COMP 4304 / 6934 – Assignment 1 (1%)

MEMORIAL UNIVERSITY

Due: 11:59pm, Jan. 26, 2024

Learning Objectives

The goal of this assignment is to become acquainted with Jupyter notebooks, pandas and Matplotlib.

Data

Download the data sets, `plastic-production.csv` and `river-plastics.csv`, from Brightspace.

The first data set contains information on global plastic production over the past 30 years. The total production is given, along with production by various industrial sectors.

The second data set contains information about plastic waste on a per country basis. Provided is the total mismanaged plastic waste (MPW) given in metric tons per year, and the MPW that is emitted into the oceans (ME), also in metric tons per year. Geographic details of each country are provided, such as its area (in km^2), length of coast lines (km), and annual rainfalls (mm / year).

The data comes from the following articles about plastic pollution:

- Plastic Pollution, *Our World in Data*, https://ourworldindata.org/plastic-pollution
- More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean, *Science*, https://www.science.org/doi/10.1126/sciadv.aaz5803
- *OECD Stats (Organisation for Economic Co-operation and Development*, https://stats.oecd.org/viewhtml.aspx?datasetcode=PLASTIC_USE_10

Instructions

Using the provided data sets, create a Jupyter notebook to answer the following questions.

You may only use the pandas and Matplotlib libraries. That is, pandas and Matplotlib should be the only libraries imported into your notebook to complete this assignment.

Question 1: (20 pts)

Using pandas `groupby()`, calculate the sum of plastic production from all sectors on a yearly basis for the `plastic-production.csv` data set.

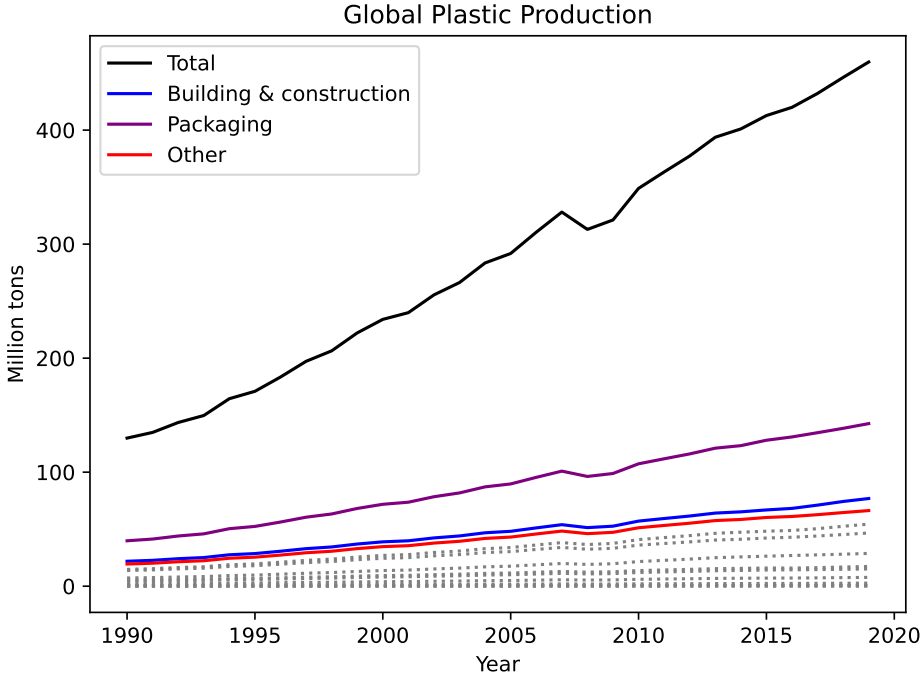
Confirm that the calculated sums equal the total production for all years provided in the data set (within an absolute precision of 1e-3).

Question 2: (40 pts)

Global plastic production has steadily increased and more than tripled in the last 30 years. The plot below shows this increase in total production (solid black line), along with the increasing plastic production by various sectors.

Using the `plastic-production.csv` data set, re-create the plot using Matplotlib's **functional interface**. All elements should be replicated as shown. In particular, ensure that,

- the legend is in the top left and that the ordering of the labels is correct,
- the colours of the lines are correct (black, purple, blue, red, grey),
- the labeled lines are solid lines and unlabelled lines are dotted lines, and
- the axes labelling and plot title are present.

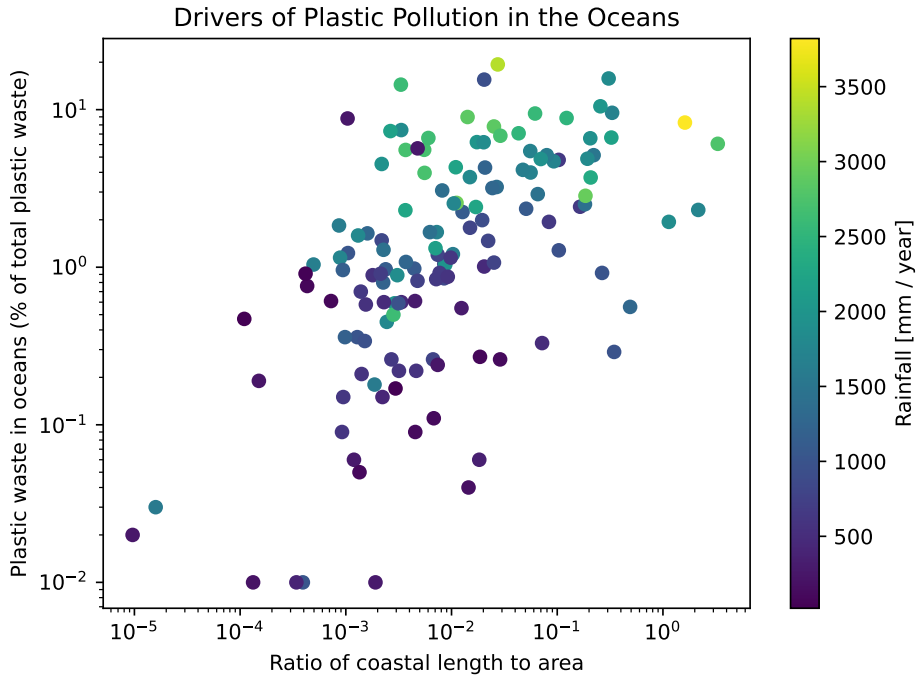


Question 3: (40 pts)

Approximately 1–2 million tons of plastic waste enter the oceans annually, and rivers bring a significant portion of this plastic waste to the ocean. The plot below shows how a country's coastal length, area and rainfall affect its percentage of plastic waste that is emitted to the oceans. Each point represents a country that has a coast line. Countries with a higher ratio of coastal length to their land area, and that receive higher annual rainfalls, tend to have a larger percentage of their mismanaged plastic waste end up in the oceans.

Using the `river-plastics.csv`, re-create the plot below using Matplotlib's **object-oriented interface** (explicitly created Figure and Axes objects). Ensure that all elements are replicated, such as,

- the x-axis is the ratio of coastal length to area,
- the y-axis is the percentage of plastic waste that is emitted into the oceans,
- points are coloured according to the country's annual rainfall,
- a colourbar is added to the figure with the default colour map ('viridis') and the correct label (hint: `fig.colorbar()`),
- the x- and y-axes use log scale, and
- the axes labels and plot title are present.



Submission

Submit your Jupyter notebook (`.ipynb`) through Brightspace.

Late submissions will be subject to a 10% penalty for each hour past the deadline.

Attribution

Submissions should include an attribution section indicating any sources of material, ideas or contribution of others to the submission.

Submissions must represent your independent work.

You are encouraged to use any resources to help with your solution, but your solution must represent independent work. If your submitted work includes unacknowledged collaboration, code materials, ideas or other elements that are not your original work, it may be considered plagiarism or some other form of cheating under MUN general regulations 6.12.4.2 (4.12.4.2 for graduate students) and academic penalties will be applied accordingly.

Avoid academic penalties by properly attributing any contribution to your submission by others, including internet sources and classmates. This will also help distinguish what elements of the submission are original. You may not receive full credit if your original elements are insufficient, but you can avoid penalties for plagiarism or copying if you acknowledge your sources.

Github

I encourage you to store and version your work on GitHub. It is good practice to do so as everyone uses git in the real world.

However, **it is a requirement that git repositories containing assignment material be private.** University regulations (undergraduate 6.12.4.2 and graduate 4.12.4.2) consider it cheating if you allow your work to be copied. There will be zero tolerance for this.