**“US Greenhouse Gas Emissions Relative to Shifts in Energy Sources – Data Analysis from 1990 to 2017”**

**Outline**

Comparison of US CO2 emissions vs percentage change in renewables relative to all energy types

**Data Sources**

**US Population Data 1960 to 2018 – World Bank**

<https://data.worldbank.org/indicator/SP.POP.TOTL?locations=US>

**CO2 Data sourced from OECD**

<https://www.oecd-ilibrary.org/energy/data/iea-co2-emissions-from-fuel-combustion-statistics/co2-emissions-by-product-and-flow-edition-2019_751f13bf-en>

<https://cfpub.epa.gov/ghgdata/inventoryexplorer/#iallsectors/allgas/inventsect/all>

**Definition of CO2 Emissions**

<https://www.epa.gov/ghgemissions/inventory-us-greenhouse-gas-emissions-and-sinks>

**US Energy Production By Type – Department of Energy**

<https://www.eia.gov/totalenergy/data/browser/?tbl=T01.02#/?f=M>

**History of Significant US Events**

<https://www.thebalance.com/national-debt-by-year-compared-to-gdp-and-major-events-3306287>

**Team Members**

Stacy Giauque

Brian Walker

1. \* \*\*E\*\*xtract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).
2. \* \*\*T\*\*ransform: what data cleaning or transformation was required.
3. \* \*\*L\*\*oad: the final database, tables/collections, and why this was chosen.

Please upload the report to Github and submit a link to Bootcampspot.

**Data Cleanup & Analysis**

Once you have identified your datasets, perform ETL on the data. Make sure to plan and document the following:

* The sources of data that you will extract from.
* The type of transformation needed for this data (cleaning, joining, filtering, aggregating, etc).
* The type of final production database to load the data into (relational or non-relational).
* The final tables or collections that will be used in the production database.

You will be required to submit a final technical report with the above information and steps required to reproduce your ETL process.

## Project Report

At the end of the week, your team will submit a Final Report that describes the following:

* **E**xtract: your original data sources and how the data was formatted (CSV, JSON, pgAdmin 4, etc).
* **T**ransform: what data cleaning or transformation was required.
* **L**oad: the final database, tables/collections, and why this was chosen.

Please upload the report to Github and submit a link to Bootcampspot.