CSC110Y1-F Fall 2020 - Fundamentals of Computer Science 1
Course Project Proposal

Ching Chang Letian Cheng Arkaprava Choudhury Hanrui Fan

December 13, 2020

1. Part 1

An interactive report of the effect of Amazon rainforest's area on the local annual precipitation and CO_2 emission (with python prime shipping).

Ching Chang, Letian Cheng, Arkaprava Choudhury, Hanrui Fan

2. Part 2

Research Question: **To what extent do the changes in the Amazon rainforest's area affect the local annual precipitation and CO_2 emission?**

During our initial research on topics related to climate change, we found that the South American rainforest contributes 20% (Thomas, 2020) of the oxygen produced by photosynthesis on land, while the Amazon rainforest is responsible for 10% of the current greenhouse gas emissions (Melillo et al., y 20). This information was surprising to us since 20% of photosynthesis implies a lot of conversion from carbon dioxide, which is a greenhouse gas, to oxygen, whereas the 10% contribution to greenhouse gas seems to contradict that information. Due to this contradiction, we were curious about whether tree populations actually help control the climate. After some research, we learned that the effect of trees on climate change is more complex than we originally thought. There are many factors to consider, such as the carbon dioxide to oxygen conversion, the tendency to trap heat due to their dark color, reaction to form methane and ozone, and political movements revolving around tree plantations (Marshall, y 26). This led us into choosing an empirical approach —we wanted to directly observe the relationship between the change of tree population and climate change. We chose to focus our data on the Amazon rainforest not only because it is the largest rainforest on Earth (World Wildlife Fund., 2013), but also because there have been several pieces of evidence that show that the Amazon rainforest has been suffering from deforestation recently. For example, over $700,000km^2$ $(270,000mi^2)$ of Amazon rainforest had been lost since 1970, reducing its size to 80.7% of its original size, in 2018 (Butler, ry 4); There have been more than 40,000 fires in the rainforest in 2019 (Government of Brazil., er 4); and that forest exploitation in Amazon has risen for 14 consecutive months in June 2020 (Reuters, 2020). With these major pieces of evidence of deforestation correlating to the change in global and local climate, we believe that it is a relevant topic to contemporary society that should not be ignored.

3. Part 3

The data we show here is based on joint estimates by the Brazilian National Institute of Space Research and the United Nations Food and Agriculture Organization with map data provided by MapBiomas.
The data is presented as a text-based table, but we will convert it into a CSV file. The data shows the area of the Amazon rainforest in Brazil, which accounts for about 60% of the total area of the country.
For these data we did some pre-processing, we first removed items in the data that had NA (unrecorded or incomplete data), these included data up to 1985 and 2019, but we needed to record data from pre-1970 as this is the baseline data for the subsequent percentage change.
We then selected data from the start of 1988 because our team considered the absence of the change in deforestation loss data between 1978-1987 to be unrepresentative, so we filtered the 1978-1987 data as well.
At the same time, these data contain four columns including estimated remaining forest cover, average annual forest area loss, percentage of forest cover compared to 1970 area, and total forest area lost since 1970.

Sample:

Source: (Butler, ry 4)

Period	Estimated Natural Forest Cover	Deforestation (INPE)
2016	3,406,796	7,893
2017	3,399,308	6,947
2018	3,390,835	7,900

continue:

Period	Forest cover as % of pre-1970 cover	Total forest loss since 1970
2016	83.1%	693,204
2017	82.9%	700,692
2018	82.7%	709,165

4. Part 4

We first create a function that parses the `html` element of the stats on the website as a string, and converts it to a `nested array` so that it's easier to work with. This will involve using a `for loop`, `if statements`, and an `accumulator` keeping track of the data parsed so far. Using this function, we will collect our data for deforestation in the Amazon rain-forest over the past few decades. With the data now converted into a form that we can easily manipulate, we shall focus on analysing the data using our own functions.

For this project, we use smooth polynomial fitting to relate two of the variables in our `nested list`. Now, although there exist readily available functions that would do the same in the module `numpy` (SciPy community, e 29), we try implementing our own functions for the same, to test our learning from the course.

We split the mathematical algorithm for this problem using top-down design. Firstly, the main function would have two lists, `l_x, l_y`, of same length as input (for the two variables), along with an integer n (where $n \leq$ `len(l_x)`; representing degree of intended polynomial). The function body would have calls to helper functions. Note: this is only a rough outline and the exact technical details may be changed based on the results after testing the functions.

Firstly, we have a function to calculate the perpendicular distance of one point from a given polynomial. As opposed to the naive approach to the problem, we use the Newton-Raphson method *repeatedly* to estimate a solution for the derivative of the expression for the difference between the point and the polynomial, hence finding the coordinates of the foot of perpendicular, and consequently, the length of perpendicular.

The first estimate for the polynomial will be trivial, and we will then run the simplex algorithm to minimize the sum of the squares of the perpendiculars using two more helper functions, yielding the polynomial regression model. We now move towards plotting the resulting graph using the `matplotlib` library.

We also plan to write a function to calculate the coefficient of determination r_2 to check whether the graph shows an appropriate relationship between the independent variable (forest cover) and the dependent variable.

In addition to the graph, we plan to create an interactive text-based report of our data, where the user inputs a value for the independent or dependent variable, and the program will provide the corresponding dependent or independent value, coefficient of determination, or the slope of the tangent at the point, depending on which one the user asks for. The output will be text-based, and will require string concatenation, and if statements to check whether to add trivial information to the report.

The input/output model will use while loops and input prompt to keep the program interactive. We also extrapolate the data to yield predictions about future data using the interactive i/o model. Finally, we use the extrapolated data to summarize the upcoming significant years where the dependent variable will reach a certain milestone.

Technical requirement:

The Annual total CO_2 emissions dataset we have is a CSV file. In order to parse it easily, We decide to use pandas for CSV file reading.

Pandas is a python library written for data manipulation and analysis. (Pandas Development Team., 2020) In particular, it offers data structures and operations for manipulating numerical tables and time series.

Specific, we will use the function `read_csv`". This function has many parameters available for us, but we will only use `filepath_or_buffer` parameter. It takes a valid string path, which is the path to our CSV file.

This function will return a class - Pandas DataFrame that contains the data of the CSV file. Pandas DataFrame is easier for calculating dataset thanks to its various built-in functions. This way, we can access each cell of the CSV file for our later research.

References

Butler, R. (2020, January 4). Calculating deforestation figures for the amazon. *Mongabay.* https://rainforests.mongabay.com/amazon/deforestation_calculations.html.

Government of Brazil. (2020, November 4). Queimadas. *INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS.* http://queimadas.dgi.inpe.br/queimadas/portal-static/situacao-atual/.

Marshall, M. (2020, May 26). Planting trees doesn't always help with climate change. *BBC.* https://www.bbc.com/future/article/20200521-planting-trees-doesnt-always-help-with-climate-change.

Melillo, J., McGuire, A., Kicklighter, D., Moore III, B., and Vörösmarty, C. (1993, May 20). Global climate change and terrestrial net primary production. *Nature.* 363:234–240. https://www.bbc.com/future/article/20200521-planting-trees-doesnt-always-help-with-climate-change.

Pandas Development Team. (2020). pandas documentation. https://pandas.pydata.org/docs/.

Reuters, T. (2020). Brazil amazon deforestation up in june, set for worst year in over a decade. *CBC.* https://www.cbc.ca/news/world/amazon-deforestation-up-june-1.5644730.

SciPy community (2020, June 29). numpy.polyfit. https://numpy.org/doc/stable/reference/generated/numpy.polyfit.html.

Thomas, A. (2020). Biodiversity and the amazon rainforest. *Greenpeace.* https://www.greenpeace.org/usa/biodiversity-and-the-amazon-rainforest/.

World Wildlife Fund. (2013). Our world's largest rainforest: The amazon [youtube]. *Youtube.* https://www.youtube.com/watch?v=bYAZ3NWVgtc.