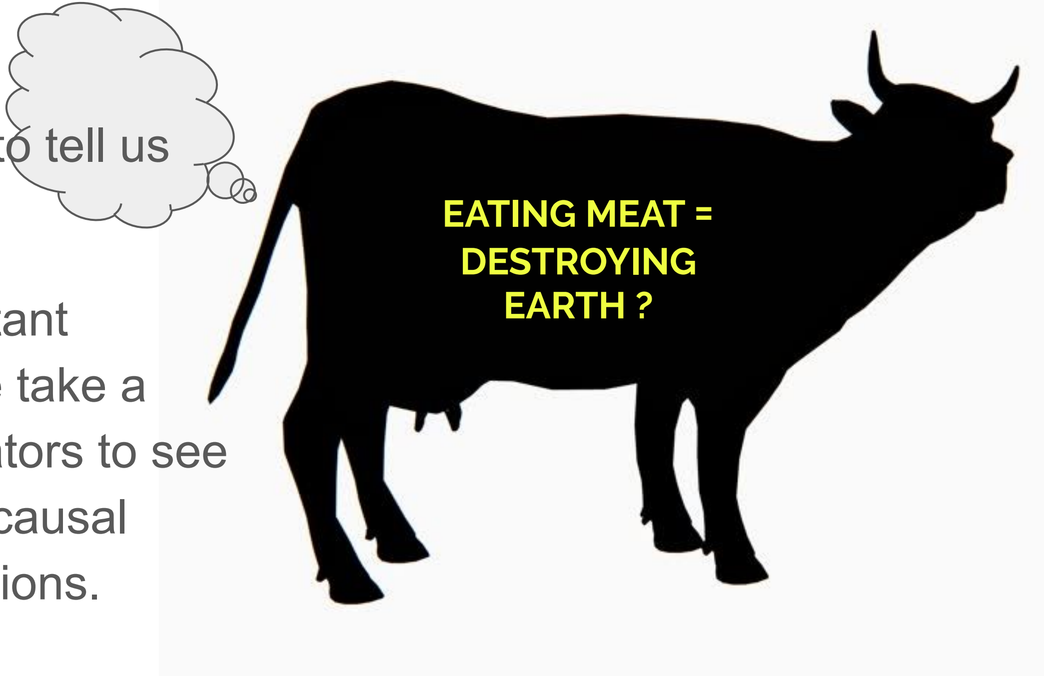
- World Development Indicators Dataset
  - **Source:** World Bank - Download (Last Updated May 2017)
  - **Description:** This dataset contains over a thousand annual indicators of economic development from hundreds of countries around the world.
  - **Focus:** This study examines a subset data focused on Environmental Topics, specifically at **emissions related** and **agriculture livestock** indicators. The emissions related and livestock indicators data set ranges from 1960s to 2010s typically. Missing for some indicators or for certain periods are not significant.
  - **Limitations:** The study did not expand beyond the World Development Indicator Datasets and the indicators provided by the dataset. Some of indicators have different duration of available data.

Motivation

For a while, many news media outcried about scientist's underestimation of cows' farts role in global warming[1].

Is it truly the case? Do the data dare to tell us eating meat is a catastrophe?

To answer one of today's most important question for the future generation, we take a look at the World Development Indicators to see what it can possibly tell us about the causal relationship between meat and emissions.



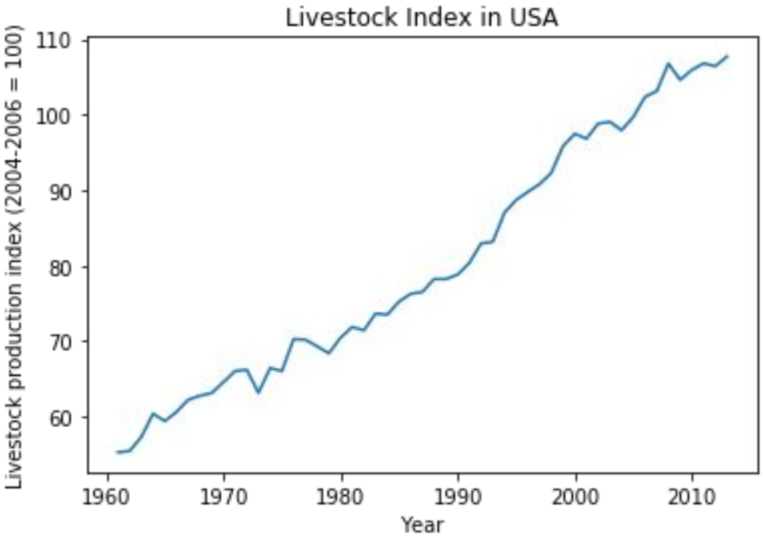
Research Question(s)

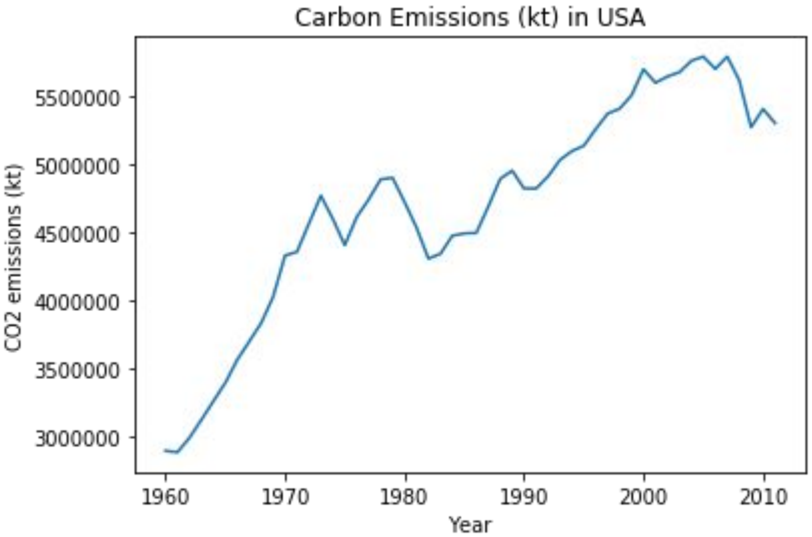
The goal of the research questions is get an idea of potential correlation.

What are the trends in the past 50 years in the U.S. of livestock production and GHG emissions? Is there a correlation between livestocks and Greenhouse Gases (GHG) emissions? How does it compares outside of the U.S.?

Findings

What are the trends of CO2 equivalent (CO2e) emissions in the past 50 years in the U.S.?

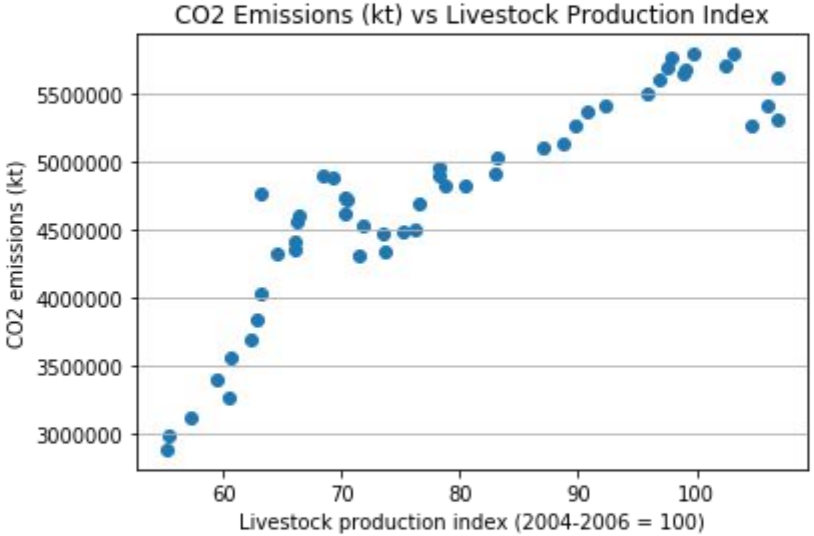




The trends from 1961 to 2012 for Livestock Index and Carbon Emissions show a positive upward trend. Carbon emissions in the U.S. between 1961 and 1981 show a higher increase that the Livestock Index trend does not, indicating there are likely other causal factors to the carbon emissions other than Livestock Index.

Findings

What is the correlation between livestocks and Greenhouse Gases (GHG) emissions in the U.S.?

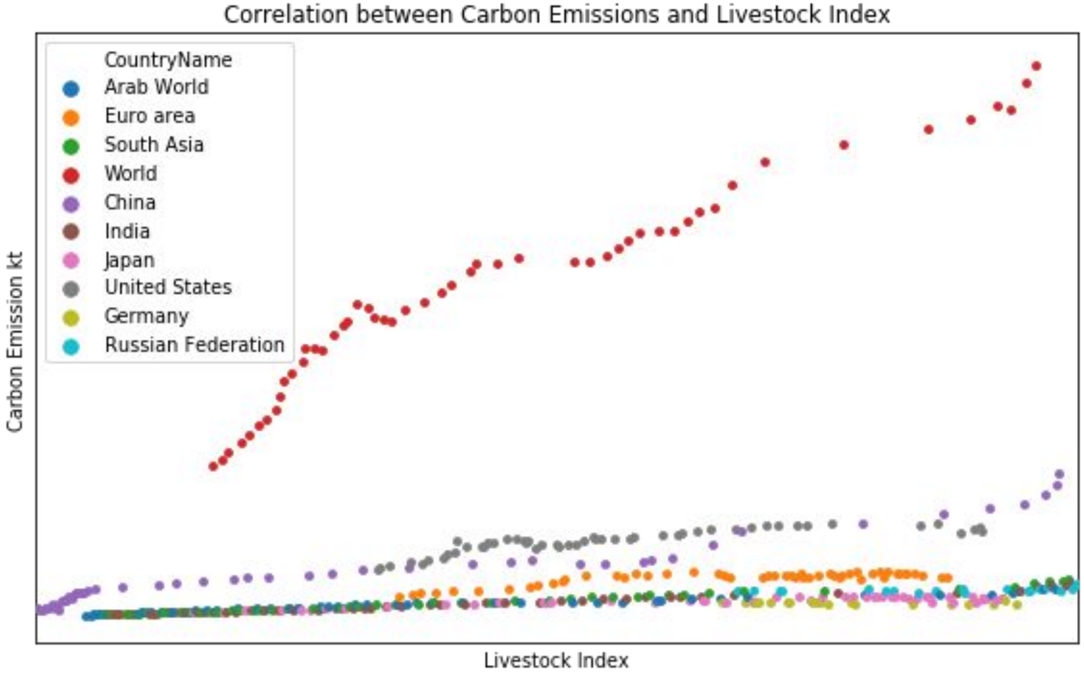


Correlation = 0.89

The correlation between Livestock Index and Carbon Emissions (during 1961 and 2012) shows a positive one.

Findings

How does it compare to top emitting countries or regions?

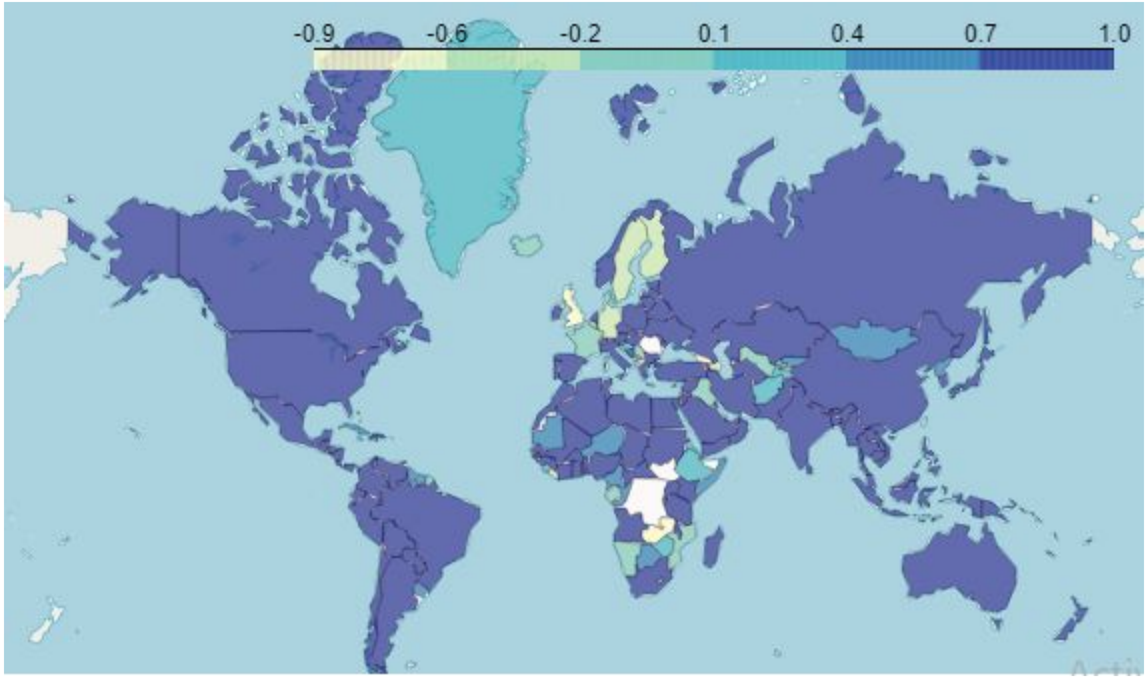


CountryName	CorrVal
South Asia	0.996246
India	0.995452
Arab World	0.986659
World	0.975142
China	0.955384
Japan	0.914964
United States	0.891535
Euro area	0.864608
Russian Federation	0.861487
Germany	-0.546637

Most of the top countries or regions show a high correlation values, except for Germany, which appears to be an outlier. Is it there are many vegans in Germany[2]? To understand why there are outliers or specifically why Germany is an outlier, we may need additional datasets and analysis.

Findings

What can we say about the rest of the world?

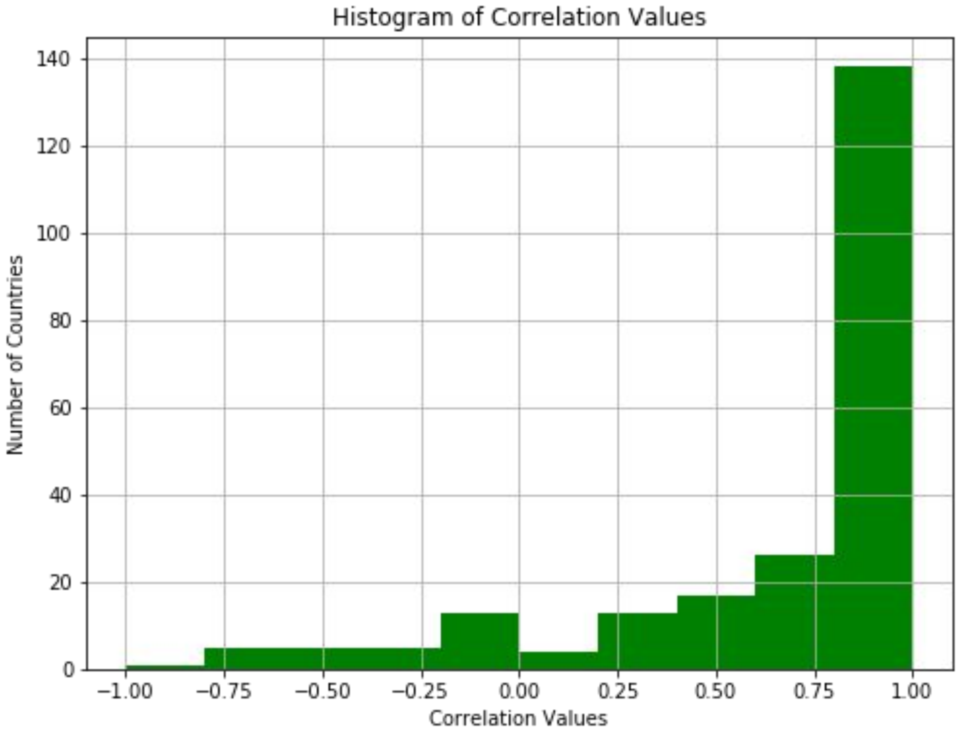


The correlation between Carbon Emissions and Livestock Index is found for every Country Name in the dataset and shown on a map. It is very interesting to see that there is a small subset of countries that have low correlation values (light blue) or even negative correlation values (yellow and green).

What are the possible that this is happening? Is it because these countries do not have livestocks or much agricultural activities?

Findings

What can we say about the rest of the world?



Since the Livestock Index is the only indicator in the dataset to directly address Livestock activities, it's possible add a different livestock indicator (like meat consumption instead of Livestock Index) to truly measure consumption behaviors effects on emissions.

All in all, we can see that distribution of the correlation values for each countries skewed toward high correlation spectrum, indicating a likely positive correlation between livestock productions and carbon emissions.

Acknowledgements

I did not have feedback from anyone for this project.

Outside of the class's resources and lessons, I used:

- Stackoverflow and other google sites
- Seaborn library

References

1. Lemonick, Sam (2017). “Scientists Underestimated How Bad Cow Farts Are”. Forbes Online Magazine. Link.

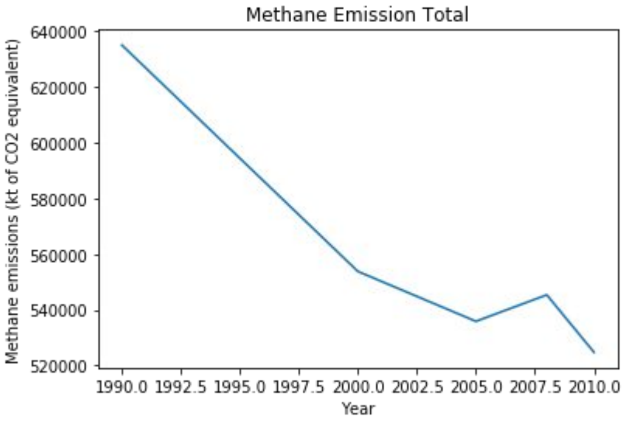
   “Livestock Emissions: Still Grossly Underestimated?”. Worldwatch Institute. 2019. Link.

2. Chiorando, Maria (2018). “Germany Dominates Global Vegan Product Market, Says Report”, Plant Based News. Link.

Additional Analysis

What if we look at methane in the U.S.?

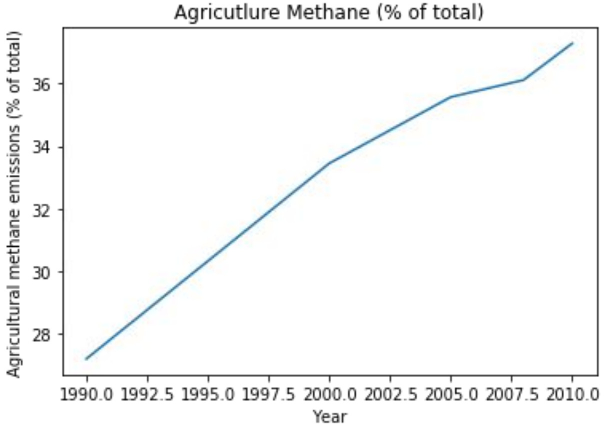
As the Livestock Index (e.g. an indicator of how much livestocks are produced annually relative to a baseline year) increases, methane emissions - particularly from the agriculture sector - are expected to increase.
At first glance, the methane emissions (1990 to 2010) seem to decrease.

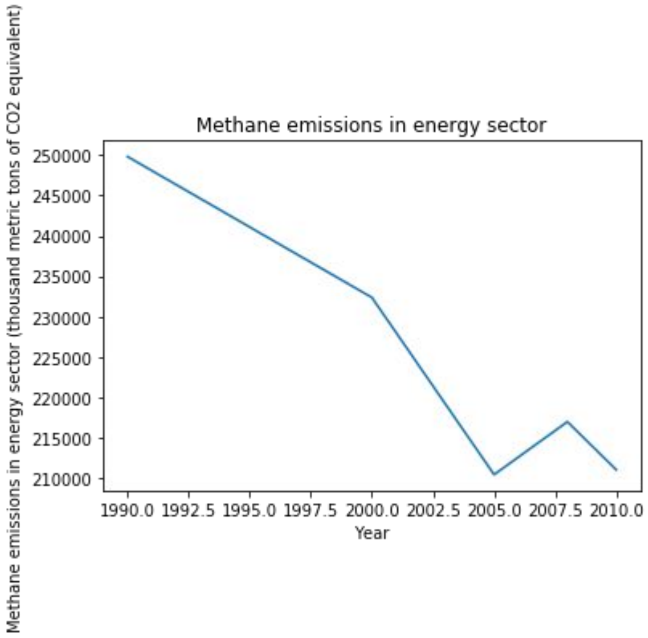


Additional Analysis

What if we look at methane in the U.S.?

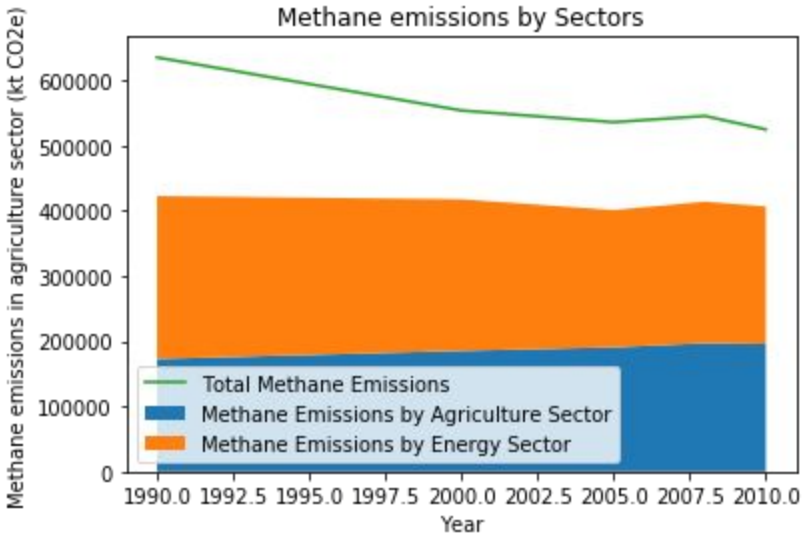
However, when looking at the breakdown by sector, agricultural methane emissions (% of total) increase over time while the methane emissions from the energy sector shows a decrease over time.





Additional Analysis

What if we look at methane in the U.S.?



When looking at the methane emissions by sectors in a stacked area chart, we can see there are gaps (or other sectors data) that were not reported that may contribute to methane declining trend.

All in all, the methane emissions from the agriculture sector shows an expected increase as livestock index increases over time.

We may need additional granular data sets to dig into the impact of cows' farts since agricultural emissions indicator is too broad. If we really care about this issue, we can ask the original data collector to include Livestock emissions (% of total or % agricultural emissions total).

Mini Project YN

June 21, 2019

```
In [2]: import pandas as pd
        import numpy as np
        import random
        import matplotlib.pyplot as plt
```

0.1 Data Exploratory of the World Development Indicators

I want to explore this dataset a bit more before coming up with a research question. I was most interested in the Environmental indicators.

```
In [3]: #This is the main spreadsheet to work from.
        data = pd.read_csv('./world-development-indicators/Indicators.csv')
        print("Data Shape = " , data.shape)

Data Shape =  (5656458, 6)
```

```
In [4]: data.head()

Out[4]:   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
```

```
In [5]: #The goal is to add topic from the Series dataset.
        series = pd.read_csv('./world-development-indicators/Series.csv')
        len(series)

Out[5]: 1345
```

In [6]: series.head()

```
Out[6]:             SeriesCode                                              Topic  \
        0       BN.KLT.DINV.CD  Economic Policy & Debt: Balance of payments: C...
        1  BX.KLT.DINV.WD.GD.ZS  Economic Policy & Debt: Balance of payments: C...
        2    BX.KLT.DINV.CD.WD  Economic Policy & Debt: Balance of payments: C...
        3    BM.KLT.DINV.GD.ZS  Economic Policy & Debt: Balance of payments: C...
        4       BN.TRF.KOGT.CD  Economic Policy & Debt: Balance of payments: C...

                                             IndicatorName ShortDefinition  \
        0  Foreign direct investment, net (BoP, current US$)             NaN
        1  Foreign direct investment, net inflows (% of GDP)             NaN
        2  Foreign direct investment, net inflows (BoP, c...             NaN
        3  Foreign direct investment, net outflows (% of ...             NaN
        4             Net capital account (BoP, current US$)             NaN

                                            LongDefinition UnitOfMeasure  \
        0  Foreign direct investment are the net inflows ...           NaN
        1  Foreign direct investment are the net inflows ...           NaN
        2  Foreign direct investment refers to direct inv...           NaN
        3  Foreign direct investment are the net inflows ...           NaN
        4  Net capital account records acquisitions and d...           NaN

          Periodicity BasePeriod  OtherNotes AggregationMethod  \
        0      Annual        NaN         NaN               NaN
        1      Annual        NaN         NaN  Weighted average
        2      Annual        NaN         NaN               Sum
        3      Annual        NaN         NaN  Weighted average
        4      Annual        NaN         NaN               NaN

                                    LimitationsAndExceptions NotesFromOriginalSource  \
        0                                                NaN                     NaN
        1  FDI data do not give a complete picture of int...                     NaN
        2  FDI data do not give a complete picture of int...                     NaN
        3                                                NaN                     NaN
        4                                                NaN                     NaN

                                           GeneralComments  \
        0  Note: Data are based on the sixth edition of t...
        1  Note: Data starting from 2005 are based on the...
        2  Note: Data starting from 2005 are based on the...
        3  Note: Data are based on the sixth edition of t...
        4  Note: Data are based on the sixth edition of t...

                                                    Source  \
        0  International Monetary Fund, Balance of Paymen...
        1  International Monetary Fund, International Fin...
        2  International Monetary Fund, Balance of Paymen...
```

```
3  International Monetary Fund, International Fin...
4  International Monetary Fund, Balance of Paymen...

                     StatisticalConceptAndMethodology  \
0                                                 NaN
1  Data on equity flows are based on balance of p...
2  Data on equity flows are based on balance of p...
3                                                 NaN
4                                                 NaN

                               DevelopmentRelevance RelatedSourceLinks  \
0                                                 NaN                NaN
1  Private financial flows - equity and debt - ac...                NaN
2  Private financial flows - equity and debt - ac...                NaN
3                                                 NaN                NaN
4                                                 NaN                NaN

  OtherWebLinks  RelatedIndicators LicenseType
0           NaN                NaN        Open
1           NaN                NaN        Open
2           NaN                NaN        Open
3           NaN                NaN        Open
4           NaN                NaN        Open
```

In [7]:
```python
indicators = data['IndicatorName'].unique().tolist()
len(indicators)
```

Out[7]:
```
1344
```

In [8]:
```python
countries = data['CountryCode'].unique().tolist()
len(countries)
```

Out[8]:
```
247
```

In [9]:
```python
#Create new data frame that will match the Topic to data table.
data_addtopic = pd.merge(data,series[['IndicatorName','Topic']],on ="IndicatorName")
data_addtopic.head()
```

Out[9]:
```
                              CountryName CountryCode  \
0                              Arab World         ARB
1                  Caribbean small states         CSS
2          Central Europe and the Baltics         CEB
3  East Asia & Pacific (all income levels)         EAS
4    East Asia & Pacific (developing only)         EAP

                                    IndicatorName IndicatorCode  Year  \
0  Adolescent fertility rate (births per 1,000 wo...   SP.ADO.TFRT  1960
1  Adolescent fertility rate (births per 1,000 wo...   SP.ADO.TFRT  1960
2  Adolescent fertility rate (births per 1,000 wo...   SP.ADO.TFRT  1960
```

```
3  Adolescent fertility rate (births per 1,000 wo...   SP.ADO.TFRT  1960
4  Adolescent fertility rate (births per 1,000 wo...   SP.ADO.TFRT  1960

        Value                        Topic
0  133.560907  Health: Reproductive health
1  162.871212  Health: Reproductive health
2   46.716752  Health: Reproductive health
3   66.015974  Health: Reproductive health
4   75.043631  Health: Reproductive health
```

In [10]:
```
# Get list of all the topic.
topic = data_addtopic['Topic'].unique().tolist()
topic
```

Out[10]:
```
['Health: Reproductive health',
 'Health: Population: Dynamics',
 'Public Sector: Defense & arms trade',
 'Environment: Emissions',
 'Infrastructure: Communications',
 'Health: Health services',
 'Social Protection & Labor: Migration',
 'Health: Mortality',
 'Private Sector & Trade: Exports',
 'Private Sector & Trade: Imports',
 'Private Sector & Trade: Total merchandise trade',
 'Economic Policy & Debt: Official development assistance',
 'Health: Population: Structure',
 'Environment: Density & urbanization',
 'Economic Policy & Debt: National accounts: US$ at current prices: Expenditure on GDI
 'Economic Policy & Debt: National accounts: US$ at current prices: Aggregate indicato
 'Environment: Energy production & use',
 'Economic Policy & Debt: National accounts: US$ at constant 2005 prices: Value added
 'Economic Policy & Debt: National accounts: Shares of GDP & other',
 'Economic Policy & Debt: National accounts: US$ at constant 2005 prices: Aggregate in
 'Economic Policy & Debt: National accounts: US$ at constant 2005 prices: Expenditure
 'Economic Policy & Debt: National accounts: US$ at current prices: Other items',
 'Economic Policy & Debt: Balance of payments: Capital & financial account',
 'Economic Policy & Debt: National accounts: US$ at current prices: Value added',
 'Infrastructure: Technology',
 'Financial Sector: Exchange rates & prices',
 'Economic Policy & Debt: National accounts: Local currency at current prices: Expendi
 'Economic Policy & Debt: National accounts: Local currency at current prices: Aggrega
 'Economic Policy & Debt: Balance of payments: Reserves & other items',
 'Economic Policy & Debt: National accounts: Local currency at constant prices: Value
 'Economic Policy & Debt: National accounts: Local currency at current prices: Value a
 'Economic Policy & Debt: National accounts: Local currency at constant prices: Expend
 'Economic Policy & Debt: National accounts: Local currency at constant prices: Other
 'Economic Policy & Debt: National accounts: Local currency at constant prices: Aggreg
```

```
'Economic Policy & Debt: National accounts: Adjusted savings & income',
'Economic Policy & Debt: Balance of payments: Current account: Transfers',
'Financial Sector: Interest rates',
'Environment: Land use',
'Environment: Agricultural production',
'Economic Policy & Debt: National accounts: Growth rates',
'Economic Policy & Debt: National accounts: Growth rates:',
'Environment: Freshwater',
'Economic Policy & Debt: National accounts: Atlas GNI & GNI per capita',
'Health: Nutrition',
'Infrastructure: Transportation',
'Environment: Natural resources contribution to GDP',
'Economic Policy & Debt: External debt: Debt ratios & other items',
'Economic Policy & Debt: External debt: Net flows',
'Economic Policy & Debt: External debt: Terms',
'Economic Policy & Debt: External debt: Commitments',
'Economic Policy & Debt: External debt: Currency composition',
'Economic Policy & Debt: External debt: Debt service',
'Economic Policy & Debt: External debt: Disbursements',
'Economic Policy & Debt: External debt: Debt outstanding',
'Economic Policy & Debt: External debt: Interest',
'Economic Policy & Debt: External debt: Amortization',
'Economic Policy & Debt: External debt: Arrears, reschedulings, etc.',
'Economic Policy & Debt: External debt: Net transfers',
'Economic Policy & Debt: External debt: Undisbursed debt',
'Financial Sector: Capital markets',
'Poverty: Poverty rates',
'Health: Disease prevention',
'Social Protection & Labor: Labor force structure',
'Social Protection & Labor: Economic activity',
'Social Protection & Labor: Unemployment',
'Private Sector & Trade: Trade indexes',
'Poverty: Income distribution',
'Public Sector: Policy & institutions',
'Private Sector & Trade: Tariffs',
'Public Sector: Conflict & fragility',
'Economic Policy & Debt: Purchasing power parity',
'Environment: Biodiversity & protected areas',
'Health: Risk factors',
'Private Sector & Trade: Private infrastructure investment',
'Public Sector: Government finance: Deficit & financing',
'Public Sector: Government finance: Expense',
'Public Sector: Government finance: Revenue',
'Private Sector & Trade: Travel & tourism',
'Financial Sector: Assets',
'Financial Sector: Monetary holdings (liabilities)',
'Education: Inputs',
'Financial Sector: Access',
```

```
 'Private Sector & Trade: Business environment',
 'Poverty: Shared prosperity',
 'Social Protection & Labor: Performance',
 'Economic Policy & Debt: Balance of payments: Current account: Goods, services & inco
 'Private Sector & Trade: Trade facilitation',
 'Economic Policy & Debt: Balance of payments: Current account: Balances']
```

In [11]:
```python
# Narrow down to looking at environmental topics.
env_topic = [x for x in topic if 'Environment' in x]
#Create a mask of the environment topics
#env_topic_mask = data_addtopic['Topic'].str.contains(env_topic)
env_data = data_addtopic['Topic'].isin(env_topic)
data_env = data_addtopic.loc[env_data]

#get list of indicator in data_env
indicators_env = data_env['IndicatorName'].unique().tolist()
indicators_env
```

Out[11]:
```
['CO2 emissions (kt)',
 'CO2 emissions (metric tons per capita)',
 'CO2 emissions from gaseous fuel consumption (% of total)',
 'CO2 emissions from liquid fuel consumption (% of total)',
 'CO2 emissions from liquid fuel consumption (kt)',
 'CO2 emissions from solid fuel consumption (% of total)',
 'Population in the largest city (% of urban population)',
 'Population in urban agglomerations of more than 1 million (% of total population)',
 'Rural population',
 'Rural population (% of total population)',
 'Urban population',
 'Urban population (% of total)',
 'Alternative and nuclear energy (% of total energy use)',
 'CO2 emissions from electricity and heat production, total (% of total fuel combustic
 'CO2 emissions from manufacturing industries and construction (% of total fuel combus
 'CO2 emissions from other sectors, excluding residential buildings and commercial anc
 'CO2 emissions from residential buildings and commercial and public services (% of tc
 'CO2 emissions from solid fuel consumption (kt)',
 'CO2 emissions from transport (% of total fuel combustion)',
 'CO2 intensity (kg per kg of oil equivalent energy use)',
 'Combustible renewables and waste (% of total energy)',
 'Electric power transmission and distribution losses (% of output)',
 'Electricity production from renewable sources, excluding hydroelectric (kWh)',
 'Energy imports, net (% of energy use)',
 'Fossil fuel energy consumption (% of total)',
 'CO2 emissions (kg per 2005 US$ of GDP)',
 'Electric power consumption (kWh per capita)',
 'Electricity production from coal sources (% of total)',
 'Electricity production from hydroelectric sources (% of total)',
 'Electricity production from natural gas sources (% of total)',
```

```
'Electricity production from nuclear sources (% of total)',
'Electricity production from oil sources (% of total)',
'Electricity production from oil, gas and coal sources (% of total)',
'Electricity production from renewable sources, excluding hydroelectric (% of total)
'Energy use (kg of oil equivalent per capita)',
'CO2 emissions from gaseous fuel consumption (kt)',
'Population in largest city',
'Population in urban agglomerations of more than 1 million',
'Rural population growth (annual %)',
'Urban population growth (annual %)',
'Agricultural land (% of land area)',
'Agricultural land (sq. km)',
'Agricultural machinery, tractors',
'Agricultural machinery, tractors per 100 sq. km of arable land',
'Arable land (% of land area)',
'Arable land (hectares per person)',
'Cereal production (metric tons)',
'Cereal yield (kg per hectare)',
'Crop production index (2004-2006 = 100)',
'Food production index (2004-2006 = 100)',
'Land area (sq. km)',
'Land under cereal production (hectares)',
'Livestock production index (2004-2006 = 100)',
'Permanent cropland (% of land area)',
'Population density (people per sq. km of land area)',
'Surface area (sq. km)',
'Arable land (hectares)',
'Renewable internal freshwater resources per capita (cubic meters)',
'Renewable internal freshwater resources, total (billion cubic meters)',
'Average precipitation in depth (mm per year)',
'Annual freshwater withdrawals, total (% of internal resources)',
'Annual freshwater withdrawals, total (billion cubic meters)',
'Water productivity, total (constant 2005 US$ GDP per cubic meter of total freshwate
'Annual freshwater withdrawals, agriculture (% of total freshwater withdrawal)',
'Annual freshwater withdrawals, domestic (% of total freshwater withdrawal)',
'Annual freshwater withdrawals, industry (% of total freshwater withdrawal)',
'Coal rents (% of GDP)',
'Forest rents (% of GDP)',
'Mineral rents (% of GDP)',
'Natural gas rents (% of GDP)',
'Oil rents (% of GDP)',
'Total natural resources rents (% of GDP)',
'Agriculture value added per worker (constant 2005 US$)',
'Access to electricity (% of population)',
'Access to electricity, rural (% of rural population)',
'Access to electricity, urban (% of urban population)',
'Access to non-solid fuel (% of population)',
'Agricultural methane emissions (% of total)',
```

```
'Agricultural nitrous oxide emissions (% of total)',
'CO2 emissions (kg per 2011 PPP $ of GDP)',
'CO2 emissions (kg per PPP $ of GDP)',
'Energy intensity level of primary energy (MJ/$2011 PPP GDP)',
'Energy related methane emissions (% of total)',
'Energy use (kg of oil equivalent) per $1,000 GDP (constant 2011 PPP)',
'Forest area (% of land area)',
'Forest area (sq. km)',
'GDP per unit of energy use (constant 2011 PPP $ per kg of oil equivalent)',
'GDP per unit of energy use (PPP $ per kg of oil equivalent)',
'Land area where elevation is below 5 meters (% of total land area)',
'Marine protected areas (% of territorial waters)',
'Methane emissions (kt of CO2 equivalent)',
'Methane emissions in energy sector (thousand metric tons of CO2 equivalent)',
'Nitrous oxide emissions (thousand metric tons of CO2 equivalent)',
'Nitrous oxide emissions in energy sector (thousand metric tons of CO2 equivalent)',
'Nitrous oxide emissions in industrial and energy processes (% of total nitrous oxide
'Other greenhouse gas emissions, HFC, PFC and SF6 (thousand metric tons of CO2 equiva
'PM2.5 air pollution, mean annual exposure (micrograms per cubic meter)',
'PM2.5 air pollution, population exposed to levels exceeding WHO guideline value (% o
'Population living in areas where elevation is below 5 meters (% of total population)
'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)',
'HFC gas emissions (thousand metric tons of CO2 equivalent)',
'PFC gas emissions (thousand metric tons of CO2 equivalent)',
'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)',
'GHG net emissions/removals by LUCF (Mt of CO2 equivalent)',
'Agricultural irrigated land (% of total agricultural land)',
'Fertilizer consumption (% of fertilizer production)',
'Fertilizer consumption (kilograms per hectare of arable land)',
'GEF benefits index for biodiversity (0 = no biodiversity potential to 100 = maximum)
'Droughts, floods, extreme temperatures (% of population, average 1990-2009)',
'Access to non-solid fuel, rural (% of rural population)',
'Access to non-solid fuel, urban (% of urban population)',
'Disaster risk reduction progress score (1-5 scale; 5=best)',
'Bird species, threatened',
'Fish species, threatened',
'Mammal species, threatened',
'Plant species (higher), threatened']
```

0.2 Narrowing the Research Question

Looking at the list of environmental indicators, I can see what are the available data and begin to wonder how some of these may correlate relative to the countries and year. I am curious about livestock's impact on the planet (broad topic).

I'm further narrowing down the data column I'm interested in: * Livestock production index (2004-2006 = 100) (This is the index calculated from the 2004-2006 baseline) * Crop production index (2004-2006 = 100) * Food production index (2004-2006 = 100) * CO2 emissions (kt) * Agricultural land (sq. km) * Land area (sq. km) * Arable land (% of land area) * Arable land (hectares) * Cereal production (metric tons) * Land under cereal production (hectares) * Agricultural methane emissions (% of total) * Agricultural nitrous oxide emissions (% of total) * Methane emissions (kt of CO2 equivalent) * Nitrous oxide emissions (thousand metric tons of CO2 equivalent) * Nitrous oxide emissions in industrial and energy processes (% of total nitrous oxide emissions) * Other greenhouse gas emissions, HFC, PFC and SF6 (thousand metric tons of CO2 equivalent) * Fertilizer consumption (kilograms per hectare of arable land) * Population density (people per sq. km of land area)

From my industry (environmental & energy consulting), my colleagues and I always wondered whether the emissions from the agriculture sector is truly accounted in the total greenhouse gases emissions since there are many indirect whether secondary or tertiery by-products that are latched on to other categories. For instance, emissions from "eating meat" are typically spread out between many categories, (not simply the "Agriculture" catetory) including energy production, refrigeration, and mobile/transportation. Meat is needed to maintain at a low temperature and typically "grown" outside of urban areas (even import/export across country lines) where they are mostly consumed.

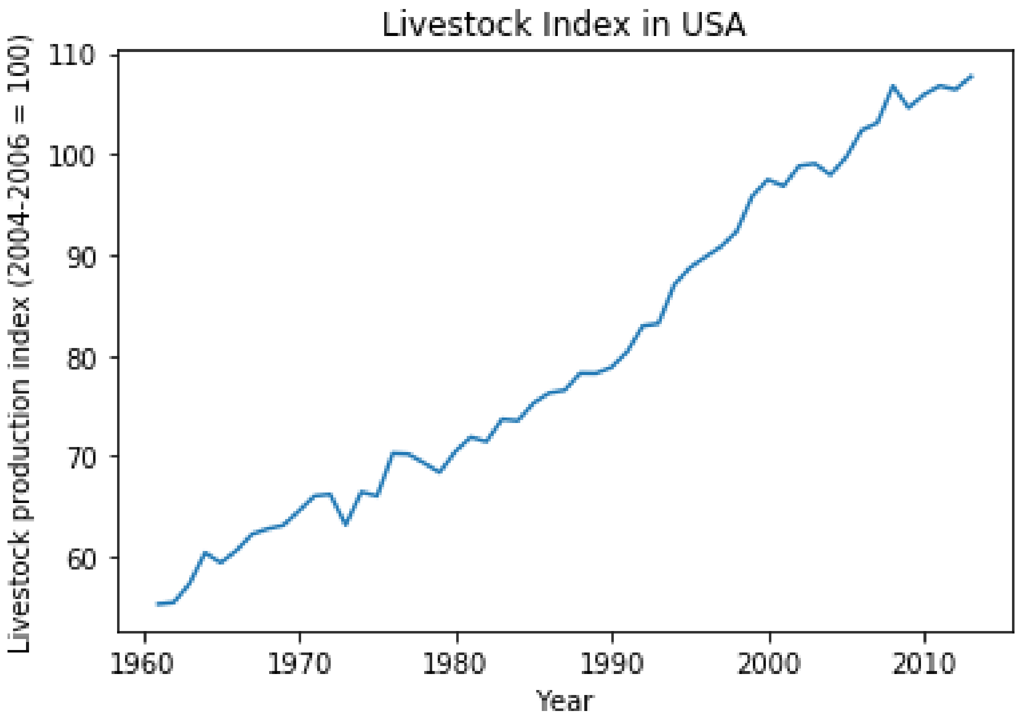
Let's look at USA alone and what the data show.

In [12]:
```
country = 'USA'
livestock_index = 'Livestock production index'

mask_livestock_index = data_env['IndicatorName'].str.contains(livestock_index)
mask_country = data_env['CountryCode'].str.contains(country)

output1 = data_env[mask_livestock_index & mask_country]
```

In [13]:
```
type(mask_country)
```

Out[13]:
```
pandas.core.series.Series
```

In [14]:
```
#plot the livestock index over time in USA
plt.plot(output1['Year'].values, output1['Value'].values)
plt.xlabel('Year')
plt.ylabel(output1['IndicatorName'].iloc[0])
plt.title('Livestock Index in USA')
plt.show()
```



In [15]: data_env.head()

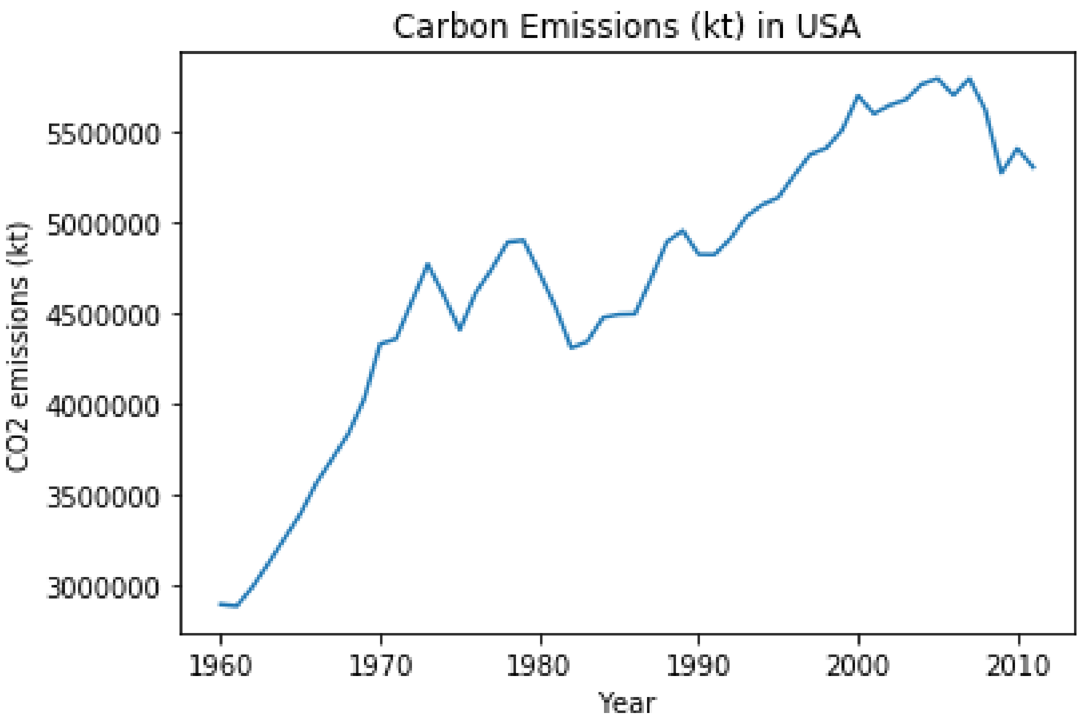
```
Out[15]:                                    CountryName CountryCode  \
         72745                               Arab World         ARB
         72746                   Caribbean small states         CSS
         72747            Central Europe and the Baltics         CEB
         72748  East Asia & Pacific (all income levels)         EAS
         72749    East Asia & Pacific (developing only)         EAP

                    IndicatorName   IndicatorCode  Year         Value  \
         72745  CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  5.956399e+04
         72746  CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  5.878201e+03
         72747  CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  4.674500e+05
         72748  CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  1.211359e+06
         72749  CO2 emissions (kt)  EN.ATM.CO2E.KT  1960  8.541146e+05

                                 Topic
         72745  Environment: Emissions
         72746  Environment: Emissions
         72747  Environment: Emissions
         72748  Environment: Emissions
         72749  Environment: Emissions
```

```
In [16]: #Let's look at carbon emission in total.
         CO2e_kt = 'CO2 emissions \(kt'
```

```
print(CO2e_kt)
mask_carbon = data_env['IndicatorName'].str.contains(CO2e_kt)
#print(data_env['IndicatorName'].str.contains(CO2e_kt).any())
output2 = data_env[mask_carbon & mask_country]
```

```
CO2 emissions \(kt
```

In [17]:
```
#plot the livestock index over time in USA
plt.plot(output2['Year'].values, output2['Value'].values)
plt.xlabel('Year')
plt.ylabel(output2['IndicatorName'].iloc[0])
plt.title('Carbon Emissions (kt) in USA')
plt.show()
```



In [18]:
```
#Check size
print("Size of output1 or Livestock Index = ", len(output1))
print("Size of output2 or CO2e_kt = ", len(output2))
```

```
Size of output1 or Livestock Index =  53
Size of output2 or CO2e_kt =  52
```

```
In [19]: #Check min and max
         print("Livestock Index Min Year = ", output1['Year'].min(), "max: ", output1['Year'].m
         print("CO2 Emission Tot Min Year = ", output2['Year'].min(), "max: ", output2['Year']
```

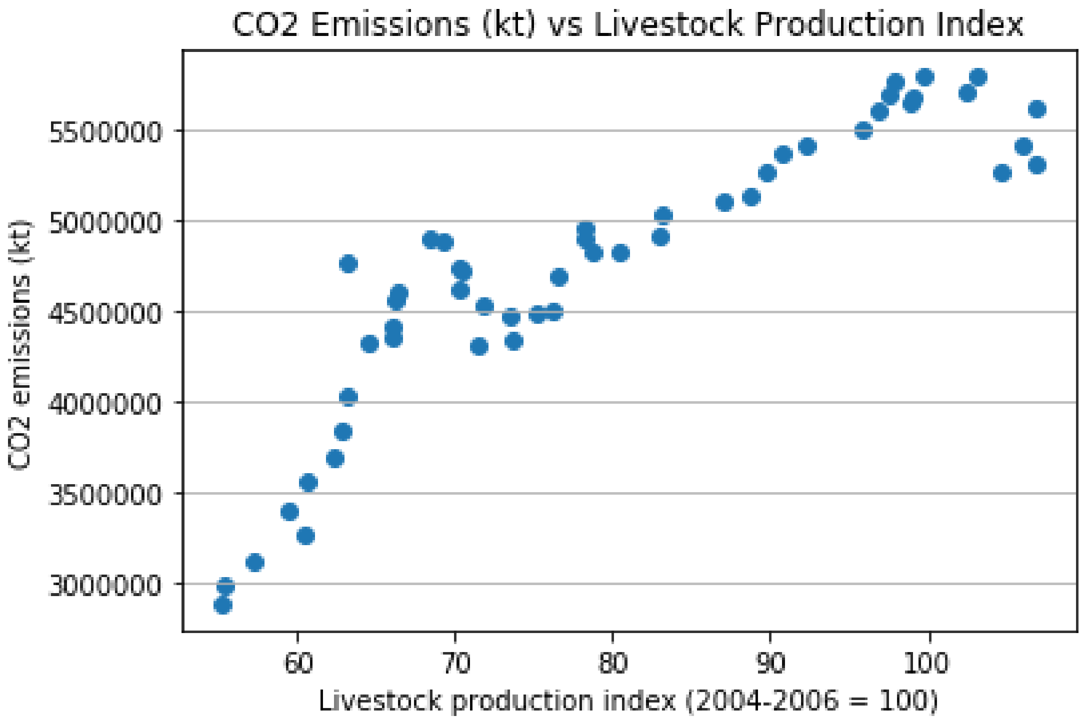
```
Livestock Index Min Year =  1961 max:  2013
CO2 Emission Tot Min Year =  1960 max:  2011
```

We need to truncate the CO2 Emissions and Livestock Index to have the same size.

```
In [20]: ls_output1_trunc = output1[output1['Year'] < 2012]
         CO2_output2_trunc = output2[output2['Year'] > 1960]
         print(len(ls_output1_trunc))
         print(len(CO2_output2_trunc))
```

```
51
51
```

```
In [21]: #Let's look at the correlation between the Livestock Index vs Carbon Emissions.
         #%matplotlib inline
         #import matplotlib.pyplot as plt

         fig, axis = plt.subplots()
         # Grid lines, Xticks, Xlabel, Ylabel

         axis.yaxis.grid(True)
         axis.set_title('CO2 Emissions (kt) vs Livestock Production Index',fontsize=12)
         axis.set_xlabel(ls_output1_trunc['IndicatorName'].iloc[0],fontsize=10)
         axis.set_ylabel(CO2_output2_trunc['IndicatorName'].iloc[0],fontsize=10)

         X = ls_output1_trunc['Value']
         Y = CO2_output2_trunc['Value']

         axis.scatter(X, Y)
         plt.show()
```




```
In [22]: np.corrcoef(ls_output1_trunc['Value'],CO2_output2_trunc['Value'])

Out[22]: array([[1.        , 0.89153464],
                [0.89153464, 1.        ]])
```

This shows a correlation of 0.89, which is a pretty high correlation.

Now, to better analyze the data, we should also look at the total livestock production (accounting for both import and export). Since this information is missing from the current dataset. Other studies such as this one[https://tind-customer-agecon.s3.amazonaws.com/5f4c6fab-9cbc-4745-8075-db48b2f3b7d8?response-content-disposition=inline%3B%20filename%2A%3DUTF-8%27%27pip09.pdf&response-content-type=application%2Fpdf&AWSAccessKeyId=AKIAXL7W7Q3XHXDVDQYS&Expires=1560843294&Signature=c shows that the import and export varies since 1972 to 1999 for Beef alone, showing more import than export historically. However, this also did not include chickens, pigs, lambs, and others.

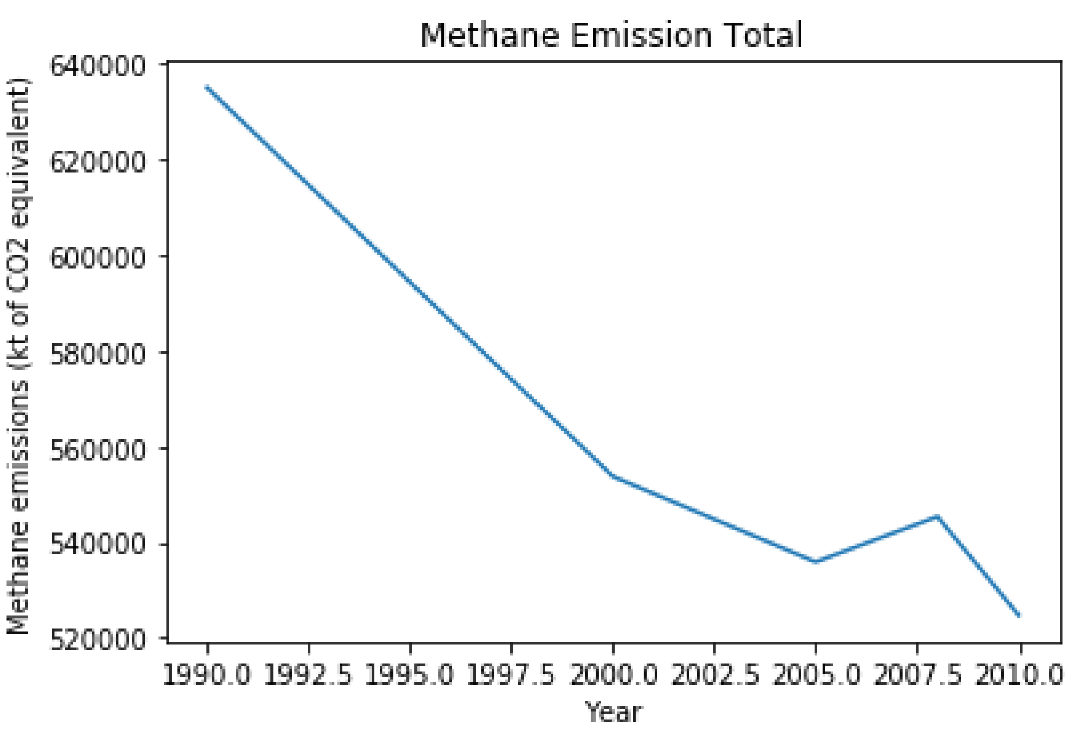
Let's take a look at other factors related to livestock production to see if they share similar trends.

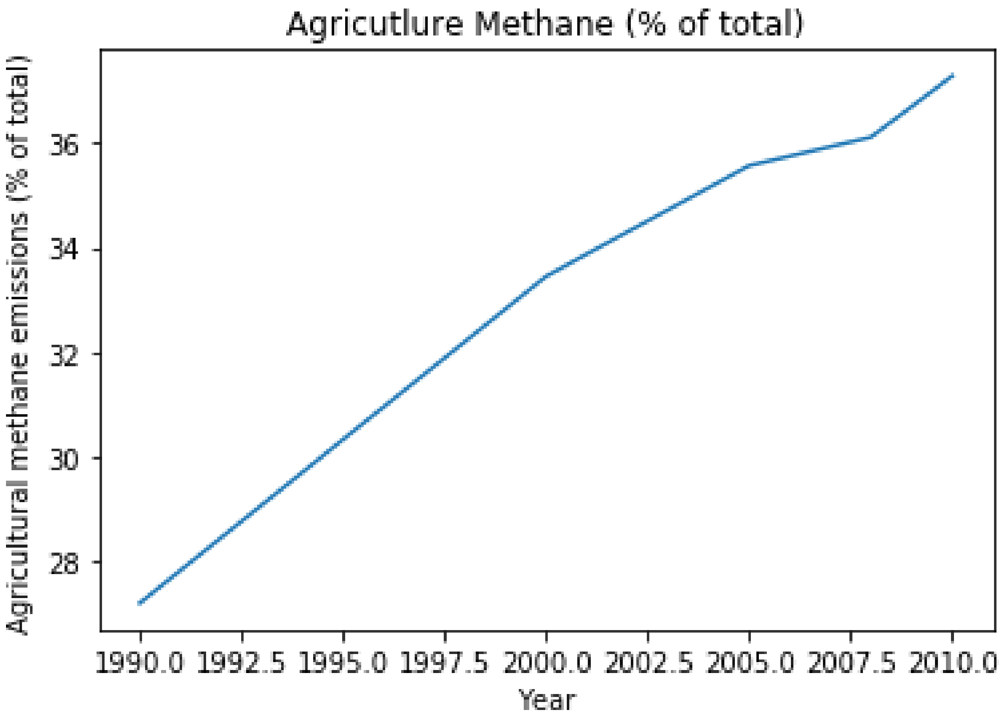
```
In [23]: methane_CO2e = 'Methane emissions \(kt of CO2 equivalent'
         #  Methane emissions (kt of CO2 equivalent)
         agmethane_per = 'Agricultural methane emissions \(% of total'
         # Agricultural methane emissions (% of total)

         mask_methane = data_env['IndicatorName'].str.contains(methane_CO2e)
         output3 = data_env[mask_methane & mask_country]
```

```
print("Methane Emission Total = ", output3['Year'].min(), "max: ", output3['Year'].ma
print(len(output3))

mask_agmethane_per = data_env['IndicatorName'].str.contains(agmethane_per)
output4 = data_env[mask_agmethane_per & mask_country]
print("Agriculture Methane % = ", output4['Year'].min(), "max: ", output4['Year'].max
print(len(output3))
```

```
Methane Emission Total =  1990 max:  2010
5
Agriculture Methane % =  1990 max:  2010
5
```

In [24]:
```
plt.plot(output3['Year'].values, output3['Value'].values)
plt.xlabel('Year')
plt.ylabel(output3['IndicatorName'].iloc[0])
plt.title('Methane Emission Total')

plt.show()

plt.plot(output4['Year'].values, output4['Value'].values)
plt.xlabel('Year')
plt.ylabel(output4['IndicatorName'].iloc[0])
plt.title('Agricutlure Methane (% of total)')
plt.show()
```

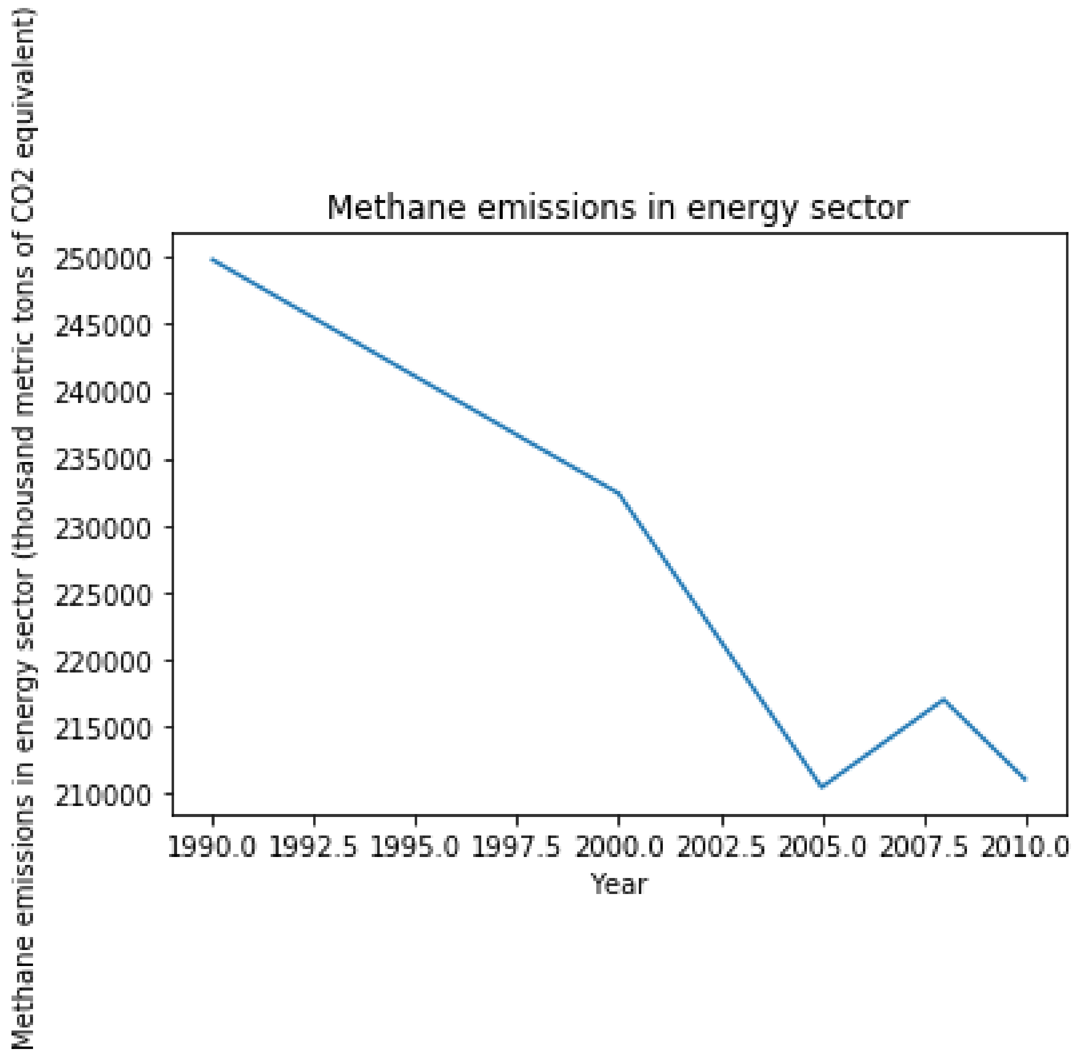




This shows that the decline of methane total emissions is not dependent on the agricultural methane sourced.

```
In [25]: #Let's look at Methane emissions in energy sector
         cereal = 'Methane emissions in energy sector'

         mask_cereal = data_env['IndicatorName'].str.contains(cereal)
         output5 = data_env[mask_cereal & mask_country]
         print("Methane emissions in energy sector = ", output5['Year'].min(), "max: ", output5
         print(len(output5))

         plt.plot(output5['Year'].values, output5['Value'].values)
         plt.xlabel('Year')
         plt.ylabel(output5['IndicatorName'].iloc[0])
         plt.title('Methane emissions in energy sector')
         plt.show()
```

```
Methane emissions in energy sector =  1990 max:  2010
5
```



It appears that the total methane emissions is impacted through the energy sector more so than the agriculture sector. Let's compare the methane emissions between agriculture and energy sector.

In [26]:
```
agmethane_ktCO2e = output3[['Year','Value']].copy()
agmethane_ktCO2e['TotalValue'] = agmethane_ktCO2e['Value'].values / 100 * output4['Va
```

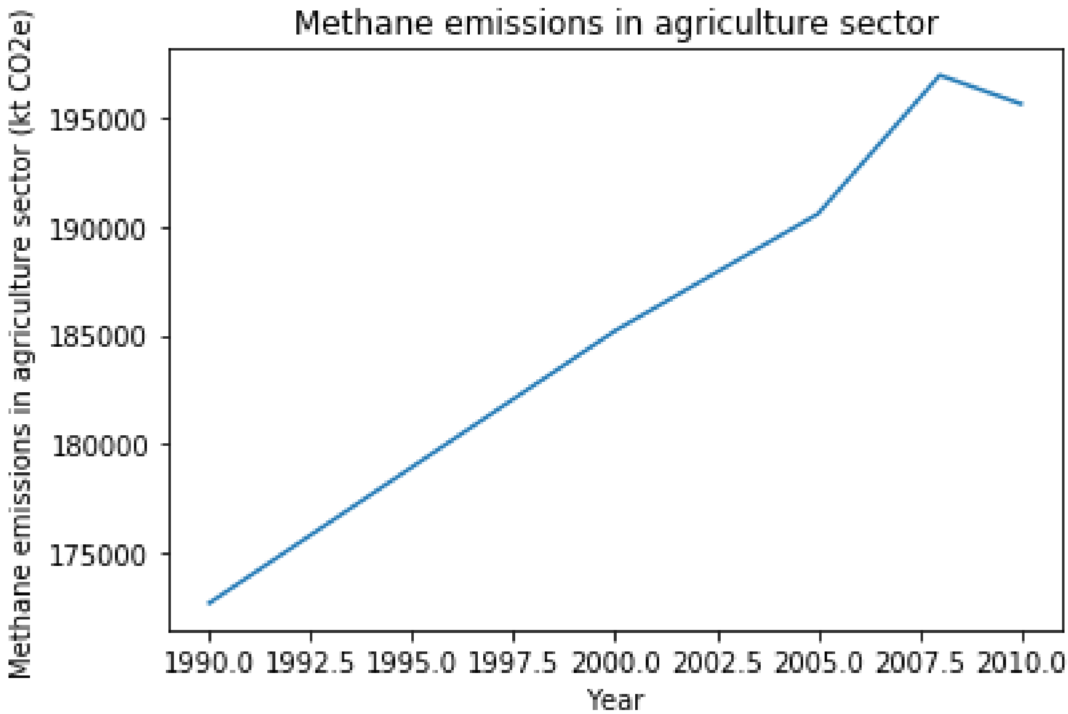
In [27]:
```
agmethane_ktCO2e.head(2)
```

Out[27]:

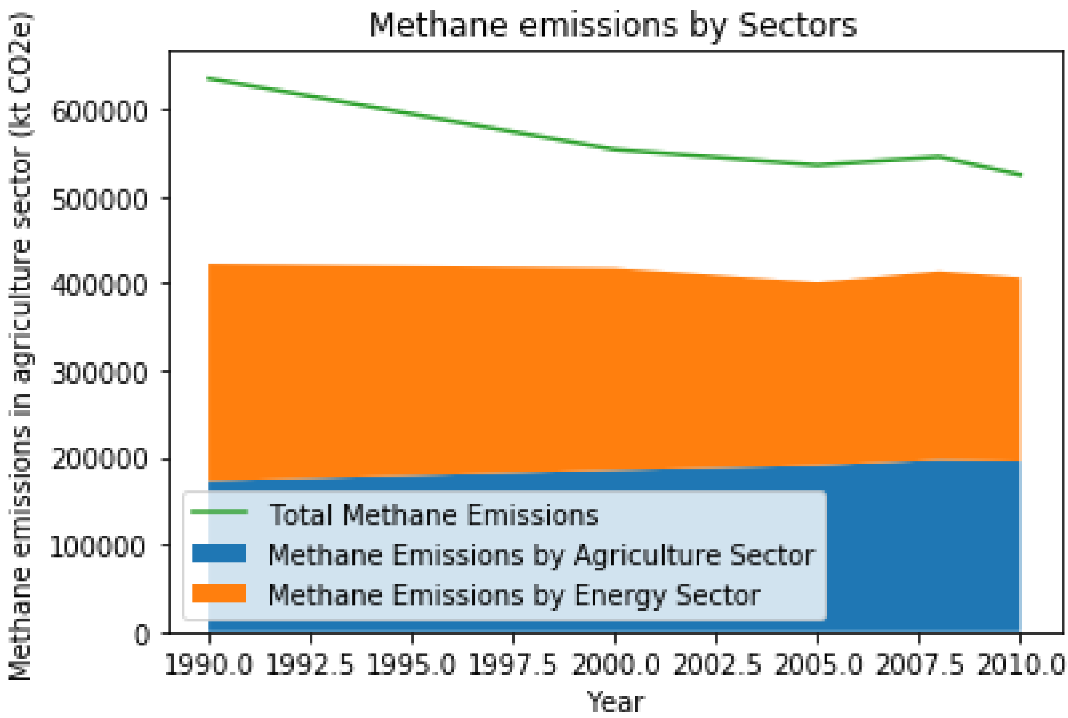
	Year	Value	TotalValue
4465227	1990	635108.2	172714.8
4465393	2000	553739.7	185198.6

In [28]:
```
plt.plot(agmethane_ktCO2e['Year'].values, agmethane_ktCO2e['TotalValue'].values)
plt.xlabel('Year')
plt.ylabel('Methane emissions in agriculture sector (kt CO2e)')
```

```
plt.title('Methane emissions in agriculture sector')
plt.show()
```




```
In [246]: plt.stackplot(agmethane_ktCO2e['Year'].values, agmethane_ktCO2e['TotalValue'].values
          plt.plot(output3['Year'].values, output3['Value'].values)
          plt.xlabel('Year')
          plt.ylabel('Methane emissions in agriculture sector (kt CO2e)')
          plt.title('Methane emissions by Sectors')
          plt.legend(['Total Methane Emissions','Methane Emissions by Agriculture Sector','Met
          plt.show()
```



This shows that there are methane emissions from other sources (not from agriculture or the energy sector) that are not reported as a separate category from the Indicators list.

While news articles may associate "cow farts" to global warming, the latest methane trends and contribution is driven more so by the energy and other sectors.

Now, let's look at Nitrious Dioxide, which is another potent GHG gases.

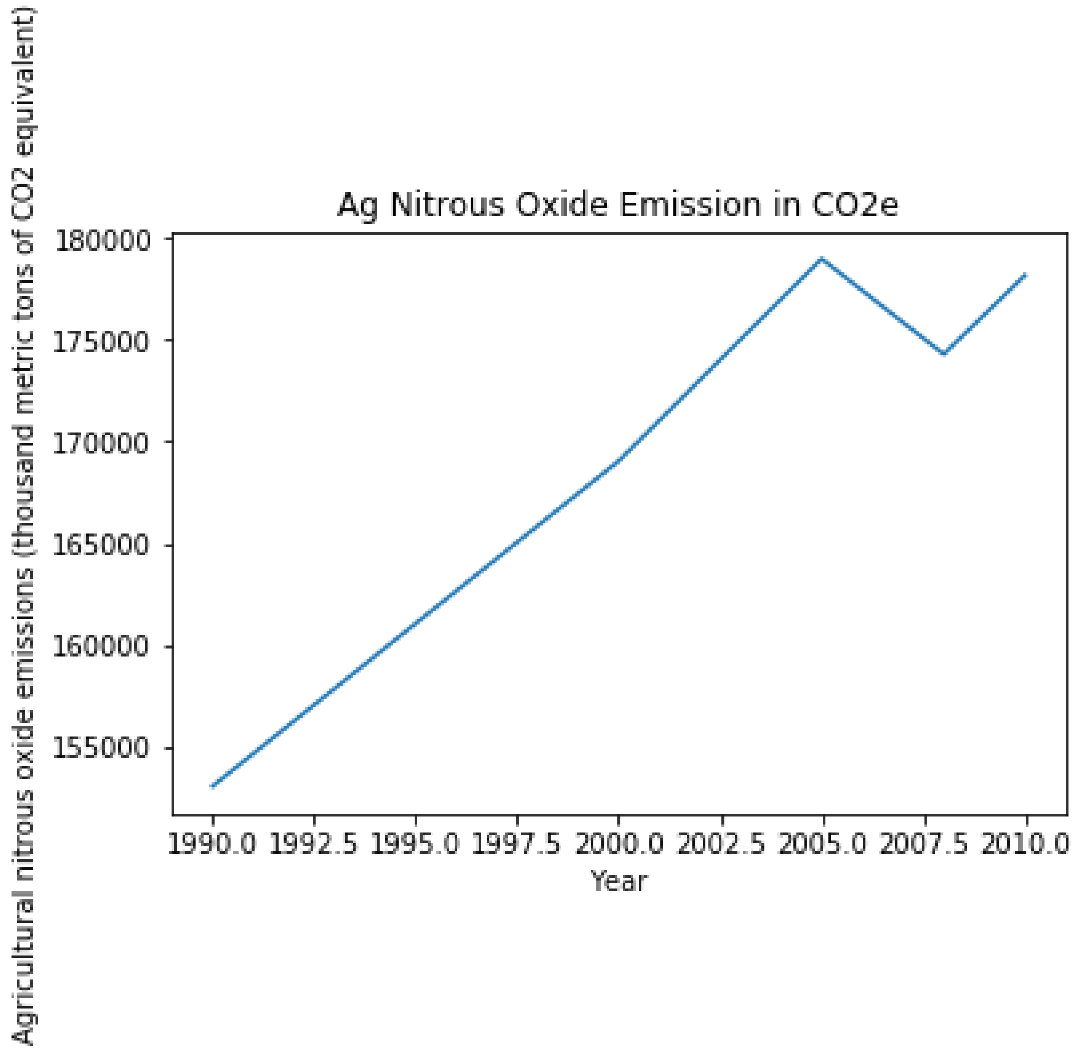
- Agricultural nitrous oxide emissions (% of total)
- Nitrous oxide emissions (thousand metric tons of CO2 equivalent)
- Nitrous oxide emissions in industrial and energy processes (% of total nitrous oxide emissions)

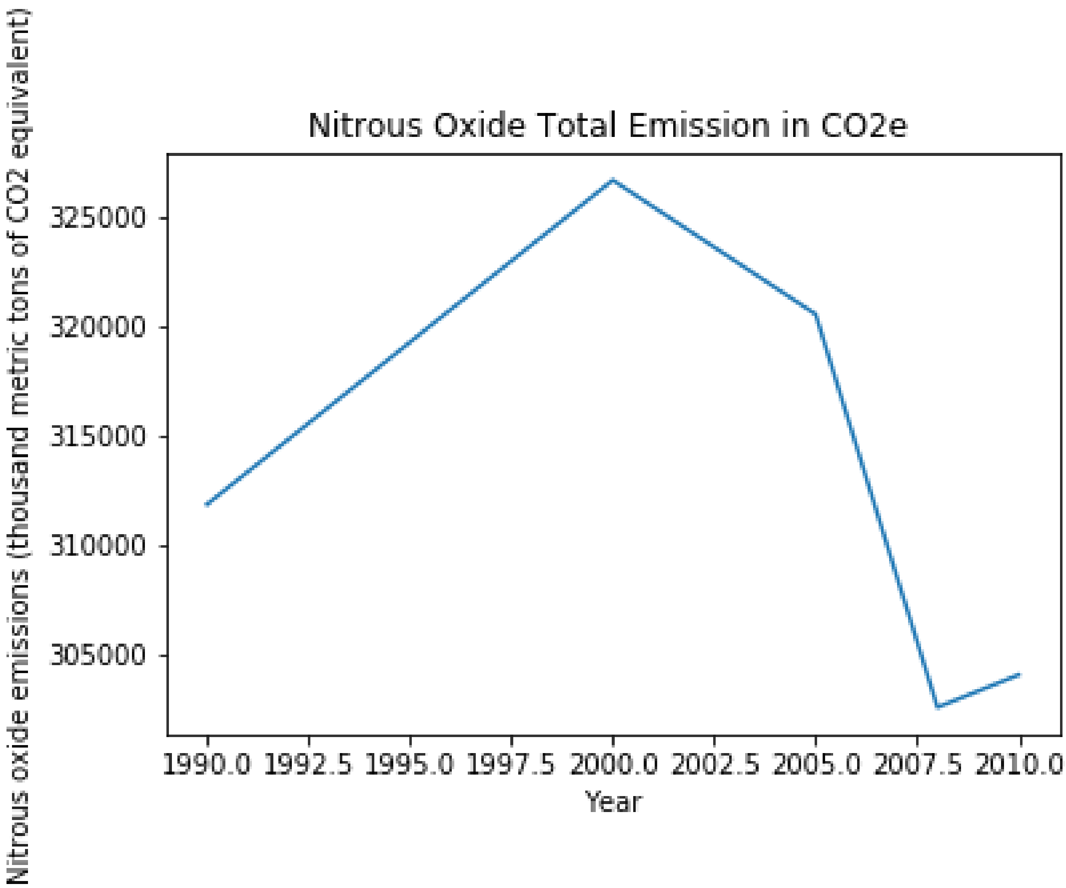
```
In [30]: NO_per = 'Agricultural nitrous oxide emissions \(thousand'
         NO_tot = 'Nitrous oxide emissions \(thousand metric tons of CO2 equivalent'
         NO_indandenergy = 'Nitrous oxide emissions in industrial and energy processes'

         mask_NO_per = data_env['IndicatorName'].str.contains(NO_per)
         o_NO_per = data_env[mask_NO_per & mask_country]
         print("Ag Nitrous Oxide Emission in Percent = ", o_NO_per['Year'].min(), "max: ", o_N
         print(len(o_NO_per))

         mask_NO_tot = data_env['IndicatorName'].str.contains(NO_tot)
         o_NO_tot = data_env[mask_NO_tot & mask_country]
         print("Nitrous Oxide Total Emission in CO2e = ", o_NO_tot['Year'].min(), "max: ", o_N
         print(len(o_NO_tot))
```

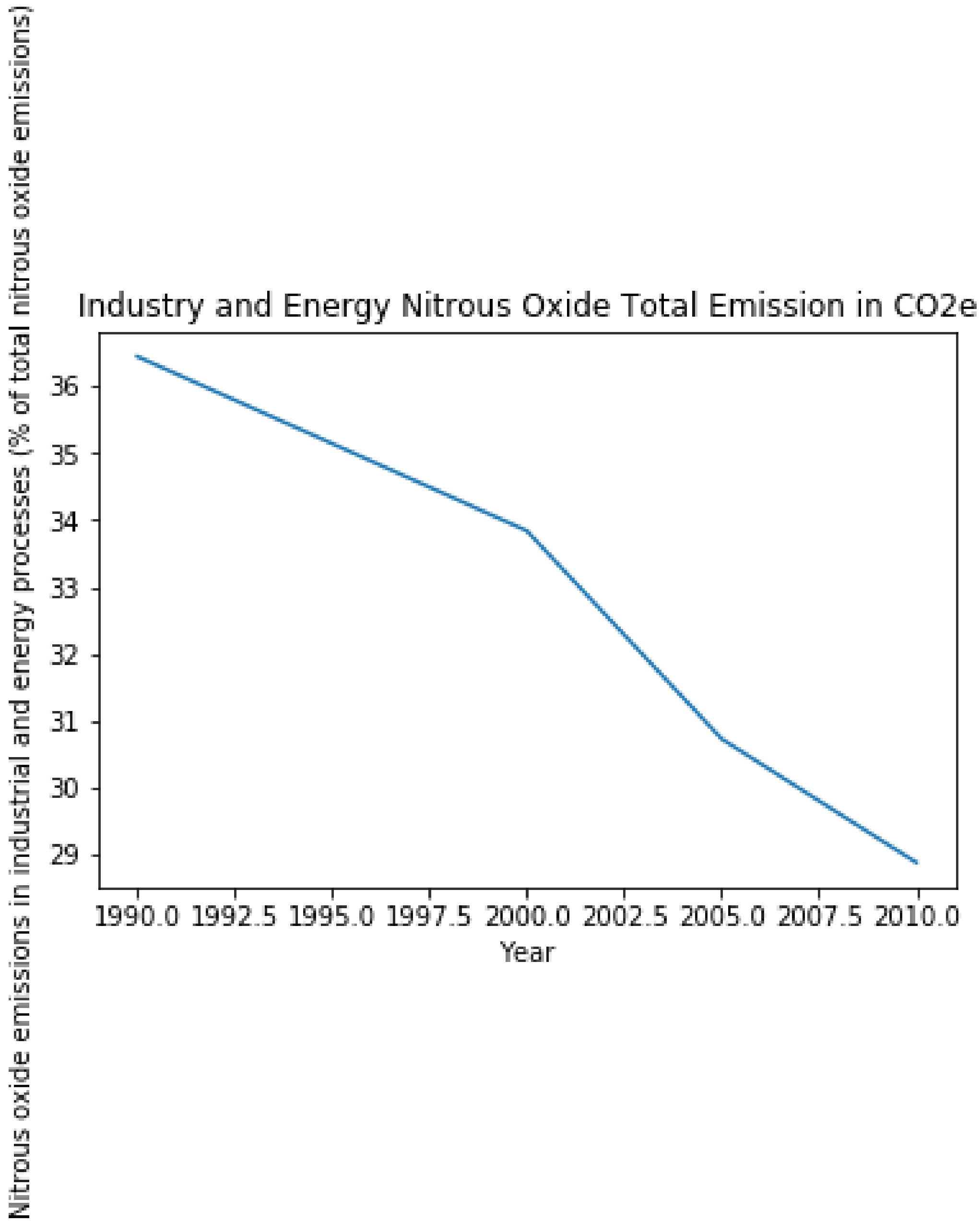
```
mask_NO_indandenergy = data_env['IndicatorName'].str.contains(NO_indandenergy)
o_NO_indandenergy = data_env[mask_NO_indandenergy & mask_country]
print("Industry and Energy Nitrous Oxide Total Emission in CO2e = ", o_NO_indandenerg
print(len(o_NO_indandenergy))
```

```
Ag Nitrous Oxide Emission in Percent =  1990 max:  2010
5
Nitrous Oxide Total Emission in CO2e =  1990 max:  2010
5
Industry and Energy Nitrous Oxide Total Emission in CO2e =  1990 max:  2010
5
```

In [31]:
```
plt.plot(o_NO_per['Year'].values, o_NO_per['Value'].values)
plt.xlabel('Year')
plt.ylabel(o_NO_per['IndicatorName'].iloc[0])
plt.title('Ag Nitrous Oxide Emission in CO2e')
plt.show()

plt.plot(o_NO_tot['Year'].values, o_NO_tot['Value'].values)
plt.xlabel('Year')
plt.ylabel(o_NO_tot['IndicatorName'].iloc[0])
plt.title('Nitrous Oxide Total Emission in CO2e ')
plt.show()

plt.plot(o_NO_indandenergy['Year'].values, o_NO_indandenergy['Value'].values)
plt.xlabel('Year')
plt.ylabel(o_NO_indandenergy['IndicatorName'].iloc[0])
plt.title('Industry and Energy Nitrous Oxide Total Emission in CO2e')
plt.show()
```


Ag Nitrous Oxide Emission in CO2e
Agricultural nitrous oxide emissions (thousand metric tons of CO2 equivalent)
180000
175000
170000
165000
160000
155000
1990.0
1992.5
1995.0
1997.5
2000.0
2002.5
2005.0
2007.5
2010.0
Year


Nitrous Oxide Total Emission in CO2e
Nitrous oxide emissions (thousand metric tons of CO2 equivalent)
325000
320000
315000
310000
305000
1990.0
1992.5
1995.0
1997.5
2000.0
2002.5
2005.0
2007.5
2010.0
Year



Nitrous Oxide Emissions show similar trend as Methane.

0.3 Positive Correlation Across the Globe

Is this consistent across different countries for CO2 emissions? Let's look at the correlation factor in a map form.

The goal: 1) For each country, find the max and min year into table. 2) Use data_env data frame. 3) Bring the carbon onto another column. 4) Create a new table with the country and year. 5) Filter out the year: 1961 to 2012

Already existed: mask_livestock_index and mask_carbon

```
In [32]: data_carbon = data_env[mask_carbon]
         data_livestock = data_env[mask_livestock_index]

In [33]: #create new key, CountryName and Year
         data_carbon['CountryNameYear'] = data_carbon['CountryName'].map(str) + ' ' + data_carl
         data_livestock['CountryNameYear'] = data_livestock['CountryName'].map(str) + ' ' + da

C:\Users\-\Anaconda3\lib\site-packages\ipykernel_launcher.py:2: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html

C:\Users\-\Anaconda3\lib\site-packages\ipykernel_launcher.py:3: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html
  This is separate from the ipykernel package so we can avoid doing imports until


In [34]: #rename column
         data_livestock.rename(columns={'Value': 'Livestock Index'}, inplace=True)
         data_carbon.rename(columns={'Value': 'Carbon Emission kt'}, inplace = True)

C:\Users\-\Anaconda3\lib\site-packages\pandas\core\frame.py:4025: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: http://pandas.pydata.org/pandas-docs/stable/indexing.html
  return super(DataFrame, self).rename(**kwargs)


In [35]: data_cls = pd.merge(data_carbon,data_livestock[['CountryNameYear','Livestock Index']]

In [36]: data_cls.head()

Out[36]:                                 CountryName CountryCode       IndicatorName  \
         0                                Arab World         ARB  CO2 emissions (kt)
         1                    Caribbean small states         CSS  CO2 emissions (kt)
         2  East Asia & Pacific (all income levels)         EAS  CO2 emissions (kt)
         3    East Asia & Pacific (developing only)         EAP  CO2 emissions (kt)
         4                                 Euro area         EMU  CO2 emissions (kt)

             IndicatorCode  Year  Carbon Emission kt                   Topic  \
         0  EN.ATM.CO2E.KT  1961        6.515110e+04  Environment: Emissions
         1  EN.ATM.CO2E.KT  1961        8.804467e+03  Environment: Emissions
         2  EN.ATM.CO2E.KT  1961        1.042288e+06  Environment: Emissions
```

```
3  EN.ATM.CO2E.KT  1961        6.240318e+05  Environment: Emissions
4  EN.ATM.CO2E.KT  1961        1.190487e+06  Environment: Emissions

                                CountryNameYear  Livestock Index
0                                  Arab World 1961        20.525848
1                        Caribbean small states 1961        26.634305
2  East Asia & Pacific (all income levels) 1961        12.690781
3    East Asia & Pacific (developing only) 1961         6.979030
4                                   Euro area 1961        58.581630
```

In [251]:
```
top10_raw = data_cls.groupby('CountryName')['Carbon Emission kt'].nlargest(3).sum(le
avoid_list = ['Low & middle income', 'Lower middle income', 'Middle income', 'High i
             'East Asia & Pacific (developing only)', 'Europe & Central Asia (all in
             'European Union','Latin America & Caribbean (all income levels)','North
             'Middle East & North Africa (developing only)','Latin America & Caribbe
#top10 = top10_raw[~top10_raw['CountryName'].isin(avoid_list)]
#top10
```

In [252]:
```
top10_countries = top10_raw.index.values
type(top10_countries)
top10_c = top10_countries.tolist()
if avoid_list in top10_c: top10_c.remove(avoid_list)
top10_c1 = [x for x in top10_c if x not in avoid_list]
```

In [253]:
```
top10_c1
```

Out[253]:
```
['World',
 'China',
 'United States',
 'Euro area',
 'South Asia',
 'India',
 'Russian Federation',
 'Arab World',
 'Japan',
 'Germany']
```

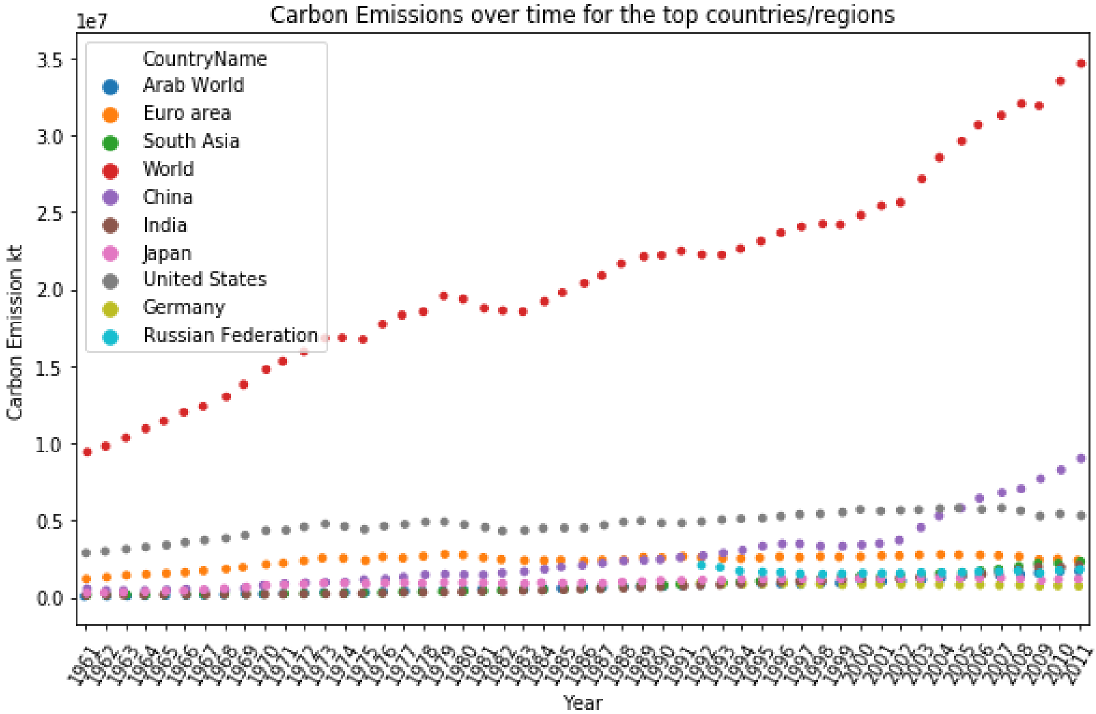
In [254]:
```
#mask_top5 = data_cls['CountryName'].str.contains(top10_c)

filterMesh = data_cls['CountryName'].isin(top10_c1)
country1_data = data_cls.loc[filterMesh]
len(country1_data)
```

Out[254]:
```
449
```

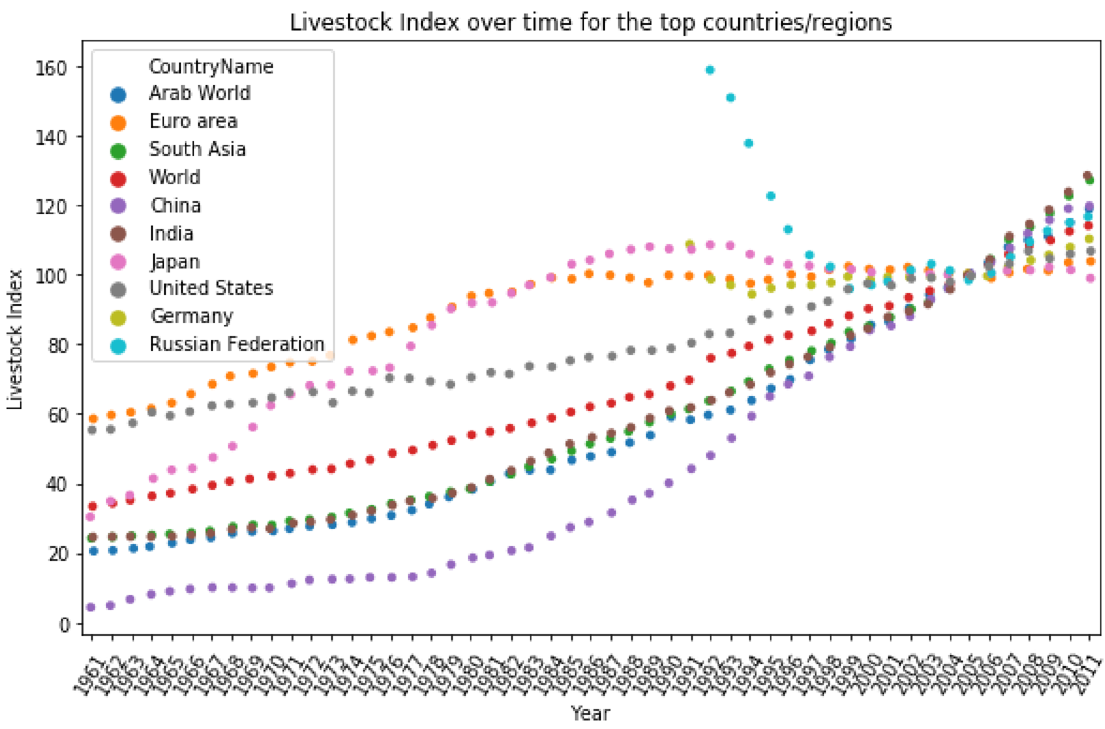
In [41]:
```
import seaborn as sns
```

In [255]:
```
axis.yaxis.grid(True)
plt.figure(figsize=(10,6),dpi=70)
```

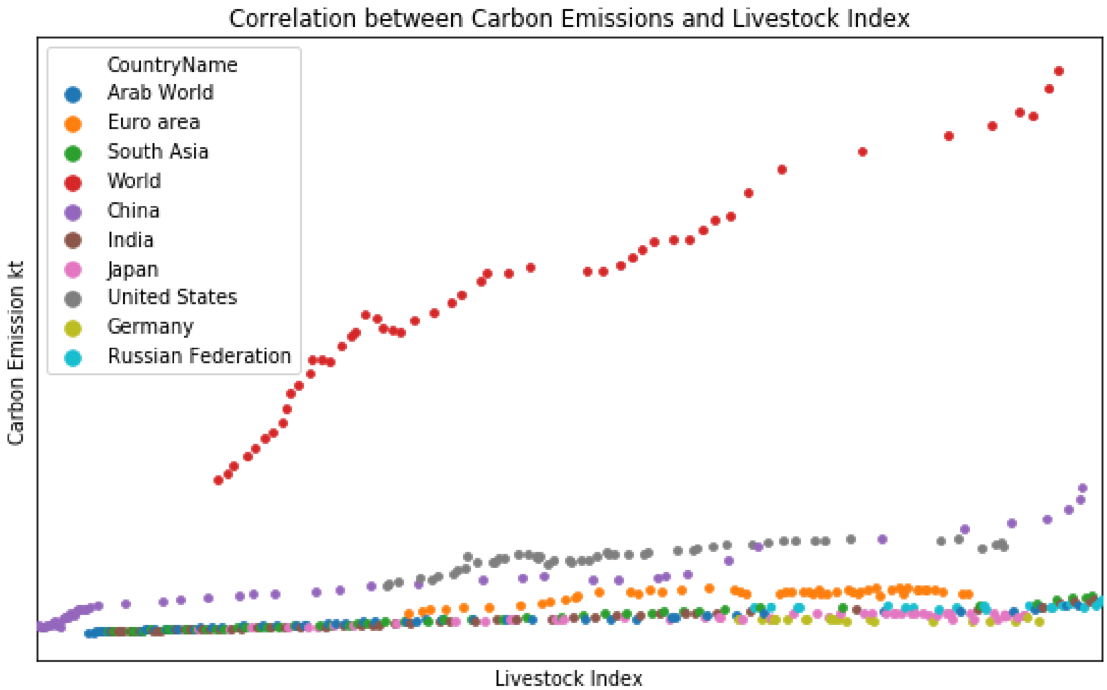
```
plt.xticks(np.arange(1960,2020,10), rotation=60)
plt.tick_params(which='minor',length=10)
plt.title('Carbon Emissions over time for the top countries/regions')
axi = sns.stripplot(x='Year', y='Carbon Emission kt', data=country1_data, hue='Count
#plt.xticks(rotation=60)
```




```
In [256]: plt.figure(figsize=(10,6),dpi=70)
          plt.xticks(np.arange(1960,2020,10))
          plt.xticks(rotation=60)
          plt.title('Livestock Index over time for the top countries/regions')
          sns.stripplot(x='Year', y='Livestock Index', data=country1_data, hue='CountryName')

Out[256]: <matplotlib.axes._subplots.AxesSubplot at 0x1f78aca8048>
```




```
In [257]: plt.figure(figsize=(10,6),dpi=70)
          sns.stripplot(x='Livestock Index', y='Carbon Emission kt', data=country1_data, hue='C
          plt.yticks([])
          plt.xticks([])
          plt.title("Correlation between Carbon Emissions and Livestock Index")

Out[257]: Text(0.5, 1.0, 'Correlation between Carbon Emissions and Livestock Index')
```




```
In [45]: np.corrcoef(data_cls['Livestock Index'], data_cls['Carbon Emission kt'])[0, 1]

Out[45]: 0.022196825264866528

In [47]: data_cls.groupby('CountryName')[['Livestock Index','Carbon Emission kt']].corr()

Out[47]:                                        Livestock Index  Carbon Emission kt
         CountryName
         Afghanistan         Livestock Index           1.000000            0.243237
                             Carbon Emission kt        0.243237            1.000000
         Albania             Livestock Index           1.000000           -0.176606
                             Carbon Emission kt       -0.176606            1.000000
         Algeria             Livestock Index           1.000000            0.948630
                             Carbon Emission kt        0.948630            1.000000
         Angola              Livestock Index           1.000000            0.877841
                             Carbon Emission kt        0.877841            1.000000
         Antigua and Barbuda Livestock Index           1.000000           -0.188851
                             Carbon Emission kt       -0.188851            1.000000
         Arab World          Livestock Index           1.000000            0.986659
                             Carbon Emission kt        0.986659            1.000000
         Argentina           Livestock Index           1.000000            0.916898
                             Carbon Emission kt        0.916898            1.000000
         Armenia             Livestock Index           1.000000            0.892630
                             Carbon Emission kt        0.892630            1.000000
         Australia           Livestock Index           1.000000            0.940243
```

```
                     Carbon Emission kt        0.940243           1.000000
Austria              Livestock Index           1.000000           0.820021
                     Carbon Emission kt        0.820021           1.000000
Azerbaijan           Livestock Index           1.000000          -0.250122
                     Carbon Emission kt       -0.250122           1.000000
Bahamas, The         Livestock Index           1.000000          -0.158869
                     Carbon Emission kt       -0.158869           1.000000
Bahrain              Livestock Index           1.000000           0.769745
                     Carbon Emission kt        0.769745           1.000000
Bangladesh           Livestock Index           1.000000           0.986036
                     Carbon Emission kt        0.986036           1.000000
Barbados             Livestock Index           1.000000           0.914791
                     Carbon Emission kt        0.914791           1.000000
...                                                 ...                ...
Ukraine              Livestock Index           1.000000           0.962555
                     Carbon Emission kt        0.962555           1.000000
United Arab Emirates Livestock Index           1.000000           0.949820
                     Carbon Emission kt        0.949820           1.000000
United Kingdom       Livestock Index           1.000000          -0.560015
                     Carbon Emission kt       -0.560015           1.000000
United States        Livestock Index           1.000000           0.891535
                     Carbon Emission kt        0.891535           1.000000
Upper middle income  Livestock Index           1.000000           0.975258
                     Carbon Emission kt        0.975258           1.000000
Uruguay              Livestock Index           1.000000           0.502872
                     Carbon Emission kt        0.502872           1.000000
Uzbekistan           Livestock Index           1.000000          -0.095289
                     Carbon Emission kt       -0.095289           1.000000
Vanuatu              Livestock Index           1.000000           0.550094
                     Carbon Emission kt        0.550094           1.000000
Venezuela, RB        Livestock Index           1.000000           0.960215
                     Carbon Emission kt        0.960215           1.000000
Vietnam              Livestock Index           1.000000           0.965961
                     Carbon Emission kt        0.965961           1.000000
West Bank and Gaza   Livestock Index           1.000000           0.421471
                     Carbon Emission kt        0.421471           1.000000
World                Livestock Index           1.000000           0.975142
                     Carbon Emission kt        0.975142           1.000000
Yemen, Rep.          Livestock Index           1.000000           0.954895
                     Carbon Emission kt        0.954895           1.000000
Zambia               Livestock Index           1.000000          -0.630333
                     Carbon Emission kt       -0.630333           1.000000
Zimbabwe             Livestock Index           1.000000           0.285804
                     Carbon Emission kt        0.285804           1.000000

[454 rows x 2 columns]
```

```
In [52]: data_country = data_cls.groupby('CountryName')[['Livestock Index','Carbon Emission kt
```

```
In [85]: data_country_df = pd.DataFrame(data_country.values,data_country.index, columns=['Corr
```

```
In [86]: data_country_df
```

```
Out[86]:               CountryName          level_1   CorrVal
         0            Afghanistan  Livestock Index  0.243237
         1                Albania  Livestock Index -0.176606
         2                Algeria  Livestock Index  0.948630
         3                 Angola  Livestock Index  0.877841
         4        Antigua and Barbuda  Livestock Index -0.188851
         5             Arab World  Livestock Index  0.986659
         6              Argentina  Livestock Index  0.916898
         7                Armenia  Livestock Index  0.892630
         8              Australia  Livestock Index  0.940243
         9                Austria  Livestock Index  0.820021
         10            Azerbaijan  Livestock Index -0.250122
         11          Bahamas, The  Livestock Index -0.158869
         12               Bahrain  Livestock Index  0.769745
         13            Bangladesh  Livestock Index  0.986036
         14              Barbados  Livestock Index  0.914791
         15               Belarus  Livestock Index  0.689344
         16               Belgium  Livestock Index -0.123015
         17                Belize  Livestock Index  0.824802
         18                 Benin  Livestock Index  0.838583
         19               Bermuda  Livestock Index -0.788773
         20                Bhutan  Livestock Index  0.836944
         21               Bolivia  Livestock Index  0.958913
         22  Bosnia and Herzegovina  Livestock Index  0.745259
         23              Botswana  Livestock Index  0.585566
         24                Brazil  Livestock Index  0.967688
         25     Brunei Darussalam  Livestock Index  0.405880
         26              Bulgaria  Livestock Index  0.895944
         27          Burkina Faso  Livestock Index  0.936150
         28               Burundi  Livestock Index  0.313439
         29            Cabo Verde  Livestock Index  0.789659
         ..                   ...              ...       ...
         197            Swaziland  Livestock Index  0.215325
         198               Sweden  Livestock Index -0.402177
         199          Switzerland  Livestock Index  0.702987
         200    Syrian Arab Republic  Livestock Index  0.941789
         201           Tajikistan  Livestock Index  0.319078
         202             Tanzania  Livestock Index  0.878476
         203             Thailand  Livestock Index  0.953034
         204          Timor-Leste  Livestock Index -0.001147
         205                 Togo  Livestock Index  0.914880
         206                Tonga  Livestock Index  0.884595
         207      Trinidad and Tobago  Livestock Index  0.912612
         208              Tunisia  Livestock Index  0.979853
```

```
209                  Turkey  Livestock Index  0.973415
210            Turkmenistan  Livestock Index  0.956560
211                  Uganda  Livestock Index  0.889811
212                 Ukraine  Livestock Index  0.962555
213    United Arab Emirates  Livestock Index  0.949820
214          United Kingdom  Livestock Index -0.560015
215           United States  Livestock Index  0.891535
216     Upper middle income  Livestock Index  0.975258
217                 Uruguay  Livestock Index  0.502872
218              Uzbekistan  Livestock Index -0.095289
219                 Vanuatu  Livestock Index  0.550094
220           Venezuela, RB  Livestock Index  0.960215
221                 Vietnam  Livestock Index  0.965961
222      West Bank and Gaza  Livestock Index  0.421471
223                   World  Livestock Index  0.975142
224             Yemen, Rep.  Livestock Index  0.954895
225                  Zambia  Livestock Index -0.630333
226                Zimbabwe  Livestock Index  0.285804

[227 rows x 3 columns]
```

```
In [120]: test_index = data_country_df[data_country_df['CountryName'] == 'United States'].inde
          print("Double check US correlation: " , data_country_df.loc[test_index,'CorrVal'].va
```

```
Double check US correlation:  [0.89153464]
```

```
In [258]: #Get the top countries/region table
          top10_c1
```

```
Out[258]: ['World',
           'China',
           'United States',
           'Euro area',
           'South Asia',
           'India',
           'Russian Federation',
           'Arab World',
           'Japan',
           'Germany']
```

```
In [283]: mask_top10 = data_country_df['CountryName'].isin(top10_c1)
          data_top10c = data_country_df[mask_top10]
          data_top10c.reset_index()
          data_top10c[['CountryName','CorrVal']].sort_values('CorrVal',ascending=False)
```

```
Out[283]:            CountryName   CorrVal
          187         South Asia  0.996246
          94               India  0.995452
```

```
5            Arab World  0.986659
223               World  0.975142
39                China  0.955384
102               Japan  0.914964
215       United States  0.891535
63            Euro area  0.864608
171  Russian Federation  0.861487
76              Germany -0.546637
```

In [134]:
```
# Map the correlation on the map using Folium
import folium
country_geo = 'geo/world-countries.json'
```

In [123]:
```
data_livestock.head()
```

Out[123]:
```
                                          CountryName CountryCode  \
2168017                                    Arab World         ARB
2168018                          Caribbean small states         CSS
2168019  East Asia & Pacific (all income levels)         EAS
2168020    East Asia & Pacific (developing only)         EAP
2168021                                     Euro area         EMU

                                         IndicatorName   IndicatorCode  Year  \
2168017  Livestock production index (2004-2006 = 100)  AG.PRD.LVSK.XD  1961
2168018  Livestock production index (2004-2006 = 100)  AG.PRD.LVSK.XD  1961
2168019  Livestock production index (2004-2006 = 100)  AG.PRD.LVSK.XD  1961
2168020  Livestock production index (2004-2006 = 100)  AG.PRD.LVSK.XD  1961
2168021  Livestock production index (2004-2006 = 100)  AG.PRD.LVSK.XD  1961

         Livestock Index                                 Topic  \
2168017        20.525848  Environment: Agricultural production
2168018        26.634305  Environment: Agricultural production
2168019        12.690781  Environment: Agricultural production
2168020         6.979030  Environment: Agricultural production
2168021        58.581630  Environment: Agricultural production

                                          CountryNameYear
2168017                                   Arab World 1961
2168018                         Caribbean small states 1961
2168019  East Asia & Pacific (all income levels) 1961
2168020    East Asia & Pacific (developing only) 1961
2168021                                    Euro area 1961
```

In [124]:
```
# Need to add Country Code to data_country_df
data_corrcountry_df = pd.merge(data_country_df,data_livestock[['CountryName','Country
```

In [127]:
```
data_corrcountry_df = data_corrcountry_df.drop_duplicates(subset='CountryName')
```

In [128]:
```
data_corrcountry_df
```

Out[128]:

	CountryName	level_1	CorrVal	CountryCode
0	Afghanistan	Livestock Index	0.243237	AFG
53	Albania	Livestock Index	-0.176606	ALB
106	Algeria	Livestock Index	0.948630	DZA
159	Angola	Livestock Index	0.877841	AGO
212	Antigua and Barbuda	Livestock Index	-0.188851	ATG
265	Arab World	Livestock Index	0.986659	ARB
318	Argentina	Livestock Index	0.916898	ARG
371	Armenia	Livestock Index	0.892630	ARM
412	Australia	Livestock Index	0.940243	AUS
465	Austria	Livestock Index	0.820021	AUT
518	Azerbaijan	Livestock Index	-0.250122	AZE
559	Bahamas, The	Livestock Index	-0.158869	BHS
612	Bahrain	Livestock Index	0.769745	BHR
665	Bangladesh	Livestock Index	0.986036	BGD
718	Barbados	Livestock Index	0.914791	BRB
771	Belarus	Livestock Index	0.689344	BLR
812	Belgium	Livestock Index	-0.123015	BEL
837	Belize	Livestock Index	0.824802	BLZ
890	Benin	Livestock Index	0.838583	BEN
943	Bermuda	Livestock Index	-0.788773	BMU
996	Bhutan	Livestock Index	0.836944	BTN
1049	Bolivia	Livestock Index	0.958913	BOL
1102	Bosnia and Herzegovina	Livestock Index	0.745259	BIH
1143	Botswana	Livestock Index	0.585566	BWA
1196	Brazil	Livestock Index	0.967688	BRA
1249	Brunei Darussalam	Livestock Index	0.405880	BRN
1302	Bulgaria	Livestock Index	0.895944	BGR
1355	Burkina Faso	Livestock Index	0.936150	BFA
1408	Burundi	Livestock Index	0.313439	BDI
1461	Cabo Verde	Livestock Index	0.789659	CPV
...	...	...	...	...
9911	Swaziland	Livestock Index	0.215325	SWZ
9964	Sweden	Livestock Index	-0.402177	SWE
10017	Switzerland	Livestock Index	0.702987	CHE
10070	Syrian Arab Republic	Livestock Index	0.941789	SYR
10123	Tajikistan	Livestock Index	0.319078	TJK
10145	Tanzania	Livestock Index	0.878476	TZA
10198	Thailand	Livestock Index	0.953034	THA
10251	Timor-Leste	Livestock Index	-0.001147	TMP
10304	Togo	Livestock Index	0.914880	TGO
10357	Tonga	Livestock Index	0.884595	TON
10410	Trinidad and Tobago	Livestock Index	0.912612	TTO
10463	Tunisia	Livestock Index	0.979853	TUN
10516	Turkey	Livestock Index	0.973415	TUR
10569	Turkmenistan	Livestock Index	0.956560	TKM
10591	Uganda	Livestock Index	0.889811	UGA
10644	Ukraine	Livestock Index	0.962555	UKR

```
10666    United Arab Emirates  Livestock Index  0.949820         ARE
10719          United Kingdom  Livestock Index -0.560015         GBR
10772           United States  Livestock Index  0.891535         USA
10825     Upper middle income  Livestock Index  0.975258         UMC
10878                 Uruguay  Livestock Index  0.502872         URY
10931              Uzbekistan  Livestock Index -0.095289         UZB
10953                 Vanuatu  Livestock Index  0.550094         VUT
11006           Venezuela, RB  Livestock Index  0.960215         VEN
11059                 Vietnam  Livestock Index  0.965961         VNM
11112      West Bank and Gaza  Livestock Index  0.421471         WBG
11149                   World  Livestock Index  0.975142         WLD
11202             Yemen, Rep.  Livestock Index  0.954895         YEM
11255                  Zambia  Livestock Index -0.630333         ZMB
11308                Zimbabwe  Livestock Index  0.285804         ZWE

[227 rows x 4 columns]
```

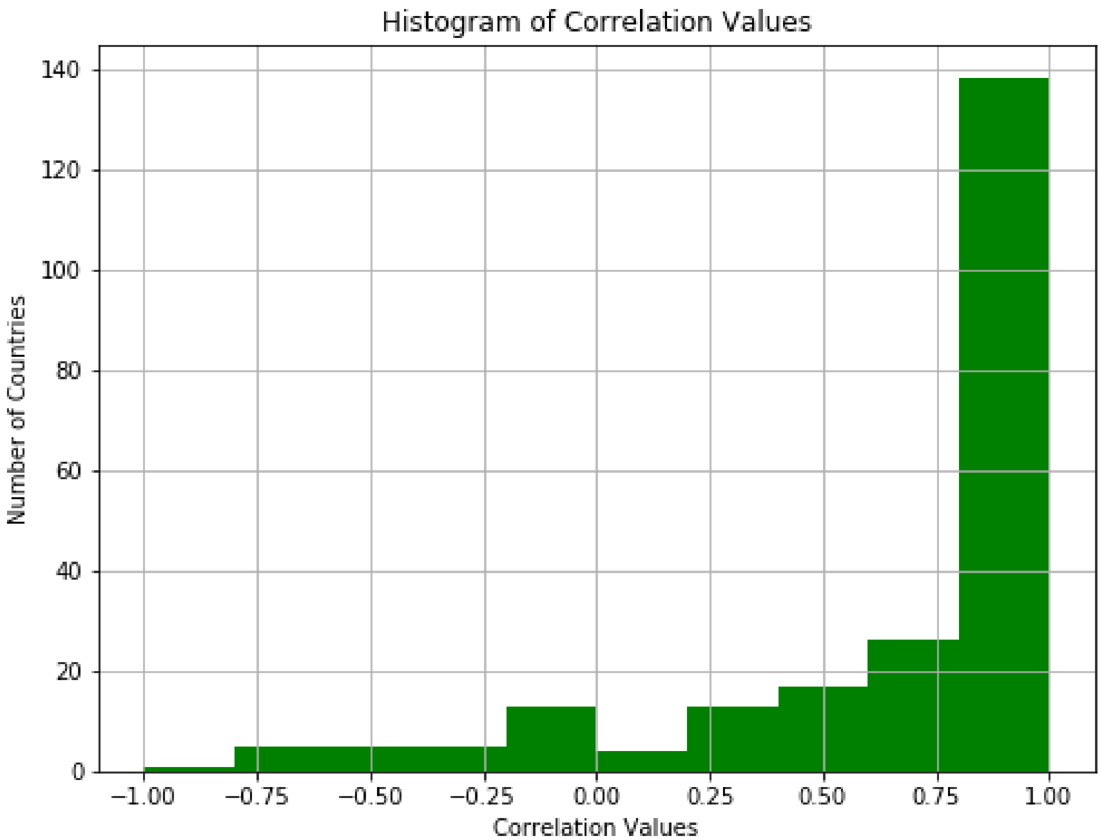
In [244]:
```python
# Setup a folium map at a high-level zoom @Alok - what is the 100,0, doesn't seem li
map = folium.Map(location=[-25, 0], zoom_start=1.5)



# choropleth maps bind Pandas Data Frames and json geometries.  This allows us to qu
map.choropleth(geo_data=country_geo, data=data_corrcountry_df,
             columns=['CountryCode', 'CorrVal'],
             key_on='feature.id',
             nan_fill_color='white',
             fill_color='YlGnBu', fill_opacity=0.7, line_opacity=0.2)

# Create Folium plot
map.save('plot_data.html')

# Import the Folium interactive html file
from IPython.display import HTML
HTML('<iframe src=plot_data.html width=900 height=650></iframe>')
```

```
C:\Users\-\Anaconda3\lib\site-packages\IPython\core\display.py:689: UserWarning: Consider using
  warnings.warn("Consider using IPython.display.IFrame instead")
```

Out[244]: <IPython.core.display.HTML object>

In [298]:
```python
# the histogram of the data
plt.figure(figsize=(8,6))
plt.hist(data_corrcountry_df['CorrVal'].values, 10, density=False, facecolor='green'
        range=(-1,1))

plt.xlabel('Correlation Values')
plt.ylabel('Number of Countries')
```

```
plt.title('Histogram of Correlation Values')
plt.grid(True)

plt.show()
```




```
In [ ]:
```