

Project Statement

The goal of this project is to try and predict the future area of the Amazonian rainforest, and to analyze and discover what factors have an impact on it. We will be focusing primarily on the deforestation information to predict the future land area. On top of it we will be considering other environmental and economic factors, such as rainfall and GDP, respectively.

Preliminary EDA

When looking for data, we will start by seeking out any form of common sense connection, such as with rainfall. Clearly this has an impact on the growth of vegetation, i.e. the rainforest. However, the deforestation of the rainforest is occurring because it's lucrative to do so. What is causing these changes, and will these factors be able to help predict the trends in the rainforest's coverage?

Datasets:

- http://terrabrasilis.dpi.inpe.br/download/dataset/legal-amz-prodes/vector/yearly_deforestation.zip - Terrabrasilis amazon deforestation dataset from 2008 to 2021
- <https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=BR> – Brazil's GDP dataset from 1960 to 2021
- [World Development Indicators | DataBank \(worldbank.org\)](https://data.worldbank.org/indicator/SL.LS.OV.XD?locations=BR) – Brazil's labor force from 1990 to 2021
- <https://zenodo.org/record/4428308#.ZA0-6XbMI2x> – Brazil's rainfall erosivity dataset.

The list of datasets may grow as time carries on, as we will continue to brainstorm other sets of data that could possibly have an impact on the rainforests deforestation.

Tools:

We will utilize ML and data exploration tools such as Sklearn, matplotlib and other libraries that are needed for reading the data files, pandas, numpy, just to name a few. We will be utilizing Jupyter Notebook for this project.

Research questions:

- What are the trends in these datasets mentioned and how would it be in the future?
- What correlations exists among these factors and how does it impact the Brazilian rainforest?