Database for CO_2 emission

SIHAN CAO* and ADITYA BIYANI*, College of Engineering, Carnegie Mellon University, USA

Carbon-dioxide is the most commonly occurring gas in the atmosphere. Due to the increase in anthropogenic activities, there is an imbalance in nature and human activities. Carbon-dioxide emissions have become the primary source for this imbalance and a shift in global climate cycle. Countries all around the globe are taking initiatives to become carbon-negative. Technologies are being developed for carbon capture and sequestration. Thus, it has become all the more important to monitor this activity now than ever. This study involves creating a database for CO2 emissions for every state in the United States of America over the span of 3 decades from 1980 to 2017. The database helps in analysing the carbon intensity, energy related emissions and carbon intensity change with the change in economy.

Additional Key Words and Phrases: datasets, neural networks, gaze detection, text tagging

ACM Reference Format:
Sihan Cao and Aditya Biyani. 2020. Database for CO_2 emission. 1, 1 (December 2020), 6 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 MOTIVATION

The population of the United States has grown by an average of 35% over 3 decades since 1930 (Kelvin et al., 2020) (see Figure 1). Although the United States contributes to only 4.3 % of the total world population, it has been responsible for 90% of CO2 emissions along with Europe and China. The emissions started to increase after the industrial revolution and remained unchecked for decades until the 21st century and the growing awareness for environment protection and reduction in emissions.

The total emissions as seen in figure 2 have increased 2 fold in the past two decades to an average of alarming 500 MMT. Since, the USA contributes to over 33% of the total emissions, it is important to understand the source of these emissions in the country and how it has contributed to the country's growing economy and population.

*Both authors contributed equally to this research.

Authors' address: Sihan Cao, sihancao@andrew.cmu.edu; Aditya Biyani, abiyani@andrew.cmu.edu, College of Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, Pennsylvania, USA, 15289.

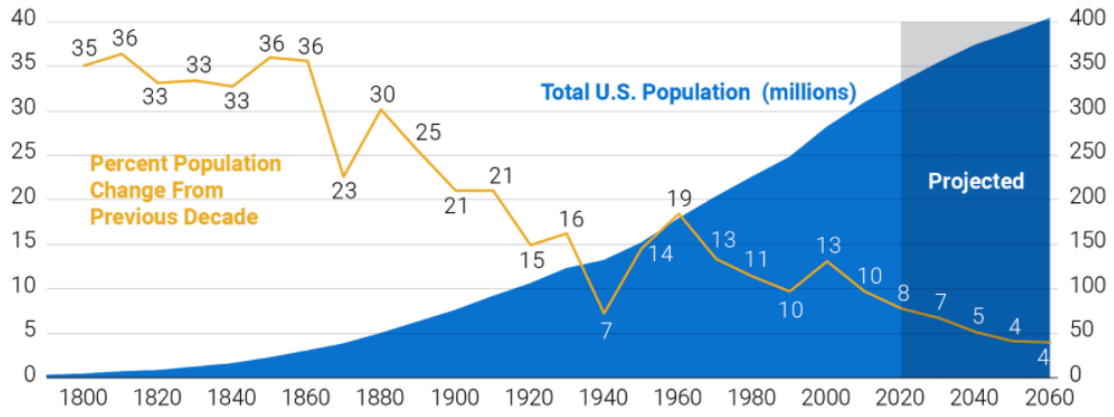

Fig. 1. Source: U.S. Census Bureau, decennial censuses, and vintage 2018 population estimates

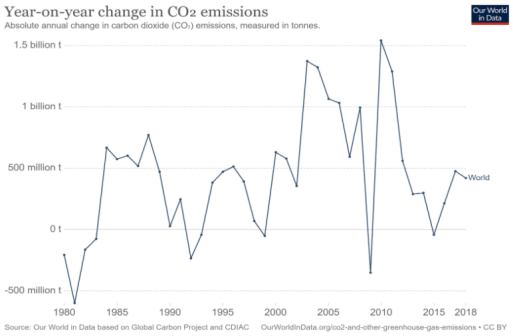

Fig. 2. CO_2 emissions change year in year

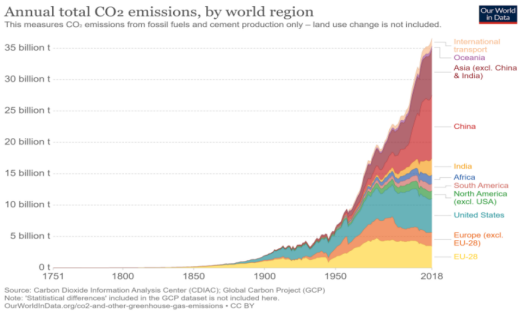

Fig. 3. CO_2 emission, world wide

2 SPECIFIC OBJECTIVE

In this project, we hope to build the database can integrate the information about CO_2 emission on the EIA website together. So we can easily compare and analyze the data between originally from different states, and find the pattern, e.g. investigate the relationship between how total energy related emission is influenced by industrial emission along with time. The task part include some examples that we can take advantage of this database.

3 DATA SOURCE

The data is sourced from the U.S. Energy Information Administration (Administration Energy Information, 2020). The tables used to create our database are as follows:

- Energy related emissions and per capita energy related emissions - emissions released at the locations where fossil fuels were used. They vary significantly from state to state on both absolute and per capita basis. (Units: MMT and MMT per million)

- Energy Intensity by state: Is the amount of energy produced by every state per unit the GDP of the country. (Units: thousand Btu per chained 2009 dollar of GDP)
- Carbon intensity of energy supply per state: For every unit of Energy produced how much carbon dioxide emissions were recorded. (Units: kilograms of energy-related carbon dioxide per million Btu)
- Carbon intensity of the economy by state: It is the amount of CO2 emissions recorded for every dollar of the US economy generated (Units: metric tons of energy-related carbon dioxide per chained 2009 million dollars of GDP).
- Energy Emissions by sector: The energy related emissions are due to use fossil fuels like coal, natural gas and petroleum in different sectors like Transportation,
- How has carbon intensity changed with the carbon economy? Residential, Commercial Industrial and Electric Power. (Units: MMT)

4 TASK

To design the database, a ER diagram was developed:

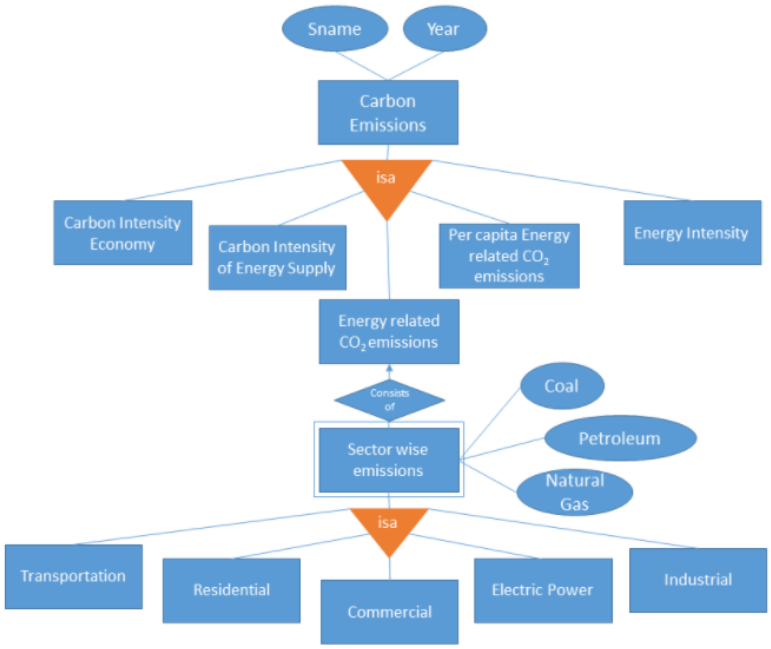

Fig. 4. ER diagram

To build the database, we first use pandas package in Python language to process the data, including transform and classify it into different sector category. Then we use the import wizard in Mysql Workbench to create the data table, change the column table use query (see the "type transfer.sql"),and use query to define the primary key and foreign key for each table (see the detials in "primary key.sql" and "foreign key.sql").

Then, we mainly excute query to test the performance of the database as below:

- Use query to obtain different sectors data for industrial and compare it to the energy related data, see the query script in "select.sql"
- How has energy related consumption evolved with the increase in population over 3 decades?
- How do different sectors contribute to energy consumption in a particular state?

5 SUMMARY

First query try to compare the different sectors' contributions to the total CO_2 emission, from 1990 to 2017. This example is for California only, but can be changed for any state conveniently:

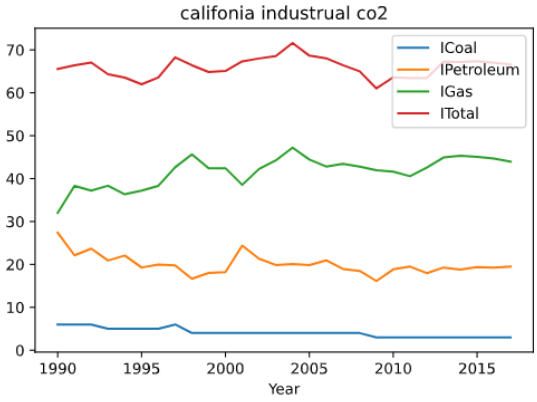

Fig. 5. Different sectors' contribution to California CO_2 emission, from 1990 to 2017

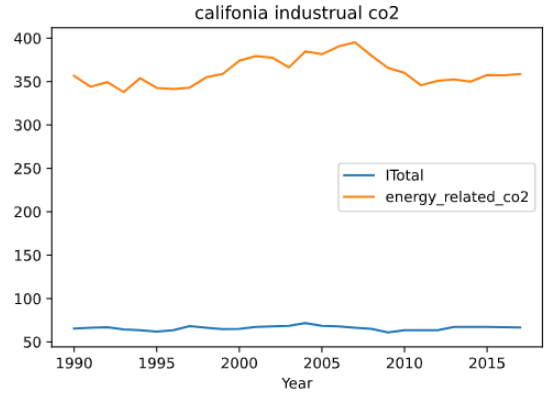

Fig. 6. Industrial sector CO_2 emission versus total CO_2 emission of California

The query costs 0.11 second and returned 28 messages.And it combines the information between different two tables.

The second query can be found in "consump_popu.sql". For this query, the output was imported into excel for visualization, see result in fig7. Per capita energy consumption for the state of pennsylvania over three decades has decreased from 21 MMT/ million to 15 MMT/ million. This can be explained because the major contributors to CO2 emissions were the cities of Philadelphia and Pittsburgh (aka as the steel capital of USA) after the industrial revolution and shift of steel industry from Pittsburgh, these cities became more habitable and less manufacturing industry loveable. Thus, the emissions reduced. Trends like these can be plotted for other states to analyse their history of emissions as the population grew.

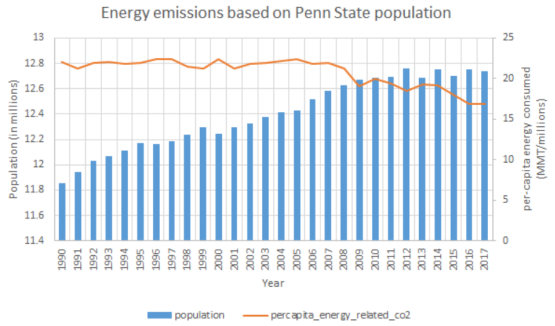

Fig. 7. population - percapita relation

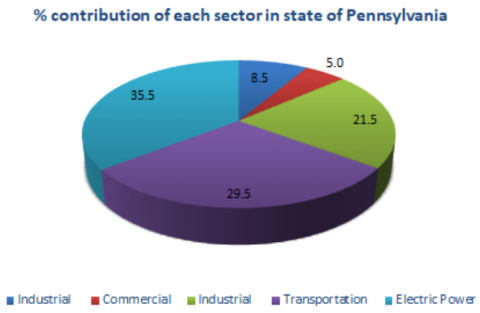

Fig. 8. Contribution from Pennsylvania

The final query match the last task in task section, and the details could be found in "intensity.sql". The result is shown as below: This is a stacked bar chart which compares the amount of carbon emitted per thousand BTU of energy

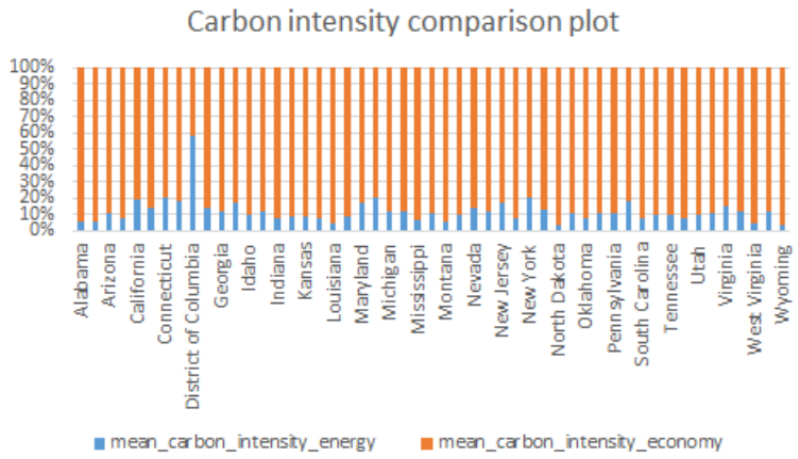

Fig. 9. Intensity energy versus economy

consumed and per USD of economy generated. The District of Columbia has the highest carbon intensity per energy whereas the state of Wyoming has the highest carbon intensity economy. In general most states have performed well over three decades to improve the carbon economy.

6 SCHEMA DESCRIPTION

The following part shows the table schema:

all_commercial(Year, Sname,CCoal, CPetroleum, CGas, CTotal)
all_electric(Year, Sname,ECoal, EPetroleum, EGas, ETotal)
all_industrial(Year, Sname,ICoal, IPetroleum, IGas, Itotal)
all_residential(Year, Sname,RCoal, RPetroleum, RGas, Rtotal)
all_transportation(Year, Sname,TCoal, TPetroleum, TGas, Ttotal)

carbon_emissions(Year, Sname, energy related co2, percapita energy related co2, energy intensity, carbon intensity energy supply, carbon intensity economy)