

COM6018 Assignment 1

Exploring Global Carbon Emissions Data

Due: 15:00, Friday 3rd November 2023

Copyright © 2023 Jon Barker, University of Sheffield. All rights reserved.

26th October 2023. v1.1

Table of Contents

- 1. Introduction
- 2. The Data
- 3. The Tasks
- 4. Assessment
- 5. Marks
- 6. Submission

1. Introduction

The assignment provides you with a dataset that describes the carbon dioxide emissions from fossil fuels and cement production for countries around the world. It then asks you to use this data to answer a series of questions. For each question you will need to produce an illustrative figure and write a short paragraph to describe the figure and answer the question. You will need to think carefully about how to use the data and how to design the figure to make your answer clear.

Your assignment report will take the form of a Jupyter notebook. You are provided with a template notebook that contains the questions and some code to get you started. You will need to add code to process the data and produce the figures. You will also need to add text to describe the figures and answer the questions. You will then submit the completed notebook to Blackboard. You will be allowed to use any modules in the Python standard library, plus the following packages: `numpy`, `pandas`, `matplotlib`, `seaborn`.

Further details of the data, the questions, the assessment criteria and the submission process are given below.

2. The Data

The dataset for this assignment is the CO₂ and Greenhouse Gas Emissions dataset from Our World in Data. The data is available from the following GitHub repository, <https://github.com/owid/co2-data>. You will need to download the file `owid-co2-data.csv` from this repository. Store this file in the same directory where you store your notebook.

The dataset contains records that describe the CO2 emissions for a country (or region) in a given year. The dataset contains a very large number of fields (80 in total). There is another csv file in the repository, `owid-co2-codebook.csv`, that describes the fields in the dataset. You should read this file carefully to make sure you understand the meaning of the fields. Understanding these fields correctly will be important for answering the questions in this assignment. You can get a further understanding of the data by reading the accompanying article on the Our World in Data website, <https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions>.

3. The Tasks

The questions are presented in the template notebook but are also listed below for convenience. Note, for each question think carefully about the best type of plot to use, e.g., a line graph, a scatter plot, a pie chart, a box plot, a violin plot, etc.

Question 1

How have CO2 emissions grown over time?

The dataset contains entries for countries grouped into four income groups: low, lower-middle, upper-middle and high. i) How have the total annual CO2 emissions of each income group varied over the years 1950 to 2021? ii) How has the per capita CO2 emissions (i.e., the emissions per person) of these groups varied over the same period. Comment on how the contribution of each income group is different when measured in terms of total emissions and per capita emissions.

Question 2

Which countries have the highest CO2 emissions?

Compare the share of global carbon emissions for the top 5 emitting countries as a proportion of the total world emissions. Make separate plots for the years 1960, 1990 and 2020. Make a similar set of plots but normalised in a way that accounts for the population of each country. Note, all the plots should appear in the same figure.

Comment on how the top 5 emitting countries have changed over time and how the top 5 emitting countries change when you normalise for population.

Question 3

Do countries with high GDP always have high per capita CO2 emissions?

Plot the per capita CO2 emissions against the GDP per capita for each country. Design your plot so that the size of the marker is proportional to the population

of the country. Only consider countries with a population of at least 5 million people.

Compare these plots for the year 1978, 1998 and 2018. Comment on the relationship between the two variables and how it has changed over time.

Question 4

Has the distribution of wealth across countries become more or less equal over time?

Make a plot that compares the distribution of GDP per capita across the countries in the world at 10-yearly intervals from 1950 to 2020. Comment on how the distribution has changed over time. In particular, does it appear that wealth inequality has increased or decreased over time?

4. Assessment

The assignment is worth 40% of the module mark and will be marked out of 40.

For each of the 4 questions you will need to write some code to produce an appropriate figure. You will then need to write a short paragraph that describes the key points of the figure and answers the question. Each question is worth 10 marks.

The marks for each question will be awarded after considering the following criteria:

- **code:** is the code concise, clear and well documented? Does it make good use of the available libraries?
- **plots:** is the plot well-designed? Is it well labelled? Is it easy to interpret? Does it display the data in a way that addresses the question?
- **text:** does the text describe the plot clearly and correctly? does the text provide a good answer to the question? Are conclusions drawn correctly from the plot?

5. Marks

The question mark will reflect the overall quality of the answer after considering the three aspects above. Marks can be interpreted using a scale consistent with the degree's pass, merit and distinction grade boundaries.

- 0 - No attempt made
- 1-4 - Unsatisfactory, an attempt has been made but it is incomplete or has serious deficiencies in one or more of the criteria.
- 5 - Satisfactory, a minimum passing mark
- 6 - Good, work consistent with a Merit level grade.
- 7 - Very good, work consistent with a Distinction level grade.
- 8-9 - Excellent work, at least one criterion is met to the highest standard.

- 10 - Exceptional work, all criteria are met to the highest standard.

6. Submission

Please name your notebook using the following convention:

`COM6018-assignment1-<your-student-username>.ipynb`

For example, if your username is ‘ac1jpb’ then your notebook should be named `COM6018-assignment1-ac1jpb.ipynb`

You must submit your completed notebook to the assignment 1 submission area on Blackboard. Multiple submissions are allowed up to the deadline. Only the latest submission will be marked.

The assessors will be **running your notebook** to check the code runs correctly. The notebook will be run in a directory that contains a copy of the `owid-co2-data.csv` datafile. So, make sure that your data file is stored in the same directory as the notebook when you test it. If your notebook does not run correctly then you will be in danger of losing marks.

The assignment is due by 15:00 on Friday, 3rd November. Standard lateness penalties will be applied.

:warning: Academic Misconduct

The University takes academic misconduct very seriously. Academic misconduct includes plagiarism, collusion and fabrication of data. You should read the University’s policy on academic misconduct to ensure you understand what is and what is not acceptable. The policy is available from <https://www.sheffield.ac.uk/new-students/unfair-means>. In particular, this is an individual assignment and any suspected collusion will be investigated. Do not share your code with anyone else.

Copyright © 2023 Jon Barker, University of Sheffield. All rights reserved.